5th Meeting
Salt Lake City, 2018

Edited by Frédéric Noo

Sponsored by

GE Healthcare

SIEMENS Healthineers

Canon

Welcoming Address

Dear Colleague,

It is a pleasure to welcome you in Salt Lake City for the fifth edition of *The International Conference on Image Formation in X-ray Computed Tomography*, also known as "The CT meeting".

The idea for this meeting came in 2010, following a wish from many of us to have a venue where CT scientists could meet together to discuss in depth all aspects that impact the image formation process in CT. These aspects include dose evaluation and dose reduction strategies, non-linearity effects and compensation schemes for these effects, image reconstruction algorithms, spectral decomposition, dynamic effects, geometrical calibration, phase-contrast physics, deep learning, and image quality assessment.

Pleasantly, the meeting this year remains strong, as it attracted more than 100 submissions. The breadth of the topics being covered is amazing. I am grateful to all authors for submitting their work. Similarly to the previous meetings, oral presentations have been allocated significant time to allow for in-depth discussion between the attendees. Also, comfortable poster sessions have been planned, and attractive locations have been selected to promote discussion during the meals.

The meeting could not be a success without a great scientific committee. I would like to take the opportunity here to once again thank the following scientists, for their support, as well as for their help with the evaluation of submitted abstracts, which was critical in shaping the scientific program:

Cristian Badea	Duke University, NC, USA
Kirsten Boedeker	Canon Medical Research USA, Inc.
Yoram Bresler	University of Illinois at Urbana-Champaign, IL, USA
John Boone	UC Davis Medical Center, CA, USA
Guang-Hong Chen	University of Wisconsin-Madison, WI, USA
Dianna Cody	MD Anderson Cancer Center, TX, USA
Mini Das	University of Houston, TX, USA
Michel Defrise	Vrije Universiteit of Brussels, Belgium
Bruno De Man	GE Global Research Center, NY, USA
Jeff Fessler	University of Michigan, MI, USA
Thomas Flohr	Siemens Healthcare GmbH, Germany
Taly Gilat-Schmidt	Marquette University, WI, USA
Stephen Glick	Office of Science and Engineering Labs, CDRH, FDA, MD, USA
Michael Grass	Philips Research Laboratories, Hamburg, Germany
Scott Hsieh	David Geffen School of Medicine at UCLA, CA, USA
Xun Jia	UT Southwestern Medical Center, TX,USA
Marc Kachelriess	German Cancer Research Center (DKFZ), Heidelberg, Germany
Andrew Karellas	University of Arizona, AZ, USA

Thomas Koehler	Philips Research Laboratories, Hamburg, Germany
Hiroyuki Kudo	University of Tsukuba, Japan
Patrick La Riviere	University of Chicago, IL, USA
Tobias Lasser	Technical University Munich, Germany
Guenter Lauritsch	Siemens Healthcare GmbH, Germany
Ke Li	University of Wisconsin-Madison, WI, USA
Hongbin Lu	Fourth Military Medical University, Xi□an, china
Andreas Maier	University of Erlangen-Nuremberg, Germany
Cynthia McCollough	Mayo Clinic, MN, USA
Michael McNitt-Gray	David Geffen School of Medicine at UCLA, CA, USA
Klaus Mueller	Stony Brook University, NY,USA
Johan Nuyts	Katholieke Universiteit, Belgium
Jed Pack	GE Global Research Center, NY, USA
Xiaochuan Pan	University of Chicago, IL, USA
Norbert Pelc	Stanford University, Stanford, CA, USA
Cyril Riddell	GE Healthcare, Buc, France
Ken Sauer	University of Notre Dame, IN, USA
Emil Sidky	University of Chicago, IL, USA
Jeff Siewerdsen	Johns Hopkins University, MD, USA
Joseph W Stayman	Johns Hopkins University, MD, USA
Karl Stierstorfer	Siemens Healthcare GmbH, Germany
Katsuyuki Taguchi	Johns Hopkins University, MD, USA
Xiangyang Tang	Emory University, GA, USA
Srinivasan Vedantham	University of Arizona, AZ, USA
Wenli Wang	Avant Tomography Consulting, LLC, USA
Adam Wang	Varian Medical Systems, CA, USA
Bruce Whiting	University of Pittsburgh, PA, USA
Yuxian Xing	Tsinghua University, Beijing, China
Jingyan Xu	Johns Hopkins University, MD, USA
Jong Chul Ye	Korea Advanced Inst. of Science and Technology (KAIST), South Korea
Lifeng Yu	Mayo Clinic, MN, USA
Zhou Yu	Canon Medical Research USA, Inc., IL
Stanislav Zabic	United Imaging, OH, USA

Once again, we were also fortunate to receive generous support from

- ✓ Siemens Medical Solutions, USA
- ✓ Canon Medical Systems USA
- ✓ GE Healthcare

This financial support is essential to accommodate a lower registration fee for graduate students. This year, we have more than 150 attendees out of which a third are students.

Finally, I am thankful to my research team including Zijia Guo, Viktor Haase, Mehmet “Bugra” Oktay for providing me with a strong hand; and I am grateful for the help from my colleague, Larry Zeng, without whom I could not run such a conference.

I wish you all a pleasant meeting.

Frederic Noo, Ph.D.
General Chair
Utah Center for Advanced Imaging Research (UCAIR)
Department of Radiology and Imaging Sciences
University of Utah
E-mail: Frederic.Noo@hsc.utah.edu

Table of contents

Session M1: Model-based iterative reconstruction I
Moderated by Fessler J (University of Michigan) & La Riviere P (University of Chicago)

Session M2: The four best scored abstracts
Moderated by Kachelrieß M (German Cancer Research Center) & Taguchi K (Johns Hopkins University)

Session M3: Novel data acquisition concepts
Moderated by Chen G-H (University of Wisconsin-Madison) & McCollough C (Mayo Clinic)

CBCT

Dynamic CT

Data truncation and consistency

Industrial CT

Session T5: Limited data CT
Moderated by Sidky E (University of Chicago) & Xu J (Johns Hopkins University)

Session W1: Model-based iterative reconstruction II
Moderated by Zabic S (United Imaging) & Boursier Y (Aix Marseille University)

Session W2: Photon counting CT
Moderated by Flohr T (Siemens Healthcare, GmbH) & Brown K (Philips Healthcare)

Session W3: Posters -- Analytical methods, MBIR, spectral CT, deep learning
Moderated by Riddell C (GE Healthcare) & Rit S (Lyon INSA, France)

Analytical methods

MBIR

Spectral CT

Deep learning

pDART: Discrete algebraic reconstruction using a polychromatic forward model

Nathanaël Six, Jan De Beenhouwer and Jan Sijbers

***Abstract*—The discrete algebraic reconstruction technique (DART) is a tomographic method to reconstruct images from X-ray projections in which prior knowledge of the object materials' attenuation values is exploited. In monochromatic X-ray CT (e.g., synchrotron), DART has been shown to lead to high quality reconstructions, even with a low number of projections. However, most X-ray sources are polychromatic, leading to beam hardening effects, which significantly degrade the performance of DART. To reduce beam hardening artefacts, we developed an algorithm, pDART, that exploits sparsity in the attenuation values using DART and simultaneously accounts for the polychromaticity of the X-ray source. The results show that pDART leads to a vastly improved segmentation on simulated polychromatic data.**

***Index Terms*—Computed tomography, discrete tomography, DART, polychromaticity, beam hardening**

I. Introduction

COMPUTED tomography is a widely used technique for non-invasive imaging of a sample. Classical reconstruction techniques, e.g. Filtered Backprojection, assume a linear acquisition model and a large number of projections taken from a full angular view. When these requirements are not met, artefacts arise in the reconstruction.

One of the main sources of reconstruction artefacts is undersampling. When only few projections are available, the system of equations that needs to be solved for the reconstruction becomes severely underdetermined. By including prior knowledge about the sample, one can reduce this underdetermination, effectively decreasing the space of possible solutions. One of the ways this can be achieved is Discrete Tomography (DT), in which the grey levels in the image corresponding to the attenuation values of the object materials are assumed to be known a priori. This reduces the reconstruction from a continuous problem to a discrete problem. The *Discrete Algebraic Reconstruction Technique* or DART algorithm is one of the possible DT algorithms and has proven to be effective if the sample consists of few materials [1].

The DART algorithm relies on the same linearized acquisition model as most classical techniques; it assumes that the log-corrected normalized projection is the sum of attenuation values along a ray. This is an adequate model for a monoenergetic source, but not for a polyenergetic one. Due to this model mismatch, the reconstructed image will show beam hardening artefacts as the effective energy of the spectrum shifts upwards [2]. Beam hardening artefacts and the inability to clearly select a correct grey level for each material, lead to inaccurate reconstructions when DART is used in combination with polychromatic projections. One of the current methods for reducing beam hardening artefacts consists of placing metal filters in front of the source, to pre-harden the beam. However, the pre-hardened spectrum is still polychromatic and the filtering decreases the number of photons available for imaging.

We propose to combine the DART framework, with the *polychromatic Simultaneous Algebraic Reconstruction Technique* (pSART) introduced in [3], as the underlying reconstruction method. The pSART algorithm has previously successfully been combined with a Total Variation approach in [4]. This technique reconstructs a single grey level image, corresponding to a simulated monochromatic reconstruction, while using a polychromatic forward model.

In this paper, a short overview of DART and pSART is given, to clarify the models and algorithms that were used. Next, we explain the combined pDART algorithm and where it differs from the normal DART scheme. Finally, different discrete reconstructions from simulated polychromatic data are compared to demonstrate the improvements of the proposed method.

II. Method

A. Discrete reconstruction: DART

The DART algorithm is an efficient reconstruction method for discrete tomography. A complete discussion of DART can be found in [1]. We briefly recall the different steps in the DART algorithm:

1) **Initial reconstruction**: An algebraic reconstruction method (ARM) is performed to obtain an initial reconstruction $\boldsymbol{V}_0$.
2) **Segmentation**: Segment $\boldsymbol{V}_0$ based on the a priori knowledge of the grey levels and corresponding threshold values.
3) **Masking**: Choose a set of fixed and free pixels. The free pixels encompass all boundary pixels [1] and a small percentage of randomly chosen pixels.
4) **Reconstruction**: Perform a number of iterations of the ARM, on the set of free pixels only. The fixed

N. Six, J. De Beenhouwer and J. Sijbers are with Vision Lab, an imec research group at the University of Antwerp, Belgium (email: nathanael.six@uantwerpen.be)

[1] A pixel is called a *boundary pixel* if it has a neighbouring pixel with a different grey value.

pixels are kept at their segmented value. This gives us a reconstruction $\boldsymbol{V}$.

5) **Smoothing**: Apply smoothing to weaken the harsh fluctuations that can occur due to the segmentation.
6) **Iteration**: Set $\boldsymbol{V}_0 = \boldsymbol{V}$ and repeat steps 2-5 until a stopping criterion is met.

B. Polychromatic SART (pSART)

The pSART algorithm [3] is based on the same iterative scheme as SART [5]. In matrix form, the k^{th} update step can be expressed as:

$$\boldsymbol{V}^{(k+1)} = \boldsymbol{V}^{(k)} - \boldsymbol{C}\boldsymbol{A}^{\top}\boldsymbol{R}\left(\text{Proj}(\boldsymbol{V}^{(k)}) - \boldsymbol{s}\right), \qquad (1)$$

with $\boldsymbol{A}$ the system matrix of the acquisition geometry, $\boldsymbol{C}$ and $\boldsymbol{R}$ diagonal matrices with the inverse sums of the columns and rows of $\boldsymbol{A}$, respectively, Proj the forward projection operator, and $\boldsymbol{s}$ the sinogram. In the case of SART, $\text{Proj}(\boldsymbol{V}) = \boldsymbol{A}\boldsymbol{V}$. In pSART, however, a polychromatic forward projection is used instead.

The polychromatic projection P_{L_r} along the ray L_r[2] is modelled as:

$$P_{L_r} = \int_{\epsilon_{min}}^{\epsilon_{max}} I(\epsilon) \exp\left(-\int_{L_r} \mu(\boldsymbol{x}, \epsilon)\, \mathrm{d}\boldsymbol{x}\right) \mathrm{d}\epsilon, \qquad (2)$$

with μ the function giving the attenuation coefficient of the object at a point $\boldsymbol{x}$ and an energy level ϵ and I the X-ray source spectrum [2]. This model can be discretized over N_e total energy levels and N_m different materials to achieve the following form:

$$\hat{P}_{L_r} = \sum_{\epsilon=1}^{N_e} w(\epsilon) \exp\left(-\sum_{m=1}^{N_m} l_{r,m}\mu_m(\epsilon)\right). \qquad (3)$$

Here $w(\epsilon)$ is the weight for each energy level, $l_{r,m}$ the distance that ray L_r travels through material m, and $\mu_m(\epsilon)$ the attenuation coefficient of material m at energy ϵ. By log-correcting and normalizing the projection $\hat{P}_{L_r}$ from Eq. (3) we have:

$$\text{Proj}(\boldsymbol{V})_r = -\log\left(\frac{\hat{P}_{L_r}}{\sum_{\epsilon=1}^{N_e} w(\epsilon)}\right). \qquad (4)$$

The extra prior information needed for the pSART algorithm involves the spectrum: N_e and $w(\cdot)$ and the different materials in the sample: $\mu_m(\cdot)$ and N_m. Note that this knowledge about materials is also a prerequisite of the DART algorithm.

For the forward projection of $\boldsymbol{V}^{(k)}$, the line-lengths $l_{r,m}^{(k)}$ need to be calculated as follows. First, a reference energy ϵ_{ref} is chosen, which provides reference attenuation values $\mu_m = \mu_m(\epsilon_{ref})$. In each step, the reconstruction represents the monoenergetic attenuation map at energy level ϵ_{ref} [3]. Next, for each pixel v in the current reconstruction $\boldsymbol{V}^{(k)}$, with value $t_v \in [\mu_i, \mu_{i+1}]$, the fractions in the following decomposition are calculated:

$$t_v = \frac{\mu_{m+1} - t_v}{\mu_{m+1} - \mu_m}\mu_m + \frac{t_v - \mu_m}{\mu_{m+1} - \mu_m}\mu_{m+1}. \qquad (5)$$

[2] L_r is the ray from the source to the rth pixel on the detector.

These percentages are grouped, per material, in a mask $\boldsymbol{M}_m^{(k)}$. This essentially models each pixel v as the mixture of two materials. The $l_{r,m}^{(k)}$ are now found as the values in the product $\boldsymbol{A}\boldsymbol{M}_m^{(k)}$.

C. Polychromatic DART (pDART)

We propose a new algorithm based on the principles of DART and pSART that combines the benefits of both methods. As pSART returns a single grey value image, it is suited for combination with DART. However, due to the non-linearity of the polychromatic model and the assumptions made in the pSART model, changes to the masking step of DART and the material selection of pSART are needed.

Choosing pSART as the underlying reconstruction method implies that a weight vector and a matrix containing the energy dependent attenuation values and reference values have to be specified. This prior knowledge allows us to make a natural choice for the grey levels: the pSART algorithm already uses the reference values to decide which materials each pixel consists of and to reconstruct a single image which represents a monochromatic reconstruction at the reference energy level [3]. For this reason, the grey levels for segmentation are chosen to be the same as the reference attenuation values.

In the original DART approach, the forward projection of the fixed voxels is subtracted from the sinogram [1]. However,

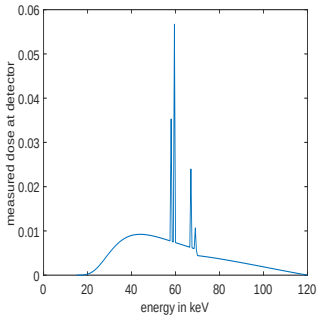
$$\boldsymbol{s} - \text{Proj}(\boldsymbol{V}_{fixed}) \qquad (6)$$

as the new sinogram data to the pSART algorithm will not be equal to a reconstruction of the reduced system of equations. As a result, running the DART algorithm with the conventional settings will give rise to divergent behaviour. Because of the fundamental non-linearity of the polychromatic problem, the reconstruction on the reduced set of pixels cannot be performed in an analogue way to the monochromatic version. Instead, the line lengths through each material are calculated once in every DART iteration. The material locations are known, since the image is segmented. These same fixed numbers will be added to the calculated line lengths in each iteration of pSART. This way the pixel values are effectively fixed in the pSART iterations. However, the fixed values can no longer be grouped together in an updated sinogram.

To pass more information about the segmented interfaces to the pSART algorithm, a different interpolation is performed for boundary pixels. Instead of interpolating between the two closest reference values, interpolation is performed between the maximal and minimal attenuating material found within the neighbourhood of the boundary pixel.

For the implementation of the pDART algorithm, we implemented pSART in Matlab and used the DART framework from the ASTRA toolbox [6].

(a) Phantom image with three materials.



(b) Polychromatic spectrum of a 120keV source.

Figure 1: Phantom (a) and employed spectrum (b).

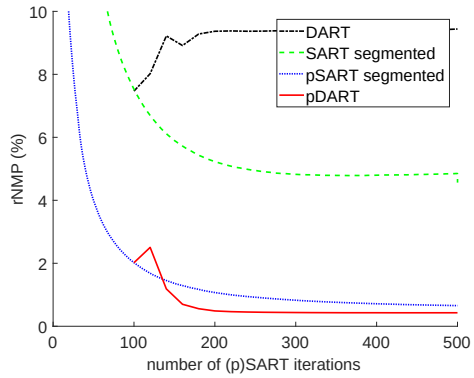


Figure 2: Plot of rNMP vs. (p)SART-iterations on 150 angles.

III. Experiments

To validate the proposed algorithm, simulation experiments were set up with an image quality phantom shown in Fig.1a. This phantom consists of three simulated materials on a vacuum background: water, cortical bone and titanium. The attenuation values at different energy levels were used from [7]. Different parallel beam polychromatic sinograms of the phantom were created, based on the weights and energy levels of a realistic spectrum, shown in Fig.1b. Sinograms were created with a variable number of angles, ranging from 2 to 500, equally spaced in the interval $[0, \pi)$. The forward model defined in Eq. (4) was used for the polychromatic projections. To limit inverse crime, the phantom was defined on a higher resolution grid than the reconstruction. To study the performance under noisy conditions, Poisson noise was added to the projections with 50 angles, based on different source intensities I_0, ranging from 5000 to 100000 photons per detector pixel.

From the simulated polychromatic sinograms, images were reconstructed with the following methods: FBP segmented with Otsu's method, pSART segmented, SART segmented, DART with manually optimized grey levels[3] and pDART. The segmentations of SART and pSART were performed with the same global threshold as DART and pDART, respectively.

All pDART reconstructions were performed with 1% random free pixels, 100 initial and 20 inner iterations of pSART. The same settings were used for DART. The percentage of free pixels was chosen as advised in [1]. When testing the influence of the variable number of angles, the error was measured after 300 (p)SART iterations, which corresponds to 10 (p)DART iterations.

For the noisy reconstructions we performed less (p)SART iterations per iteration of (p)DART: only 50 initial and 4 inner iterations, to prevent overfitting to noise. In this case, the error was measured after 210 (p)SART iterations. Due to the lower amount of initial and inner iterations, this corresponds to 40 (p)DART iterations. For the polychromatic reconstruction methods, the spectrum shown in Fig. 1b was rebinned to 18 bins. The reference energy level was set at 120keV, as higher energy levels lead to faster convergence [3].

All parameters (e.g. percentage of free pixels, number of inner iterations and reference energy level) were either chosen empirically or in accordance with the literature.

All reconstructions and projections were generated using the ASTRA toolbox [6].

[3]Out of the different available attenuation values, at the different energy levels, the one with the lowest reconstruction error was chosen.

IV. Results

To quantify the performance of pDART against segmented FBP, DART, segmented SART and segmented pSART, we computed the *relative number of mismatched pixels* or *rNMP* [1]. The effects of undersampling and beam hardening were very strong for FBP, so we have omitted FBP from all plots to improve visibility.

In Fig.2, the rNMP of the different methods is shown in function of the number of (p)SART iterations, for 150 projection angles between 0 and π. The rNMP for (p)SART was calculated at every iteration, whereas the rNMP for (p)DART was calculated at every DART iteration, i.e. every 20 (p)SART iterations. From Fig.2, one can observe that pDART converges more quickly than pSART. The original DART algorithm is outperformed by the other methods, except for FBP.

Next, the effect of varying amounts of projection angles on the rNMP was studied. These results are shown in Fig.3. Again, pDART outperforms the other methods in terms of mismatched pixels, reaching a lower stable rNMP and reaching this convergence point at a lower number of projections. This suggests that pDART benefits from both the beam hardening correction property of pSART and the imposed discreteness.

Lastly, the robustness to noise of the new method was studied. In Fig.4, the rNMP is plotted as a function of the beam intensity (represented by the photon count per detector element in the absence of attenuation). These simulations were performed for 50 projection angles.

Even in the presence of noise, pDART performs the best out of the different studied methods. However, comparing Fig. 4 with Fig. 3 shows that the influence of noise on our proposed method is still substantial.

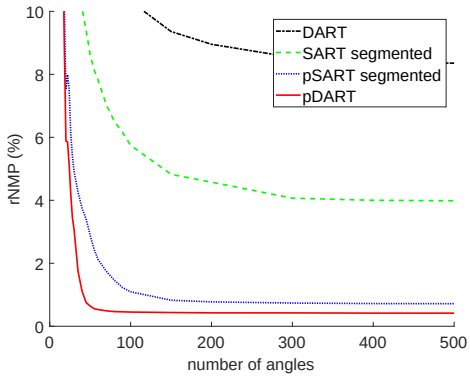


Figure 3: Plot of rNMP vs. number of angles, equally spaced in $[0, \pi)$.

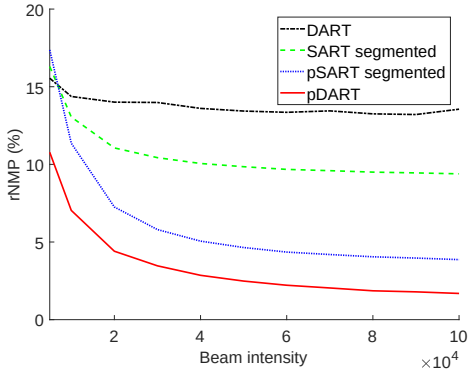
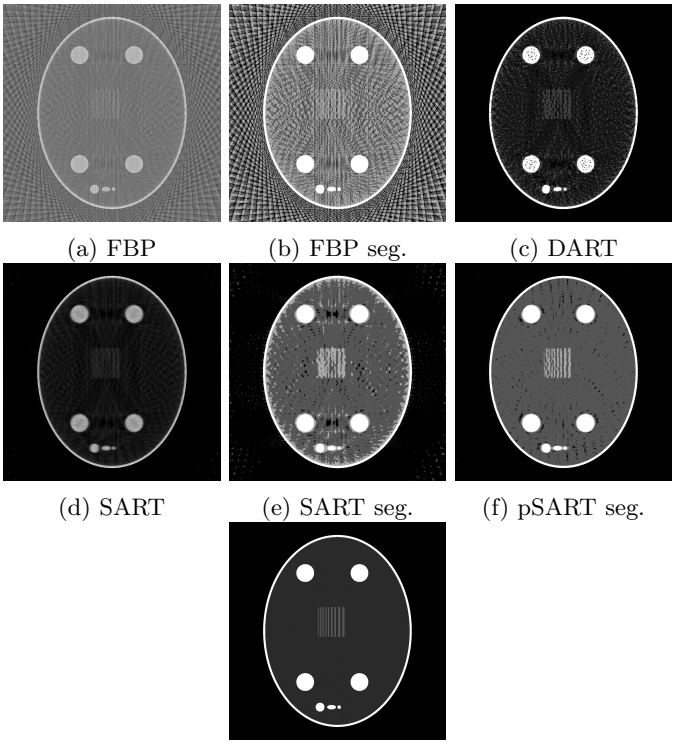


Figure 4: Plot of rNMP vs. different Poisson noise levels at 50 projection angles.

Finally, a comparison of the different reconstruction methods in the case of 50 noiseless projections can be found in Fig. 5. From this figure it can be observed that the two polychromatic methods show the least reconstruction artefacts, with the other methods showing beam hardening artefacts and streaks due to the low amount of projection angles. Both pSART and pDART have some mismatched black pixels in the interior, but the effect is less pronounced in the pDART reconstruction. Furthermore, it can be observed that pDART is the only algorithm, in our comparison, that accurately reconstructs the bar phantom in the center.

V. Conclusion

Many objects consist of a limited number of materials. This prior knowledge can be exploited in the reconstruction of images from X-ray projection data using discrete tomography. Current discrete tomography methods (such as DART), however, do not account for polychromaticity of X-ray sources, leading to various reconstruction artefacts. In this paper, pDART was proposed, a discrete tomography method that exploits sparseness in attenuation values, while taking a polychromatic projection model into account. Simulation experiments revealed that pDART results in substantially improved image reconstruction quality compared to DART or segmented versions of FBP, SART or pSART for polychromatic X-ray data.



(a) FBP (b) FBP seg. (c) DART

(d) SART (e) SART seg. (f) pSART seg.

(g) pDART

Figure 5: Comparison of the different reconstruction techniques with 50 projection angles.

Acknowledgements

This research is funded by the FFG Bridge Early Stage project no. 851249: ADAM as well as the Fund for Scientific Research-Flanders (FWO, grant nr S004217N: MetroFlex). The authors would like to thank Arjan den Dekker for his insightful comments.

References

[1] K. J. Batenburg and J. Sijbers, "DART: a practical reconstruction algorithm for discrete tomography," *IEEE Transactions on Image Processing*, vol. 20, no. 9, pp. 2542–2553, 2011.

[2] T. M. Buzug, *Computed tomography: from photon statistics to modern cone-beam CT*. Springer Science & Business Media, 2008.

[3] Y. Lin and E. Samei, "An efficient polyenergetic SART (pSART) reconstruction algorithm for quantitative myocardial CT perfusion," *Medical Physics*, vol. 41, no. 2, 2014.

[4] T. Humphries, J. Winn, and A. Faridani, "Superiorized algorithm for reconstruction of CT images from sparse-view and limited-angle polyenergetic data," *Physics in Medicine & Biology*, vol. 62, no. 16, pp. 6762–6783, 2017.

[5] A. C. Kak and M. Slaney, *Principles of computerized tomographic imaging*. SIAM, 2001.

[6] W. van Aarle, W. J. Palenstijn, J. Cant, E. Janssens, F. Bleichrodt, A. Dabravolski, J. De Beenhouwer, K. J. Batenburg, and J. Sijbers, "Fast and flexible X-ray tomography using the ASTRA toolbox," *Optics Express*, vol. 24, no. 22, pp. 25 129–25 147, 2016.

[7] J. H. Hubbell and S. M. Seltzer, "Tables of X-ray mass attenuation coefficients," https://www.nist.gov/pml/x-ray-mass-attenuation-coefficients, National Inst. of Standards and Technology-PL, Gaithersburg, Tech. Rep., 1995.

Joint image reconstruction for multiphase CT

Jingyan Xu and Frédéric Noo

***Abstract*—Joint image reconstruction for multiphase CT aims to jointly reconstruct a sequence of CT data sets acquired following contrast injection. Multiphase CT scans are not acquired simultaneously. There exists small organ shifts and organ boundary misalignment among different phases due to inter-scan patient breathing. Existing multi-channel image reconstruction algorithms with joint total variation (TV) type regularizers introduce artifacts at misaligned organ boundaries. We propose a multi-channel regularizer that is robust to organ misalignment. The regularizer is based on the infimal convolution between a joint TV and a separable TV. It can work like a joint TV, a separable TV, and a continuum between them depending on the local image properties, thus incorporating the shared information if it exists. The proposed regularizer is used in a constrained image denoising problem. The optimization problem is solved using the linearized alternating direction method of multipliers. Our numerical studies demonstrate that denoising using the infimal convolution regularizer can improve upon joint TV when misalignments are present, while maintaining the advantages of joint TV when misalignments are absent.**

I. Introduction

Multiphase CT scans are common in abdominal CT. Following contrast injection, multiple (2-4) acquisitions are performed at different time points to capture the variation of contrast enhancement in the organs of interest. Instead of reconstructing each phase independently as is now done in the clinics, joint image reconstruction for multiphase CT aims to reconstruct the multiple data sets together, and use the shared information among the multiple acquisitions to improve image quality and reduce radiation dose.

Unlike other multi-acquisition schemes such as multi-energy photon counting CT, the multiphase images are not acquired simultaneously. The sequence of scans is often finished within minutes while the patient remains on the table. During each acquisition, the patient is instructed to hold his/her breath to minimize motion; between acquisitions, the patient is instructed to breath. Thus the motion in each CT image is negligible, but motion between scans causes small organ shifts or misalignment. An example set of a patient 2-phase (arterial and venous) CT images is shown in Fig. 1. The organ shift is obvious around the spleen, stomach, and liver boundaries.

Joint multi-channel image reconstruction has been proposed for multi-energy CT, simultaneous PET-MR, longitudinal PET, multi-parameter MR, and color image processing. Many of these methods rely on the assumption that there exist common boundaries among the different images and design regularizers that leverage such information. As our preliminary data will show, the existing multi-channel regularizers can introduce artifacts to multiphase CT, as the organ boundaries are not perfectly aligned. This motivates us to improve the robustness of joint image reconstruction as applied to multiphase CT.

J Xu is with Deparment of Radiology, Johns Hopkins University. F Noo is with Department of Radiology and Imaging Sciences, the University of Utah. This project was supported by the National Cancer Institute of the National Institutes of Health under R21CA211035.

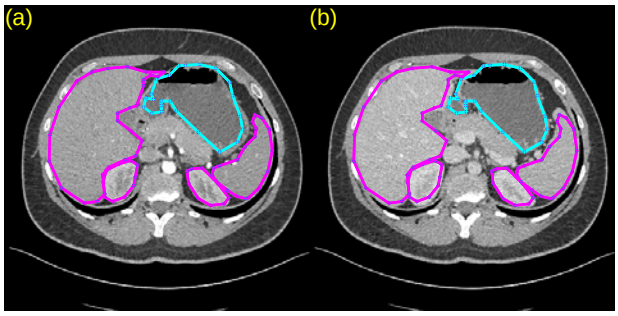


Fig. 1: The images at the arterial phase (a) and the venous phase (b) were acquired $\sim$ 38 sec apart. The organ boundaries at the arterial phase (magenta) and the venous phase (cyan) are overlaid onto each other. The misalignments are observed near the stomach, spleen, and liver with kidney interface. (C, W) = (10, 400) HU.

II. Preliminaries

For notational convenience, we mostly use joint two-phase and 2-D reconstruction as our running example. The same idea can be extended to many-phases and 3-D reconstruction in a straightforward manner.

A. Framework of joint multi-channel image reconstruction

The joint image reconstruction problem can be formulated as the following constrained minimization problem.

$$\min_{x_1,x_2} \quad J(x_1,x_2) \tag{1a}$$

$$\text{subject to} \left\| \begin{array}{c} A_1x_1 - y_1 \\ A_2x_2 - y_2 \end{array} \right\| \leq \epsilon \left\| \begin{array}{c} y_1 \\ y_2 \end{array} \right\| \tag{1b}$$

where $A_{1,2}$ is the forward projection operator for the two scans; $x_{1,2}$ are the reconstructed images at phase 1 and 2, *e.g.*, the arterial and the venous phase; $y_{1,2}$ are the projection data, ϵ is a percent error tolerance on data fitting. The objective function to be minimized is a joint regularizer for x_1 and x_2. We use the joint total variation (TV) type regularizer as an example.

$$J(x_1,x_2) = \sum_i \|T_i(x_1,x_2)\|_U \tag{2}$$

where T_i is the joint image gradient tensor at voxel i, *i.e.*

$$T_i(x_1,x_2) = \left[\begin{array}{cc} [Dx_1]_i & [Dx_2]_i \end{array} \right]$$

and $\|T\|_U$ denotes a unitarily invariant matrix norm of T, and $Dx_{1,2}$ denotes the image gradient, $[Dx_{1,2}]_i$ the gradient

evaluated at voxel i, a column vector composed of the neighboring voxel differences at i of image 1 (or 2) along each dimension. In (2), we consider both the Frobenius norm and the Nuclear norm in this work, but other matrix norms have also been proposed for multi-channel reconstruction, see, *e.g.*, [1], [2]. The joint TV regularizers encourage joint sparsity in the multiple images.

B. Infimal convolution

The infimal convolution (inf-conv) of two closed, proper, convex functions f_1, f_2: $\mathrm{R}^n \to \mathrm{R} \cup \{+\infty\}$, is defined as

$$f(s) = \inf_t f_1(t) + f_2(s-t) = \inf_{t+t'=s} f_1(t) + f_2(t') \quad (3)$$

If f_1 is coercive[1] and f_2 is bounded below, then from [3, Proposition 12.11, 12.14], $f(s)$ is also closed, proper, convex. Further, there exists t for any s, such that $f(s) = f_1(t) + f_2(s-t)$, *i.e.*, the infimal convolution is exact. In this case, we may also write $f(s) = f_1(s) \boxdot f_2(s)$.

Infimal convolution is a fundamental convexity preserving operation for functions [3, page 167]. For imaging applications, infimal convolution has been used to combine regularizers to encourage different image properties [4]. For example, the total generalized variation (TGV) is the infimal convolution of the first order and the higher-order differences. The second order TGV has been found to reduce the stair-casing artifacts that is well known to TV. The infimal convolution has also been used to combine spatial regularization and temporal regularization for dynamic imaging [5], [6].

As a more familiar example, the Huber function, defined by

$$h(s) = \begin{cases} s^2/(2\delta) & |s| \le \delta \\ |s| - \delta/2 & |s| > \delta \end{cases} \quad (4)$$

is the infimal convolution of $|s|$ and $s^2/(2\delta)$, *i.e.*,

$$h(s) = |s| \boxdot s^2/(2\delta) = \inf_t \{|t| + \frac{(s-t)^2}{2\delta}\}. \quad (5)$$

In imaging applications, s usually represents intensity differences. The parameter δ can be regarded as a combination parameter between the two terms. For a fixed δ, if the image difference s is large, the Huber function is similar to $|s|$; when s is small, $h(s)$ behaves like a quadratic function. Thus the parameter δ determines which term in the infimal convolution plays a more dominant role at each voxel.

III. Method

The organ shifts in multiphase CT create artifacts if we directly apply the joint TV regularizers. An ideal regularizer should be robust to organ shifts: if organ shifts exist, apply separable (channel-by-channel) regularization, and if they do not, apply the joint regularization. In other words, a desirable property for a robust regularizer is that it should adapt to the local image content, instead of applying the joint or the individual regularization invariably.

In this work, such local adaptation is accomplished by the infimal convolution of a joint multi-channel penalty and a separable channel-by-channel penalty. A combination parameter between the two terms, similar to δ in (4), (5), determines which term, the joint or the separable penalty, plays a major role at each voxel.

Denote by $X = [\cdots x_j \cdots]$, $j = 1, \cdots, p$, x_j is the vectorized image at phase j, *i.e.*, X is the Casorati matrix of the multiphase images, the proposed regularizer can be written as

$$R(X;\beta) = \inf_{Z+W=DX} J(Z) + \beta S(W) \quad (6)$$

where $J(\cdot)$ denotes a joint penalty term, and $S(\cdot)$ denotes a separable penalty term, $\beta > 0$ is a parameter that determines the relative contribution of the two. For a 2-phase problem, (6) is simplified to

$$R(x_1, x_2; \beta) = \inf_{\binom{z_1+w_1=Dx_1}{z_2+w_2=Dx_2}} \|z_1, z_2\|_U + \beta(\|w_1\| + \|w_2\|) \quad (7)$$

Letting $d_{1,i} = [Dx_1]_i$, $d_{2,i} = [Dx_2]_i$, $d_1 = \{[Dx_1]_i\}$, $d_2 = \{[Dx_2]_i\}$, the regularizer in (7) can be rewritten as

$$R(x_1, x_2; \beta) = \min_{z_1, z_2} \sum_i \{\|z_{1,i}, z_{2,i}\|_U + \beta(\|d_{1,i} - z_{1,i}\| + \|d_{2,i} - z_{2,i}\|)\} \quad (8)$$

the unknowns are $z_1 = \{z_{1,i}\}$, $z_2 = \{z_{2,i}\}$, where $z_{1,i}$, $z_{2,i}$ are column vectors of the same size as the gradient vector $[Dx_1]_i$ and $[Dx_2]_i$.

For the Huber function, the single voxel minimization problem (5) has an analytic solution (a soft thresholding of s), which provides insight into the properties of the Huber function as a regularizer. Similarly, the special 2-phase problem (8) consists of separable single voxel problems that can be solved for simple matrix norms. The solution provides insights into the properties of the regularizer (6).

Omitting the voxel index i (and with some abuse of notation), the single voxel minimization problem involved in (8) is

$$R_0(d_1, d_2) \triangleq \min_{z_1, z_2} \ \|z_1, z_2\|_U + \beta(\|d_1 - z_1\| + \|d_2 - z_2\|) \quad (9)$$

where d_1, d_2, z_1, z_2 are column vectors of length m, $m \ge 2$. For example, d_1 and d_2 can be the horizontal and vertical differences of image 1 and image 2 (at a certain voxel i), then $m = 2$. The minimization problem (9) is convex but not strictly convex. For some critical values of β, the minimizer z_1, z_2 may not be unique; but any local minimizer is a global minimizer.

Below we characterize the solution to (9) for two choices of the unitarily invariant matrix norm: (a) the Frobenius norm, and (b) the Nuclear norm. The details of our analysis are omitted due to space limit. Again, such a characterization is important to understand the properties of the inf-conv regularizers and informative to select the parameter β. Table I highlights our findings when $d_1 \neq 0$, $d_2 \neq 0$, and $\|d_1\| \ge \|d_2\|$.

The solution z_1, z_2 depends on β for both the Frobenius norm and the Nuclear norm. It is easy to see that z_1 and z_2 must lie in the subspace of d_1 and d_2 (otherwise their

[1] $f_1(s) \to \infty$ as $\|s\| \to \infty$.

TABLE I: The solution to the single voxel problem (9) as a function of the input vectors $d_1 \neq 0$, $d_2 \neq 0$, and $\|d_1\| \geq \|d_2\|$, for using (a) the Frobenius norm and (b) the Nuclear norm.

(a)	$\beta \leq \beta_0$	$\beta_0 < \beta < 1$	$\beta \geq 1$
	$z_1 = 0$ $z_2 = 0$	$z_1 = \min\{r\|d_2\|, \|d_1\|\}\hat{d}_1$ $z_2 = d_2$ $r = \beta/\sqrt{1-\beta^2}$	$z_1 = d_1$ $z_2 = d_2$
	$\beta_0 = \sqrt{1/2}$		

(b)	$\beta \leq \beta_0(\theta)$	$\beta_0(\theta) < \beta < 1$	$\beta \geq 1$
	$z_1 = 0$ $z_2 = 0$	(i) $z_1 \parallel z_2, \|z_1\| = \|z_2\|$ (ii) $z_1 \parallel z_2 = d_2$, $\|z_1\| > \|d_2\|$	$z_1 = d_1$ $z_2 = d_2$
	$\beta_0(\theta) = 1/(\sqrt{1+\|\cos\theta\|})$		

The angle θ is defined as $\theta = \arccos\langle \hat{d}_1, \hat{d}_2 \rangle$. $\hat{d}_i = d_i/\|d_i\|$, $i = 1, 2$.

projection onto the subspace of $\{d_1, d_2\}$ would have smaller objective values). When $\beta \geq 1$, $R_0(d_1, d_2) = \|d_1, d_2\|_U$, *i.e.*, similar to a joint TV regularizer; when $\beta \leq \beta_0$, $R_0(d_1, d_2) = \beta(\|d_1\| + \|d_2\|)$, *i.e.*, similar to a separable TV regularizer. For intermediate values of β, the solution z_1, z_2 induces a decomposition of the input vectors d_1, d_2, such that z_1, z_2 are assigned to the joint penalty, and the remainders, $d_1 - z_1$ and $d_2 - z_2$, are assigned to the separable penalty. These observations are common to both the Frobenius norm and the Nuclear norm.

For the Frobenius norm, z_1 and z_2 will maintain the direction of their respective input d_1 and d_2. Further, the assignment z_1, z_2 to the joint penalty is such that the "shorter" input vector, say $\|d_1\| \geq \|d_2\|$, is always saturated $z_2 = d_2$, and no more than a certain extension factor $r(\beta)$ is assigned to its partner z_1, $\|z_1\| = \min\{r\|d_2\|, \|d_1\|\}$, in the joint penalty.

A difference between the Frobenius norm and the Nuclear norm is that the Frobenius norm is orientation independent, while the Nuclear norm is orientation dependent. Here by orientation we refer to the relative orientation (angle θ) between the column vectors that make up a matrix. This difference leads to differences in the solution z_1, z_2, and different properties of the regularizer $R_0(d_1, d_2)$. The orientation dependency of the Nuclear norm is reflected in z_1, z_2 in the following aspects.

1) The threshold $\beta_0(\theta)$ is orientation dependent. The larger the separation angle θ between d_1, d_2, the higher the threshold. For two pairs of vectors d_1 and d_2 that differ only in the relative orientation, the pair with a larger separation angle θ will more likely be subject to the separable TV.
2) The solutions, z_1 and z_2, are always parallel to each other; this shared direction may not (case (i)) or may (case (ii)) be one of the input directions. The solutions also maintain the sign of their correlation.[2] In other words, if $\langle d_1, d_2 \rangle \lesseqgtr 0$, then $\langle \hat{z}_1, \hat{z}_2 \rangle = \mp 1$; Case (ii) happens when two conditions are met (1) the inputs d_1, d_2 are very "well aligned;" specifically, if $d_1 \cdot d_2 > 0$, the condition is $(d_1 - d_2) \cdot \hat{d}_2 / \|d_1 - d_2\| \triangleq \cos\theta_1 \geq 1/\sqrt{2}$. and (2) $\beta \geq \beta_1 = 1/(\sqrt{2}\cos\theta_1)$. Note that $\beta_0 = 1/(\sqrt{2}\cos(\theta_0/2))$. As $\theta_1 \geq \theta_0 \geq \theta_0/2$, we always have $\beta_1 \geq \beta_0$.

We point out that for intermediate range of $\beta_0(\theta) < \beta < 1$, the decomposition is such that the matrix formed by z_1, z_2 is of rank-1, *i.e.*, z_1 and z_2 are always aligned.

These observations echo the desirable properties for a robust regularizer in which either the joint or the separable penalty plays a major role depending on the local image content.

We develop a generic algorithm for multiphase image reconstruction, to minimize the joint regularizer (6) or (7) with the constraint (1b). Our algorithm uses the linearized form of the alternating direction method of multipliers (ADMM). Two other methods were used for comparison: (a) conventional joint TV as in (1) and (b) separable TV, which is simply

$$\min_{x_1, x_2} \ \mathrm{TV}(x_1) + \mathrm{TV}(x_2)$$
$$\text{subject to} \ \ \|A_1 x_1 - y_1\| \leq \epsilon \|y_1\|, \ \|A_2 x_2 - y_2\| \leq \epsilon \|y_2\|$$

IV. Numerical studies

Our numerical studies focused on three issues: (1) qualitative and (2) quantitative effect of anatomical misalignment on joint image denoising; and (3) mitigating the detrimental effects of misalignment by using the inf-conv regularizer.

To answer question (1), the patient data (Fig. 1) was denoised using separable TV, joint TV with the Nuclear norm, and joint TV with the Frobenius norm for various error tolerance values ϵ [Fig. 2]. Using the separable TV as a reference for artifacts appearance, we notice that organ misalignment creates artifacts for both the joint TV penalties. At low to intermediate error tolerance values, these artifacts might be mistaken for noise; but the artifacts resemble the misaligned anatomy in the other phase, which becomes more evident at higher error tolerance values. At very high error tolerance values, both the artifacts and important signal features will eventually disappear.

Comparing the two joint denoising methods, we observe that at the same error tolerance value, using the Frobenius norm is more prone to artifacts than using the Nuclear norm.

To provide a more quantitative answer, we created hybrid patient data by adding two sets of 20 signals of size 3x3 pixels and contrast 40 HU to the arterial image of the same patient (Fig. 3). Signals of Set 1 were located at the organ misaligned region, Set 2 at the aligned region.

The hybrid patient data were then denoised using the separable TV, the joint TV with both the Nuclear and the Frobenius norm, and the inf-conv regularizer between the joint and the separable TV (for a variety of β values, cf.(7)).

We defined a two-phase contrast-to-noise (CNR) as the quantitative figure of merit. The contrast was defined as the mean intensity difference between the signal in the arterial phase (for either Set 1 or Set 2) and the background in the venous phase. The noise was defined as the summed noise in the signal and the background. Written symbolically,

$$\mathrm{CNR}(i,j) = \frac{\mathrm{mean}[s(i,j)] - \mathrm{mean}[b(i,j=2)]}{\sqrt{\mathrm{var}[s(i,j)] + \mathrm{var}[b(i,j=2)]}} \qquad i = 1, \cdots, 20, \quad j = 1, 2 \quad (10)$$

[2] When the inputs are perpendicular, *i.e.*, $d_1 \cdot d_2 = 0$, $\beta_0(\theta) = 1$. We have separable TV if $\beta < 1$ and joint TV if $\beta > 1$; when $\beta = 1$, the solution $z_1 = \gamma d_1$, $z_2 = \gamma d_2$, for $0 \leq \gamma \leq 1$.

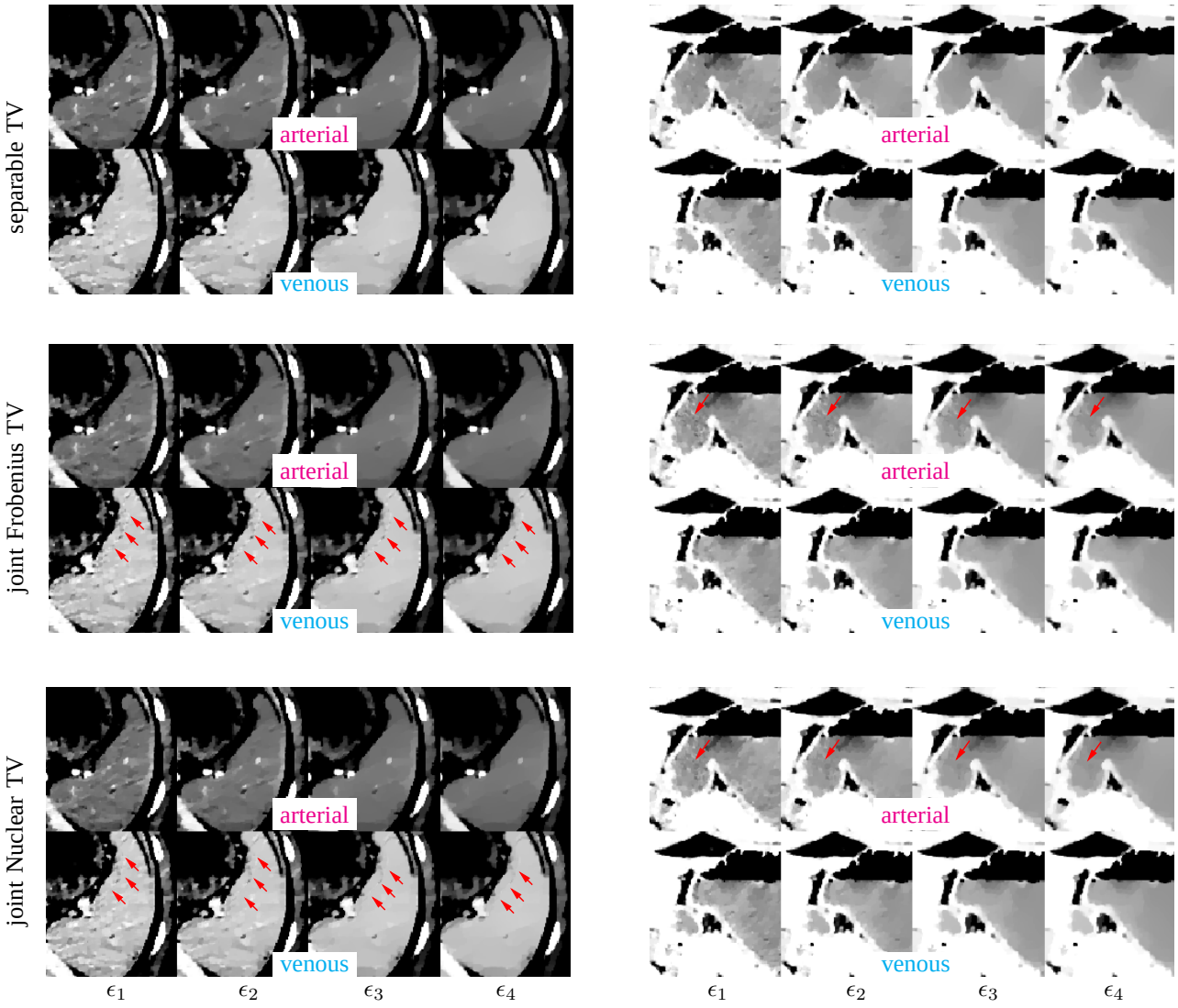


Fig. 2: Image artifacts at misaligned organ boundaries for joint TV denoising using the nuclear norm and the Frobenius norm. Image denoising using the separable (channel-by-channel) TV is shown for reference. The error tolerance in these examples were 2%, 2.2%, 2.5%, 2.8% for ϵ_1, ϵ_2, ϵ_3, ϵ_4, respectively. Misalignment artifacts were more observable in images using the joint Frobenius norm than the joint Nuclear norm. The display window for the 1st column (spleen) is (C, W) = (66, 100). for the 2nd column (stomach) is (C, W) = (0, 100).

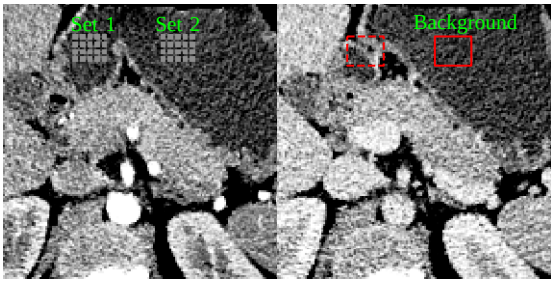
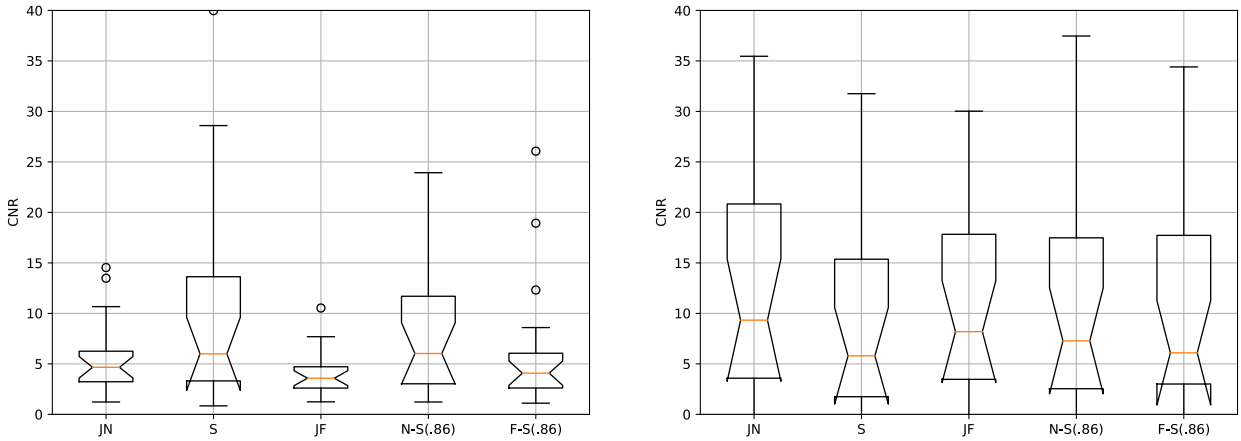


Fig. 3: Hybrid data generation. We added two sets of 20 signals of size 3x3 pixels and contrast of 40 HU to the arterial image at the marked locations. Signals of Set 1 are at the arterial-venous misaligned region (the dashed box), Set 2 at the arterial-venous aligned region. For both sets of signals, the CNR calculation used the same background area (divided into 20 3x3 cells, not shown) marked by the solid (red) box. (C, W) = (66, 200) HU.

where i indexes the 20 signals of size 3x3, and j indexes the set location, $j = 1$ (the misaligned) and $j = 2$ is the aligned region. Both the mean and the variance in (10) were calculated over 3x3 pixels.

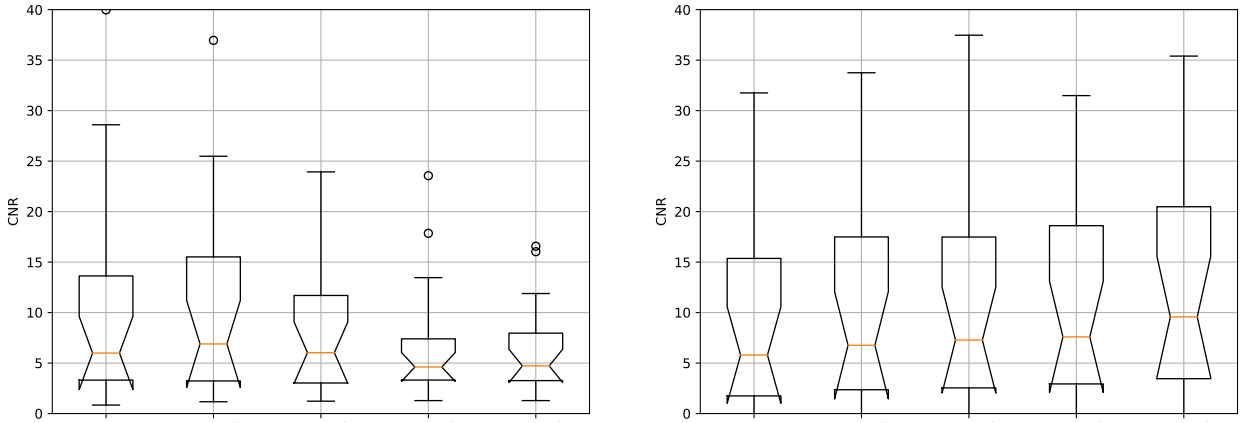
Among the four error tolerance values of $2.0\%, 2.2\%, 2.5\%, 2.8\%$ that we used in Fig. 2, we found that the CNR in (10) peaked at $\epsilon = 2.5\%$. Therefore this value was used for comparing different methods. The (notched) boxplots of the CNRs for the two sets (two locations) of 20 signals are shown in Fig. 4 for the different methods.

Comparing joint denoising and separable denoising (Fig. 4(a), the first three columns), we notice that the misaligned organ boundaries in Set 1 reduces the CNR, and that the separable TV achieved higher CNR than either of the joint denoising methods. When there was no misalignment (Set 2, Fig. 4(b)), the joint denoising methods achieve higher CNR than the separable TV. Comparing between the two joint



(a) Set 1 (misaligned locations) (b) Set 2 (aligned locations)

Fig. 4: Boxplots of CNR using different image denoising methods. The labels are: J-N: joint TV with Nuclear norm; S: separable TV; J-F: joint TV with Frobenius norm; N-S(β): infimal convolution between nuclear norm TV and separable TV, with parameter β; F-S(β): infimal convolution between Frobenius norm TV and separable TV, with parameter β.



(a) Set 1 (misaligned locations) (b) Set 2 (aligned locations)

Fig. 5: Boxplots of CNR using infimal convolution between joint nuclear TV and separable TV for different parameters β. As $\beta \uparrow 1$, N-S(β) approaches the performance of JN in Fig. 4. As $\beta \downarrow 0.707$, N-S(β) approaches the performance of S (the first column). The labels are: S: separable TV; N-S(β): infimal convolution between nuclear norm TV and separable TV, with parameter β.

denoising methods, using the Nuclear norm performed better than using the Frobenius norm. These observations support and strengthen the qualitative statements we made about artifacts level between the two methods.

For proper weighting parameters β, the inf-conv regularizer can maintain the advantages of the joint TV when there is no organ misalignment, while improving upon joint TV when there is.

Further, comparing between the two inf-conv regularizers, N-S and F-S, *i.e.*, infimal convolution between joint nuclear TV and separable TV (N-S), and joint Frobenius TV and separable TV (F-S), we observe that N-S achieved higher CNR, consistent with the performance of the two joint TV penalties.

By adjusting the weighting parameter β in (7), the infimal convolution achieves a continuum between joint TV and separable TV. For a 2-phase problem, as $\beta \uparrow 1$, it behaves more like joint TV; and $\beta \downarrow 0.707$, it behaves more like separable TV. These general trends can be observed from the CNR boxplots in Fig. 5 for some sample β values.

V. Discussion

We did not use computer simulations in the evaluation as it might be difficult to simulate realistic anatomical variations, contrast uptake, and organ misalignment. Instead, our numerical studies were based on patient data. The strength of this approach is the realism in data generation. But the lack of ground truth and noise variations may limit the statistical power of our quantitative statements. Future studies should supplement the current work with experimental data acquired under controlled conditions.

VI. Conclusions

We evaluated the performance of joint total variation (TV) penalty for multiphase CT when there exists organ shifts

and boundary misalignment among the different phases. We observed image artifacts at the misaligned organ boundaries for the two joint TV penalties we studied (joint nuclear norm TV and joint Frobenius norm TV). The image artifacts might be mistaken for increased noise for low or intermediate penalty parameters (ϵ in this work). The statistical noise combined with anatomical error caused by organ misalignment negatively affect the performance of joint TV denoising.

A simple approach to maintain the advantages of joint TV while avoiding anatomical artifacts is to use the separable TV when misalignments are present, and use the joint TV when they are not. Building on this idea, we proposed to use the infimal convolution between the joint TV and separable TV to achieve a continuum between the two. The infimal convolution can adapt to the local image content, incorporate the common edge information when it exists, and resort to separable TV when organ misalignments are present.

Our preliminary evaluation using a two-phase patient data set demonstrated that the infimal convolution regularizer can improve upon joint TV when misalignments are present, while maintaining the advantage of joint TV over separable TV when they are not. Moreover, it may be important to adapt the weighting parameter β in the infimal convolution to truly achieve "the best of both worlds."

REFERENCES

[1] B. Goldluecke, E. Strekalovskiy, and D. Cremers, "The natural vectorial total variation which arises from geometric measure theory," *SIAM Journal on Imaging Sciences*, vol. 5, no. 2, pp. 537–563, 2012.

[2] D. S. Rigie, A. A. Sanchez, and P. J. L. Riviere, "Assessment of vectorial total variation penalties on realistic dual-energy ct data," *Physics in Medicine & Biology*, vol. 62, no. 8, p. 3284, 2017.

[3] H. H. Bauschke and P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*, vol. 408. Springer, 2011.

[4] A. Chambolle and P.-L. Lions, "Image recovery via total variation minimization and related problems," *Numerische Mathematik*, vol. 76, no. 2, pp. 167–188, 1997.

[5] M. Holler and K. Kunisch, "On infimal convolution of TV-type functionals and applications to video and image reconstruction," *SIAM Journal on Imaging Sciences*, vol. 7, no. 4, pp. 2258–2300, 2014.

[6] M. Schloegl, M. Holler, A. Schwarzl, K. Bredies, and R. Stollberger, "Infimal convolution of total generalized variation functionals for dynamic MRI," *Magnetic resonance in medicine*, vol. 78, no. 1, pp. 142–155, 2017.

Statistical Image Reconstruction with Sample-Based Beam-Hardening compensation for X-ray CT

C. Martinez, J. A. Fessler, M. Desco, and M. Abella

***Abstract*— CT images are often affected by beam-hardening artifacts due to the polychromatic nature of the X-ray spectra. These artifacts appear in the image as cupping in homogeneous areas and streaks between dense parts in heterogeneous samples.**

This paper proposes a new statistical reconstruction method for X-ray CT based on Poisson statistics, taking into account the non-linearities caused by beam hardening. To avoid needing knowledge of the X-ray spectrum, the method obtains the 2D beam-hardening function using information provided by the acquired data itself.

Evaluation using simulations showed beam hardening artifact reductions similar to those achieved with conventional post-processing techniques while avoiding noise and artifacts in low-dose studies.

***Index Terms*—Beam-hardening, CT, artifacts, penalized-likelihood, streaks, polychromatic.**

INTRODUCTION

The beam hardening effect in computed tomography derives from the polychromatic nature of the radiation produced by X-ray tubes. Due to the energy dependence of mass attenuation coefficients, low energy photons are preferably absorbed, causing a shift of the mean energy of the X-ray beam to higher values. This effect leads to two main artifacts in uncorrected reconstructed images: cupping in homogeneous regions and streaks between dense areas in heterogeneous regions [1].

Several strategies can be found in the literature to compensate for this effect. Physical filters are generally used to pre-harden the beam before reaching the sample, but this is not enough to remove the artifacts. Another method implemented in most commercial scanners is the water-linearization, based on a prior calibration with a water-equivalent phantom. This method models the object as homogeneous and corrects only cupping artifact [2]. To correct also streaks, Nalcioglu et al. [3] proposed a method that requires knowledge of the spectrum, the linear attenuation coefficients and the thickness of soft tissue and bone traversed estimated by means of a preliminary reconstruction. Joseph et al. [4] proposed a similar idea modeling the corrected data with a second-order polynomial dependent on the bone traversed thickness. However, the optimum parameters for this model could be exactly obtained only with a complete characterization of the spectrum. This need of knowledge of the spectra was avoided in [5, 6], also based on a linear combination of basis images to correct streaks. The coefficients of this linear combination are obtained iteratively maximizing the flatness of the soft tissue areas, which could reduce the soft-tissue contrast. Cupping correction is achieved using the water-linearization method, which needs a calibration step. We recently proposed two new methods extending the water-linearization to a 2D linearization [7, 8]. However, they require a good bone and soft tissue segmentation which may hinder their use in low-dose studies.

To deal with low-dose studies, Elbrakri et al. presented a statistical method that requires knowledge of the spectrum [9, 10]. This is avoided in [11, 12] with a simplified statistical algorithm that parametrizes the beam-hardening function following the model proposed by Joseph and Spital [4].

This paper presents a variation of [12] that replaces the approximation functions with the real measured line integrals of bone and soft tissue of the sample as proposed in [8].

MATERIALS AND METHODS

A. Forward model

We model the measurements as independently distributed Poisson random variables [13] contaminated by extra background counts, primarily scatter:

$$Y_i \sim Poisson\left\{ \overline{Y_i} \right\}, i = 1, \dots, N \quad (1)$$

with

$$\overline{Y_i} = \int I_i(\varepsilon) e^{-\int_{L_i} \mu(\varepsilon) dl} d\varepsilon + r_i \quad (2)$$

where $\mu(\varepsilon)$ is the attenuation coefficient at each energy ε, the integral in the exponent is taken over the line L_i

Manuscript received April 2, 2018. This work was partially supported by the Ministry of Economy, Industry and Competitiveness (DPI2016-79075-R) and ISCIII-FIS grants DTS17/00122, co-financed by ERDF (FEDER) Funds from the European Commission, "A way of making Europe". The CNIC is supported by the Ministry of Economy, Industry and Competitiveness (MEIC) and the Pro CNIC Foundation, and is a Severo Ochoa Center of Excellence (SEV-2015-0505).

C. Martinez, M. Abella and M. Desco are with the Dpto. Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid and the Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, España (e-mail: mabella@ing.uc3m.es, cmartinez@hggm.es).

J. A. Fessler is with Electrical Engineering and Computer Science department, The University of Michigan, Ann Arbor, MI 48109-2122, USA.

M. Abella and M. Desco are with the Centro Nacional de Investigaciones Cardiovasculares Carlos III (CNIC), Madrid, España (e-mail: manuel.desco@cnic.es).

M. Desco is with the Centro de investigación en red salud mental (CIBERSAM), Madrid, España. (e-mail: desco@hggm.es).

followed by the ray, and $I_i(\varepsilon)$ is the incident intensity and the term r_i accounts for mean scatter and mean other background signals for the *i*-th ray.

Following [4, 14] we model the attenuation coefficient in Eq. 2 at each pixel *j* as:

$$\mu_j(\varepsilon) = \sum_{k=1}^{K} mac_k(\varepsilon) f_k^j \rho_j \quad (3)$$

where mac_k is the mass attenuation coefficient of the material *k*, ρ the density and f_k^j is a unitless fraction that describes the contribution of the material *k* to attenuation in the pixel *j*. We assume *K*=2, i.e., the object contains only soft tissue (*ST*) and bone (*B*). The contribution of each tissue type to the line integral along the *i*-th ray is:

$$t_{ST}(\rho) = \sum_{j=1}^{p} a_{ij} f_{ST}^j(\rho_j)\rho_j \quad (4)$$

$$t_B(\rho) = \sum_{j=1}^{p} a_{ij} f_B^j(\rho_j)\rho_j \quad (5)$$

where a_{ij} are the elements of the system matrix. Here we allow the unitless fraction (f_k^j) to only be 1 or 0, i.e., the pixels do not contain mixtures of tissues. Eq. 2 for the expected value of the measured data along the path *i* results in:

$$\overline{Y_i}(\rho) = \int I_i(\varepsilon) e^{-\int_{L_i} \mu(\varepsilon) dl} d\varepsilon + r_i = \\ = I_i e^{-F\left(t_{ST}^i(\rho), t_B^i(\rho)\right)} + r_i \quad (6)$$

where

$$I_i \equiv \int I_i(\varepsilon) d\varepsilon \quad (7)$$

and the beam-hardening function, *F*, is:

$$F(t_{ST}, t_B) = -\log \int \frac{I(\varepsilon)}{I} e^{-mac_{ST}(\varepsilon) t_{ST} - mac_B(\varepsilon) t_B} \, d\varepsilon \quad (8)$$

dropping the dependence on ray *i* for simplicity.

B. Beam-Hardening function

The beam-hardening function $F(t_{ST}, t_B)$ could be analytically calculated from a known spectrum, but often this information is not available. To avoid assuming spectrum knowledge, the proposed method determines $F(t_{ST}, t_B)$ experimentally using the acquired data following the process outlined in Fig. 1.

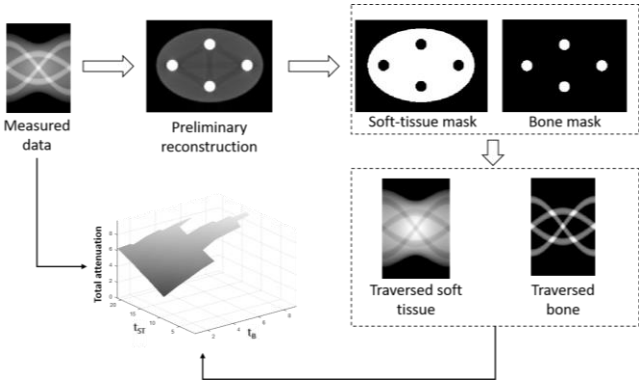


Fig. 1. Workflow for the generation of the beam-hardening function.

First bone and soft-tissue masks are obtained by thresholding a preliminary reconstructed image. These two masks are then multiplied by the density of each tissue and projected, which will be the *x* and *y* axis, where *x* corresponds to t_{ST} and *y* corresponds to t_B. The value in the original projection will be the *z* axis, corresponding to $F(t_{ST}, t_B)$.

The generated $F(t_{ST}, t_B)$ will not cover the whole space, since a specific acquisition will not have all possible combinations of soft tissue and bone (Fig. 2, left). To completely characterize the beam-hardening function, we "extrapolate" the incomplete function *F* using a quadratic function (Fig. 2, right).

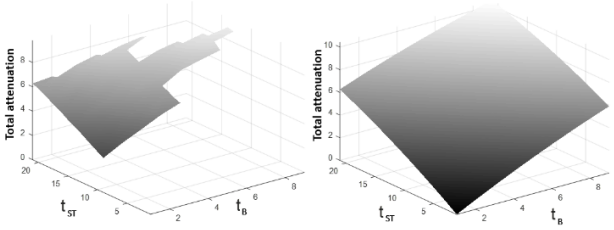


Fig. 2. Measured (left) and extrapolated (right) beam-hardening function.

C. Cost function

The negative log-likelihood for independent Poisson measurement is:

$$L(\rho) = -\sum_{i=1}^{N} h_i\left(F\left(t_{ST}(\rho), t_B(\rho)\right)\right) \quad (9)$$

where

$$h_i(d) = -Y_i \log(I_i e^{-d} + r_i) + I_i e^{-d} + r_i \quad (10)$$

Since minimizing $L(\rho)$ is generally an ill-posed problem, regularization is included by adding a penalty term to control how much the object ρ departs from our assumptions about image properties. In this work, we use a 3D roughness penalty function with the convex edge-preserving Huber potential. The resulting penalized cost function is:

$$\Phi(\rho) = L(\rho) + \beta R(\rho) \quad (11)$$

where β is a scalar that controls the tradeoff between the data-fit and penalty terms.

D. Algorithm

We derive an iterative algorithm based on separable quadratic surrogates using the principles of optimization transfer [14], resulting in the following update:

$$\rho^{n+1} = \rho^n - D^{-1} \nabla \Phi(\rho^n) \quad (12)$$

where *D* is a diagonal matrix that influences the rate of convergence. We originally designed *D* to ensure that the algorithm monotonically decreases the cost function. As in [9], in practice we choose the elements of *D* approximately by using the precomputed curvature:

$$d_j = \left(mac_{ST}^2(\varepsilon_{eff}) + mac_B^2(\varepsilon_{eff})\right) \sum_{i=1}^{N} a_{ij} \sum_j a_{ij} Y_i \quad (13)$$

where the effective $mac(\varepsilon_{eff})$ values for each tissue are approximated using the derivative of beam-hardening function at (0,0).

Artifacts in the preliminary reconstruction may hinder the segmentation to obtain the bone and soft-tissue masks, resulting in an erroneous beam-hardening function. To tackle this problem, new bone and soft-tissue masks are obtained from the solution at each iteration and are used to recalculate the beam-hardening function.

EVALUATION

Preliminary evaluation used simulations of a 2D phantom with two cortical bone inserts having density 1.9 gr/cm^3, one trabecular bone insert with density 1.5 gr/cm^3 and one adipose-tissue insert with density 0.9 gr/cm^3, inside of a soft-tissue ellipse with density 1.06 gr/cm^3 (Fig. 3).

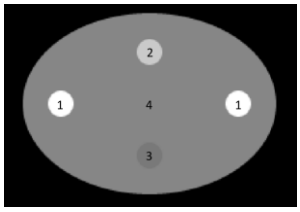


Fig. 3. Test phantom with two cortical bone inserts (1), one trabecular bone insert (2), and one adipose-tissue insert (3) inside of a soft-tissue ellipse (4).

Four polyenergetic X-ray datasets were generated using MIRT*(http://www.eecs.umich.edu/~fessler/code/index.html)* with a 50 kVp spectrum and 0.1 mm aluminum filtration, typically used in preclinical studies. The number of counts per ray were 10^5 and 10^6 to simulate low-SNR and high-SNR scenarios respectively. For each scenario, we obtained 45 and 180 projections in a span of 180 degrees with a matrix size of 256×256 pixels and 0.1×0.1 mm pixel size.

The data were reconstructed with FBP, with FBP corrected by the free calibration method (fCM) proposed in [8] and by the proposed statistical algorithm.

RESULTS

Fig. 4 shows the segmented soft-tissue mask along different iterations for the low-SNR scenario with 45 projections. The mask in the first iteration has holes due to the streaks from beam-hardening effect and low sampling, which are removed in subsequent iterations.

Fig. 4. Soft tissue mask in iteration 1 (left), 5 (center) and 15 (right).

Fig. 5 shows the results for 180 projections. The result of FBP corrected with fCM shows a good compensation of beam-hardening artifacts but it fails to correct streaks with high noise (low SNR scenario), while the proposed method results in good quality reconstructions in high-SNR and low-SNR scenarios.

When the number of projection is reduced to 45 (Fig. 6) FBP+fCM eliminates the beam hardening artifacts in the high-SNR scenario, but the reconstruction is contaminated with streaks due to the lack of projections. For the low-SNR scenario, FBP+fCM also fails to compensate the streaks due to beam-hardening effect, because of the wrong segmentation used to create the beam-hardening function. The proposed method eliminates the beam hardening and the streaks in all cases.

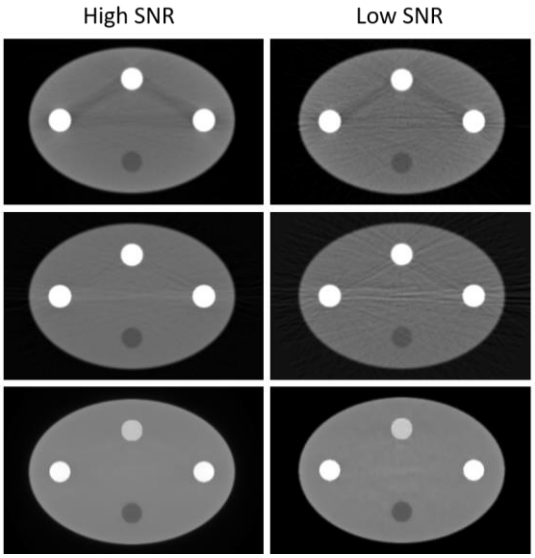


Fig. 5. Results for the 180 projections datasets using FBP (top), FBP + fCM (center) and the proposed algorithm (bottom).

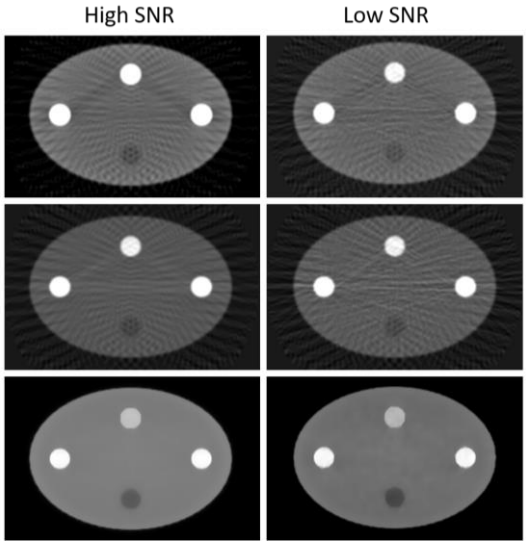


Fig. 6. Results for the 45 projections datasets using FBP (top), FBP +fCM (center) and the proposed algorithm (bottom).

CONCLUSIONS

We present a new statistical reconstruction algorithm that includes beam-hardening correction without needing any spectrum knowledge or correction parameters optimization.

The method models the polychromatic effect via a beam-hardening function determined from the acquired data and a segmentation of bone and soft-tissue masks, which are

iteratively improved. Results on simulated data show a reduction of streaks due to both beam-hardening effect and low number of projections.

REFERENCES

[1] J. F. Barrett and N. Keat, "Artifacts in CT: Recognition and Avoidance," *RadioGraphics*, vol. 24, pp. 1679-1691, 2004.

[2] R. A. Brooks and G. D. Chiro, "Beam hardening in x-ray reconstruction tomography," *Phys. Med. Biol.*, vol. 21, pp. 390-8, 1976.

[3] O. Nalcioglu and R. Y. Lou, "Post-reconstruction Method for Beam Hardening in Computerised Tomography," *Phys. Med. Biol.*, vol. 24, pp. 330-40, 1979.

[4] P. M. Joseph and R. D. Spital, "A method for correcting bone induced artifacts in computed tomography scanners," *J. Comp. Assisted Tomo.*, vol. 2, pp. 100-8, 1978.

[5] Y. Kyriakou, E. Meyer, D. Prell, and M. Kachelrieß, "Empirical beam hardening correction (EBHC) for CT," *Med Phys*, vol. 37, pp. 5179-5187, 2010.

[6] S. Schüller, S. Sawall, K. Stannigel, M. Hülsbusch, J. Ulrici, E. Hell, *et al.*, "Segmentation-free empirical beam hardening correction for CT," *Med Phys*, vol. 42, pp. 794-803, 2015.

[7] C. Martinez, C. De Molina, M. Desco, and M. Abella, "Simple method for beam-hardening correction based on a 2D linearization function," in *The 4th International Meeting on Image Formation in X-Ray Computed Tomography (CTMeeting 2016)*, Bamberg, Germany, 2016.

[8] C. Martínez, C. de Molina, M. Desco, and M. Abella, "Calibration-free method for beam-hardening compensation: preliminary results," in *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, Atlanta, GA, 2017.

[9] I. A. Elbakri and J. A. Fessler, "Statistical Image Reconstruction for Polyenergetic X-Ray Computed Tomography," *IEEE Trans. Med. Imaging*, vol. 21, pp. 89-99, 2002.

[10] I. A. Elbakri and J. A. Fessler, "Segmentation-free statistical image reconstruction for polyenergetic x-ray computed tomography with experimental validation," *Phys. Med. Biol.*, vol. 48, pp. 2453-77, 2003.

[11] S. Srivastava and J. A. Fessler, "Simplied statistical image reconstruction algorithm for polyenergetic X-ray CT," presented at the IEEE Nuclear Science Symposium Conference Record, 2005.

[12] M. Abella and J. A. Fessler, "A new statistical image reconstruction algorithm for polyenergetic X-ray CT," *Proceedings of the 2009 IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 165-8, 2009.

[13] H. Erdogan and J. A. Fessler, "Monotonic algorithms for transmission tomography," *IEEE Trans. Med. Imaging*, vol. 18, pp. 801-14, 1999.

[14] R. E. Alvarez and A. Macovski, "Energy-selective reconstruction in x-ray computerized tomography," *Phys. Med. Biol.*, vol. 21, pp. 733–44, 1976.

Model Based Iterative Reconstruction With Spatially Adaptive Sinogram Weights for Wide-Cone Cardiac CT

Amirkoushyar Ziabari, Dong Hye Ye, Lin Fu, Somesh Srivastava, Ken D. Sauer, Jean-Baptiste Thibault, and Charles A. Bouman

***Abstract*—With the recent introduction of CT scanners with large cone angles, wide coverage detectors now provide a desirable scanning platform for cardiac CT that allows whole heart imaging in a single rotation. On these scanners, while half-scan data is strictly sufficient to produce images with the best temporal resolution, acquiring a full 360 degree rotation worth of data is beneficial for wide-cone image reconstruction at negligible additional radiation dose. Applying Model-Based Iterative Reconstruction (MBIR) algorithm to the heart has already been shown to yield significant enhancement in image quality for cardiac CT. But imaging the heart in large cone angle geometry leads to apparently conflicting data usage considerations. On the one hand, in addition to using the fastest available scanner rotation speed, a minimal complete data set of 180 degrees plus the fan angle is typically used to minimize both cardiac and respiratory motion. On the other hand, a full 360 degree acquisition helps better handle the challenges of missing frequencies and incomplete projections associated with wide-cone half-scan data acquisition. In this paper, we develop a Spatially Adaptive sinogram Weights MBIR algorithm, deemed SAW-MBIR, that is designed to achieve the benefits of both half-scan and full-scan reconstructions in order to maximize temporal resolution performance over the heart region while providing stable results over the whole volume covered with the wide-area detector. Spatially-adaptive sinogram weights applied to each projection measurement in SAW-MBIR are designed to selectively perform back projection from the full-scan and half-scan portion of the sinogram based on both projection angle and reconstructed voxel location. We demonstrate with experimental results of SAW-MBIR applied to whole-heart cardiac CT clinical data that overall temporal resolution performance matches half-scan results while full volume image quality compares positively to the standard MBIR reconstruction of full-scan data.**

***Index Terms*—MBIR, Selective Back Projection, Cardiac CT, Temporal Resolution, Spatially-Adaptive**

I. Introduction

CARDIAC CT reconstruction requires high temporal resolution to capture the moving heart. To achieve high temporal resolution in the reconstructed image, it is desirable to reconstruct from limited-view angle projections, i.e. a minimal complete dataset (180 + the fan angle), also referred to as half-scan, instead of a full-scan acquisition with 360 degrees of data [1], [2]. Although half-scan is sufficient for image reconstruction in the mid-plane, it results in incomplete projections and missing spatial frequencies for image planes under larger cone angles. Without explicit corrections, analytical algorithms such as filtered back projection (FBP) may lead to artifacts when performing reconstruction from half-scan data at high cone angles [3], [4].

On the other hand, wide detector apertures and large cone angles are advantageous in cardiac CT as they allow the acquisition of the whole heart in a single rotation, with reduced overall scan time over the target volume and higher X-ray tube efficiency [5]. While most conventional clinical CT platforms with smaller detectors rely only on half-scan data for heart imaging, wide-area detectors allow full-scan acquisitions with 360 degree projections with negligible additional radiation dose. In this geometry, spatially-dependent sinogram weighting can help conserve temporal resolution performance over the heart region. In analytical methods such as FBP, this is relatively straightforward because the reconstruction at each spatial location can be carried out in a closed form equation, independent of other spatial locations, as long as the weights are calculated to sum to a constant over redundant projection data contributing to the same location. However, large cone angles result in missing data and inconsistent projections which can lead to image distortions without explicit compensation [5].

Model Based Iterative Reconstruction (MBIR) affords additional flexibility in data handling and has been shown to perform better than analytical methods in terms of noise, artifact reduction, and spatial resolution performance [6]–[8], thus allowing significant reduction in patient radiation dose [1], [8], [9]. But specifically for cardiac imaging, MBIR needs to balance temporal resolution performance with image spatial resolution, noise, and uniformity. Spatially-dependent sinogram weighting can be employed to control the contributions from full-scan and half-scan data sets depending on voxel location in MBIR as well, as long as the algorithm provides for consistent problem formulation that allows stable convergence to the minimum of the cost function.

Individual voxel locations or sub-regions of the cardiac volume may be properly reconstructed with spatially-dependent sinogram weighting in MBIR when back projecting a single full-scan wide-cone dataset which treats differently locations within the heart area vs. outside the heart. Parker weights [10] can for instance be used for back projection of half-scan data over the heart region, whereas regions outside the heart are back projected using the corresponding full-scan sinogram weights. In this case, the algorithm needs to deal with an

A. Ziabari, D. Ye, and C. A. Bouman (corresponding author) are with the ECE Department, Purdue University, IN 47907 USA (Emails: aziabari@purdue.edu; yed@purdue.edu; Bouman@purdue.edu.)

S. Srivastava and J.-B. Thibault are with GE Healthcare Technologies, Waukesha, WI 53188 USA (Emails: someshsrivastava@ge.com; Jean-Baptiste.Thibault@med.ge.com)

L. Fu is with GE Global Research, Niskayuna, NY, 12309 USA (Email: fulin@ge.com)

K. D. Sauer is with the EE Department, University of Notre Dame, Notre Dame, IN 46556-5637 USA (Email:sauer@nd.edu).

unmatched forward/backward projection pair (or dual system matrix) [11]. Such approaches have previously been investigated to reduce artifacts [12] as well as to accelerate the convergence of IR algorithms [11], [13], [14]. However, in these past works, the back projection operator does not contain any specific information about the locations of both the measurements and the reconstructed voxels.

In this paper, we propose the SAW-MBIR algorithm that uses spatially-adaptive sinogram weights to perform selective back projection from different sub-regions or voxels in the reconstructed image, and apply it to wide-cone angle cardiac CT. In a single iteration, SAW-MBIR selectively performs half-scan back projection over the heart region and full-scan back projection over the rest of the volume in order to address the challenges normally associated with incomplete and missing data from the half scan geometry in regions with high cone angles. Experimental results demonstrate that the SAW-MBIR algorithm achieves consistent temporal resolution with half scan MBIR reconstruction over the heart region, and good image quality consistent with full-scan MBIR reconstruction in the rest of the reconstructed volume, all in a consistent algorithm framework operating with a single reconstruction.

II. SAW-MBIR

A. Theoretical Formulation

The objective of this work is to develop an MBIR formulation with spatially-adaptive sinogram weighting (SAW-MBIR) for selective back projection of cardiac full-scan CT data. In this approach, separate back projections are performed from the sinogram residual corresponding to different sub-regions in the reconstructed image volume. Subsequently, the back projection results are weighted using a mask applied to the spatial location of each sub-regions. When used over half-scan and full-scan projection regions, this supports the reconstruction of an image volume with good temporal resolution over the central region while using full sampling to reconstruct the rest of the volume.

To explain the idea, we use gradient descent to find the solution to the problem:

$$y = Ax \tag{1}$$

where $A \in \Re^{M \times N}$ is the system matrix, $x \in \Re^{N}$ is the unknown vector of the image, and $y \in \Re^{M}$ is the vector of sinogram measurements. The corresponding maximum *a posteriori* (MAP) cost function is:

$$f(x) = \frac{1}{2}||y - Ax||^2_W + \Phi(x) \tag{2}$$

The norm in the first term is the data fidelity term, where W is a diagonal statistical weighting matrix that models the noise [8] and $\Phi(x)$ is the regularization function. We can then write the gradient descent update equation as follows:

$$g(x) = \nabla f(x) = A^T W(Ax - y) + \nabla\Phi(x) \tag{3}$$

$$x^{(k+1)} = x^k - \alpha g(x), \tag{4}$$

where k denotes the iteration number and α is the step size. Here, A^T is the matched back projection operator to the forward projection system matrix A.

For a cardiac CT scan that includes projections from the full 360 degree rotation scan, we denote y the complete set of projections (full-scan), and consider the limited view angle measurements y_h (half-scan) as a subset of the full-scan measurements. We note $y_{h'}$ the complement of y_h, which includes all the rest of the projections from outside the limited-view angles. We define:

$$y \equiv \begin{bmatrix} y_h \\ y_{h'} \end{bmatrix} \tag{5}$$

We also partition the image x into two regions of x_m, which corresponds to the target region (here, the heart region) and should be back projected from half-scan measurement data, and $x_{m'}$, which corresponds to the background region.

$$x \equiv \begin{bmatrix} x_m \\ x_{m'} \end{bmatrix} \tag{6}$$

Consistently, we partition the A matrix into four sub-matrices:

$$A \equiv \begin{bmatrix} A_{hm} & A_{hm'} \\ A_{h'm} & A_{h'm'} \end{bmatrix} \tag{7}$$

A_{hm} and $A_{hm'}$ are the sub-matrices that project x_m and x'_m into the half-scan measured sinogram y_h, respectively. On the other hand, $A_{h'm}$ and $A_{h'm'}$ project x_m and x'_m into the region outside the half-scan measured sinogram y'_h.

We can rewrite equation (1) as follows:

$$\begin{bmatrix} y_h \\ y_{h'} \end{bmatrix} = \begin{bmatrix} A_{hm} & A_{hm'} \\ A_{h'm} & A_{h'm'} \end{bmatrix} \begin{bmatrix} x_m \\ x_{m'} \end{bmatrix} \tag{8}$$

Using half-scan projections is advantageous in improving the temporal resolution relative to the full-scan data. The corresponding back projection matrix can be written as:

$$A^T_{half} \equiv \begin{bmatrix} A^T_{hm} & 0 \\ A^T_{hm'} & 0 \end{bmatrix} \tag{9}$$

It is challenging to reconstruct image slices at high cone angles from only half-scan data because of incomplete and missing projections in the half-scan dataset. Instead, one may benefit from using full-scan projections outside the central primary region of interest, in conjunction with half-scan data only for the center region, in order to both improve temporal resolution at the center and maintain high image quality throughout.

Algorithm 1 SAW-MBIR

$y \leftarrow$ measured sinogram
$x^0 \leftarrow$ FBP
$\alpha \leftarrow$ step size
For k iterations {
$g_s(x^k) = A^T_{masked} W(Ax^k - y) + \nabla\Phi(x^k)$
$x^{(k+1)} = x^k - \alpha g_s(x^k)$ }

Here, we introduce the SAW-MBIR algorithm that uses spatially-adaptive sinogram weights to perform selective back projection while retaining a consistent framework for iterative

optimization. We first define the masked back projection operator A^T_{masked}:

$$A^T_{masked} \equiv \begin{bmatrix} A^T_{hm} & 0 \\ A^T_{hm'} & A^T_{h'm'} \end{bmatrix} \tag{10}$$

Compared to A^T, by setting $A_{h'm} = 0$, A^T_{masked} decouples $y_{h'}$ from x_m when back projecting the sinogram residual to the image. The back projector is masked (i.e. set to zero) depending on the spatial location of the regions that are outside the half-scan projection data. Putting A^T_{masked} in equation (3), we get the following pseudo-gradient:

$$\begin{aligned} g(x) = g_s(x) &\equiv \begin{bmatrix} [g_s(x)]_m \\ [g_s(x)]_{m'} \end{bmatrix} + \nabla\Phi(x) \\ &= A^T_{masked} W(Ax - y) + \nabla\Phi(x), \end{aligned} \tag{11}$$

where:

$$[g_s(x)]_m \equiv \begin{bmatrix} A_{hm} \\ 0 \end{bmatrix}^T W(Ax - y) \tag{12}$$

$$[g_s(x)]_{m'} \equiv \begin{bmatrix} A_{hm'} \\ A_{h'm'} \end{bmatrix}^T W(Ax - y) \tag{13}$$

And the update equation becomes:

$$x^{(k+1)} = x^k - \alpha g_s(x^k) \tag{14}$$

Eqs. (12) and (13) above show how the proposed SAW-MBIR algorithm selectively performs half-scan back projection of locations within the mask (Eq. (12)) and full-scan back projection of locations outside the mask (Eq. (13)). A summary of the proposed method is shown in algorithm 1.

B. Application to Wide-Cone Cardiac CT

SAW-MBIR can be applied in whole-heart wide-cone cardiac CT to produce better results than standard MBIR. An illustration of the transverse view of the cardiac CT data acquisition is shown in Fig. 1a. Cone-beam projections are drawn for two opposite view angles. Standard MBIR with full-scan projection data can produce good image quality in the full volume by reducing noise and artifacts compared to analytical methods, but does not inherently produce the temporal resolution of the half-scan acquisition over the heart area typically covered in the fully-sampled (i.e. light purple) region. The SAW-MBIR algorithm introduced in section II.A. can be used to improve the temporal resolution in the heart region relative to full-scan.

We define a mask to distinguish between the area where complete projection data is available from the half-scan data and the rest of the volume. We intend to reconstruct the light purple region of Fig. 1a primarily from the half-scan data, and the dark purple region primarily from the full-scan data. Cross sections of the mask taken at the center, between the center and the edge, and at the edge, are shown in Fig. 1b, c, and d, respectively.

The mask is set to one based on the spatial location of the voxels with respect to the light and dark purple regions of Fig. 1a. Since the half-scan and full-scan back projection operators only differ in the angular range of integration, the mask is used to implement A^T_{masked} of Eq. (10). The masked back projector selectively performs back projection from the half-scan and full-scan measurements at each iteration of MBIR. Basically, the masked back projector works as if only measurements inside the region corresponding to the half-scan data are back projected by half-scan back projector of Eq. (12), and the rest of the measurements are back projected using the full-scan back projector of Eq. (13).

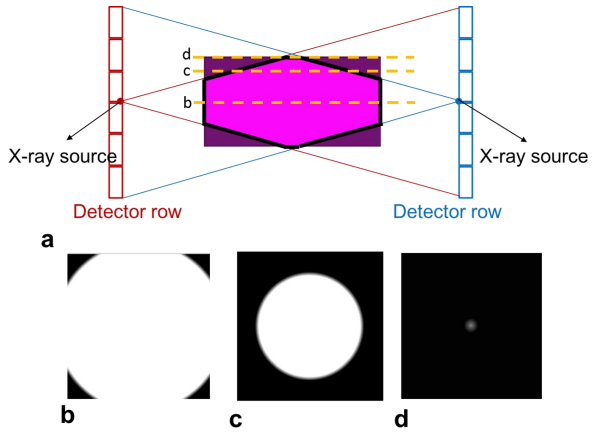


Fig. 1. a. A schematic of the transverse view of the cardiac CT data acquisition. Projections from two opposite view angles are drawn in red and blue. The desired reconstruction volume is indicated by the rectangular box at the center. A mask is defined to differentiate between half-scan and full-scan back projected regions. Our intention is to reconstruct the light purple region primarily based on the half-scan data, and the dark purple region primarily based on the full-scan data. Slices of the reconstruction mask are shown: (b) at the center, (c) between center and edge, and (d) at the edge.

III. Results and Discussions

Here we use the SAW-MBIR algorithm to improve upon the standard MBIR implementation in [15], [16] applied to cardiac CT. This approach uses a preconditioned gradient-based IR algorithm to simultaneously update all the voxels. Further, the ordered subset (OS) method [17] is used to calculate the sub-gradient for each subset at each iteration, then a preconditioner operator is used to accelerate the high frequency convergence, and the Nesterov's method [18] further reduces the number of iterations to achieve convergence. The method included a line-search step that ensures the monotonic decrease of the cost function relative to the previous estimate.

It is worth noting that normally making the statistical weighting dependent on location in the image volume would result in an inconsistent problem formulation for iterative convergence. Using a line search, however, mitigates the concern of inconsistent weights, and in turn guarantees convergence and stability. The convergence path would be influenced by both the gradient of the original cost function 3 and the pseudo-gradient 11. Intuitively, a fixed point of the iteration is reached when the two gradient vectors become orthogonal or have a negative inner product. We leave further theoretical analysis of the properties of the convergence point to future studies, so that this paper would focus on achieving the intended image quality benefits for wide-cone cardiac CT.

Clinical datasets from GE Healthcare Revolution CT scanner with 160 mm detector aperture at the isocenter are used here, with a single axial rotation covering the whole heart.

Comparisons between full-scan and half-scan MBIR as well as the proposed SAW-MBIR algorithm are shown in Figs. 2 and 3.

Fig. 2 compares the performance of SAW-MBIR against full-scan and half-scan MBIR for wide-cone cardiac CT reconstruction in the chest (Figs. 2a-c), and in the liver (Figs. 2d-f) regions. Figs. 2a-c correspond to a slice between cross sections c and d (closer to d at the edge) in Fig. 1, and Figs. 2d-f belong to a same region on the opposite side of the reconstruction volume.

The full-scan and half-scan MBIR results are shown in Figs. 2a, and d, and Figs. 2b, and e, respectively. Figs. 2c, and f belong to SAW-MBIR results. SAW-MBIR perform consistently as good as full-scan MBIR outside the heart region. Half-scan MBIR, on the other hand, shows some distortions in the regions outside the heart due to incomplete and missing projection data.

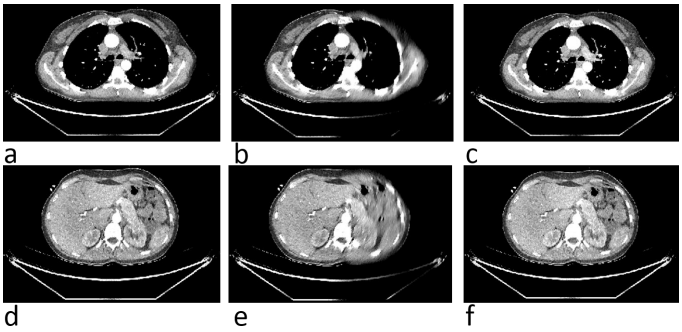


Fig. 2. Comparison between reconstructed cardiac CT images in the chest and liver regions. a, and d. Full-scan MBIR. b, and e. Half-scan MBIR. c, and f. SAW-MBIR. Window settings show the [-200 200] HU range. SAW-MBIR shows consistent image quality with full-scan outside the heart where half-scan MBIR is distorted. Panels a-c correspond to a slice between cross sections c and d (closer to d at the edge) in Fig. 1, and panels d-f are from a slice on the opposite side in the reconstructed volume.

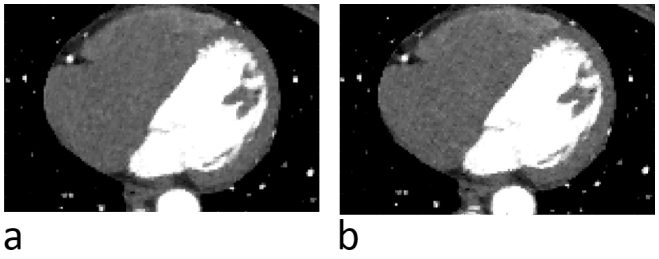


Fig. 3. Impact on temporal resolution. An expanded view of the heart region in the reconstructed center slice of the wide-cone acquisition is shown in (a) half-scan MBIR, and (b) SAW-MBIR. Window settings show the [-200 500] HU range.

We also compared reconstruction results specifically in the heart region (center slices) between half-scan, and SAW-MBIR in Figs. 3a, and b. Qualitative comparisons show close agreement between SAW-MBIR and half-scan results.

Further, the root mean squared error (RMSE) between full-scan MBIR and half-scan MBIR, as well as between full-scan MBIR and SAW-MBIR at each slice location is computed. The results are shown in Fig. 4. It is clear that in the non-heart regions (the edge slices) SAW-MBIR matches better with full-scan MBIR without the distortions of half-scan MBIR. The

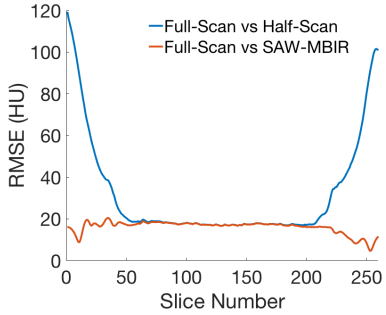


Fig. 4. RMSE at each slice. The RMSE of half-scan MBIR (blue) and SAW-MBIR (red) with respect to full-scan MBIR is illustrated. Background pixels with zero intensity were excluded.

results also suggest very good agreement between the half-scan and SAW-MBIR results in the heart region, which further verify that the SAW-MBIR maintains comparable temporal resolution as the half-scan MBIR.

It is worth noting that, to reduce half-scan artifacts while maintaining temporal resolution, Cho in [19] proposed akin approach where they use extra measurements (such as full-scan data) to modify the statistical weighting using an extrapolation scheme. However, the experimental results on numerical phantoms proved challenging to tune parameters of their model to obtain optimum solution with temporal resolution of half-scan and reduced artifacts as good as full-scan results.

IV. Conclusion

In this work, we developed the SAW-MBIR algorithm that uses spatially-adaptive weights to perform selective back projection of sinogram residuals to different sub-regions or voxels in the reconstructed image. Back projection weights may be determined based on both the measurement position and the location of the reconstructed voxel in the field of view. We examined the performance of the SAW-MBIR using whole-heart cardiac CT clinical data sets with temporal heart motion. The experimental results obtained using the SAW-MBIR demonstrate marked performance in achieving high temporal resolution in the heart region with similar image quality to standard full-scan MBIR outside the heart region. While cardiac CT is shown as an example, the method can be extended to other scan geometries or imaging modalities wherein image artifacts or image degradations may be spatially localized, for instance with scatter, low signal, or metal artifacts.

References

[1] Johan Nuyts, Bruno De Man, Jeffrey A Fessler, Wojciech Zbijewski, and Freek J Beekman, "Modelling the physics in the iterative reconstruction for transmission computed tomography.," *Physics in Medicine and Biology*, vol. 58, no. 12, pp. R63–96, 2013.

[2] Jiang Hsieh, Brian Nett, Zhou Yu, Ken D. Sauer, Jean-Baptiste Thibault, and Charles A. Bouman, "Recent advances in CT image reconstruction," *Current Radiology Reports*, vol. 1, pp. 39–51, 2013.

[3] Jed D Pack, Zhye Yin, Kai Zeng, and Brian E Nett, "Mitigating cone-beam artifacts in short-scan CT imaging for large cone-angle scans," in *The 12th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, 2013, pp. 300–303.

[4] Jie Tang, Jiang Hsieh, and Guang-Hong Chen, "Temporal resolution improvement in cardiac CT using PICCS (TRI-PICCS): performance studies.," *Medical physics*, vol. 37, no. 8, pp. 4377–4388, 2010.

[5] Clemens Maaß, Frank Dennerlein, Frédéric Noo, and Marc Kachelrieß, "Comparing short scan ct reconstruction algorithms regarding cone-beam artifact performance," in *Nuclear Science Symposium Conference Record (NSS/MIC), 2010 IEEE*. IEEE, 2010, pp. 2188–2193.

[6] Jean-Baptiste Thibault, Ken D. Sauer, Charles A Bouman, and Jiang Hsieh, "A three-dimensional statistical approach to improved image quality for multislice helical CT," *Med Phys*, vol. 34, no. 11, pp. 4526–4544, 2007.

[7] Zhou Yu, Jean-Baptiste Thibault, Charles A. Bouman, Ken D. Sauer, and Jiang Hsieh, "Fast model-based X-ray CT reconstruction using spatially nonhomogeneous ICD optimization," *IEEE Transactions on Image Processing*, vol. 20, no. 1, pp. 161–175, 2011.

[8] Dong Hye Ye, Somesh Srivastava, Jean-Baptiste Thibault, Ken D. Sauer, and Charles A Bouman, "Roi reconstruction for model based iterative reconstruction (mbir) via a coupled dictionary learning," in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2017, pp. 101322R–101322R.

[9] Tobias A Fuchs, Julia Stehli, Sacha Bull, Svetlana Dougoud, Olivier F Clerc, Bernhard A Herzog, Ronny R Buechel, Oliver Gaemperli, and Philipp A Kaufmann, "Coronary computed tomography angiography with model-based iterative reconstruction using a radiation exposure similar to chest x-ray examination," *European heart journal*, vol. 35, no. 17, pp. 1131–1136, 2014.

[10] Stefan Wesarg, Matthias Ebert, and Thomas Bortfeld, "Parker weights revisited," *Medical physics*, vol. 29, no. 3, pp. 372–378, 2002.

[11] Chris Kamphuis, Freek J. Beekman, Peter P. Van Rijk, and Max A. Viergever, "Dual matrix ordered subsets reconstruction for accelerated 3D scatter compensation in single-photon emission tomography," *European Journal of Nuclear Medicine*, vol. 25, no. 1, pp. 8–18, 1998.

[12] G. L. Zeng and G. T. Gullberg, "Frequency domain implementation of the three-dimensional geometric point response correction in spect imaging," *IEEE Transactions on Nuclear Science*, vol. 39, no. 5, pp. 1444–1453, 1992.

[13] Adam S. Wang, J. Webster Stayman, Yoshito Otake, Sebastian Vogt, Gerhard Kleinszig, and Jeffrey H. Siewerdsen, "Accelerated statistical reconstruction for C-arm cone-beam CT using Nesterov's method," *Medical Physics*, vol. 42, no. 5, pp. 2699–2708, 2015.

[14] Hanming Zhang, Lei Li, Linyuan Wang, Yanmin Sun, Bin Yan, Ailong Cai, and Guoen Hu, "Computed Tomography Sinogram Inpainting With Compound Prior Modelling Both Sinogram and Image Sparsity," *IEEE Transactions on Nuclear Science*, vol. 63, no. 5, pp. 2567–2576, 2016.

[15] Lin Fu, Zhou Yu, and Jean-Baptiste Thibault, "Space-variant channelized preconditioner design for 3D iterative CT reconstruction," in *The 12th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, 2013, pp. 205–208.

[16] Zhou Yu, Lin Fu, Debashish Pal, Jean-Baptiste Thibault, Charles A. Bouman, and Ken D. Sauer, "Nested loop algorithm for parallel model based iterative reconstruction," in *The 12th International meeting on Fully 3D Image Reconstruction in Radiology and Nuclear Medicine*, 2013, pp. 197–210.

[17] H Malcolm Hudson and Richard S Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE transactions on medical imaging*, vol. 13, no. 4, pp. 601–609, 1994.

[18] Yurii Nesterov, "How to make the gradients small," *Optima*, vol. 88, pp. 10–11, 2012.

[19] Jang Hwan Cho, *Improving statistical image reconstruction for cardiac X-ray computed tomography*, Ph.D. thesis, University of Michigan, 2014.

A Memory-efficient Algorithm for Large-scale Sparsity Regularized Image Reconstruction

Greg Ongie, Naveen Murthy, Laura Balzano, Jeffrey A. Fessler

***Abstract*—We derive a memory-efficient first-order variable splitting algorithm for convex image reconstruction problems with non-smooth regularization terms. The algorithm is based on a primal-dual approach, where one of the dual variables is updated using a step of the Frank-Wolfe algorithm, rather than the typical proximal point step used in other primal-dual algorithms. We show in certain cases this results in an algorithm with far less memory demand than other first-order methods based on proximal mappings. We demonstrate the algorithm on the problem of sparse-view X-ray computed tomography (CT) reconstruction with non-smooth edge-preserving regularization and show competitive run-time with other state-of-the-art algorithms while using much less memory.**

***Index Terms*—primal-dual algorithm, Frank-Wolfe algorithm, image reconstruction, sparse-view CT**

I. Introduction

Sparsity regularized inverse problems arising in medical imaging result in large-scale non-smooth convex optimization problems that are computationally challenging to solve. General purpose first-order algorithms for nonsmooth convex optimization, such as subgradient descent or smoothing techniques [7], while memory-efficient, converge undesirably slow for this class of problem. Instead, specialized proximal splitting algorithms, such as the primal-dual algorithm of Chambolle-Pock (PDCP) [2], represent the current state-of-the-art for these problems. However, PDCP and other proximal methods require storing and operating on one or more dual variables with dimensions potentially several times larger than the image volume to be reconstructed. Even for realistic problem sizes arising in X-ray CT reconstruction, storing and operating on these additional dual variables may be prohibitive or infeasible, limiting the scope of these algorithms in practice. This is especially true when using 3D regularization with all 26 neighboring voxels [11], or when computing on GPUs that are limited in memory relative to traditional CPUs.

To address this issue, this paper introduces a novel algorithm that has far less memory demand than previous approaches. Specifically, we focus on reconstruction via weighted least squares with a non-smooth edge-preserving regularization term. This includes the total variation semi-norm and related penalties. The proposed algorithm is based on a novel primal-dual approach. Existing first-order primal-dual approaches [2], [3] alternate between updating the primal variable with a gradient descent step and the dual variable with a projected gradient ascent step (or their proximal equivalents). The main idea of the proposed approach is to replace the projected gradient ascent step in the dual update with a step of the Frank-Wolfe algorithm [4], [6]. We show that this modification allows for substantial memory savings over standard primal-dual approaches. In particular, the algorithm requires storing at most two additional auxiliary variables with dimensions matching the primal variable. We prove convergence of the algorithm under certain assumptions on its step-sizes.

Finally, we demonstrate the proposed algorithm by reconstructing a sparse-view X-ray CT dataset. Empirically, the proposed algorithm shows competitive convergence with state-of-the-art proximal splitting methods for this problem [10], [8], but with much less memory demand.

G. Ongie, N. Murthy, L. Balzano, and J.A. Fessler are with University of Michigan, Department of Electrical Engineering and Computer Science, Ann Arbor, MI, USA. E-mail: {gongie,nnmurthy,girasole,fessler}@umich.edu. This work was supported in part by NIH Grant U01 EB018753, DARPA-16-43-D3M-FP-037, and NSF ECCS-1508943.

II. Problem Formulation

We consider the following optimization problem:

$$\min_{\boldsymbol{x}} \tfrac{1}{2}\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_{\boldsymbol{W}}^2 + \lambda\, R(\boldsymbol{x}). \tag{1}$$

Here $\boldsymbol{x} \in \mathbb{R}^n$ represents a vectorized discrete image. The first term in (1) measures the data-fit, where $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ with $m \le n$ is a matrix representing the linear measurement operator, $\boldsymbol{b} \in \mathbb{R}^m$ are the (noisy) measurements, and $\|\cdot\|_{\boldsymbol{W}}$ is the weighted ℓ^2-norm defined as $\|\boldsymbol{z}\|_{\boldsymbol{W}} = \sqrt{\boldsymbol{z}^T\boldsymbol{W}\boldsymbol{z}}$ for a fixed diagonal matrix $\boldsymbol{W} \in \mathbb{R}^{m\times m}$ with positive diagonal entries. The second term is a regularization penalty, where $\lambda > 0$ is a parameter balancing the tightness of data-fit and regularization. This work focuses on regularizers of the form $R(\boldsymbol{x}) = \varphi(\boldsymbol{D}\boldsymbol{x})$ for some *regularization transform* $\boldsymbol{D} \in \mathbb{R}^{N\times n}$ and where $\varphi(\boldsymbol{y}) := \sum_i \phi(\boldsymbol{y}_i)$ for some convex and possibly non-smooth sparsity promoting *potential function* $\phi : \mathbb{R} \to \mathbb{R}^+$. We assume the regularization transform $\boldsymbol{D}$ is a tall matrix ($N \gg n$) having block form $\boldsymbol{D} = [\boldsymbol{D}_1^T\ \boldsymbol{D}_2^T\ \cdots\ \boldsymbol{D}_\ell^T]^T$ with $\boldsymbol{D}_i \in \mathbb{R}^{N_i\times n}$. For example, if $\{\boldsymbol{D}_i\}_{i=1}^\ell$ is a collection of first-order finite-difference operators in each dimension and $\phi(x) = |x|$ then $\varphi(\boldsymbol{D}\boldsymbol{x}) = \|\boldsymbol{D}\boldsymbol{x}\|_1$ is the discrete (anisotropic) total variation (TV) of $\boldsymbol{x}$. Other choices for $\{\boldsymbol{D}_i\}_{i=1}^\ell$ include oriented higher-order finite differences [5] or a collection of pre-trained sparsifying transforms [12]. Likewise, the proposed method also generalizes to other convex potential functions ϕ, such as the Huber loss or Fair potential.

III. Primal-Dual Frank-Wolfe Algorithm

For ease of exposition we focus on the case $\varphi(\boldsymbol{y}) = \|\boldsymbol{y}\|_1$ in the remainder of this work, *i.e.*, $R(\boldsymbol{x}) = \|\boldsymbol{D}\boldsymbol{x}\|_1$. Because the ℓ^∞-norm is dual to the ℓ^1-norm, the primal problem (1)

with an ℓ^1 regularizer has an equivalent primal-dual saddle point formulation[1] given by

$$\min_{\boldsymbol{x}} \max_{\|\boldsymbol{y}\|_\infty \leq \lambda} \tfrac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_{\boldsymbol{W}}^2 + \langle \boldsymbol{D}\boldsymbol{x}, \boldsymbol{y}\rangle \tag{2}$$

Partitioning the dual vector as $\boldsymbol{y}^T = [\boldsymbol{y}_1^T \ \cdots \ \boldsymbol{y}_\ell^T]$ corresponding to the blocks $\boldsymbol{D}^T = [\boldsymbol{D}_1^T \ \cdots \ \boldsymbol{D}_\ell^T]$, we re-express the inner product in (2) as

$$\langle \boldsymbol{D}\boldsymbol{x}, \boldsymbol{y}\rangle = \sum_{i=1}^{\ell} \langle \boldsymbol{D}_i\boldsymbol{x}, \boldsymbol{y}_i\rangle = \left\langle \boldsymbol{x}, \sum_{i=1}^{\ell} \boldsymbol{D}_i^T\boldsymbol{y}_i \right\rangle.$$

Therefore, introducing the auxiliary variable $\boldsymbol{z} \triangleq \sum_{i=1}^{\ell} \boldsymbol{D}_i^T\boldsymbol{y}_i$, (2) is equivalent to the equality constrained problem:

$$\min_{\boldsymbol{x}} \max_{\substack{\|\boldsymbol{y}_i\|_\infty \leq \lambda, \\ i=1,\ldots,\ell}} \tfrac{1}{2}\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_{\boldsymbol{W}}^2 + \langle \boldsymbol{x}, \boldsymbol{z}\rangle \ \text{ s.t. } \boldsymbol{z} = \sum_{i=1}^{\ell} \boldsymbol{D}_i^T\boldsymbol{y}_i. \tag{3}$$

Below we show that the proposed algorithm only needs to maintain an estimate of the auxiliary variable $\boldsymbol{z} \in \mathbb{R}^n$, which has dimensions of the image volume, rather than the full dual variable $\boldsymbol{y} \in \mathbb{R}^N$ that is generally several times larger.

We also dualize the data-fit term by defining $g(\boldsymbol{t}) := \frac{1}{2}\|\boldsymbol{t} - \boldsymbol{b}\|_{\boldsymbol{W}}^2$ and using the identity

$$g(\boldsymbol{A}\boldsymbol{x}) = \max_{\boldsymbol{t}} \ \langle \boldsymbol{t}, \boldsymbol{A}\boldsymbol{x}\rangle - g^*(\boldsymbol{t}),$$

where g^* is the convex conjugate of g. Simple analysis yields $g^*(\boldsymbol{t}) = \frac{1}{2}\|\boldsymbol{t}+\boldsymbol{W}\boldsymbol{b}\|_{\boldsymbol{W}^{-1}}^2 - \frac{1}{2}\|\boldsymbol{b}\|_{\boldsymbol{W}}^2$. Inserting this into (3) and dropping constant terms yields the equivalent formulation

$$\min_{\boldsymbol{x}} \max_{\boldsymbol{t}} \max_{\substack{\|\boldsymbol{y}_i\|_\infty \leq \lambda, \\ i=1,\ldots,\ell}} \langle \boldsymbol{t}, \boldsymbol{A}\boldsymbol{x}\rangle - \frac{1}{2}\|\boldsymbol{t} + \boldsymbol{W}\boldsymbol{b}\|_{\boldsymbol{W}^{-1}}^2 + \langle \boldsymbol{x}, \boldsymbol{z}\rangle \tag{4}$$

subject to $\boldsymbol{z} = \sum_{i=1}^{\ell} \boldsymbol{D}_i^T\boldsymbol{y}_i$.

A. Frank-Wolfe dual update

The Frank-Wolfe (FW) algorithm [4], [6], also known as the conditional gradient method, is a projection-free approach to solving constrained problems of the form

$$\max_{\boldsymbol{y} \in \mathcal{C}} f(\boldsymbol{y}),$$

where f is a concave function and $\mathcal{C}$ is a closed, convex set. At each iteration, the FW algorithm solves for a search-direction $\boldsymbol{s}^\star$ via

$$\boldsymbol{s}^\star = \arg\max_{\boldsymbol{s} \in \mathcal{C}} \ \langle \boldsymbol{s}, \nabla f(\boldsymbol{y}^{(k)})\rangle$$

then updates $\boldsymbol{y}$ with a convex combination of the previous iterate $\boldsymbol{y}^{(k)}$ and the search direction

$$\boldsymbol{y}^{(k+1)} = (1 - \alpha_k)\boldsymbol{y}^{(k)} + \alpha_k \boldsymbol{s}^\star$$

where α_k is some iteration-dependent step-size.

If we apply one step of the FW algorithm to the dual variable $\boldsymbol{y}$ in (2) while holding the primal variable $\boldsymbol{x}$ fixed, then the function to maximize is simply the linear function $f(\boldsymbol{y}) = \langle \boldsymbol{D}\boldsymbol{x}, \boldsymbol{y}\rangle$ with $\nabla f(\boldsymbol{y}) = \boldsymbol{D}\boldsymbol{x}$ subject to $\|\boldsymbol{y}\|_\infty \leq \lambda$. The FW search-direction update in this case is

$$\boldsymbol{s}^\star = \arg\max_{\|\boldsymbol{s}\|_\infty \leq \lambda} \langle \boldsymbol{s}, \boldsymbol{D}\boldsymbol{x}\rangle = \lambda \operatorname{sign}(\boldsymbol{D}\boldsymbol{x}).$$

where $\operatorname{sign}(\cdot)$ is applied entrywise and we define $\operatorname{sign}(0) = 0$. Hence, a FW update of $\boldsymbol{y}$ has the form

$$\boldsymbol{y}^{(k+1)} = (1 - \alpha_k)\boldsymbol{y}^{(k)} + \alpha_k \, \lambda \operatorname{sign}(\boldsymbol{D}\boldsymbol{x}).$$

A key to saving memory is that the above update is separable in terms of the $\boldsymbol{y}_i$-blocks:

$$\boldsymbol{y}_i^{(k+1)} = (1 - \alpha_k)\boldsymbol{y}_i^{(k)} + \alpha_k \, \lambda \operatorname{sign}(\boldsymbol{D}_i\boldsymbol{x}), \ \text{for all } i = 1, ..., \ell.$$

Applying $\boldsymbol{D}_i^T$ to both sides above and summing over i yields

$$\boldsymbol{z}^{(k+1)} = (1 - \alpha_k)\boldsymbol{z}^{(k)} + \alpha_k \lambda \sum_{i=1}^{\ell} \boldsymbol{D}_i^T \operatorname{sign}(\boldsymbol{D}_i x) \tag{5}$$

where we define $\boldsymbol{z}^{(k)} = \sum_{i=1}^{\ell} \boldsymbol{D}_i^T\boldsymbol{y}_i^{(k)}$ for all $k \geq 0$. To save memory, we compute $\boldsymbol{z}^{(k+1)}$ incrementally, first by rescaling the current estimate by $(1 - \alpha_k)$ then by adding $\alpha_k \lambda \boldsymbol{D}_i^T \operatorname{sign}(\boldsymbol{D}_i x)$ for all $i = 1, ..., \ell$ in sequence.

B. Proximal dual update

Similar to other primal-dual approaches [2], [3], to update the dual variable $\boldsymbol{t}$ we take one step of a proximal point algorithm applied to (4) while fixing the other variables. Specifically, given the current iterates $(\boldsymbol{x}^{(k)}, \boldsymbol{t}^{(k)})$, we set

$$\boldsymbol{t}^{(k+1)} = \arg\max_{\boldsymbol{t}} \langle \boldsymbol{t}, \boldsymbol{A}\boldsymbol{x}^{(k)}\rangle - \tfrac{1}{2}\|\boldsymbol{t}+\boldsymbol{W}\boldsymbol{b}\|_{\boldsymbol{W}^{-1}}^2 - \tfrac{1}{2\sigma_k}\|\boldsymbol{t}-\boldsymbol{t}^{(k)}\|_{\boldsymbol{W}^{-1}}^2$$

where $\sigma_k > 0$ is a step-size parameter to be specified later. This has the closed form solution

$$\boldsymbol{t}^{(k+1)} = \tfrac{1}{1+\sigma_k}\boldsymbol{t}^{(k)} + \tfrac{\sigma_k}{1+\sigma_k}\boldsymbol{W}(\boldsymbol{A}\boldsymbol{x}^{(k)} - \boldsymbol{b}). \tag{6}$$

C. Primal update

Finally, we update the primal variable $\boldsymbol{x}$ via a gradient descent step (or equivalently a proximal-point step) applied to (4) with the dual variables fixed:

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \tau_k(\boldsymbol{z}^{(k+1)} + \boldsymbol{A}^T\boldsymbol{t}^{(k+1)}) \tag{7}$$

where $\tau_k > 0$ is a step-size parameter to be specified later. Inspired by [2] we include an optional over-relaxation step:

$$\overline{\boldsymbol{x}}^{(k+1)} = \boldsymbol{x}^{k+1} + \theta(\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}), \tag{8}$$

where $\theta \in [0, 1]$, and perform the dual variable updates (5) and (6) with $\overline{\boldsymbol{x}}^{(k)}$ in place of $\boldsymbol{x}^{(k)}$.

D. Algorithm summary and convergence

Algorithm 1 summarizes the proposed primal-dual Frank-Wolfe (PDFW) algorithm. Using similar analysis as in [1] we are able to prove the following convergence result for Algorithm 1 in the special case $\theta = 0$ by showing it is a particular instance of an ϵ-subgradient descent method; we omit the proof for brevity.

Theorem 1. *Let $\mathcal{X}^*$ denote the set of minimizers to* (1)*, and let $\{\boldsymbol{x}^k\}_{k=1}^{\infty}$ be the iterates generated by Algorithm 1 with $\theta = 0$.*

[1] For a general regularization penalty of the form $\varphi(\boldsymbol{D}\boldsymbol{x})$, we can derive a similar saddle-point formulation (2) by writing $\varphi(\boldsymbol{D}\boldsymbol{x}) = \max_{\boldsymbol{y}} \langle \boldsymbol{D}\boldsymbol{x}, \boldsymbol{y}\rangle - \varphi^*(\boldsymbol{y})$ where φ^* is the convex conjugate of φ.

Algorithm 1 Primal-Dual Frank-Wolfe (PDFW)

Initialize $\boldsymbol{x}^{(0)} = \overline{\boldsymbol{x}}^{(0)}, \boldsymbol{z}^{(0)} = 0 \in \mathbb{R}^n$ and $\boldsymbol{t}^{(0)} = 0 \in \mathbb{R}^m$.
Choose step-sizes $(\tau_k, \sigma_k, \alpha_k)$, $\theta \in [0, 1]$.
for $k = 0, ..., k_{max}$ **do**
$\boldsymbol{t}^{(k+1)} = \frac{1}{1+\sigma_k}\boldsymbol{t}^{(k)} + \frac{\sigma_k}{1+\sigma_k}\boldsymbol{W}(\boldsymbol{A}\overline{\boldsymbol{x}}^{(k)} - \boldsymbol{b})$
$\boldsymbol{z}^{(k+1)} = (1-\alpha_k)\boldsymbol{z}^{(k)} + \alpha_k\lambda\sum_{i=1}^{\ell}\boldsymbol{D}_i^T\text{sign}(\boldsymbol{D}_i\overline{\boldsymbol{x}}^{(k)})$
$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} - \tau_k(\boldsymbol{A}^T\boldsymbol{t}^{(k+1)} + \boldsymbol{z}^{(k+1)})$
$\overline{\boldsymbol{x}}^{(k+1)} = \boldsymbol{x}^{(k+1)} + \theta(\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)})$
end for

Suppose the iterates $\{\boldsymbol{x}^k\}_{k=1}^{\infty}$ are bounded. If the step-size sequences $\{\alpha_k\} \subset [0, 1]$, $\{\sigma_k\} \subset (0, \infty)$, and $\{\tau_k\} \subset (0, \infty)$ satisfy $\tau_k \to 0$, $\sum_{k=0}^{\infty}\tau_k = +\infty$, and

$$\sum_{j=1}^{k}\tau_{j-1}\prod_{i=j}^{k}(1-\alpha_i) \to 0, \quad \sum_{j=1}^{k}\tau_{j-1}\prod_{i=j}^{k}\frac{1}{1+\sigma_k} \to 0$$

as $k \to \infty$, then $\mathsf{dist}(\boldsymbol{x}^k, \mathcal{X}^*) \to 0$, where dist denotes the Euclidean distance of a point to a set. In particular, if the solution $\boldsymbol{x}^*$ to (1) is unique then $\boldsymbol{x}^k \to \boldsymbol{x}^*$.

The step-size conditions in Theorem 1 are satisfied, for example, when $\tau_k = O(\frac{1}{k^p})$, $0 < p \leq 1$, α_k and σ_k are constant. There are also valid choices of $\tau_k, \alpha_k, \sigma_k$ for which $\alpha_k \to 0$ and $\sigma_k \to \infty$, such as $\tau_k = O(\frac{1}{k^p})$, $\alpha_k = O(\frac{1}{k^q})$, and $\sigma_k = O(\frac{1}{\tau_k})$ with $0 < p < 1$ and $0 < q < p/2$.

Empirically, we observe improved convergence rates using $\theta = 1$ and a constant step-size $\tau_k = \tau$. However, our current proof of Theorem 1 does not extend to the case $\theta \neq 0$ nor to the case of τ_k constant, and we leave its convergence under these conditions as an open problem for future work.

E. Connections to Chambolle-Pock primal dual algorithm

Algorithm 1 is closely related to the primal-dual algorithm of Chambolle-Pock (PDCP) [2]. If we introduce an auxiliary variable $\boldsymbol{s}^{(k)} \in \mathbb{R}^N$ and replace the $\boldsymbol{z}^{(k+1)}$ update in Algorithm 1 with the alternative update

$$\boldsymbol{z}^{(k+1)} = \boldsymbol{D}^T\boldsymbol{s}^{(k+1)} := \boldsymbol{D}^T\mathsf{proj}_{\|\cdot\|_\infty \leq \lambda}(\boldsymbol{s}^{(k)} + \sigma_k\boldsymbol{D}\overline{\boldsymbol{x}}^{(k)})$$

where $\mathsf{proj}_{\|\cdot\|_\infty \leq \lambda}$ denotes Euclidean projection onto the set $\{\boldsymbol{s} : \|\boldsymbol{s}\|_\infty \leq \lambda\}$, then this modified version of Algorithm 1 coincides with PDCP applied to (4). In [2] it is shown that PDCP converges when $\sigma_k = \sigma$ and $\tau_k = \tau$ are constant and $\tau\sigma L^2 < 1$ and $\theta = 1$, where L is the operator norm of the concatenated matrix $[\boldsymbol{A}^T, \boldsymbol{D}^T]^T$.

F. Memory benefits

Table 1 summarizes the memory requirements of different first-order proximal methods for solving (2). An important feature of Algorithm 1 is that it only requires storing at most three arrays having the size of the image volume to be reconstructed. In contrast, the linearized augmented Lagrangian method (LALM) of [8] would need to store several arrays have the same size as the image plus two additional arrays of size N, the output dimension of the regularization transform. Similarly, the PDCP algorithm [2] implemented as in [10] needs to store at least one array of size N. The last column of Table 1 we illustrates the memory demand of these algorithms for the iterative reconstruction of a 3D axial CT scan as specified in the next section. The proposed PDFW algorithm requires an order of magnitude less memory for this example because it avoids having to store large auxiliary variables associated with the regularization transformed image.

	number of variables of size			total memory
	$\boldsymbol{x} \in \mathbb{R}^n$ (image)	$\boldsymbol{Dx} \in \mathbb{R}^N$ (reg. transform)	$\boldsymbol{b} \in \mathbb{R}^m$ (data)	3D CT example (in GB)
LALM [8]	4	2	2	3.02
PDCP [2], [10]	2	1	2	1.60
PDFW, $\theta = 1$	3	0	2	0.47
PDFW, $\theta = 0$	2	0	2	0.38

TABLE I
MEMORY DEMANDS OF FIRST-ORDER METHODS FOR SOLVING (1).

IV. EXPERIMENTS

Here we demonstrate the proposed PDFW algorithm's potential for sparse-view X-ray CT reconstruction. We simulate an axial CT scan of the XCAT phantom [9] of size $1024 \times 1024 \times 154$ voxels to obtain a sinogram of size $m = 888 \times 64 \times 120$ (channels $\times$ rows $\times$ views) and reconstruct on a coarser grid of size $n = 512 \times 512 \times 90$. Our reconstruction is obtained by solving (1) with regularizer $R(\boldsymbol{x}) = \|\boldsymbol{Dx}\|_1$ where $\boldsymbol{D}$ computes all finite-differences with thirteen nearest-neighbors of each voxel. We set the statistical weighting matrix $\boldsymbol{W} = \boldsymbol{I}$, and set the regularization parameter $\lambda = 4096$. We compare against two state-of-the-art first-order algorithms for solving (1): PDCP as adapted to CT reconstruction in [10], and the linearized augmented Lagrangian method [2] (LALM) of [8]. For the proposed PDFW algorithm we test two sets of step-sizes. The first set (S1) is designed to satisfy the conditions of Theorem 1:

$$\tau_k = \tfrac{2}{2+k}, \quad \sigma_k = \tfrac{1}{L^2\tau_k}, \quad \alpha_k = \left(\tfrac{2}{2+k}\right)^{0.49}, \quad \theta = 0. \quad \text{(S1)}$$

The second set (S2) uses a constant step-size $\tau_k = \tau$, violating the conditions of Theorem 1, but matches the settings proposed for the PDCP algorithm in [10] (except for the choice of α_k):

$$\tau_k = 1/L, \quad \sigma_k = 1/L, \quad \alpha_k = \tfrac{2}{2+k}, \quad \theta = 1. \quad \text{(S2)}$$

Figure 1 shows cropped images from the central transaxial plane initial filtered back projection reconstruction, reference solution, and the reconstruction obtained from the proposed PDFW algorithm with settings (S2) after 500 iterations. We obtained a reference solution to the optimization problem by running several thousands of iterations of the LALM algorithm, which reached the smallest cost among the competing algorithms. Observe that there is almost no visual difference between the reference solution and the PDFW solution after 500 iterations.

Figure 2 compares the performance of the algorithms with respect to two convergence metrics: (1) the *normalized cost*

[2] An ordered subsets variant of LALM is also presented in [8]. The proposed PDFW algorithm could also be modified to include ordered subsets updates, but is outside the scope of this work. For fair comparison, we compare against LALM without ordered subsets.

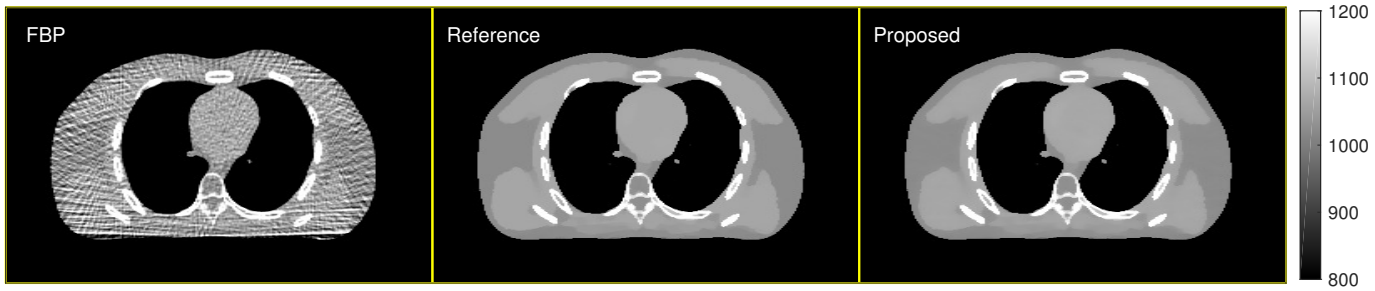


Fig. 1. Reconstructions of central transaxial plane of XCAT phantom from sparse-view measurements (left to right): filtered back projection (FBP) reconstruction, a reference solution, and a reconstruction obtained after running 500 iterations of the proposed PDFW algorithm with settings (S2). Images displayed in HU (modified so that air is 0) clipped to range $[800, 1200]$ and cropped to the region of interest.

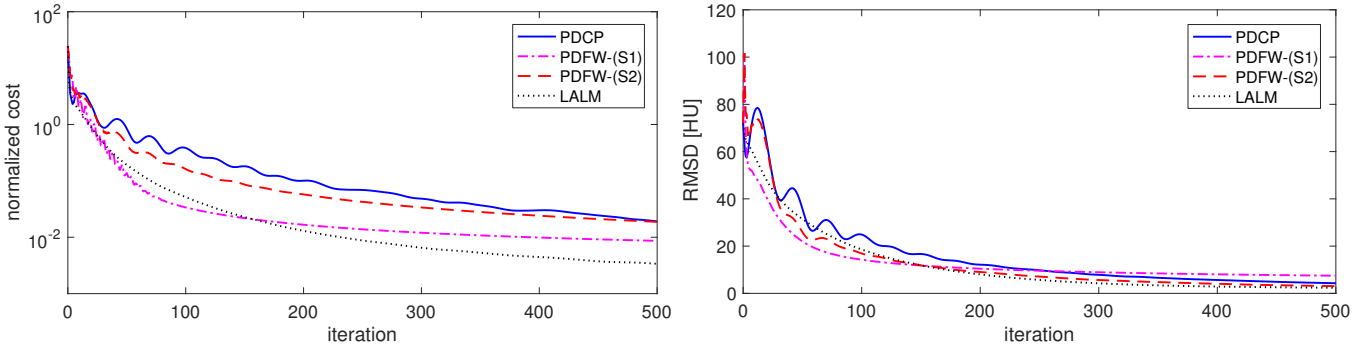


Fig. 2. Plots of convergence metrics using the proposed PDFW algorithm with settings (S1) and (S2), the PDCP algorithm [10], and LALM algorithm [8].

defined as $(f(\boldsymbol{x}^{(k)}) - f(\boldsymbol{x}^*))/f(\boldsymbol{x}^*)$ where $f(\boldsymbol{x})$ is the cost function in (1), $\boldsymbol{x}^{(k)}$ is the kth iteration of a given algorithm, and $\boldsymbol{x}^*$ is the reference solution; and (2) *the root mean square difference* (RMSD) computed as RMSD $= \sqrt{\frac{1}{|\Omega|}\sum_{\boldsymbol{i}\in\Omega} |\boldsymbol{x}_{\boldsymbol{i}}^{(k)} - \boldsymbol{x}_{\boldsymbol{i}}^*|^2}$ where Ω is the index set of voxels in a cylindrical region of interest containing the phantom anatomy. Overall, the LALM algorithm performs best in terms of the convergence metrics, reaching the lowest cost and RMSD after 500 iterations. However, we reiterate that the LALM algorithm has the highest memory demand of the compared methods (see Table I). The proposed PDFW algorithm with step-size scheme (S1) shows a fast initial decrease in the cost and RMSD, but slows in improvement after 100 iterations and has the highest RMSD after 500 iterations, indicating that the (S1) step-size scheme may yield slow asymptotic convergence. The PDFW algorithm with step-size scheme (S2) has better long-run performance in RMSD, yielding nearly the same as LALM after 500 iterations, and its reduction in normalized cost is similar to PDCP.

V. Conclusion

We introduce a memory-efficient algorithm for solving large-scale convex image reconstruction problems with transform sparse regularization based on a novel hybrid of proximal methods and the Frank-Wolfe algorithm. Our experiments demonstrate that the algorithm has competitive performance with other first-order algorithms but with substantially less memory demand. In our experiments we use all of the sinogram measurements to update the primal variable in each iteration, but the proposed algorithm could potentially be modified to incorporate ordered subsets updates similar to [8] for improved computational efficiency and faster convergence.

References

[1] S. Bonettini and V. Ruggiero. On the convergence of primal-dual hybrid gradient algorithms for total variation image restoration. *J. Math. Im. Vision*, 44(3):236–53, Nov. 2012.

[2] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Im. Vision*, 40(1):120–145, 2011.

[3] L. Condat. A primal–dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms. *J. of Opt. Theory and Applications*, 158(2):460–479, Dec 2012.

[4] M. Frank and P. Wolfe. An algorithm for quadratic programming. *Naval research logistics quarterly*, 3(1-2):95–110, 1956.

[5] Y. Hu, G. Ongie, S. Ramani, and M. Jacob. Generalized higher degree total variation regularization. *IEEE TIP*, 23(6):2423–2435, 2014.

[6] M. Jaggi. Revisiting Frank-Wolfe: Projection-free sparse convex optimization. In *ICML*, pages 427–435, 2013.

[7] Y. Nesterov. Smooth minimization of non-smooth functions. *Mathematical Programming*, 103(1):127–52, May 2005.

[8] H. Nien and J. A. Fessler. Fast X-ray CT image reconstruction using a linearized augmented Lagrangian method with ordered subsets. *IEEE TIP*, 34(2):388–99, Feb. 2015.

[9] W. P. Segars, M. Mahesh, T. J. Beck, E. C. Frey, and B. M. W. Tsui. Realistic CT simulation using the 4D XCAT phantom. *Med. Phys.*, 35(8):3800–8, Aug. 2008.

[10] E. Y. Sidky, J. H. Jorgensen, and X. Pan. Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm. *Phys. Med. Biol.*, 57(10):3065–92, 2012.

[11] J.-B. Thibault, K. Sauer, C. Bouman, and J. Hsieh. A three-dimensional statistical approach to improved image quality for multi-slice helical CT. *Med. Phys.*, 34(11):4526–44, Nov. 2007.

[12] B. Wen, S. Ravishankar, and Y. Bresler. Structured overcomplete sparsifying transform learning with convergence guarantees and applications. *Int. J. of Computer Vision*, 114(2-3):137–167, 2015.

First Experimental Validation of a Novel Concept for Dynamic Beam Attenuation in CT

Sascha Manuel Huck[1,2], Katia Parodi[2], Karl Stierstorfer[1]

***Abstract*—It has been a long-standing wish in CT (Computed Tomography) to compensate the emitted X-ray beam profile for the patient's changing attenuation.**

Taking into account the varying conditions, the image quality as well as the delivered patient dose could benefit. The latter is of particular importance as it is claimed by the principle of ALARA (as low as reasonably achievable). State of the art are so called bowtie filters. These filters, however, are *static* and constitute a compromise between different patient anatomies. Thus, bowtie filters cannot cope for changing attenuation profiles. Furthermore, pre-patient bowtie filters are a source of scatter radiation impairing image quality and increasing patient dose.

In this work, a novel concept for a *dynamic* beam attenuator (DBA) is presented. This filter is based on an array of closely spaced, highly absorbing sheets focused towards the focal spot. Controlled defocussing enables a dynamic adaption of the propagated X-ray beam to changing attenuation profiles. The concept of intensity modulation is confirmed in simulations as well as in first-time experiments in a CT scanner. Compared to a conventional bowtie filter, substantial enhancements regarding scatter-to-primary ratio, beam hardening and effective shielding of the patient from scatter radiation are confirmed. Although the presented DBA is a structured object, the first reconstructions of a water phantom are nearly free of ring artifacts.

I. Introduction

In the field of medical diagnostics, CT imaging has developed as a reliable modality for gaining diagnostic information. The price to be paid for full 3D information, however, is an increased dose compared to projectional X-ray imaging. Hence, the goal to reduce the dose has been a major motivation since the early days of CT.

The following common clinical scenario emphasizes the demand for a DBA: In this example, the object under examination is an abdomen approximated as an ellipse. The fan beam width necessary to cover the entire abdomen changes for different gantry positions, see fig. 1. Furthermore, also the depth of the abdomen changes during the gantry rotation such that the propagated X-rays are attenuated more when the tube is positioned laterally and less when the tube is in an anterior or posterior position. The relative contribution of noise compared to the total signal is higher in the center, which is typically the diagnostically relevant region, than in the periphery. Therefore, it is necessary to adapt the propagated intensity pre-patient in order to optimize dose and equalize noise. In lateral direction (see fig. 1 position (i)), a comparably narrow attenuation profile with large variations of the intensity is needed for optimal exposure. In contrast, a comparable broad attenuation profile with rather small intensity variations is required in anterior or posterior direction (see fig. 1 position (ii)). Apart from the angular adaptation various patient circumferences (pediatric or bariatric) and different body areas (abdomen, head, shoulder) require different attenuation profiles.

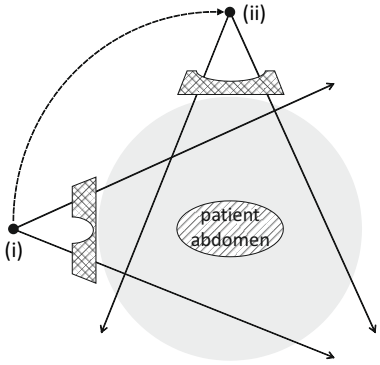


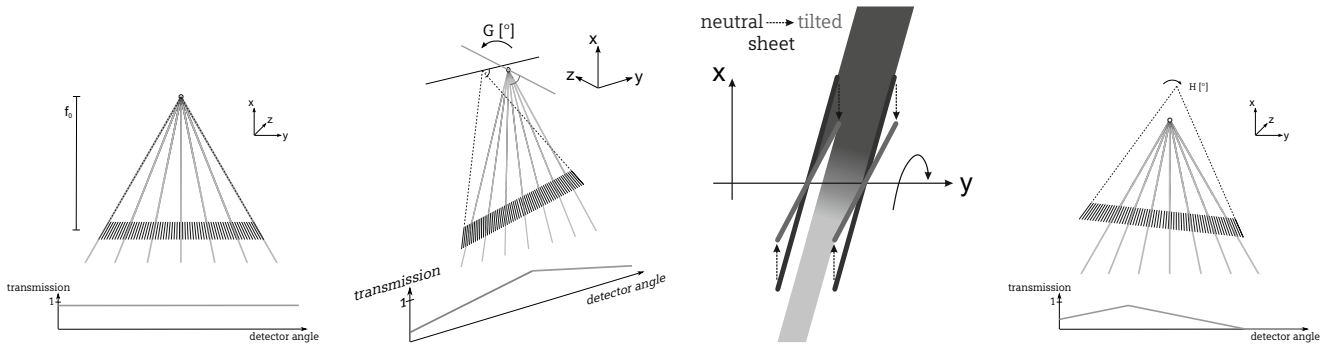
Fig. 1: Required **intensity modulation** (crosshatched) in order to compensate for the attenuation caused by the abdomen (hatched): The X-ray source (black dot) rotates around the abdomen. (i): lateral and (ii): frontal direction.

State of the art are so called bowtie filters, causing increasing attenuation towards the periphery of the fan beam. The attenuation properties are predefined by their shape. Variations are possible only by modulations of the tube voltage or current – which have their technical limitations [1, 7, 9]. Therefore, the present solution cannot cope with the desired, altering attenuation profiles required to optimize dose and noise distribution. Another drawback of the bowtie filter is that it becomes a source of scatter radiation when radiated, which impairs image quality and contributes to patient dose. Different concepts have been proposed in the past trying to overcome the mentioned difficulties [2, 4, 5, 6, 8]. None of these approaches are clinically used.

A novel concept is presented in the following which was elaborated in detail in a previous work [3]: Several advantages over the current bowtie filter are expected from this concept:

- spatial and temporal modulation of the attenuation profile according to the patient's anatomy can:

[1]Siemens Healthineers AG, Forchheim, Germany.
[2]Department of Medical Physics, Faculty of Physics, Ludwig-Maximilians-Universität München, Munich, Germany.
Contact: sascha.huck@siemens-healthineers.com



(a) (b) (c) (d)

Fig. 2: 2a: DBA in NP: The sheets (black) are focused towards the focal spot at distance f_0 covered by the fan beam (grey). The resulting transmission is constantly at almost 1 (normalized). 2b: Defocusing the DBA by a rotation G around the y-axis (perspective view). 2c: Attenuation of a passing X-ray by two non-central sheets (zoomed view). 2d: Lateral translation of the transmission profile by a rotation H around the z-axis.

  - provide a more homogeneous noise level
  - reduce the required dynamical range of the detector
- imaging of a VOI (e.g. the heart), sparing radiosensitive organs (e.g. the lung)
- enhanced image quality
  - negligible beam hardening
  - negligible scatter radiation
- reduced required space within collimator box

In particular, considerable improvements in patient dose and image quality are valuable. In order to confirm these features, a prototype has been designed and tested in simulations as well as in experimental measurements.

Conceptional Idea

The DBA is composed of an array of closely spaced, highly absorbing sheets focused towards the focal spot, see fig. 2. The structure is comparable to an anti-scatter grid (ASG), however, located pre-patient. In neutral position (NP), each sheet is parallel to its adjacent X-rays. Hence, the overall transmission is only slightly affected, see fig. 2a. For common CT scenarios, as described previously, suitable attenuation profiles can be achieved by defocusing the DBA's sheets. For this purpose, the entire DBA is rotated around its y-axis by the attenuation angle G, see fig. 2b. There is no effect on the central X-ray as the adjacent sheets are still parallel. Towards the periphery, however, the sheets are more and more tilted out of their initial plane, see zoomed view in fig. 2c. Hence, non-central sheets are increasingly defocused with respect to the focal spot. The resulting transmission is still at maximum in the center and declines towards the periphery. Depending on the attenuation angle G, the DBA is defocused accordingly realizing the desired attenuation profiles. Independent from the attenuation, the profile can be moved in φ-direction. The basis is a slight rotation of the DBA around its z-axis by the translation angle H. In tilted state, a certain, non-central sheet becomes parallel to its adjacent rays. The transmission profile moves accordingly along the beam width, see fig. 2d. These modifications, which can be deduced from reference anatomy data, build the central function of a DBA.

II. Methods and Tools

Simulation: The geometrical properties of the DBA are elaborated by an internal simulation software of *Siemens Healthcare*, Forchheim. It is based on Monte Carlo simulations taking into account the relevant (scatter) interaction processes. A user selected composition of X-ray source, prefilters, phantoms, ASG and detector can be determined and various parameters be adjusted.

Scan Mode: Unless stated otherwise, the following default parameters are applied: The X-ray tube operates at $120\,\mathrm{kV}$ anode voltage and $300\,\mathrm{mA}$ current. A standard high resolution focus with flying focal spot (FFS) in φ- and z-direction is selected and all data preprocessing options are disabled.

DBA Construction: The DBA is fixed in a *Cardan suspension* which is mounted onto the collimator box. It enables a (manual) adjustment of the DBA around two axes independently and simultaneously. The sheets of the DBA are fastened by two opposing combs. The latter guarantee the exact position and orientation of the sheets.

III. Characterization

Intensity Modulation: In experimental measurements the intensity modulation of the propagated X-ray beam is performed. In this context, the attenuation ratio is defined as the quotient I_0/I, with I_0 the measured intensity without any attenuator and I the one with attenuator. Since the outermost pixels (1 to 120 and 800 to 920) are not covered by the DBA in these measurements, this area is neglected.

The influence of the attenuation angle G on the attenuation profile is demonstrated in fig. 3. At the center of the fan beam the attenuation ratios of all alignments are about $1 : 1.5$ compared to the absence of any attenuator. Hence, the intensity modulation by the attenuation angle G does not affect the intensity in the center. In NP ($G = 0°$), the

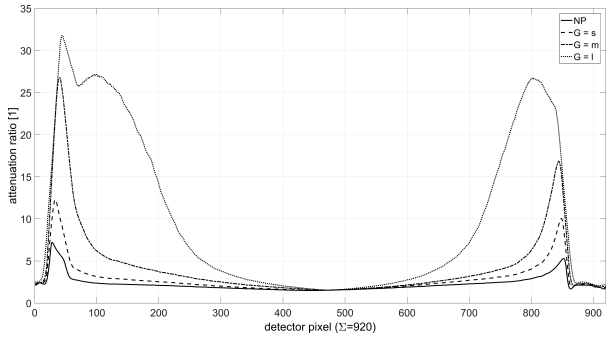


Fig. 3: Various **attenuation angles** G applied to the DBA causing different attenuation ratios.

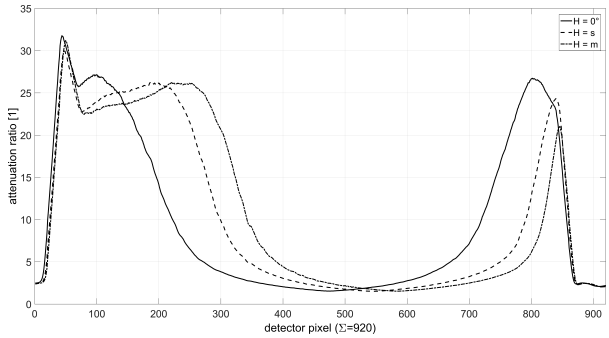


Fig. 4: Different **translation angles** H are applied to the DBA at $G = l = const.$

attenuation ratio is nearly constant. For angles $G \geq 0°$, however, the profiles diverge towards the periphery. The attenuation at a given pixel is the larger, the larger the attenuation angle (G: $s < m < l$) is – up to $1 : 27$ for the largest angle investigated.

As second modification, the lateral translation of the profile is demonstrated for three different translation angles ($H : 0° \leq s \leq m$) at $G = l$, see fig. 4. By increasing the translation angle H, the profile is shifted accordingly in φ-direction without changing the profile's shape considerably.

As a first result of the measurements, the presented concept fulfills the central function of a DBA, providing more dose in the center and less in the periphery.

Beam Hardening: In order to quantify beam hardening, the average energy $\bar{E}_i$ in detector pixel i is simulated in *MOCASSIM* for different attenuation modalities, see fig. 5. As the path length through the bowtie filter (aluminum) increases towards the periphery, the average energy ascends about $12\,\mathrm{keV}$. In contrast, the beam hardening of the DBA is considerably less, at attenuations comparable or even higher. Furthermore, in the region where the average energy $\bar{E}_i$ increases for the DBA, the absolute number of photons (dashed line) is relatively low compared to the center. These simulations confirm negligible beam hardening of the DBA.

Scatter Radiation: In simulations both attenuators are exposed to a narrow beam of 10^4 photons. Figure 6 depicts the traces of secondary photons in the entire CT system – inside and outside the CT funnel. It is apparent

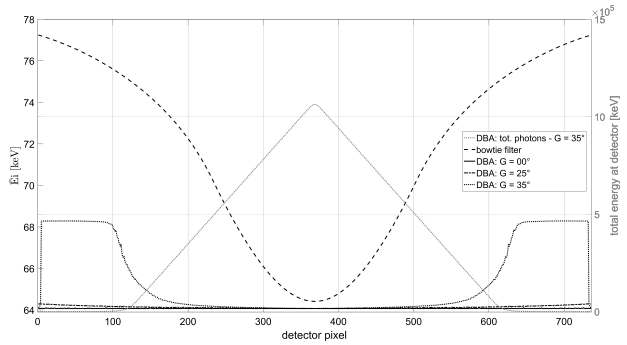


Fig. 5: The **average energy** $\bar{E}_i$ in keV for detector pixel i for different attenuation modalities.

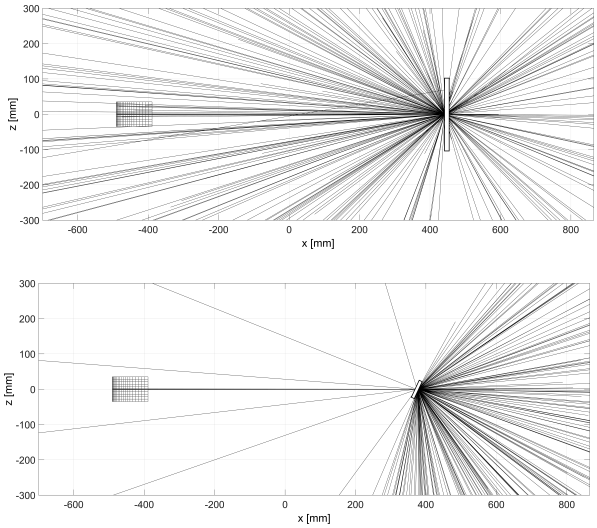


Fig. 6: **Traces** of scattered and K-emission photons exiting the bowtie filter (top) or the DBA (bottom), respectively. The focus is located on the right, the attenuator (white box) at $x = 375\,\mathrm{mm}$ and the detector on the left.

from fig. 6 (top) that photons scattered at the bowtie filter are propagating in all directions. Approximately half of which reach the inside of the CT funnel. This situation changes substantially with the DBA installed instead, fig. 6 (bottom). Scatter radiation and K-emission are also present but nearby exclusively towards the focus. Only very few photons reach the inside of the funnel. Hence, the patient is effectively shielded from scatter radiation which contributes to patient dose and impairs image quality.

Anode Voltage: The impact of the selected anode spectrum on the attenuation is investigated in measurements and quantified by a quantity κ. It is the quotient of the attenuation ratio at $80\,\mathrm{kV}$ anode voltage and the one at $120\,\mathrm{kV}$. In best case, it is $\kappa = 1.0$ along the entire fan beam. A benchmark representing an impairment of $20\,\%$ is drawn at $\kappa = 1.2$, see fig. 7. Considering the bowtie filter, only the central part is within the benchmark. Besides, the quotient rapidly ascends to $\kappa \approx 2.2$. In contrast, the quotient for the DBA is about $\kappa \approx 1.0$ or within the benchmark for a wide range at comparable ($G = 25°$) or even higher ($G = 35°$)

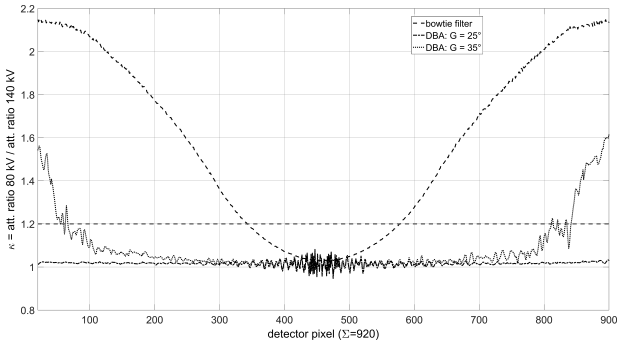


Fig. 7: The **impact of the spectrum** on the attenuation ratio.

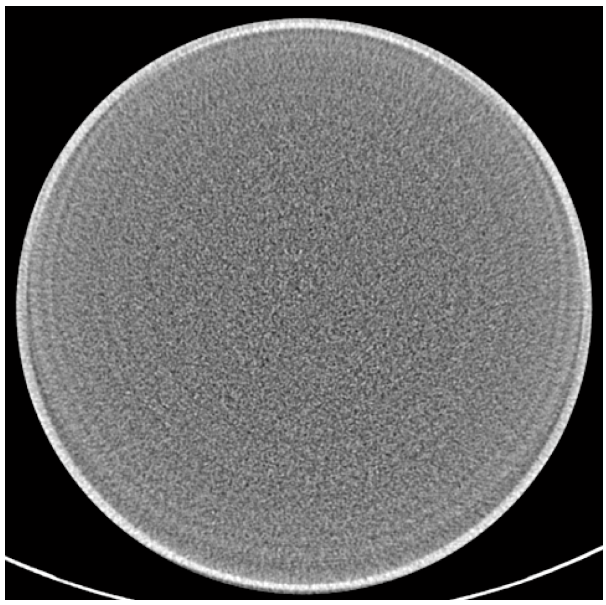


Fig. 8: Water phantom: 2D reconstruction with segmented air calibration (C: 0, W: 400).

attenuation ratio, respectively. These measurements prove that the attenuation of the DBA is hardly affected by the anode spectrum.

Phantom Scan: In addition to the previous air measurements, scans of a $30\,\mathrm{cm}$ water phantom are conducted and reconstructed. The DBA alignment is $G = 35°$ and $H = 0°$ with default scan parameters. A segmented air calibration is taken as reference measurement. The first reconstructed 2D image of the water phantom shows only few ring artifacts, see fig. 8, which could be caused by slight displacements of the sheets during a rotation and the readjustment of the Cardanian suspension between reference and phantom scan.

IV. Conclusion

In the scope of this work, a novel concept of a DBA based on metal sheets has been presented. The concept enabling changing fan beam attenuation profiles was validated in simulations and first-time measurements in a CT scanner. Intensity modulation, the core of a DBA, is required to optimize radiation exposure to different patient circumferences, body parts and gantry positions. Further, it provides a more homogeneous noise level and reduces the required dynamic range of the detector. Future steps might enable the definition of a VOI whereas radiosensitive organs could be spared. Independent of the photon energy, the DBA operates almost as a binary filter as photons either are absorbed within a metal sheet or pass the attenuator unscattered [3]. Moreover, the scatter-to-primary ratio is improved.

A number of advantages over the conventional bowtie filter arise from the DBA: Since only few scattered photons are propagated towards the examination area, the patient is effectively shielded from scatter radiation which causes additional dose and impairs image quality. Furthermore, the spectral properties of the incoming beam remain nearly unchanged when passing the DBA so that beam hardening is negligible. This may be advantageous for applications in spectral CT. A decrease in patient dose is expected and the dynamic range of the detector may also be reduced. The reconstructed water phantom images do not show pronounced ring artifacts, even if the DBA represents a periodical structure within the beam path. Precise calibration and avoidance of mechanical motion, however, are prerequisites for the sensitive setup.

Compared to other DBA concepts, this solution offers a highly advantageous, lean concept motivating further investigations and advancements.

References

[1] Michael Gies et al. “Dose reduction in CT by anatomically adapted tube current modulation. I. Simulation studies”. In: *Med. Phys.* 26.11 (1999), pp. 2235–2247.

[2] Scott S Hsieh and Norbert J Pelc. “The feasibility of a piecewise-linear dynamic bowtie filter.” In: *Med. Phys.* 40.3 (Mar. 2013), p. 031910.

[3] Sascha M Huck. “Dynamic Beam attenuation in CT Experimental Validation of a Novel Concept”. MA thesis. Ludwig-Maximilians-Universität München, 2017.

[4] Fenglin Liu et al. “Dynamic bowtie filter for cone-beam/multi-slice CT”. In: *PloS one* 9.7 (2014), p. 1.

[5] Aswin J Mathews et al. “Design of dual multiple aperture devices for dynamical fluence field modulated CT”. In: *Conference proceedings. International Conference on Image Formation in X-Ray Computed Tomography*. Vol. 2016. NIH Public Access, p. 29.

[6] Jed D Pack and Vance S Robinson. *X-ray filtration*. US Patent App. 14/793,432. July 2015.

[7] Stefan Popescu. “Method for adaptively modulating the power level of an x-ray tube of a computer tomography (CT) system”. Pat. US Patent 5,822,393. Oct. 1998.

[8] Timothy P Szczykutowicz and Charles A Mistretta. “Experimental realization of fluence field modulated CT using digital beam attenuation”. In: *Physics in medicine and biology* 59.5 (2014), p. 1305.

[9] Thomas L Toth. “Variable dose application by modulation of x-ray tube current during CT scanning”. Pat. US Patent 5,379,333. Jan. 1995.

Detectability Indices in Anisotropic X-ray Dark-Field Tomography

Theodor Cheslerean Boghiu*, Yash Sharma†*, Franz Pfeiffer†‡, and Tobias Lasser*

***Abstract*—Anisotropic X-ray Dark-field Tomography (AXDT) is a novel imaging modality aimed at the reconstruction of spherical scattering functions in every three-dimensional volume element, based on the directional X-ray dark-field contrast as measured by an X-ray grating interferometer. Using a spherical harmonics discretization, the reconstruction problem boils down to a column-block linear inverse problem. In this work, we derive a detectability index for the AXDT forward model and use it to assess several acquisition schemes. The results show that the detectability index derived here is a vital tool to design and optimize task-specific acquisition schemes for AXDT.**

I. Introduction

The X-ray dark-field contrast [1] is a relatively new contrast mechanism for X-ray imaging that is obtained using an X-ray Grating Interferometer (XGI) [2], [3]. As the name suggests, an XGI setup uses the concept of interference of wavefronts that are shifted in phase with respect to each other, in order to obtain information about the refraction and scattering of X-rays passing through a specimen. This is done by introducing three gratings between the X-ray source and the detector as shown in Figure 1. The source grating splits an incoherent wavefront into several individually coherent but mutually incoherent sources of X-rays, which is a requirement for the interferometer to work. Next, the phase grating imprints a binary phase pattern on this wavefront, resulting in spatially varying intensity modulations, known as moire fringes, at specific locations downstream of the beam. Finally, an absorption grating placed in front of the detector is used to analyze the fringes to decode three different contrasts, namely the attenuation contrast, differential phase contrast, and the dark-field contrast (which is used in this work).

The dark-field contrast originates from the small and ultra-small angle scattering of X-rays from structures of sizes in the order of a few micrometers. Owing to the unidirectional sensitivity of an XGI setup with linear gratings, the dark-field signal obtained in such a setup is anisotropic in nature and encodes information about the three dimensional scattering function [4], [5]. Wieczorek et al. [6] developed a technique to reconstruct the full three dimensional scattering function in every volume element inside an object and termed it Anisotropic X-ray Dark-field Tomography (AXDT). In order to measure the scattering function sufficiently, we need to sample the unit sphere of scattering orientations. This is achieved by placing the sample on an Eulerian cradle, which enables three axes of rotations, as shown in Figure 1. An AXDT acquisition pose is thus defined as $x = (\psi, \theta, \phi)$.

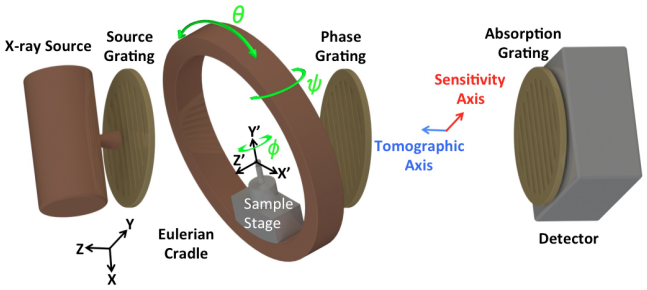


Fig. 1. Schematic of an X-ray grating interferometer (XGI) setup with an Eulerian cradle. The cradle provides three axes of rotations, y (ψ), z' (θ), and y' (ϕ). Figure by Sharma *et al.* [5] is licensed under CC BY 4.0.

Owing to the long and tedious acquisition geometries for AXDT [5], [7], there is a need to design and assess task-specific acquisition schemes. Detectability indices [8] provide a quantitative method to assess the quality of acquisition schemes in conventional CT. Fischer et al. [9] even used detectability indices to generate task-specific acquisition trajectories on-the-fly for industrial CT. In this work, we adapt an existing formulation of detectability index for the AXDT model in order to assess the reconstruction quality of different acquisition geometries. We present preliminary results on four AXDT acquisition schemes.

II. Methods

A. System model and reconstruction

AXDT aims at the reconstruction of scattering functions modeled as a field of spherical functions $\eta : \mathbb{S}^2 \times \mathbb{R}^3 \to \mathbb{R}$. We use a weighting function $h : \mathbb{S}^2 \times \mathbb{S}^2 \times \mathbb{S}^2 \to \mathbb{R}$ to model the relation between the dark-field measurement at the sample pose x, the tomographic vector $t(x)$ and the sensitivity vector $s(x)$. The tomographic vector and sensitivity vector at a pose $x = (\psi, \theta, \phi)$ are defined as:

$$\begin{aligned} s(x) &= R(x) \cdot S, \\ t(x) &= R(x) \cdot T, \end{aligned} \tag{1}$$

*Computer Aided Medical Procedures, Department of Informatics, Technical University of Munich, Germany
†Chair of Biomedical Physics, Department of Physics and School of BioEngineering, Technical University of Munich, Germany
‡Department of Diagnostic and Interventional Radiology, Klinikum rechts der Isar, Technical University of Munich, Germany
Contact: lasser@in.tum.de

where $R(x) \in \mathbb{R}^{3\times3}$ is the Euler rotation matrix for the pose x, $S \in \mathbb{S}^2$ is the setup sensitivity, and $T \in \mathbb{S}^2$ is the direction of beam propagation as shown in Figure 1 (here, $\cdot$ denotes standard matrix-vector multiplication).

By discretizing the spherical functions η and h using the basis of real-valued spherical harmonics truncated to maximum degree 4, we obtain the AXDT forward model [6]:

$$d(x) \approx \exp\left(-\frac{1}{4\pi}\sum_{k=0}^{4}\sum_{m=-k}^{k} h_k^m\big(s(x),t(x)\big)\int_{T(x)} n_k^m(r)dr\right) \tag{2}$$

where k is the degree and m is the order of the spherical harmonics. Standard discretization of line integrals over the X-ray paths $T(x)$ yields the system matrix P, which, when combined with diagonal matrices W_k^m containing the $h_k^m\big(s(x),t(x)\big)$, produce the following column-block linear system [6]:

$$\begin{aligned} p &= \sum_{k=0}^{K}\sum_{m=-k}^{m=k} W_k^m P \eta_k^m \\ &= \underbrace{\begin{pmatrix} W_0^0 P & \cdots & W_K^{-K} P & \cdots & W_K^K P \end{pmatrix}}_{=:\,\mathcal{B}} \begin{pmatrix} \eta_0^0 \\ \vdots \\ \eta_K^{-K} \\ \vdots \\ \eta_K^K \end{pmatrix}, \end{aligned} \tag{3}$$

where $p = -\log(d)$, and d is the vector containing the $d(x)$ for all poses x. We shorten (3) to:

$$\exp(-\mathcal{B}\eta) = d, \tag{4}$$

where $\mathcal{B}$ is the column-block linear operator defined in (3), η is the stacked vector containing all spherical harmonics coefficients, and d contains the dark-field measurements. Please note that η is a stacked vector containing 15 three-dimensional volumes, each representing the spherical harmonics coefficient of a specific degree and order (for details see [6]). Thus, the linear operator $\mathcal{B}$ is 15 times larger than the one for a corresponding traditional X-ray CT, making the forward and the back-projection computationally expensive.

We reconstruct η using a least-squares approach as in [10]:

$$\arg\min_{\eta} L(\eta,d), \qquad L(\eta,d) := \big\|\exp(-\mathcal{B}\eta) - d\big\|_2^2. \tag{5}$$

B. Task-based performance

In order to formulate our performance measure, approximations of useful imaging quality metrics are needed. In the following, we will focus on two performance metrics, namely the local impulse response (λ_j) of the system and the local covariance (σ_j), which we will later use to compute the modulation transfer function (MTF_j) and the noise-power spectrum (NPS_j).

The computation of the local impulse response has been previously derived in [11]. For the locality constraint, here we use a composite Dirac impulse vector e_j which contains 15 Dirac impulses, each one corresponding to the same spatial location in one of the volumes of spherical harmonics coefficients. Using $\epsilon(\eta) := \exp(-\mathcal{B}\eta)$ this results in

$$\lambda_j(\eta) = [-\nabla^{20}L(\eta,d)]^{-1}\nabla^{11}L(\eta,d)\nabla^{10}\epsilon(\eta)e_j, \tag{6}$$

where vector multiplication is element-by-element, and the different gradient operators are $\nabla^{10} = \partial/\partial\eta$, $\nabla^{20} = \partial^2/\partial\eta^2$ and $\nabla^{11} = \partial^2/\partial\eta\partial d$. In order to obtain an approximation of the local covariance, we use the findings from [12]:

$$\begin{aligned} \sigma_j(\eta) \approx [-\nabla^{20}L(\eta,d)]^{-1}[\nabla^{11}L(\eta,d)]\,\sigma(d) \\ [\nabla^{11}L(\eta,d)]^T[-\nabla^{20}L(\eta,d)]^{-1}e_j, \end{aligned} \tag{7}$$

where the covariance of the measurements $\sigma(d)$ can be approximated by the actual measurements d. Using (5), we have

$$\nabla^{10}L(\eta,d) = -2\mathcal{B}^T\epsilon(\eta)\big(2\epsilon(\eta)-d\big), \tag{8}$$

$$\nabla^{20}L(\eta,d) = 2\mathcal{B}^T\epsilon(\eta)\big(2\epsilon(\eta)-d\big)\mathcal{B}, \tag{9}$$

$$\nabla^{11}L(\eta,d) = -2\mathcal{B}^T\epsilon(\eta), \tag{10}$$

$$\nabla^{10}\epsilon(\eta) = \mathcal{B}^T\epsilon(\eta). \tag{11}$$

Putting (8) to (11) into (6) and (7), we obtain:

$$\begin{aligned} \lambda_j(\eta) = \big[-2\mathcal{B}^T\epsilon(\eta)\big(2\epsilon(\eta)-d\big)\mathcal{B}\big]^{-1} \\ -2\mathcal{B}^T\epsilon(\eta)\mathcal{B}^T\epsilon(\eta)e_j, \end{aligned} \tag{12}$$

$$\begin{aligned} \sigma_j(\eta) = \big[-2\mathcal{B}^T\epsilon(\eta)\big(-2\epsilon(\eta)-d\big)\mathcal{B}\big]^{-1}\big[-2\mathcal{B}^T\epsilon(\eta)\big]d \\ \big[-2\mathcal{B}^T\epsilon(\eta)\big]^T\big[-2\mathcal{B}^T\epsilon(\eta)\big(2\epsilon(\eta)-d\big)\mathcal{B}\big]^{-1}e_j. \end{aligned} \tag{13}$$

Using (12) and (13), we can approximate the modulation transfer function (MTF) and the noise-power spectrum (NPS)

$$\begin{aligned} MTF_j(\eta) \approx \mathcal{F}\{\lambda_j(\eta)\} = \\ \frac{\mathcal{F}\big\{-2\big(\mathcal{B}^T\epsilon(\eta)\big)^2 e_j\big\}}{\mathcal{F}\big\{-2\mathcal{B}^T\epsilon(\eta)\big(2\epsilon(\eta)-d\big)\mathcal{B}e_j\big\}}, \end{aligned} \tag{14}$$

$$\begin{aligned} NPS_j(\eta) \approx \mathcal{F}\{\sigma(\eta_j)\} = \\ \frac{\mathcal{F}\big\{4\mathcal{B}^T\epsilon(\eta)d\epsilon(\eta)\mathcal{B}e_j\big\}}{\big|\mathcal{F}\big\{2\mathcal{B}^T\epsilon(\eta)\big(-2\epsilon(\eta)-d\big)\mathcal{B}e_j\big\}\big|^2}. \end{aligned} \tag{15}$$

with element-by-element divisions and where the operator $\mathcal{F}$ represents the three-dimensional Fourier transform for each separate spherical component degree and order.

Having defined the MTF and the NPS, we can estimate the detectability index for a specific task using a non-prewhitening matched filter observer [13]:

$$\delta_j^2 = \frac{\big[\int_{\mathbb{R}^4}(MTF_j\mathcal{W}_{ROI})^2\,df^4\big]^2}{\int_{\mathbb{R}^4} NPS_j(MTF_j\mathcal{W}_{ROI})^2\,df^4} \tag{16}$$

where vector multiplications are element-by-element. The frequency template $\mathcal{W}_{ROI}$ should match the Fourier transform of the expected signal of a region of interest (ROI). As

a result, the frequencies corresponding to the specific ROI inside the MTF will be assigned higher weights.

The benefit of using a non-prewhitening observer model is, that it performs very similar to a human observer [13]. In conclusion, the performance predictor δ_j^2 is dependent on the acquisition scheme ($\mathcal{B}$), the specific given task ($\mathcal{W}$), its location (subscript j), the reconstructed volume η and the dark-field measurements d.

C. Acquisition Scheme

AXDT reconstruction aims at the reconstruction of a spherical function in every volume element. Owing to this, the measurements need to sample the unit sphere of orientations $q \in \mathbb{S}^2$. Sharma et al. [7] presented a method to design acquisition schemes for AXDT by extending the concept of circular tomographic trajectories to the unit sphere, and demonstrated their compatibility with the AXDT model. The fundamental idea of an AXDT acquisition scheme is to design several tomographic trajectories to measure the component of the scattering function along a sensitivity orientation $q \in \mathbb{S}^2$.

In order to do so, we begin with a standard CT acquisition scheme comprising of $N \in \mathbb{N}$ equally spaced points between the angles $[0, \Phi)$:

$$X(\Phi, N) := \left\{ x = (0, 0, \phi);\ \phi \in \left\{ 0, \frac{\Phi}{N}, \ldots, \Phi - \frac{\Phi}{N} \right\} \right\}. \tag{17}$$

However, every pose of such a scheme measures a different component of the scattering function, which is not desirable for tomographic reconstruction, as detailed by Sharma et al. [7]. Therefore, we define the scheme $Y(X, q) := \{y = T(x, q)\ \forall\ x\ \in\ X\}$, where the transformation $T(x, q)$ transforms the pose $x := (0, 0, \phi)$ to the pose $y := (\psi, \theta, \phi)$ such that $s(y) = q$, see (1). Figure 2(a) shows such a scheme.

Using the above concept, we design an acquisition scheme Z that measures L scattering orientations: $Z(X) = \{Y(X, q_l);\ l = 1, \ldots, L\}$. Here, we choose a t-design with $L = 28$ points distributed uniformly over one half of the unit sphere. A scheme $Z(X(180°, 25))$ is shown in Figure 2(b). As a last step, we select poses from the scheme Z that can actually be measured in the XGI setup, resulting in Figure 2(c). Please note that all poses cannot be measured because the Eulerian cradle intercepts the X-ray beam for higher values of ψ.

III. Results and Discussion

We investigated the performance of the model observer to illustrate the results of the proposed formulation. We measured a circular thermoplastic short fiber moulding part composed of fibers that are 7μm thick and 200μm long, at a resolution of approximately 0.32 mm using four different acquisition schemes:

1) $A = Z(X(180°, 25))$
2) $B = Z(X(135°, 25))$
3) $C = Z(X(90°, 25))$
4) $D = Z(X(45°, 25))$

All of the above schemes are cropped to contain only the measurable poses as shown in Figure 2(c-f). We reconstructed the scattering functions as spherical harmonics coefficients up to degree 4. Next, we extracted the main structure orientation in every voxel using the Funk-radon transform, as presented by Wieczorek et al. [6]. Figure 3 shows a three dimensional vector in every voxel for one slice of the result obtained using the scheme A. The structure orientations reveal a unique feature in this slice, which is highlighted with a red box. The feature is a weld-line, which is not visible in conventional CT at such a coarse resolution.

Next, we computed the voxel-wise detectability index over a region representing the weld-line using the four different acquisition schemes. The forward and back-projections and all the matrix-vector and vector-vector Hadamard products have been implemented in our in-house developed CampRecon Toolbox [14], while the Fourier transforms have been performed in Matlab. We used the weld-line feature to define W_{ROI} in (16), see also the red-colored vectors in Figure 5 for an illustration. The average detectability index over the whole region of interest for the four schemes is shown as a bar graph in Figure 4. Additionally, we show the feature obtained using each of the four schemes in Figure 5.

We observe a decreasing trend of the detectability index value from scheme A to scheme D, which is also clearly reflected in the detectability of the feature visually in Figure 5. This is also the expected outcome, since the scheme A is an ideal scheme as presented by Sharma et al. [7], while schemes B to D contain limited angle trajectories for all scattering orientations, with the solid angle of the unit sphere covered by the schemes decreasing from B to D.

Here we only cover a specific formulation of the AXDT reconstruction problem, which leaves much room for improvement. Equation (4) is only the simplest of several statistical models covered by Schilling et al. [10], and equation (16) is also only a part of the methods presented by Gang et al. [13]. In the future, we aim to fully integrate all the different reconstruction models into our detectability prediction algorithm and to also investigate the performance of the different observer models for AXDT.

In conclusion, we have adapted an existing observer model based method of measuring image quality performance from conventional CT to the novel Anisotropic X-ray Dark-field Tomography (AXDT) imaging modality. Our results definitely show a very positive trend in correctly classifying object-specific trajectories into optimal and suboptimal ones.

Acknowledgements

The authors thank O. Focke (Faserinstitut Bremen), S. Zabler (Fraunhofer EZRT) and M. Willner (MITOS GmbH) for providing the sample used in this study. Financial support was provided by the Munich-Centre for Advanced Photonics (MAP), DFG Gottfried Wilhelm Leibniz program and the Karlsruhe Nano Micro Facility (KNMF, www.kit.edu/knmf).

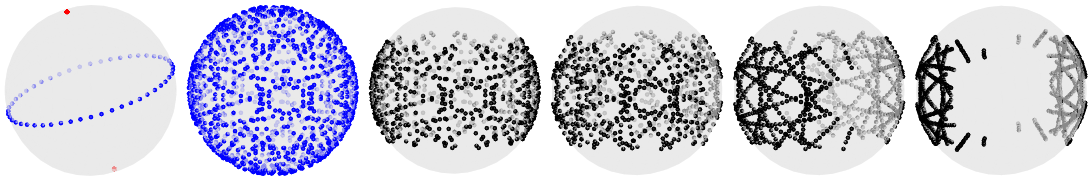


Fig. 2. Spherical representation of acquisition schemes. The blue and red points represent the tomographic and sensitivity vectors, respectively, for a given acquisition geometry. (a) Blue points representing the points $\pm t(x)$ for a scheme $Y\left(X(180°, 25), [0.15, 0.94, -0.3]^T\right)$. All the poses measure a unique component of scattering, denoted by the red point. (b) Scheme $Z\left(X(180°, 25)\right)$ (c) Poses out of (b) that can be measured in the setup shown in Figure 1, also renamed to scheme A. (d-f) Schemes B, C, and D. We can observe that these schemes are analogous to limited angle trajectories in standard CT.

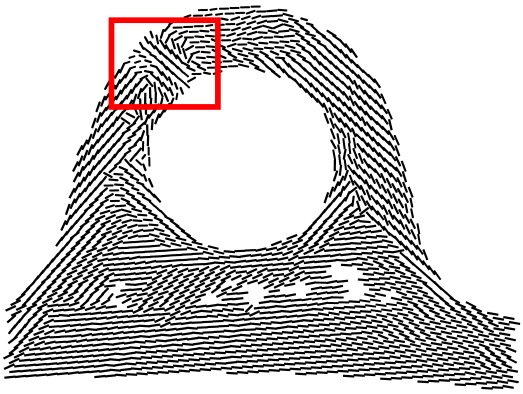


Fig. 3. Three dimensional structure orientations in a single slice out of the AXDT result obtained using the scheme A. The feature inside the red box reveals a weld-line which cannot be seen using conventional CT at this resolution of 0.32 mm.

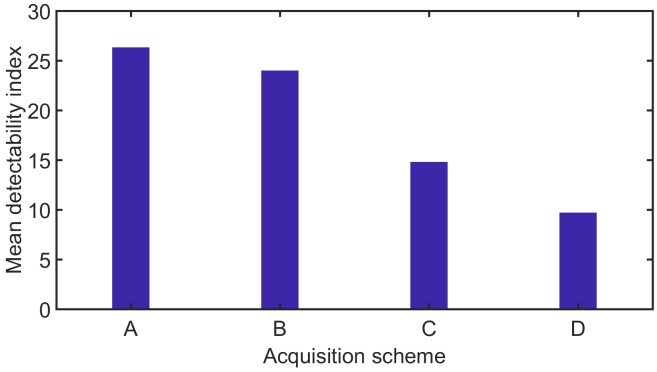


Fig. 4. The mean detectability index measure for each different acquisition scheme A-D over the ROI

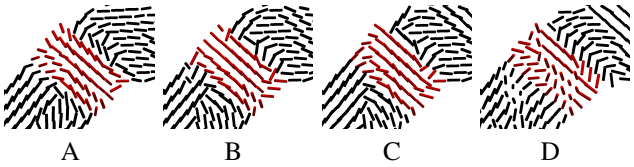


Fig. 5. Region of interest shown in Figure 3 for the schemes A to D.

References

[1] F. Pfeiffer, M. Bech, O. Bunk, P. Kraft, E. F. Eikenberry, C. Brönnimann, C. Grünzweig, and C. David, "Hard-X-ray dark-field imaging using a grating interferometer." *Nature materials*, vol. 7, no. 2, pp. 134–137, 2008.

[2] F. Pfeiffer, T. Weitkamp, O. Bunk, and C. David, "Phase retrieval and differential phase-contrast imaging with low-brilliance X-ray sources," *Nature Physics*, vol. 2, no. 4, pp. 258–261, 2006.

[3] A. Momose, S. Kawamoto, I. Koyama, Y. Hamaishi, K. Takai, and Y. Suzuki, "Demonstration of x-ray talbot interferometry," *Japanese Journal of Applied Physics*, vol. 42, no. 7B, p. L866, 2003.

[4] T. H. Jensen, M. Bech, O. Bunk, T. Donath, C. David, R. Feidenhans'l, and F. Pfeiffer, "Directional x-ray dark-field imaging," *Physics in Medicine and Biology*, vol. 55, no. 12, p. 3317, 2010.

[5] Y. Sharma, M. Wieczorek, F. Schaff, S. Seyyedi, F. Prade, F. Pfeiffer, and T. Lasser, "Six dimensional x-ray tensor tomography with a compact laboratory setup," *Applied Physics Letters*, vol. 109, no. 13, p. 134102, 2016.

[6] M. Wieczorek, F. Schaff, F. Pfeiffer, and T. Lasser, "Anisotropic x-ray dark-field tomography: A continuous model and its discretization," *Phys. Rev. Lett.*, vol. 117, p. 158101, Oct 2016.

[7] Y. Sharma, F. Schaff, M. Wieczorek, F. Pfeiffer, and T. Lasser, "Design of Acquisition Schemes and Setup Geometry for Anisotropic X-ray Dark-Field Tomography (AXDT)," *Scientific Reports*, vol. 7, no. 1, p. 3195, 2017.

[8] J. W. Stayman and J. H. Siewerdsen, "Task-based trajectories in iteratively reconstructed interventional cone-beam ct," *Proc. 12th Int. Meet. Fully Three-Dimensional Image Reconstr. Radiol. Nucl. Med*, pp. 257–260, 2013.

[9] A. Fischer, T. Lasser, M. Schrapp, J. Stephan, and P. B. Noel, "Object specific trajectory optimization for industrial x-ray computed tomography," *Scientific Reports*, vol. 6, no. 4, Jan 2016.

[10] N. Schilling, M. Wieczorek, and T. Lasser, "Statistical models for anisotropic x-ray dark-field tomography," in *Proc. Int'l Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine (Fully3D)*, Jun. 2017.

[11] J. A. Fessler and W. L. Rogers, "Resolution properties of regularized image reconstruction methods," of EECS, Univ. of Michigan, Ann Arbor, MI, Tech. Rep., 1995.

[12] J. A. Fessler, "Mean and variance of implicitly defined biased estimators (such as penalized maximum likelihood): applications to tomography," *IEEE Transactions on Image Processing*, vol. 5, no. 3, pp. 493–506, Mar 1996.

[13] G. J. Gang, J. Lee, J. W. Stayman, D. J. Tward, W. Zbijewski, J. L. Prince, and J. H. Siewerdsen, "Analysis of fourier-domain task-based detectability index in tomosynthesis and cone-beam ct in relation to human observer performance," *Medical Physics*, vol. 38, no. 4, pp. 1754–1768, 2011.

[14] M. Wieczorek, J. Vogel, and T. Lasser, "Camprecon a software framework for linear inverse problems," Tech. Rep., 01 2014.

Simultaneous reconstruction and separation in a spectral CT framework with a proximal variable metric algorithm

Souhil Tairi and Sandrine Anthoine and Christian Morel and Yannick Boursier

***Abstract*—In this paper, we propose a method to simultaneously separate and reconstruct the physical components of an object observed in the context of spectral Computed Tomography. Our contribution considers the underlying polychromatic model of the X-ray beam and combines it with a prior on the components of the object to reconstruct. The simultaneous separation and reconstruction of components is done by solving a non-convex ill-posed inverse problem with a variable metric proximal algorithm. Promising results are shown on 3D real data acquired on the micro-CT PIXSCAN prototype with a simple regularization that encompasses the positivity of the quantities of interest.**

I. Introduction

Spectral Computed Tomography (spectral CT) is a new imaging modality that relies on the recent development of photon-counting detectors. It enhances the classical tomographic measurements (which give information about spatial localization of the mass attenuation of the sample) with energy-dependent (i.e., spectral) attenuation properties that are obtained by counting photons in different energy bins [1]. The spectral information can be used to discriminate different physical phenomena such as Compton or photo-electric effects, different components such as soft tissues, bones, or contrast agents... This technology opens the way for using X-ray tomography as a functional imaging tool, which would naturally improve for example the study of biomedical or clinical data. - Although introduced as early as 1976 by Alvarez and Macovski [2], a growing interest for spectral aspect in CT has started more than a decade ago, with the advent of dual-energy CT, which can separate two materials in specific conditions [1], [3]. The development of photons-counting detectors has enabled to go a step further and simultaneously reconstruct of a larger number of material maps by exploiting a larger set of spectral measurements.

The reconstruction problem from spectral CT measurements is two-fold: one issue is to disentangle the information coming from each material, the other is to reconstruct spatial maps from tomographic measurement. While the later issue corresponds to the classical tomographic reconstruction problem in CT, the first one is a source separation problem that exploits the spectral information available in the data. Both problems are separately non-trivial by essence, they raise questions pertaining to the class of ill-posed inverse problems. Naturally, the full spectral CT problem can be tackled by either solving both problems sequentially or simultaneously.

Sequential methods have been proposed first. Among these, some first handle the reconstruction problem and then proceed to the separation in object space. This is the case for example for the pioneering K-edge methods [1], [4]. They leverage on the spectral profile of high-Z elements such as contrasts agents which have a characteristic spectral discontinuity (K-edge) to approximate the spectral separation by simple addition/subtraction. A more detailed attention to the physics of the acquisition lead others to consider the spectral separation first, estimating thus the contribution of each material in the measurement space (sinograms) and then proceed to the spatial reconstruction of each map [5], [6]. Although these latest methods are quite effective because they handle the two problems separately and thus divide the computational complexity, they might be limited by the errors perpetrated from one step to the next. Simultaneous treatment of the spectral CT has been explored as well, and our contribution lies in this line of research. The global separation and reconstruction problem can be cast as the minimization of a functional. The spectral integration on the detectors yields a non-convex data fidelity term, that is usually penalized with a priori regularizations. The non-convexity has been handled first by using quadratic majorization on smooth functionals [7]. Lately, a more involved proximal scheme called MOCCA has been developed by Barber and colleagues [8] that is able to incorporate non-smooth regularization. Our work follows the same line of research. Similarly, we propose a proximal scheme to handle the simultaneous separation and reconstruction of spectral CT data with non-smooth a-priori on the object. The main difference is the algorithm itself: while MOCCA is a primal-dual scheme inspired by [9] that has been demonstrated on 2D data only, our contribution is based on a variable-metric primal algorithm [10], [11] and we show in this contribution its applicability on 3D synthetic and real data.

The rest of the paper is organized as follows. The next section establishes the physical forward model as well as the modelization of attenuation maps. From this, we deduce the inverse problem at stake and present the variable metric algorithm that relies on a metric specifically designed for the

S. Tairi, C; Morel and Y. Boursier are with Centre de Physique des Particules de Marseille, Aix Marseille Univ, CNRS/IN2P3, CPPM, Marseille, France. S. Anthoine is with Institut de Mathématiques de Marseille, Aix Marseille Univ, CNRS, Centrale Marseille, Marseille, France.

spectral CT problem. The last section presents on results on real data.

II. Forward model

A. Acquisition model

A Computerized-Tomography (CT) scan is a set of measurements obtained by shining a rotating X-Ray light, modulated by spectral filters through an object. In the polychromatic (i.e. spectral) setting, the p-th measurement depends on

- the position of the source-detector couple with respect to the object, this is the line of sight $\mathcal{L}^p$,
- the absorption coefficients of the object $\mu(l, E)$ for l on the line of sight $\mathcal{L}^p$,
- the X-ray source beam $I_0(E)$,
- the filters attenuation factor $Fi(E)$,
- the detectors efficiency $De(E)$.

It is quantified by the Beer-Lambert law:

$$y^p = \int_{\mathbb{R}^+} I_0(E) Fi(E) De(E) e^{-\int_{\mathcal{L}^p} \mu(l,E) dl} dE \quad (1)$$

The polychromatic CT reconstruction is the estimation of the map $\mu : \mathbb{R}^3 \times \mathbb{R}^+ \to \mathbb{R}^+$ from the knowledge of $I_0(E)$, $Fi(E)$ and $De(E)$ and a set of measurements y^p obtained by i) rotating the source/detector couple around the object, ii) changing the filters and iii) modifying the efficiency energy spectrum of the detector (*e.g.* by setting an energy discrimination threshold).

Using P spatial configurations $\mathcal{L}^p$, Q spectral filters Fi and R settings per detector efficiency De, one obtains a set of $M = PQR$ measurements. Denoting $I_0(E)Fi^q(E)De^r(E) = f^m(E)$ the total spectral inputs, the m-th measurement reads:

$$y^m = \int_{\mathbb{R}^+} f^m(E) e^{-\int_{\mathcal{L}^m} \mu(l,E) dl} dE \quad (2)$$

B. Absorption maps model

Recovering the full spatio-spectral absorption map might still be unrealistic with a finite number of measurements. However different components in a object, such as soft tissues, bones, etc., have a different spectral signature. A natural model for the absorption map is thus to consider it as the sum of the contribution of each of its components (say K in total): $\mu(l, E) = \sum_{k=1}^K \mu^k(l, E)$. The spectral signature is physically independent of the spatial location of a component, leading to:

$$\mu(l, E) = \sum_{k=1}^K \mu^k(l, E) = \sum_{k=1}^K a^k(l) \sigma^k(E), \quad (3)$$

with $a^k(l)$ the concentration of component k at point l, and $\sigma^k(E)$ its interaction cross section.

The perfect measurements now read:

$$y^m = \int_{\mathbb{R}^+} f^m(E) e^{-\sum_{k=1}^K \sigma^k(E) \int_{\mathcal{L}^m} a^k(l) dl} dE \quad (4)$$

C. Discretization

Let us discretize the energy E in N bins, and the 3D-volume where the object lives in D voxels and denote by $S[p, d]$ the contribution of the d-th voxel in the integral on $\mathcal{L}^m$. Each column of the matrices F and A respectively contain the spectral inputs f^m and concentration maps a^k and each row of the matrix Σ contains the interaction cross section σ^k. The forward discretized model reads:

$$Y[p, m] = \sum_{n=1}^N F[n, m] e^{-\sum_{k=1}^K \sum_{d=1}^D S[p,d] A[d,k] \Sigma[k,n]}$$
$$Y = e^{-SA\Sigma} F, \quad (5)$$

with $Y \in \mathbb{R}^{P \times M}$, $F \in \mathbb{R}^{N \times M}$, $S \in \mathbb{R}^{P \times D}$, $A \in \mathbb{R}^{D \times K}$, and $\Sigma \in \mathbb{R}^{K \times N}$.

D. Inverse problem formulation

The measurements Y we have are in fact noisy versions of the perfect measurement described by Eq. (5). In our case the noise is caused by the counting rate which induces a Poisson noise. Our goal is to recover the matrix A from the knowledge of the noisy measurements Y, the experiment settings S, F and the spectral signatures Σ.

We propose to tackle this inverse problem by minimizing

$$J(A) = G(A) + R(A) \quad (6)$$

where

$$G(A) = \sum_{p,m} \{e^{-SA\Sigma} F\}_{p,m} - Y[p, m] \log(\{e^{-SA\Sigma} F\}_{p,m}) \quad (7)$$

is the discrepancy measure naturally given by the negative log-likelihood corresponding to the Poisson noise model. $R(A)$ is a nonsmooth regularization made of a term constraining the nonnegativity on all elements ($A[d, k] \geq 0\,, \forall\,(d, k)$) and a regularization term for each map $A[\cdot, k] \overset{def}{=} A_k$:

$$R(A) = \chi_{\{\geq 0\}}(A) + \sum_{k=1}^K \lambda_k R_k(A_k) \quad (8)$$

where $\chi_{\{\geq 0\}}(A) = 0$ if $A[d, k] \geq 0$, for all d, k and $\chi_{\{\geq 0\}}(A) = +\infty$ otherwise. The spatial regularization $R_k(A_k)$ are chosen according to the components for example a sparsity term such as an l_1 regularization or a total variation penalization.

III. Reconstruction method

Generally speaking, the regularization term $R(A)$ is non-differentiable but convex. On the other hand $G(A)$ is $\mathcal{C}^\infty$ but not convex. However, it can be shown that it possesses Kurdyka-Lojasiewicz property [12], [13]. We have to handle a non-convex minimization problem, but for which proximal descent methods - originally developed in the convex case - have recently been successfully extended [14]. In order to apply this methodology to real and three-dimensional spectral tomographic data, it is essential that the proposed

algorithm converges quickly. We propose here to use a *variable metric proximal scheme* [11], [10], which, by incorporating a well-chosen metric, enables to clearly accelerate the convergence of the proximal descent.

A. Variable Metric Forward-Backward Algorithm

Given the result of iteration t, denoted by A^t, the *variable metric forward-backward* (VMFB) algorithm proceeds in three steps to calculate A^{t+1}:

f1) We choose a metric $\mathbf{M}^t$ which is a bijection of $\mathbb{R}^{D\times K}$.

f2) We calculate a gradient descent step on G preconditioned by $\mathbf{M}^t$:

$$B^t = A^t - \tfrac{1}{L}\left\{\mathbf{M}^t\right\}^{-1}\nabla G(A^t) \tag{9}$$

f3) Then a proximal point is calculated according to the metric $\mathbf{M}^t$:

$$\begin{aligned} A^{t+1} &= \mathbf{prox}_{L\mathbf{M}^t,\,R}(B_k) \\ A^{t+1} &\overset{\text{def}}{=} \underset{y}{\mathbf{argmin}}\{\tfrac{1}{L}R(y) + \tfrac{1}{2}\|y - B^t\|^2_{\mathbf{M}^t}\} \end{aligned} \tag{10}$$

where $L \geq L_G$, with L_G the Lipschitz constant of the gradient ∇G. Note that the metric $\mathbf{M}^t$ needs to be in the set of symmetrical positive definite matrices $\mathbb{S}^+$ and the calculation of the proximal point (10) is not analytical here, we approximate it by an iterative accelerated schema (FISTA) [15].

B. Metric Design

The convergence rate of the algorithm depends greatly on the second order information incorporated in the metric $\mathbf{M}^t$. This information is available in the Hessian of $G(A)$:

$$\mathbf{H}G(A) = 2(\Sigma \otimes S^T)\ \Phi(A)\ (\Sigma^T \otimes S) \tag{11}$$

with

$$\Phi(A) = \mathbf{Q}_1(A)(FF^T \otimes \mathbf{I}_K)\mathbf{Q}_1(A) + \mathbf{Q}_2(A) \tag{12}$$

where $\mathbf{I}_K$ is the identity matrix and $\otimes$ the Kronecker product. $\mathbf{Q}_1(A)$ and $\mathbf{Q}_2(A)$ are diagonal matrices:

$$\begin{aligned} \mathbf{Q}_1(A) &= \mathbf{diag}(e^{-SA\Sigma}) \\ \mathbf{Q}_2(A) &= \mathbf{diag}((Y - e^{-SA\Sigma}F)F^T \odot -e^{-SA\Sigma}) \end{aligned}$$

where $\odot$ denotes the Hadamard product. The Hessian is not necessarily positive definite because of the residual term $\mathbf{Q}_2$. We design a metric inspired by it that will also have the two following advantages: i) it is easily invertible (needed for step **f1**)), ii) it does not depend on A so it is the same metric for each iteration of the algorithm. Dropping the residual term and dependence on A, we propose to use the following metric:

$$\mathbf{M}_1 = 2(\Sigma FF^T\Sigma^T \otimes S^T S) \tag{13}$$

$\mathbf{M}_1$ not only contains the needed second order information, it is also easy to inverse. Indeed $\Sigma FF^T\Sigma^T$ is a small matrix of size $(K \times K)$ and the operator $S^T S$ is a natural candidate [16] within the monochromatic framework. Its inverse, called the cone filter, is easy to implement because it is diagonal in the Fourier basis in the case of parallel beam.

C. Step Choice

The constant L is an important factor of the speed of convergence of the algorithm. It depends on the norm of chosen metric, in our case:

$$L = \|\mathbf{M}_1\| = 2\,\|\Sigma FF^T\Sigma^T\|\,\|S^T S\| \tag{14}$$

where the norm $\|.\|$ is defined as:

$$\|X\| = \max_{\lambda}\{|\lambda| \in \mathbb{R} \mid \lambda \,\text{eigenvalue of}\, X\}.$$

IV. Results on real data

We assess the performance of the proposed algorithm on real data acquired on the micro-CT PIXSCAN prototype equipped with the XPAD3 hybrid pixel detector. We used a twisted phantom given by three rubber pipes, 16 cm long of 3 mm inner diameter. The pipes were filled with the three following solutions: (i) $AgNO_3$ at a concentration of 118 mg/mL of Silver, (ii) Iomeron at 95 mg/mL of Iodine and (iii) $CuSO_4$ at 1.0 M. We used the X-ray energy spectra given by a Mo tube operated at 50 kVp and a Cu filter 100 μm thick. A CT scan corresponds to 360 projections (10 s/frame) spaced by a 1 degree rotation step. Three scans were performed at the three energy thresholds: (i) 21.0 keV, (ii) 25.5 keV and (iii) 33.2 keV. The photon statistics in the flat field depends strongly on the applied threshold. The reconstructed volume is of size $582 \times 30 \times 582$.

The dictionary Σ has been filled in with three tabulated spectral signatures (from the NIST XCOM database), (i) Iomeron 350 in solution at 50 mg/mL in Iodine, (ii) $AgNO_3$ in solution at 40 mg/mL in Ag and (iii) plastic Vinyltoluene. The F matrix has been filled using the output of a GATE Monte Carlo simulation of the XPAD3 detector response, taking into account charge sharing.

Results have been obtained after 200 iterations of the proposed algorithm, and 200 iterations of the proximal operator have been done in the inner loop. Fig. 1 shows a transverse slice though the twisted phantom and demonstrates the spectral identification of the two contrast agents, namely Silver and Iodine. Some artefacts remains with in particular some cross-talks between Iodine and Plastic components. Further work in progress include finer estimation of the parameters of the micro-CT scanner and adding a spatial regularization via Total Variation.

Acknowledgement

This work was partly funded by the French program "Investissement d'Avenir" run by the 'Agence Nationale pour la Recherche' ; the grant reference is 'Infrastructure d'avenir en Biologie Santé - ANR-11-INBS-0006'

References

[1] J. P. Schlomka, E. Roessl, R. Dorscheid, S. Dill, G. Martens, T. Istel, C. Bäumer, C. Herrmann, R. Steadman, G. Zeitler, A. Livne, and R. Proksa, "Experimental feasibility of multi-energy photon-counting k-edge imaging in pre-clinical computed tomography," *Physics in Medicine and Biology*, vol. 53, no. 15, p. 4031, 2008. [Online]. Available: http://stacks.iop.org/0031-9155/53/i=15/a=002

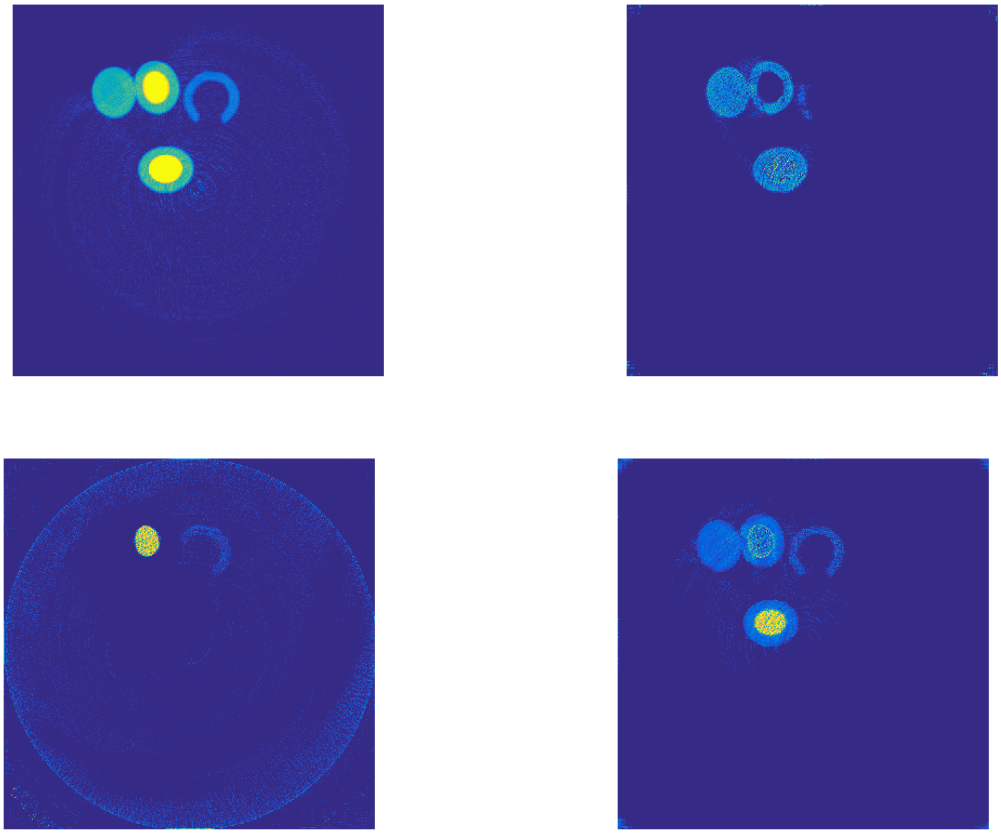


Fig. 1. From left to right and top to bottom: a FDK standard monochromatic reconstruction of the 21 keV acquisition; the plastic component estimated by the proposed algorithm; the Silver component; the Iodine component.

[2] R. E. Alvarez and A. Macovski, "Energy-selective reconstructions in x-ray computerised tomography," *Physics in Medicine and Biology*, vol. 21, no. 5, p. 733, 1976. [Online]. Available: http://stacks.iop.org/0031-9155/21/i=5/a=002

[3] P. Sukovic and N. H. Clinthorne, "Penalized weighted least-squares image reconstruction for dual energy x-ray transmission tomography," *IEEE Transactions on Medical Imaging*, vol. 19, no. 11, pp. 1075–1081, Nov 2000.

[4] F. Cassol Brunner, M. Dupont, C. Meessen, Y. Boursier, H. Ouamara, A. Bonissent, C. Kronland-Martinet, J.-C. Clemens, F. Debarbieux, and C. Morel, "First K-Edge Imaging With a Micro-CT Based on the XPAD3 Hybrid Pixel Detector," *IEEE Transactions on Nuclear Science*, vol. 60, pp. 103 – 108, 2013. [Online]. Available: http://hal.in2p3.fr/in2p3-00799983

[5] M. Dupont, Y. Boursier, A. Bonissent, F. Galland, F. Cassol Brunner, and C. Morel, "Component Separation for Spectral X-Ray Imaging Using the XPAD3 Hybrid Pixel Camera," in *2013 IEEE Nuclear Science Symposium and Medical Imaging Conference*, Oct. 2013, poster. [Online]. Available: http://hal.in2p3.fr/in2p3-00956664

[6] N. Ducros, J. F. P. J. Abascal, B. Sixou, S. Rit, and F. Peyrin, "Regularization of nonlinear decomposition of spectral x-ray projection images." *Med Phys*, vol. 44, pp. e174–e187, 2017 Sep 2017. [Online]. Available: https://hal-hcl.archives-ouvertes.fr/hal-01391538/document

[7] Y. Long and J. A. Fessler, "Multi-material decomposition using statistical image reconstruction for spectral ct," *IEEE Transactions on Medical Imaging*, vol. 33, no. 8, pp. 1614–1626, Aug 2014.

[8] T. G. Schmidt, R. F. Barber, and E. Y. Sidky, "A spectral ct method to directly estimate basis material maps from experimental photon-counting data," *IEEE Transactions on Medical Imaging*, vol. 36, no. 9, pp. 1808–1819, Sept 2017.

[9] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 120–145, May 2011. [Online]. Available: https://doi.org/10.1007/s10851-010-0251-1

[10] S. Bonettini, I. Loris, F. Porta, M. Prato, and S. Rebegoldi, "On the convergence of a linesearch based proximal-gradient method for nonconvex optimization," *Inverse Problems*, 2017, accepted. [Online]. Available: https://arxiv.org/abs/1605.03791

[11] E. Chouzenoux, J.-C. Pesquet, and A. Repetti, "Variable metric forward–backward algorithm for minimizing the sum of a differentiable function and a convex function," *Journal of Optimization Theory and Applications*, vol. 162, no. 1, pp. 107–132, 2014. [Online]. Available: http://dx.doi.org/10.1007/s10957-013-0465-7

[12] S. Lojasiewicz, "Une propriété topologique des sous-ensembles analytiques réels," *Editions du centre National de la Recherche Scientifique*, pp. 87–89, 1963.

[13] K. Kurdyka, "On gradients of functions definable in o-minimal structures," *Annales de l'institut Fourier*, vol. 48, no. 3, pp. 769–783, 1998. [Online]. Available: http://eudml.org/doc/75302

[14] H. Attouch, J. Bolte, and B. F. Svaiter, "Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward–backward splitting, and regularized gauss–seidel methods," *Mathematical Programming*, vol. 137, no. 1, pp. 91–129, 2013. [Online]. Available: http://dx.doi.org/10.1007/s10107-011-0484-9

[15] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, 2009.

[16] A. Perelli, M. A. Lexa, A. Can, and M. E. Davies, "Denoising message passing for x-ray computed tomography reconstruction," *CoRR*, vol. abs/1609.04661, 2016. [Online]. Available: http://arxiv.org/abs/1609.04661

Consistency of Fan-beam Projections of a Translating Object Along an Arc of a Circle

Thomas Boulier, Rolf Clackdoyle, Jérôme Lesaint, Laurent Desbat

***Abstract*—In this article, we compute data consistency conditions (DCCs) in fan-beam geometry for the case of a translating object irradiated by an x-ray source moving along a short arc of a circle. The DCCs are thus a generalization of those computed in [2], where the object remained still. In a second part, we use the DCCs in order to retrieve parameters of the object velocity. These results are illustrated by numerical examples.**

I. Introduction

In the field of CT-imaging, data consistency conditions (DCCs) use the redundancy of the data in order to establish relations that must be fulfilled between projections. These conditions are called *full* when they are not only necessary but also sufficient. One of the best known DCCs are the Helgason-Ludwig (H-L) conditions [3], [4], which are full for parallel projections. Full DCCs are also known for the case of fan-beam projections with source position taken along a straight line [1].

In [2], the fan-beam DCCs were modified to handle the case of a source moving along an arc of a circle. Briefly, all rays passing through the same point along a "virtual" source line between the two extreme positions of the source are gathered to form a virtual fan-beam projection. The DCCs for fan-beam projections along a line from [1] could then be applied; numerical examples showed a situation where an artificial detector attenuation was added, and the DCCs were invoked to recover the unknown attenuation coefficient. The case of a moving object has been tackled in [7], where the authors used parallel-beam H-L conditions for the case of an object undergoing rigid body motion while irradiated by a fan-beam source along a circular trajectory. Those conditions allowed the authors to recover the parameters of the the movement. The parameters were subsequently incorporated into the reconstruction procedure, in order to suppress the artifacts caused by such movement. However, using the parallel-beam H-L conditions means that the fan-beam projections had to cover all rays in the plane so that a fan-to-parallel rebinning could be achieved. The rebinning was achieved mathematically; however, the requirement of at least a standard fan-beam shortscan of $180°$ plus fan-angle was required. In our work, we are concerned with the opposite extreme of a short arc of fan-beam measurements.

T. Boulier, R. Clackdoyle, J. Lesaint and L. Desbat are with the TIMC-IMAG laboratory, CNRS UMR 5525 and Université Grenoble Alpes (e-mail: thomas.boulier@polytechnique.edu).

The authors would like to thank Simon Rit for his help using RTK.

This work is partially supported by the Agence Nationale de la Recherche (France), project “DROITE”, number ANR-12-BS01-0018, and FUI project "3D4Carm".

In this article, we will suppose that an object is translating while illuminated by a fan-beam source which is moving along an arc of a circle. With an adapted reference frame, we will use as in [2] a change of variables. Then, in a second part, we will show how the DCCs can be used in order to identify the translation velocity.

II. Theory

A. Problem under consideration

Let us begin with some notation and definitions. We will consider an object in $\mathbb{R}^2$ to be imaged in a fan-beam geometry with sources on an arc of circle with center O and radius R_0 (see Figure 1, top). The object is identified with its density function $\mathbf{x} \mapsto \mu(\mathbf{x}) \in \mathcal{C}_c^\infty(\mathbb{R}^2)$. The angular velocity of the

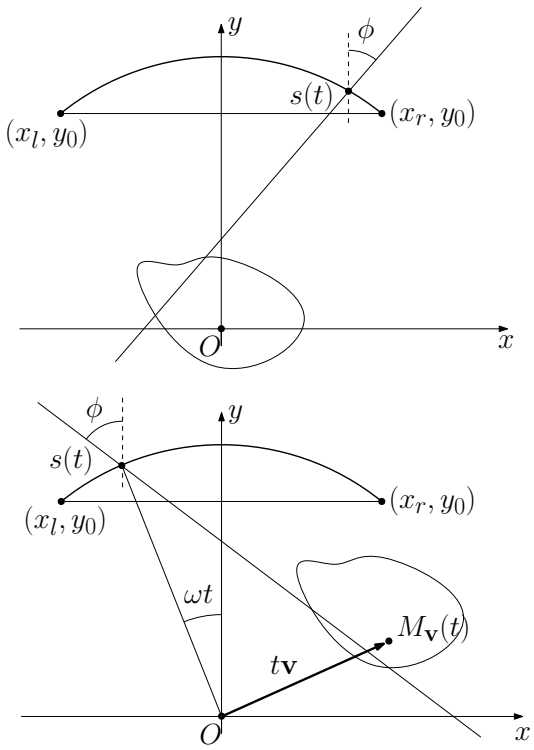


Fig. 1. Problem under consideration. The source point $s(t)$ follows the arc of circle depicted in bold. The circle has center O and radius R_0. The object is continually undergoing a translation during the movement of the source.

source will be denoted ω, and the time t will range from $-T/2$ to $T/2$, where $T > 0$. Hence, if we use $s(t)$ to denote the position of the source at time t, we have

$$s(t) = (-R_0 \sin(\omega t), R_0 \cos(\omega t)) \,. \tag{1}$$

Furthermore, we will let $s(T/2) = (x_l, y_l)$ (resp. $s(-T/2) = (x_r, y_r)$) denote the extreme left (resp. right) position of the source. Since $y_l = y_r = R_0 \cos(\omega T/2)$, we will call the common value y_0. In the following, we will assume that supp(μ) lies in the half-space $\{y < y_0\}$, and that $y_0 > 0$ (i.e.$0 < \omega T < \pi$, the total arc length is less than π).

We will suppose that at any time t, rays are simultaneously emitted from the source $s(t)$ with each ray at angle ϕ ranging from $-\pi/2$ to $\pi/2$. With this setup in mind, we can define the operator modeling the acquired data from the object.

Definition 1. *The* fan-beam projection data *of an object with density function μ is a function $(t, \phi) \mapsto \mathcal{F}\mu(t, \phi)$ defined by*

$$(\mathcal{F}\mu)(t, \phi) = \int_0^{+\infty} \mu\left(s(t) + l\left[\sin\phi, -\cos\phi\right]\right) dl, \quad (2)$$

where $t \in [-T/2, T/2]$, $\phi \in [-\pi/2, \pi/2]$ and $s(t)$ is given by (1). The operator $\mu \mapsto \mathcal{F}\mu$ is called the fan-beam projection operator.

Now let us suppose that the object is translating along a straight line with a constant velocity vector $\mathbf{v} = (v_1, v_2) \in \mathbb{R}^2$ (see Figure 1, bottom). In other words, if we let $M_{\mathbf{v}}(t)$ denote its center of mass at any time t, we have

$$M_{\mathbf{v}}(t) = M_{\mathbf{v}}(0) + (t + T/2)\,\mathbf{v} \quad (3)$$

The density function of the object now depends on both the spatial variable $\mathbf{x} \in \mathbb{R}^2$ and the time t. If we denote the time-varying object by $\mu_{\mathbf{v}}$, we have

$$\mu_{\mathbf{v}}(t, \mathbf{x}) = \mu\left(\mathbf{x} - M_{\mathbf{v}}(t)\right). \quad (4)$$

In this regard, the fan-beam projection data will be modified in the following way.

Definition 2. *The* fan-beam projection data of a translating object *with density function μ and velocity vector $\mathbf{v}$ is given by*

$$(\mathcal{F}_{\mathbf{v}}\mu)(t, \phi) = (\mathcal{F}\mu_{\mathbf{v}})(t, \phi). \quad (5)$$

The aim of this note is to derive data consistency conditions (DCCs) from (5), in order to retrieve the velocity vector $\mathbf{v}$ from the knowledge of a single element of the range of $\mathcal{F}_{\mathbf{v}}$.

B. Derivation of DCCs

In order to derive DCCs, we will first change our reference frame, from (O, x, y) to $(M(t), x', y')$, so that the origin is the center of mass of the object at any time t, and the line between the start point and the end point of the source is still parallel to the x'-axis (see Figure 2). In other words, we are performing the following change of variables

$$(x, y) \leftrightarrow (x', y') = \mathcal{R}_\beta\left((x, y) - M_{\mathbf{v}}(t)\right), \quad (6)$$

where $\mathcal{R}_\beta$ rotation by β. Note that the translation depends on t, but a single global rotation is applied. The angle β is depicted in Figure 2 (top) and is given by

$$\beta = \arctan\left(\frac{Tv_2}{2R_0 \sin(\omega T/2) + Tv_1}\right). \quad (7)$$

Note that in this equation, the denominator can be equal to zero. This situation can occur in particular cases when $v_1 < 0$. We have studied those particular cases, both in terms of physical meaning and numerical implications. For the sake of simplicity, we will assume that v_1 is not too negative (e.g. $v_1 > -2R_0/T$) to ensure that the denominator is non-zero.

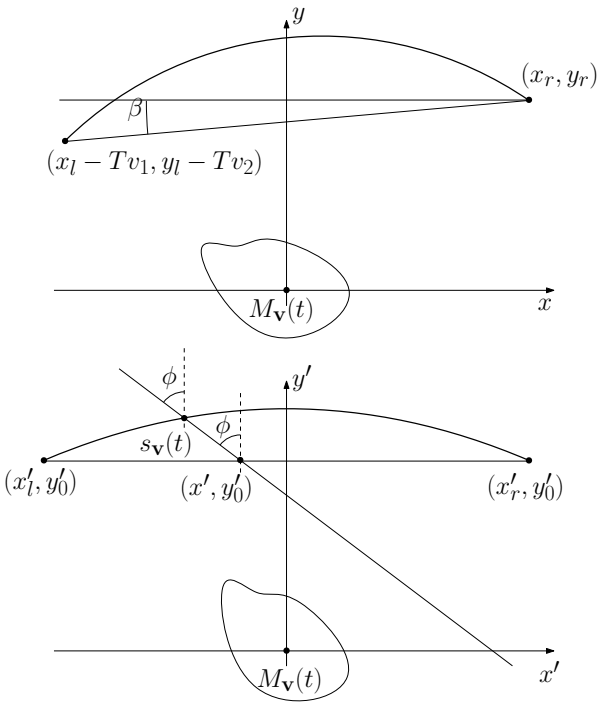


Fig. 2. Change of reference frame: the object is now at center of the coordinate system. Top: only translation of the center of frame; bottom: after rotation of angle β. The virtual source position (x', y'_0) is defined in terms of $s_{\mathbf{v}}(t)$ and ϕ; see equation (10).

In this new reference frame, the coordinates of the source position are given by $s_{\mathbf{v}}(t) = \mathcal{R}_\beta\left(s(t) - M_{\mathbf{v}}(t)\right)$. With this in mind, the data are given by the following formula

$$(\mathcal{F}_{\mathbf{v}}\mu)(t, \phi) = \int_0^{+\infty} \mu \circ \mathcal{R}_{-\beta}\left(s_{\mathbf{v}}(t) + l\left[\sin(\phi + \beta), -\cos(\phi + \beta)\right]\right) dl, \quad (8)$$

In other words, we are now dealing with a fixed object whose density function is given by $\mu \circ \mathcal{R}_{-\beta}$ irradiated by a source following an arc of a cycloid in the frame $(M(t), x', y')$ (see Figure 2, bottom).

Here, the extreme points $s_{\mathbf{v}}(-T/2)$ and $s_{\mathbf{v}}(T/2)$ have the same y'-coordinate, y'_0. We will call x'_l (resp. x'_r) the x'-coordinates of $s_{\mathbf{v}}(T/2)$ (resp. $s_{\mathbf{v}}(-T/2)$).

Now we define what we call the *virtual fan-beam projection* from a point (x', y'_0).

Definition 3. *For any point x' between x'_l and x'_r, and for any angle $\phi \in [-\pi/2, \pi/2]$, the* virtual fan-beam projection *of the object μ is defined by*

$$\left(\tilde{\mathcal{F}}\mu\right)(x', \phi') = \int_0^{+\infty} \mu \circ \mathcal{R}_{-\beta}\left((x', y'_0) + l\left[\sin\phi', -\cos\phi'\right]\right) dl. \quad (9)$$

This is called virtual *since it does not correspond to an actual position of the source.*

In the following lemma, we will make the connection between the virtual fan-beam projection and the fan-beam projection of the translating object.

Lemma 1. *Let us fix a time $t \in [-T/2, T/2]$ and an angle $\phi \in [-\pi/2, \pi/2]$. Let us define*

$$x' = s_{1,\mathbf{v}}(t) + \tan\phi\,(s_{2,\mathbf{v}}(t) - y_0'), \tag{10}$$

where, for any time t, $(s_{1,\mathbf{v}}(t), s_{2,\mathbf{v}}(t))$ are the coordinates of $s_{\mathbf{v}}(t)$. Then, we have $\left(\tilde{\mathcal{F}}\,(\mu \circ \mathcal{R}_{-\beta})\right)(x', \phi + \beta) = (\mathcal{F}_{\mathbf{v}}\mu)\,(t, \phi)$.

The idea behind the proof of lemma 1 is to note that between x' and $s_{\mathbf{v}}(t)$, the integral of μ is equal to zero since the support is assumed to remain in the half-space $\{y < y_0\}$. Hence, instead of starting the integration from x' in (9), we can start from $s_{\mathbf{v}}(t)$, which will give us (8).

We can now define the DCCs for our problem.

Theorem 1. *Let us fix a density function μ. For any integer n, there exists a function $(t, x') \mapsto W_n^{(\mathbf{v})}(t, x')$ such that the function $B_n^{(\mathbf{v})}$ defined by*

$$B_n^{(\mathbf{v})}(x') = \int_{-T/2}^{T/2} (\mathcal{F}_{\mathbf{v}}\mu)\left(t, \lambda^{(\mathbf{v})}(t, x')\right) W_n^{(\mathbf{v})}(t, x')dt, \tag{11}$$

is a polynomial of degree n. In the formula above, the angle $\lambda^{(\mathbf{v})}(t, x')$ is defined by

$$\lambda^{(\mathbf{v})}(t, x') = \arctan\left(F^{(\mathbf{v})}(t, x')\right), \tag{12}$$

where $F^{(\mathbf{v})}(t, x')$ is defined as the fraction A/B with

$$A = x' + \cos\beta\,(R_0\sin(\omega t) + (t + T/2)\,v_1) + \sin\beta\,(R_0\cos(\omega t) - (t + T/2)\,v_2) \tag{13}$$

and

$$B = \cos\beta\,(R_0\cos(\omega t) - (t + T/2)\,v_2) - \sin\beta\,(R_0\sin(\omega t) + (t + T/2)\,v_1) - y_0' \tag{14}$$

Moreover, it is possible to derive $W_n^{(\mathbf{v})}(t, x')$ analytically using the following formula

$$W_n^{(\mathbf{v})}(t, x') = \tan^n\left(\lambda^{(\mathbf{v})}(t, x')\right)\cos\left(\lambda^{(\mathbf{v})}(t, x')\right)\frac{\partial\lambda^{(\mathbf{v})}}{\partial t}(t, x') \tag{15}$$

We only have room here for an outline of the proof. The idea is to change variables in the following formula, which is known from [1] to be a polynomial in x'

$$\int_{\pi/2}^{-\pi/2} \tilde{\mathcal{F}}\mu(x', \phi)\frac{\tan^n\phi}{\cos\phi}d\phi. \tag{16}$$

The change of variable occurs between ϕ in (16) and t in (11) by using the definition in (12).

Although the formula for $W_n^{(\mathbf{v})}(t, x')$ is complicated, we observe that in the case $v_1 = v_2 = 0$, we obtain the formula (7) in [2] since

$$F^{((0,0))}(t, x') = \frac{x' + R_0\sin(\omega t)}{R_0\cos(\omega t)} \tag{17}$$

III. Numerical simulations

A. Principles

Let us suppose that we have the projections $(\mathcal{F}_{\mathbf{v}}\mu)\,(t, \phi)$. Moreover, we suppose that T is known and that the velocity $\mathbf{v}$ is constant (although unknown) during the interval $[-T/2, T/2]$. In order to recover $\mathbf{v}$, we can perform the following optimization procedure. Since $B_n^{(\mathbf{v})}(x')$ in equation (11) is supposed to be a polynomial of degree less than or equal to n, we can minimize

$$\mathcal{J}(v) = \|\left(\mathrm{res}B_n^{(\mathbf{v})}\right)\|^2 \tag{18}$$

with respect to $\mathbf{v}$, where res is the residual of the projection onto the space of polynomials of degree n or less, and n is the polynomial degree to be taken into account. This procedure will give us the velocity $\mathbf{v}$ using only the knowledge of the data $(\mathcal{F}_{\mathbf{v}}\mu)\,(t, \phi)$.

B. Application

The object under consideration is an ellipse of uniform density, whose axis lengths are 30 and 15 millimeters respectively, and making an angle of $45°$ with respect to the x-axis. The source is rotating around the object with radius $R_0 = 600\,\mathrm{mm}$, with angular velocity $\omega = 1\,\mathrm{rad}\cdot\mathrm{s}^{-1}$. The detector is a plane situated at a distance of $600\,\mathrm{mm}$ from the origin. The computations of the forward problem were performed using `simpleRTK`, a Python wrapping of `RTK` [6].

With the simulated velocity of the ellipse given by $v_1 = 0.3\,\mathrm{mm}\cdot\mathrm{s}^{-1}$, and $v_2 = 0\,\mathrm{mm}\cdot\mathrm{s}^{-1}$ we obtain the sinogram depicted in Figure 3.

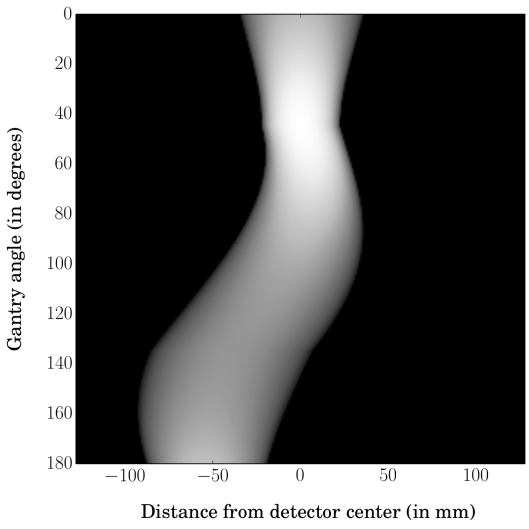


Fig. 3. Sinogram of a moving ellipse, with semi-axis $a = 30$ and $b = 15$, translating along the x-axis with velocity $\mathbf{v} = (0.3, 0)$. It is irradiated by a source at distance $R_0 = 600$, rotating with angle velocity $\omega = 1$. Gantry angle represents the angular position of the source, with respect to the x-axis ($= \omega t + \pi/2$); see Figure 1. The object translation occurs between gantry angles $45°$ and $90°$.

With this configuration in mind, the functions $x' \mapsto B_n^{(\mathbf{v})}(x')$ for $n = 0, 1, 2, 3$ were calculated from the simulated sinogram, and are illustrated in Figure 4. Note that these expressions for $B_n^{(\mathbf{v})}$, follow the predicted pattern of n^{th}-degree

polynomials, as can be seen by the fact that root mean square errors (RMSEs) between the actual values of $B_n^{(\mathbf{v})}(x)$ and their best polynomial approximations are low.

With this velocity $\mathbf{v} = (0.3, 0)$ assumed to be unknown, we performed the minimization of the cost function $\mathcal{J}(v)$ defined in (18), with a polynomial degree of $n = 2$. For this purpose, we used Powell's conjugate direction method [5], which does not require differentiation of the cost function.

We studied the influence of the time during which the object was translating, i.e.the angle α of the arc traced by the source $s(t)$ when the object was translating. As an example, for the sinogram depicted in Figure 3 we have $\alpha = 90°$. Table I summarizes all the results, where $\hat{v}$ stands for the estimated value for the velocity v_1. We observe that there is a limit below which the accuracy decreased dramatically.

TABLE I
RESULTS OF OPTIMIZATION

α	$\hat{v}$	$\|v_1 - \hat{v}\|$
95	0.300	$5.55 \cdot 10^{-17}$
90	0.300	$1.11 \cdot 10^{-16}$
85	0.300	$1.11 \cdot 10^{-16}$
84	0.300	0
83	0.300	0
82	0.376	$7.60 \cdot 10^{-2}$
81	0.415	0.115
80	0.436	0.136
75	0.436	0.136

IV. CONCLUSION

In this work, we have proposed a way to recover the velocity parameters of a translating object in fan-beam CT from a set of projections restricted to an arc of the circular trajectory. The method uses data consistency conditions (DCCs), adapted from [2]. Setting the origin of the reference frame at a point which remains still with respect to the center of mass of the object, we are in fact dealing with DCCs in the case of an arc of a cycloid. Numerical examples show that these conditions work well in the case of an object translating in a direction which is parallel to the x-axis. Further work is in progress to recover the same results for a general translation.

REFERENCES

[1] R. Clackdoyle. Necessary and sufficient consistency conditions for fanbeam projections along a line. *IEEE Transactions on Nuclear Science*, 60(3):1560–1569, 2013.

[2] R. Clackdoyle, M. Defrise, L. Desbat, and J. Nuyts. Consistency of fanbeam projections along an arc of a circle. In *The 13th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, pages 415–419, 2015.

[3] S. Helgason. *The Radon transform*. 1980.

[4] D. Ludwig. The Radon transform on euclidean space. *Communications on Pure and Applied Mathematics*, 19(1):49–81, 1966.

[5] M. Powell. An efficient method for finding the minimum of a function of several variables without calculating derivatives. *The computer journal*, 7(2):155–162, 1964.

[6] S. Rit, M.V. Oliva, S. Brousmiche, R. Labarbe, D. Sarrut, and G.C. Sharp. The reconstruction toolkit (RTK), an open-source cone-beam CT reconstruction toolkit based on the insight toolkit (itk). In *Journal of Physics: Conference Series*, volume 489, page 012079. IOP Publishing, 2014.

[7] H. Yu and G. Wang. Data consistency based rigid motion artifact reduction in fan-beam CT. *IEEE Transactions on Medical Imaging*, 26(2):249–260, 2007.

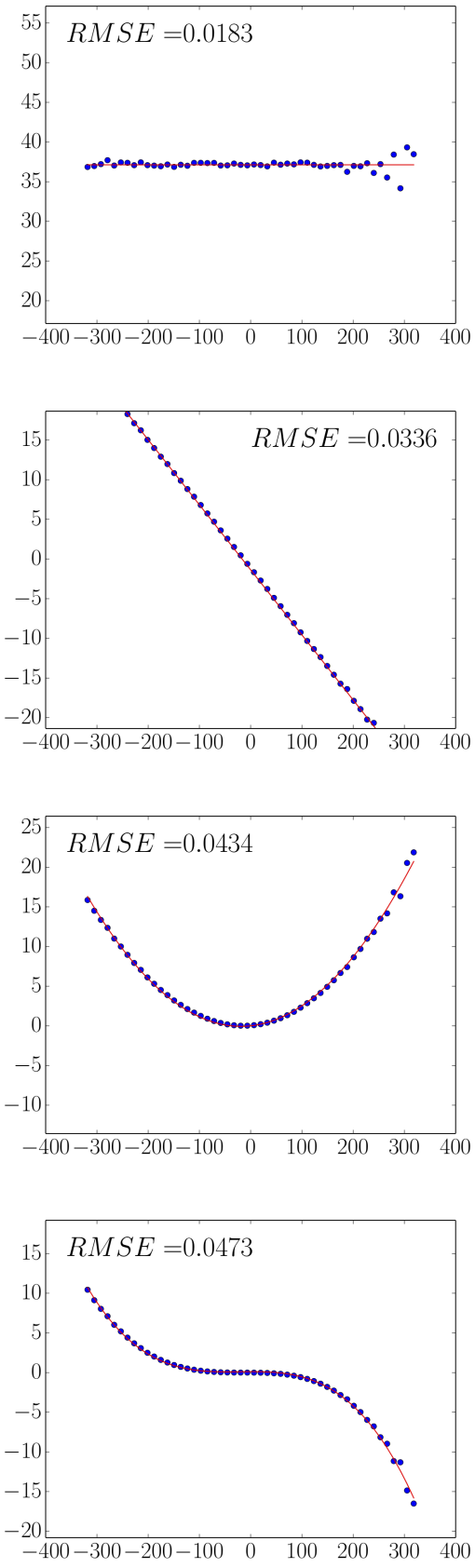


Fig. 4. From top to bottom, functions $x' \mapsto B_n^{(\mathbf{v})}(x')$ for $n = 0, 1, 2, 3$ respectively. Blue points are the actual data, while red lines are the best polynomial approximations. RMSE stands for root mean square error. The x-axis are expressed in millimeters.

Design and Evaluation of a Prototype High–Throughput Micro–CT System for In–Vivo Small Animal Imaging

Jan Kuntz, Carsten Funck, Joscha Maier, Marc Kachelrieß, Stefan Sawall

***Abstract*—Preclinical in–vivo micro–computed tomography (micro–CT) is a challenging task due to the high cardiac and respiratory rates of small animals. Commercially available micro–CT system focus on either a high spatial resolution or a high temporal resolution. This enables applications like metastasis imaging in resting bones with high spatial resolution or imaging of extensive lung tumors with a high temporal resolution. However, it prohibits the visualization of small, rapidly moving structures, e.g. micro–metastases of the lung. To visualize such structures usually long scan times are required to account for the demanding high spatial and temporal resolution and the consequently high image noise, thus preventing to examine large cohorts of animals. To overcome these issues we herein present the design and first in–vivo evaluations of a novel, gantry–based, high–throughput micro–CT system providing highest spatial and temporal resolution within only seconds of scan time.**

I. Introduction

Over the last decades micro–computed tomography (micro–CT) has evolved as an essential tool in preclinical research. Ex–vivo measurements with spatial resolutions exceeding 10 µm are readily available in practice enabling e.g. the evaluation of trabecular structures in bones [1], micro-metastases or skeletal phenotyping [2]. In–vivo imaging of small animals, in particular of large cohorts, remains challenging for a variety of reasons. First, as expected by allometric estimates, small animals show high respiratory rates of up to 300 respirations per minute (rpm) and high cardiac rates of up to 600 beats per minute (bpm) [3]. Second, structures of interest, e.g. tumors, metastases or vessels, usually exhibit sizes of only a few micrometers [4]. Third, scan times are limited to no more than a few minutes as animals cannot be held under anesthesia for longer time periods. To overcome these issues at least partially, most micro–CT systems focus on either imaging with a high spatial resolution and low temporal resolution, e.g. sufficient to visualize metastases of resting bones, or vice versa, e.g. sufficient to visualize extensive lung tumors. For example, systems with a spatial resolution of 32 µm and a detector frame rate of only 7 fps have been described [5].

Dr. Jan Kuntz, Carsten Funck, Joscha Maier, Prof. Dr. Marc Kachelrieß, and Dr. Stefan Sawall: Division of X–Ray Imaging and CT, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, Heidelberg, Germany. Dr. Jan Kuntz, Carsten Funck, Prof. Dr. Marc Kachelrieß, and Dr. Stefan Sawall: Medical Faculty, Ruprecht-Karls-University, Im Neuenheimer Feld 672, Heidelberg, Germany. Joscha Maier: Department of Physics and Astronomy, Ruprecht-Karls-University, Im Neuenheimer Feld 226, Heidelberg, Germany.

Corresponding author: j.kuntz@dkfz.de

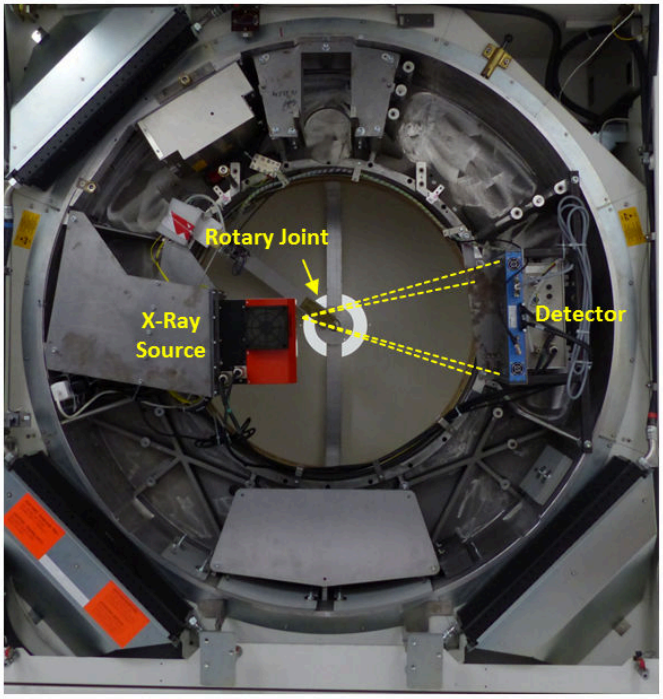


Fig. 1. Photo of our micro–CT showing the micro–focus x–ray source, the flat detector and the rotary joint mounted in the gantry. The yellow, dashed lines indicate the cone–beam geometry of the system and the green dashed lines enclose the rotating part of the gantry while all other components remain stationary.

In contrast, systems with a spatial resolution of only 210 µm but a detector frame rate of 100 fps exist as well [6]. Thus, high spatial resolution and high temporal resolution seem to be mutually exclusive and the imaging of small rapidly moving structures, e.g. micro–metastases of the lung, appears to be impossible. This is further complicated by the fact that a high spatial resolution typically results in elevated image noise. A consequent increase in tube output, however, inevitably results in an increase in the focal spot size, thus diminishing spatial resolution and a trade–off has to be found. Furthermore, a limited tube power will increase measurement time, usually in the order of minutes, and thus increase the burden to the animals due to anesthesia. This applies in particular if phase–correlated images are desired and external gating devices, e.g. ECG–electrodes, have to be attached to the animals which is a time–consuming process prior to any data acquisition. Similar reasoning applies to an increase in

temporal resolution by an increase in detector frame rate or an increase in spatial resolution by decreasing the detector pixel size. To overcome these issues we herein present the design and first in–vivo evaluation of a novel custom–made, gantry–based prototype micro–CT system that provides a high spatial and high temporal resolution with scan times in the order of seconds allowing for a high throughput and hence the in–vivo imaging of large cohorts.

II. Materials and Methods

A. Mechanical and Electrical Design

The proposed high–throughput micro–CT prototype is based on a refurbished clinical CT gantry originating from a Somatom Volume Zoom clinical CT (Siemens Healthcare, Forchheim, Germany) with a bore diameter of about $50\,\mathrm{cm}$, a rotating part with a diameter of about $162\,\mathrm{cm}$ and a total size of $199 \times 240 \times 90\,\mathrm{cm}^3$ (see figure 1). This mechanical framework allows for rapid rotations with a speed of more than 1 revolution per second and provides sufficient space and payload to house all components with the desired geometric relations, providing a sufficient magnification. While the system currently only contains a single x–ray source and detector, further upgrades are possible. In particular, the field of measurement (FOM) can be increased using a dual detector setup or a second x–ray source–detector combination can be added to obtain a dual–source CT, allowing for a reduction in scan time or for dual energy examinations. Furthermore, the rotating gantry allows placing the sample or animal under investigation in a prone position in the center of rotation. This enables easy injection of drugs like contrast agents, the administration of anesthesia and prevents a compression of the animal which is often observed in table–top micro–CT systems, i.e. micro–CT systems with stationary x–ray source and detector and a rotating sample. To reduce vibrations and mechanical inaccuracies due to imbalances of the rotating part and in consequence the loss of spatial resolution, the system is balanced using a commercially available balancing system (Caroba Balancer Pro, Präzisionsmaschinenbau Bobertag, Kaiserslautern, Germany). The power supply for all components in the gantry is realized using the slip rings where both, $230\,\mathrm{V}$ AC and $24\,\mathrm{V}$ DC, are transmitted to the rotating part. Furthermore, safety-related signals, e.g. x–ray interlock and the RS232 used to control the x–ray source, are also routed over the slip rings. Further control signals, 10 GBit Ethernet and all detector data, e.g. up to $400\,\mathrm{MB/s}$ in 1×1 binning mode, is transmitted using a optical rotary joint with 20 channels (Spinner GmbH, Munich, Germany). I.e., all detector signals are converted from a Camera Link interface to optical fiber, coupled out using the rotary joint and converted to Camera Link again before they are fed to the frame grabber. The usage of high bandwith optical transmission enables the extension of the system, e.g. with a photon–counting detector, and allows for the transmittance of even higher data rates, e.g. if multiple energy bins are available at high frame rates.

B. X–Ray Source and Detector

The x–ray source is a micro–focus transmission tube (L10951, Hamamatsu Photonics K. K., Shimokanzo, Iwata City, Japan), providing a tube voltage of up to $110\,\mathrm{kV}$ and a tube current of up to $800\,\mu\mathrm{A}$ at $50\,\mathrm{W}$. The power–dependent focal spot size varies between $15\,\mu\mathrm{m}$ at $6\,\mathrm{W}$ and $80\,\mu\mathrm{m}$ at $50\,\mathrm{W}$. The x–ray detector used is a Dexela 2923 MAM CMOS detector (Perkin Elmer, A Varex Company, Salt Lake City, USA), equipped with a $150\,\mu\mathrm{m}$ high–resolution CsI scintillator. The detector provides a matrix size of 3888×3072 pixels with a size of $74.8 \times 74.8\,\mu\mathrm{m}^2$. To account for the high cardiac and respiratory rates of small animals, the detector achieves a frame rate of up to $85\,\mathrm{fps}$, $70\,\mathrm{fps}$ and $26\,\mathrm{fps}$ in 4×4, 2×2 and 1×1 binning modes, respectively. The x–ray source is mounted into the gantry with a focus–isocenter–distance R_F of $90\,\mathrm{mm}$ while the detector–isocenter–distance R_D is $500\,\mathrm{mm}$. Thus, the FOM has a diameter of about $40\,\mathrm{mm}$. Given these components the spatial resolution Δ of the system can be estimated using

$$\Delta^2 = \left(\frac{R_\mathrm{D}}{R_\mathrm{F} + R_\mathrm{D}} W_\mathrm{F}\right)^2 + \left(\frac{R_\mathrm{F}}{R_\mathrm{D} + R_\mathrm{F}} W_\mathrm{D}\right)^2, \qquad (1)$$

with W_F and W_D being the used focus size and the detector aperture, respectively [7]. Thus, the ideal resolution is about $49\,\mu\mathrm{m}$, $28\,\mu\mathrm{m}$, and $20\,\mu\mathrm{m}$ for the said binning modes assuming a focal–spot size of $20\,\mu\mathrm{m}$. For the smallest available focus size the ideal resolution would be $46\,\mu\mathrm{m}$, $24\,\mu\mathrm{m}$, and $13\,\mu\mathrm{m}$. This seems adequate for in–vivo experiments with high temporal and spatial resolution. The geometric calibration of the system currently is performed on a per–acquisition basis using Meng's method, i.e. an intrinsic calibration procedure using the acquired rawdata themselves to estimate the used geometry [8].

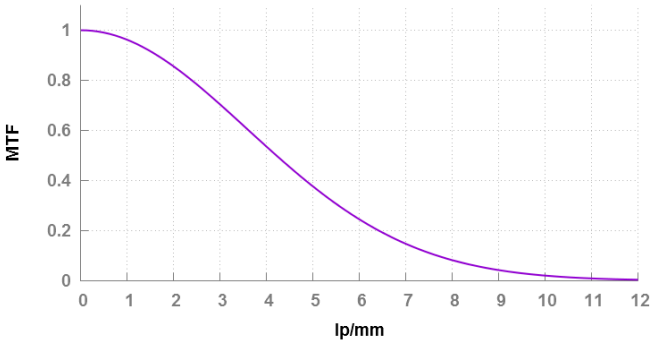


Fig. 2. Modulation transfer function (MTF) of the system using a scan protocol with 4×4 binning mode, $60\,\mathrm{kV}$ and $800\,\mu\mathrm{A}$.

C. Animal Experiments

While a variety of phantom measurements have been performed to evaluate the system, e.g. the spatial resolution and MTF were measured using a resolution pattern and the CT–value stability has been investigated using a water phantom, we herein will focus on in–vivo measurements and illustrate the high–throughput capabilities of the system. All animal experiments were approved by the local committee on animal welfare (G-256/15) and were performed in accordance with the guidelines issued by the Federation of European Laboratory

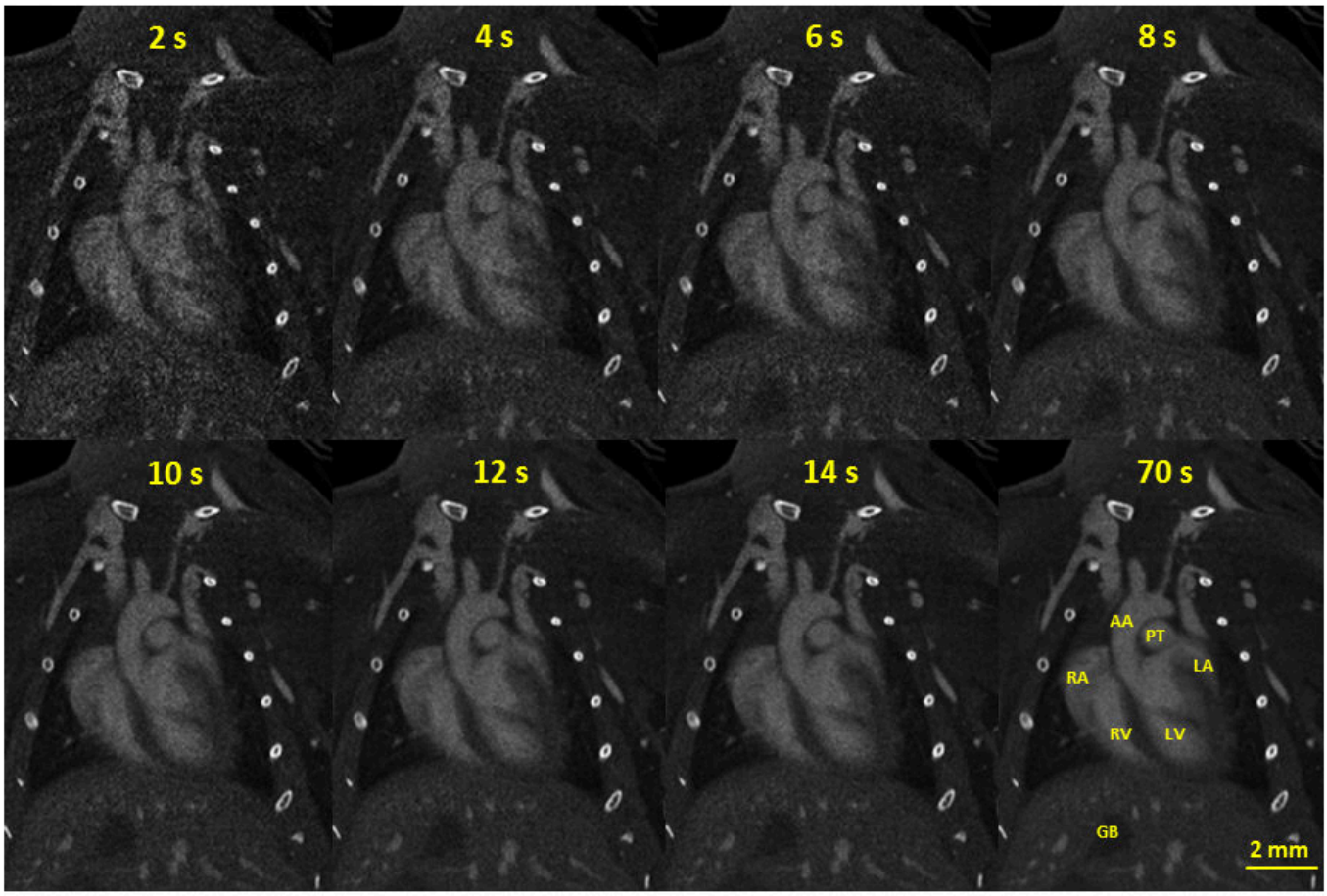


Fig. 3. Coronal slices of reconstructions of mouse one obtained after different measurement times, i.e. after 2 s to 14 s in steps of 2 s and after 70 s. The anatomical structures are labelled as follows: AA–aortic arch, RA–right atrium, LA–left atrium, RV–right ventricle, LV–left ventricle, PT–pulmonary trunk, GB–gallbladder. (C/W = 0 HU/1500 HU)

Animal Science Associations (FELASA) [9]. In particular, four mice with cardiac rates of 338 ± 78 bpm and respiratory rates of 141 ± 34 rpm were administered with 100 µL of a blood pool contrast agent (ExiTron nano 12000, nanoPET Pharma, Berlin, Germany) and measured using a tube voltage of 60 kV, a tube current of 800 µA in 4×4 binning. To limit the amount of angular blurring due to the detector frame rate of 11.7 ms the rotation speed of the gantry was limited to 4 s per revolution. All animals were scanned for 70 s with a total amount of 5,950 projections. However, to illustrate the image quality achievable with short scan times we also reconstruct fractions of the data corresponding to a scan time of 2 s to 14 s in steps of 2 s or 180° of data, respectively. All images were reconstructed using a high–performance implementation of the Feldkamp–algorithm (RayConStruct®–IR, RayConStruct, Nürnberg, Germany) on a sufficiently sized grid with a voxel size of 49 µm [10].

III. Results

A. Spatial Resolution

Figure 2 shows the modulation transfer function estimated using a line pattern phantom for the scan protocol used throughout this abstract. We observe a value of $\rho_{50\,\%} = 4.0$ lp/mm, $\rho_{10\,\%} = 7.5$ lp/mm, and $\rho_{2\,\%} = 10.0$ lp/mm. Note that the ideal resolution estimated using eq. (1) is slightly better than what was measured herein. This most likely is caused by the fact that eq. (1) does not consider angular blurring, mechanical inaccuracies of the gantry and the influence of the reconstruction.

B. Animal Experiments

Figure 3 shows the reconstructions of one of the mice for a variety of acquisition ranges. In particular, reconstruction results are illustrated for fractions of the acquired data corresponding to measurement times of 2 s to 14 s in steps of 2 s and a reconstruction after a measurement time of 70 s. Note that the unsharpness observed in certain parts of the image is caused by the fact that the shown reconstructions are not phase–correlated, i.e. show an average of all cardiac and respiratory motion states of the animal. The reconstruction obtained after 2 s shows moderate image quality with a noise level of about 450 HU measured in soft tissue and does not satisfy diagnostic image quality criteria. However, all major anatomical structures, e.g. the aortic arch or the pulmonary trunk, are visible and no streaks are apparent due to the high frame rate of the detector and the dense angular sampling.

The reconstruction obtained after a measurement time of only 4 s already shows sufficient diagnostic quality, i.e. all major anatomical structures and small vessels, e.g. small liver veins, can be easily identified and the noise level of 350 HU is well within the typical noise range of micro–CT examinations. Image noise gradually decreases with increasing measuring time as expected, reaching a noise level of 190 HU after 10 s. All major anatomical structures are labeled in the reconstruction obtained after 70 s for reference, which is nearly free of noise. Considering the fact that all reconstructions have been performed using a standard image reconstruction algorithm indicates that the usage of e.g. iterative reconstruction algorithms might allow for a further decrease in measurement time while maintaining the desired spatial resolution.

IV. Conclusion and Discussion

We herein illustrated the high–speed capabilities of a novel, gantry–based micro–CT system. A scan time of only four seconds results in images with sufficient diagnostic quality for most cases, thus minimizing stress to the animal caused by anesthesia. Hence, we estimate that at least 30 mice could be examined per hour, assuming animal preparation can be achieved in 2 min or less, e.g. by anesthetizing one animal while the other is imaged. This will allow for the imaging of large cohorts of animals emphasizing the high–throughput capabilities of the developed system. Longer scan times might be used to acquire data suitable for a phase–correlated reconstruction or for scans with a higher spatial resolution. However, recent publications illustrated that the data acquired within a few seconds already might be sufficient for cardiac– and respiratory–gated reconstructions, if sophisticated iterative methods [3], [11], [12] or motion–compensation methods are used [13], [14] and a sufficient number of cardiac and respiratory cycles are acquired. A further decrease in measurement time can be achieved if the time consuming installation of an ECG is replaced by software–based gating methods [6], [15]. While the preliminary results presented herein are promising towards a high–throughput micro–CT scanner with satisfying spatial and temporal resolution, future research will be concerned with a dose optimization of this system. This is particularly relevant in scans with higher spatial resolution, e.g. employing the 2×2 binning mode of the detector with a theoretical spatial resolution of up to 24 µm, and will be required to allow not only for the imaging of large cohorts but for the repeated imaging of these cohorts in longitudinal studies.

Acknowledgements

The high–speed image reconstruction RayConStruct®–IR was provided by RayConStruct GmbH, Nürnberg, Germany. Parts of this work were supported by the Deutsche Forschungsgemeinschaft (DFG) under grant SA 2776/1-1. We thank nanoPET Pharma GmbH, Berlin, Germany for providing their latest blood pool agent and MEDSER Medical Services, Heusenstamm, Germany for their excellent support concerning the refurbished clinical CT gantry.

References

[1] F. Peyrin, M. Salome, P. Cloetens, A. M. Laval-Jeantet, E. Ritman, and P. Ruegsegger, "Micro-CT examinations of trabecular bone samples at different resolutions: 14, 7 and 2 micron level," *Technol Health Care*, vol. 6, no. 5-6, pp. 391–401, 1998.

[2] J. Silvent, A. Akiva, V. Brumfeld, N. Reznikov, K. Rechav, K. Yaniv, L. Addadi, and S. Weiner, "Zebrafish skeleton development: High resolution micro-CT and FIB-SEM block surface serial imaging for phenotype identification," *PLoS ONE*, vol. 12, no. 12, p. e0177731, 2017.

[3] S. Sawall, F. Bergner, R. Lapp, M. Mronz, M. Karolczak, A. Hess, and M. Kachelrieß, "Low-dose cardio-respiratory phase-correlated cone-beam micro-CT of small animals," *Medical Physics*, vol. 38, no. 3, pp. 1416–1424, 2011.

[4] E. L. Ritman, "Current status of developments and applications of micro-CT," *Annu Rev Biomed Eng*, vol. 13, pp. 531–552, 2011.

[5] X. Yang, Y. Meng, Q. Luo, and H. Gong, "High resolution in vivo micro-CT with flat panel detector based on amorphous silicon," *J Xray Sci Technol*, vol. 18, no. 4, pp. 381–392, 2010.

[6] J. Dinkel, S. H. Bartling, J. Kuntz, M. Grasruck, A. Kopp-Schneider, M. Iwasaki, S. Dimmeler, R. Gupta, W. Semmler, H.-U. Kauczor, and F. Kiessling, "Intrinsic gating for small–animal computed tomography: A robust ECG–less paradigm for deriving cardiac phase information and functional imaging," *Circulation: Cardiovascular Imaging*, vol. 1, no. 3, pp. 235–243, 2008.

[7] W. A. Kalender, *Computed Tomography*. Wiley & Sons, 2005, ISBN 3–89578–216–5, 2nd Edition.

[8] Y. Meng, H. Gong, and X. Yang, "Online geometric calibration of cone-beam computed tomography for arbitrary imaging objects," *IEEE Trans Med Imaging*, vol. 32, no. 2, pp. 278–288, 2013.

[9] M. Mahler Convenor, M. Berard, R. Feinstein, A. Gallagher, B. Illgen-Wilcke, K. Pritchett-Corning, and M. Raspa, "FELASA recommendations for the health monitoring of mouse, rat, hamster, guinea pig and rabbit colonies in breeding and experimental units," *Lab. Anim.*, vol. 48, no. 3, pp. 178–192, 2014.

[10] L. Feldkamp, L. Davis, and J. Kress, "Practical cone–beam algorithm," *Journal of the Optical Society of America*, vol. 1, no. 6, pp. 612–619, 1984.

[11] L. Ritschl, F. Bergner, C. Fleischmann, and M. Kachelrieß, "Improved total variation-based CT image reconstruction applied to clinical data," *Phys Med Biol*, vol. 56, no. 6, pp. 1545–1561, 2011.

[12] L. Ritschl, S. Sawall, M. Knaup, A. Hess, and M. Kachelrieß, "Iterative 4D cardiac micro-CT image reconstruction using an adaptive spatio-temporal sparsity prior," *Phys Med Biol*, vol. 57, no. 6, pp. 1517–1525, 2012.

[13] M. Brehm, P. Paysan, M. Oelhafen, and M. Kachelrieß, "Artifact-resistant motion estimation with a patient-specific artifact model for motion-compensated cone-beam CT," *Med Phys*, vol. 40, no. 10, p. 101913, 2013.

[14] M. Brehm, S. Sawall, J. Maier, S. Sauppe, and M. Kachelrieß, "Cardiorespiratory motion-compensated micro-CT image reconstruction using an artifact model-based motion estimation," *Med Phys*, vol. 42, no. 4, pp. 1948–1958, 2015.

[15] J. Kuntz, J. Dinkel, S. Zwick, T. Bauerle, M. Grasruck, F. Kiessling, R. Gupta, W. Semmler, and S. H. Bartling, "Fully automated intrinsic respiratory and cardiac gating for small animal CT," *Phys Med Biol*, vol. 55, no. 7, pp. 2069–2085, 2010.

Performance Evaluation of a Piecewise-linear Dynamic Attenuator

Picha Shunhavanich and Norbert J. Pelc

***Abstract*—Photon counting detectors (PCDs) offer spectral imaging capability and improved image quality but can suffer from limited count rate. One approach to reducing the required count rate, a piecewise-linear dynamic bowtie filter (Hsieh and Pelc, *Med Phys* 2013), provides flexible flux control by moving its wedge components, positioned at different fan angles, along the axial direction to adjust the thickness in an axial slice. We previously implemented a full-size piecewise-linear attenuator for a table-top system. Here, our bowtie performance in minimizing a required maximum count rate for PCDs is investigated. As compared to a conventional bowtie filter, simulation results suggest an average reduction of 6.77x in max count rate and 1.71x in dose for the tested chest, abdomen, and shoulder studies.**

***Index Terms*—dynamic bowtie filter, max count rate, photon counting detector**

I. Introduction

Photon counting detectors (PCDs) offer the possibility of improved imaging capabilities including spectral detection, increased spatial resolution, higher detective quantum efficiency, and lower electronic noise. An inherent challenge to deploying PCDs is their slow count rate, resulting in lost counts and pulse pileup. A promising approach to reduce this effect is to modulate the x-ray flux with a dynamic bowtie filter, reducing the flux of photons incident on the detectors. Unlike conventional filters, the attenuation of dynamic bowtie filters can be adjusted dynamically according to patient attenuation as the CT scanner rotates.

Among proposed bowtie designs [1-6], the piecewise-linear attenuator [1-2] offers a flexible smooth profile and had promising results with practical motion control. It consists of multiple wedges, producing a piecewise-linear thickness along fan angle, with the attenuation profile present in an axial slice controlled by moving the wedges in the axial direction.

This work was supported in part by the National Institutes of Health Grant U01EB017140 and in part by Anandamahidol Foundation.

P. Shunhavanich is with Bioengineering and Radiology Departments, Stanford University, Stanford, CA 94305 USA (email: picha@stanford.edu)

N. J. Pelc is with Bioengineering and Radiology Departments, Stanford University, Stanford, CA 94305 USA (email: pelc@stanford.edu)

Previously, we implemented a full-size piecewise-linear attenuator [7]. In this work, our bowtie performance in minimizing a required maximum count rate for PCDs is investigated in simulation.

II. Bowtie implementation

The detail of our implemented piecewise-linear bowtie filter can be found in [7]. Briefly, it has 13 triangular stainless steel wedges, each with thickness linearly decreasing from 6mm to 0mm toward the tip, as shown in Fig. 1a. The bowtie cross-section is displayed in Fig. 1b, and its perspective front view is shown in Fig. 1c. Wedges are aligned in this way to produce a piecewise-linear thickness profile vs. fan angle while avoiding collisions.

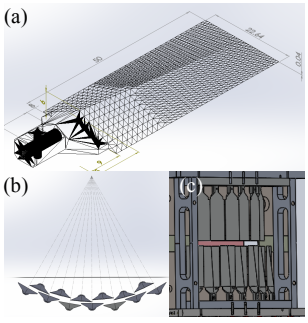


Fig. 1. Bowtie wedge design (a), bowtie cross-section (b), and bowtie in perspective front view (c).

Fig. 2 is a photo of the actual bowtie. It was designed for a table-top system that is being assembled and will include a PCD. The aspect explored here is how the wedge and mA modulation should be controlled to minimize peak count rate, and the expected benefit.

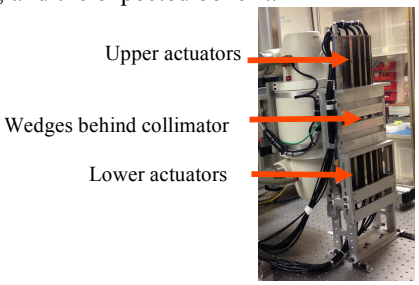


Fig. 2. The implemented full-size piecewise-linear attenuator.

III. Optimization

Tube current (mA) modulation and bowtie wedge positions were determined by minimizing dynamic range, (ratio of maximum to minimum detected counts). With a

controlled mean count, this is equivalent to minimizing maximum count rates and was employed due to its accessible formula. One approach to minimizing dynamic range while ensuring uniform noise level is to make the total attenuation of bowtie plus object flat. Therefore, an ideal target attenuation profile for the bowtie can be computed from subtracting the object sinogram from its maximum value. The optimization algorithm employed is similar to our previous work [7], with more detail on variable bounds described here.

We seek to minimize the difference between the ideal target bowtie attenuation profile, f_{target}, and the actual bowtie profile, f_{actual}, with the problem written as

$$d(\beta,\gamma) = f_{target}(\beta,\gamma) - r(\beta) - f_{actual}(\beta,\gamma)$$
$$r(\beta) = -\log\big(mA(\beta)\big)$$

$$\text{Minimize} \quad \max\big(d(\beta,\gamma)\big) - \min\big(d(\beta,\gamma)\big) + \alpha\sqrt{\frac{1}{n}\sum d(\beta,\gamma)^2} \tag{1}$$

where β and γ are view and fan angles, respectively. $d(\beta,\gamma)$ is the difference between the original target $f_{target}(\beta,\gamma)$, adjusted by the mA modulation term $mA(\beta)$, and the actual bowtie line integral $f_{actual}(\beta,\gamma)$. The objective function is the sum of log dynamic range, $\max\big(d(\beta,\gamma)\big) - \min\big(d(\beta,\gamma)\big)$, and α-weighted root mean square error, $\sqrt{\frac{1}{n}\sum d(\beta,\gamma)^2}$. n is the number of sinogram pixels within the object boundary, and α was experimentally chosen to be 10.

A. mA modulation

mA modulation helps lowering the required bowtie attenuation and preventing tube overheating. In this simulation, we assume the same bowtie dimension as our actual implementation, which has ~5.5 mm of stainless steel as maximum attenuation.

Assuming the mA modulation is sinusoidal as a function of view angle with π period, it can be written as

$$mA(\beta) = \frac{a-1}{2}\sin\left(2\beta - \frac{\pi}{2}\right) + \frac{a+1}{2}$$

where a is a ratio of peak mA along x axis to y axis, and this function is a scaled version of the actual mA.

To calculate a probable range of a, we first calculated the desired range of $r(\beta)$ that minimizes the log dynamic range, $L = \max\big(d(\beta,\gamma)\big) - \min\big(d(\beta,\gamma)\big)$. Let c_{max} be the max line integral of the implemented bowtie (~3.82),

$$0 \le f_{actual}(\beta,\gamma) \le c_{max}.$$

If the adjusted target, $f_{target}(\beta,\gamma) - r(\beta)$, exceeds c_{max} by x (e.g. c_{max}+x) or if it is negative (e.g. –x) both result in the same increase (e.g. x) in the log dynamic range L. Therefore, the optimal $r(\beta)$ could be within the following bounds,

$$r_{Lower}(\beta) = v(\beta) - \max\left(0, v(\beta) - (u(\beta) - c_{max})\right)$$
$$r_{Upper}(\beta) = v(\beta) - \min\left(0, v(\beta) - (u(\beta) - c_{max})\right)$$

where $u(\beta) = \max_{\gamma} f_{target}(\beta,\gamma)$, and $v(\beta) = \min_{\gamma} f_{target}(\beta,\gamma)$ are max and min target in view angle β, respectively.

Therefore, the possible range of an optimized a is

$$1 \le a \le \exp\Big(\max\big(r_{Upper}(\beta)\big) - \min\big(r_{Lower}(\beta)\big)\Big)$$

B. Wedge positions

In opitimization, the bowtie profile f_{actual} was assumed to be piecewise-linear in line integral with formula following Ref. [1]. The optimzation variables are line integral at the center of bowtie wedge for each wedge in each view.

C. Implementation

To solve the optimization problem (1), different values of a spanning the possible range were tested, and for each, the problem (1) was cast as a convex optimization problem. The CVX convex optimization package [8, 9] was used.

IV. Simulation

A. Scanner

The CT scanner was assumed to have a fan-beam geometry and either a gadolinium oxysulfide energy integrating or a photon counting detector with 1mm^2 elements. The X-ray rube was assumed to have 120kVp with spectrum computed from [10, 11]. The transmission profile of the conventional bowtie was retrieved from CATSIM software.

B. Testing methods

Due to inaccessibility of the vendor scanner projection data, the projection data were reproduced in simulation. Forward projection of chest, abdomen, and shoulder CT images were used as test data. These images were acquired by scanning a Kyoto Kagaku lung cancer screening CT phantom LSCT0001 with GE Discovery CT750 HD, using a clinical protocol for routine imaging (Noise index = 37). The mA along x and y axes selected by the scanner were recorded.

Variance maps of reconstructed images were estimated with the method described in [12]. The computed peak variances were assumed to be the clinically required image noise level that the scans need to obtain.

To evaluate the performance of our piecewise-linear bowtie as compared to the conventional filter, an optimized mA modulation (with weighted sum of max count rate and dose as objective) was applied for the conventional filter, and the proposed optimization for mA modulation and wedge positions (Section III) were employed to the piecewise-linear bowtie. Normalized to achieve the clinical image peak variance, the max count rate and a surrogate for dose (sum of energy incident on objects) were computed and compared.

V. RESULTS

The computed noise maps (or images of noise) are displayed in Fig. 3, with more uniformity observed for the piecewise-linear bowtie (Fig. 3d, e, f). The resulting max count rate for our piecewise-linear bowtie filter are 11.18, 7.62, and 13.08 Mcps/mm^2 (million counts per second per square millimeter) for chest, abdomen, and shoulder, respectively, which are ~6.77 times lower than those for the conventional filter on average, as shown in Table I. Our bowtie filter also reduces the dose by a factor of 2.14, 1.41, and 1.57 for chest, abdomen, and shoulder, respectively, as compared to the conventional bowtie filter.

Table I. Maximum count rate (Mcps/mm^2)

	Conventional bowtie	Piecewise-linear bowtie
Chest	53.79	11.18
Abdomen	55.06	7.62
Shoulder	108.34	13.08

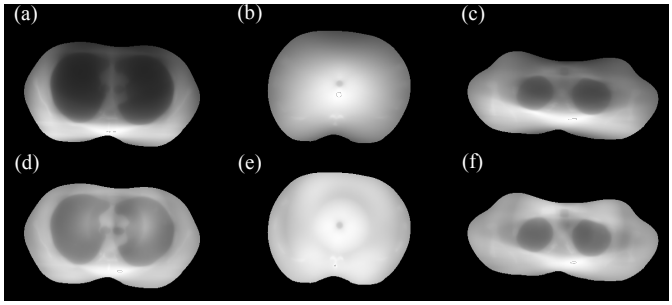


Fig. 3. Noise maps (images of noise) of chest, abdomen, and shoulder phantoms for the conventional bowtie filter (a, b, c) and for the piecewise-linear bowtie filter (d, e, f).

VI. CONCLUSION

For the same peak noise level and scan time, our piecewise-linear bowtie filter could lower the detector maximum count rates to around 10 Mcps/mm^2, which is a reasonable functioning level for current photon counting detectors. This could be improved further by increasing the maximum bowtie wedge thickness, especially for more dose reduction. However, the trade-off between max wedge thickness and the precision control of wedge attenuation should be taken into consideration, due to potential issues with actuator position control.

REFERENCES

[1] S. S. Hsieh and N. J. Pelc, "The feasibility of a piecewise-linear dynamic bowtie filter," *Medical physics*, 40(3), 031910 (2013).

[2] S. S. Hsieh, M. V. Peng, C. A. May., P. Shunhavanich, D. Fleischmann, and N. J. Pelc, "A prototype piecewise-linear dynamic attenuator," *Physics in medicine and biology*, *61*(13), 4974 (2016).

[3] T. P. Szczykutowicz, and C. A. Mistretta, "Design of a digital beam attenuation system for computed tomography: Part I. system design and simulation framework," *Medical physics*, 40(2), 021905 (2013).

[4] T. P. Szczykutowicz, and C. A. Mistretta, "Design of a digital beam attenuation system for computed tomography. Part II. performance study and initial results," *Medical physics*, 40(2), 021906 (2013).

[5] J. W. Stayman, A. Mathews, W. Zbijewski, G. Gang, J. Siewerdsen, S. Kawamoto, I. Blevis, and R. Levinson, "Fluence-field modulated x-ray CT using multiple aperture devices," *Medical Imaging 2016: Physics of Medical Imaging*, 9783, 97830X, International Society for Optics and Photonics (2016).

[6] A. J. Mathews, G. Gang, R. Levinson, W. Zbijewski, S. Kawamoto, J. H. Siewerdsen, and J. W. Stayman, "Experimental evaluation of dual multiple aperture devices for fluence field modulated x-ray computed tomography," *Medical Imaging 2017: Physics of Medical Imaging*, 10132, 101322O, International Society for Optics and Photonics (2017).

[7] P. Shunhavanich, N. R. Bennett, S. S. Hsieh, and N. J. Pelc, "Implementation of a Piecewise-linear Dynamic Attenuator," *Medical Imaging 2018: Physics of Medical Imaging*, 10573, 105730T, International Society for Optics and Photonics (2018).

[8] M. Grant and S. Boyd, CVX: MATLAB software for disciplined convex programming, version 2.1, http://cvxr.com/cvx (June 2015).

[9] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," *Recent advances in learning and control*, 95-110 (2008).

[10] Online tool for the simulation of X-ray Spectra (available URL: https://www.oem-xray-components.siemens.com/x-ray-spectra-simulation).

[11] J. M. Boone and J. A. Seibert, "An accurate method for computer-generating tungsten anode x-ray spectra from 30 to 140 kV," *Medical physics*, *24*(11), 1661-1670 (1997).

[12] D. A. Chesler, S. J.Riederer, and N. J. Pelc, "Noise due to photon counting statistics in computed x-ray tomography," *Journal of computer assisted tomography*, 1(1), 64-74 (1977)

Implementation and Assessment of Dynamic Fluence Field Modulation with Multiple Aperture Devices

Grace J. Gang, Andrew Mao, Jeffrey H. Siewerdsen, and J. Webster Stayman

***Abstract*—This work reports experimental results of dynamic fluence field modulation (FFM) using a dual multiple aperture devices (MAD) system. MAD filters use Moiré patterns produced by relative motions between two sets of thin, highly attenuating tungsten bars of varying widths and spacings. Each MAD was affixed to a linear actuator and installed on an experimental cone-beam CT bench. Phantom-specific FFM profiles were designed based on a flat detected fluence and a minimum mean variance objective and realized through a combination of MAD translations and pulse width modulation at a constant tube current. To properly correct for gains associated with the MAD filters, a correction algorithm was designed to account for focal spot shifts during scanning, as well as spectral effects from incomplete blockage of x-rays by the tungsten bars. The FFM designs were demonstrated in an elliptical phantom (25.8 $\times$ 14.1 cm) and a CatPhan module with wires and cylindrical inserts. Variance and noise power spectrum (NPS) analysis was performed on the resulting reconstructions. While conventional gain correction produced reconstructions with high frequency ring artifacts in axial slices, the proposed correction algorithm effectively removed such artifacts while preserving phantom details. Fluence field designs for the elliptical phantom were achieved using relative MAD motions over a 0.44 mm range, and measured beam profiles closely approximated the theoretically computed target profiles. The noise properties of the resulting reconstructions behave as expected: a flat detected fluence criterion yields nearly isotropic NPS and more homogeneous variance across the reconstruction as compared to an unmodulated scan; the minimum mean variance FFM results in lower mean variance compared to both the unmodulated and flat detected fluence acquisition at approximately matched total bare-beam fluence. These results suggest that a dual-MAD filter is an effective approach to provide fluence and image quality control in CT and can potentially accommodate a wide range of phantoms and design objectives.**

***Keywords*—dynamic fluence field modulation, multiple aperture device, corrections, image quality**

I. Introduction

Dynamic fluence field modulation (FFM) has been proposed as a dose reduction strategy in computed tomography (CT) by adaptively changing x-ray beam shape as a function of both the horizontal detector locations (u) and for each projection angle (θ). [1] Compared to static bowties, FFM not only accommodates varying cross sections of the patient throughout the scan, but also helps to avoid dose penalties for miscentered patients. [2], [3] Various hardware designs have been proposed to achieve such modulations. [4]–[6] Our group has proposed multiple aperture devices (MADs) consisting of a series of highly attenuating thin bars of varying widths and spacings along the u direction to control the amount of x-ray blockage locally. [6] The relative motion between the two of such devices produces Moiré patterns that effect a wide range of beam modulation profiles. The dual-MAD system designed in previous work [7] has been installed on an experimental cone-beam CT (CBCT) bench. This work details the implementation of MAD-based FFM design, calibration and correction of MAD data, and initial image quality assessments in physical phantom reconstructions.

This work is supported by NIH grant U01EB018758. G. J. Gang, A. Mao, J. W. Stayman, and J. H. Siewerdsen are with the Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD 21205 USA (e-mail: grace.j.gang, amao, web.stayman, jeff.siewerdsen@jhu.edu).

II. Methods

A. MADs Design and Setup

Dual MAD filter design has been reported in previous work. [7] Each MAD in this study consists of 137 bars with widths and spacings designed to produce modulation patterns that flatten fluence behind a 30 cm by 20 cm QRM phantom (QRM GmbH, Morehendorf, Germany). The resulting bar widths range from 0.2 mm to 0.8 mm, and spans an effective area of 135 mm $\times$ 10 mm. Each MAD is 2 mm thick with the bars focused towards the x-ray source. The spacing between the two MADs is 1 cm from center-to-center. The MADs were manufactured from tungsten powder through a powder bed laser melting process (Smit-Rontgen Facility, Best, Netherlands). Photographs of the two MADs and setup of the dual-MAD system on an experimental CBCT bench is shown in Fig.1. Each MAD was attached to a linear actuation stage capable of translation control with up to 0.002 mm precision. In this work, we focus on imaging scenarios where the objects are well-centered, i.e., the beam center remains stationary while the beam width/shape changes throughout the scan. Such modulations were achieved by keeping MAD0 stationary and translating MAD1 relative to MAD0. For miscentered

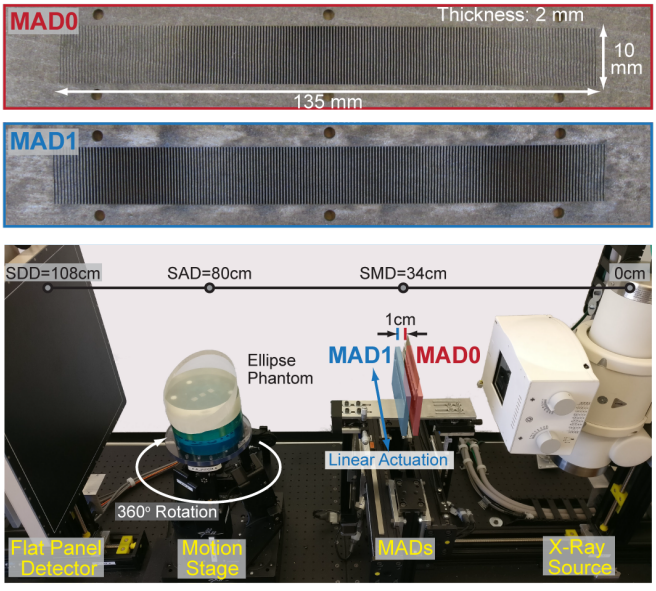


Fig. 1: Photos of the two MADs and setup of the dual-MAD system on an experimental CBCT bench.

objects, both MADs can be translated simultaneously to change both the beam center and beam width.

B. Phantom-Specific Dynamic FFM Design

Dynamic FFM is achieved through a combination of beam shape modulation through MADs and overall amplitude modulation through pulse width modulation at a constant tube current. For a particular phantom, a target FFM was first calculated by approximating the phantom as an ellipse based on measurements of the lateral and anterior-posterior (AP) diameters. The line integrals (denoted as l) of the ellipse were then computed using a linear forward projector matching the test bench geometry. Phantom-specific FFM was computed based on particular design objectives. Two example FFM objectives investigated in this work are given by the following equation when $\alpha = 1.0$ and $\alpha = 0.5$ [8]:

$$I_0(u^*, \theta) = \frac{\exp\{\alpha l(u^*, \theta)\}}{\sum_\theta \exp\{\alpha l(u^*, \theta)\}} I_0^{tot} \tag{1}$$

where I_0 denotes bare-beam fluence and u^* represents detector locations behind the object. The $\alpha = 1$ case corresponds to a flat-field criterion, i.e., uniform signal behind the object and across projections.This FFM yields uniform variance in a filtered-backprojection (FBP) reconstruction. The $\alpha = 0.5$ case minimizes the mean variance over all voxels in an FBP reconstruction.

To identify the MAD1 actuation positions and pulse widths that best approximate the target FFM, a MAD calibration scan was first performed in air through a range of actuation positions at 0.002 mm intervals. The optimal MAD1 position for each projection was identified by minimizing a RMSE criterion between the normalized target and achievable profiles. X-ray pulse width was then calculated as the scale factor required by the central detector element to achieve the target profile.

C. Corrections of Data Acquired with MADs

The presence of MADs introduce a gain term (denoted g_M) in the forward model in addition to typical gains associated with x-ray beam shape and detector sensitivity. A straightforward ***simple correction*** method would involve acquiring gain scans with MADs at the corresponding actuation positions in air and dividing them from the data. However, observations indicate that such simple correction is insufficient due to two physical effects. First, subtle shifts or changes in the focal spot between scans and/or during the same scan introduce mismatch between the g_M of the gain scans and data. Second, x-rays at the very edge of the tungsten bars are not completely blocked but exhibit significant beam hardening. Both effects results in residual high-frequency patterns in MAD projection data, which consequently appear as high frequency ring artifacts in reconstructed images.

To account for these two physical effects, we formulate the following modified data forward model:

$$y = t(\theta) g_B(u, v) g_D(u, v) g_M(u, \Delta_\theta) e^{-\kappa(u, \Delta_\theta) l(u, v, \theta)} \tag{2}$$

where $t(\theta)$ represents pulse width modulation as a function of projection angles, g_B and g_D in the detector domain (u, v) are gains associated with the intrinsic beam shape and detector sensitivity, respectively. The MAD gain/transmissivity, g_M, presumed constant along v and depends on the actuation position Δ. The term κ accounts for spectral effects as a result of MADs and carries the same dependencies as g_M. For investigations in this work, a portion of the MAD-filtered beam was acquired in air and was preserved for retrospective focal spot monitoring.

To recover the line integrals for reconstruction, we devised a ***full model correction*** method using the following calibration scans:

- Air gain calibration: Air scan without MADs for a range of pulse widths, 30 repeats at each pulse width.

$$A = t(\theta) g_B(u, v) g_D(u, v) \tag{3}$$

- MAD gain calibration: Air scan with MADs at a range of actuation positions, 5 repeats at each position, at a nominal pulse width, t_0. The $*$ indicates potential mismatch between g_M of the calibration and data (due to focal spot shift).

$$M = t_0 g_B(u, v) g_D(u, v) g_M^*(u, \Delta_\theta) \tag{4}$$

To match g_M and g_M^*, an ensemble of 1D blur kernels was estimated to minimize the RMSE between local segments of M/A and y/A across each projection (in the reserved bare-beam region), i.e., in a local sense,

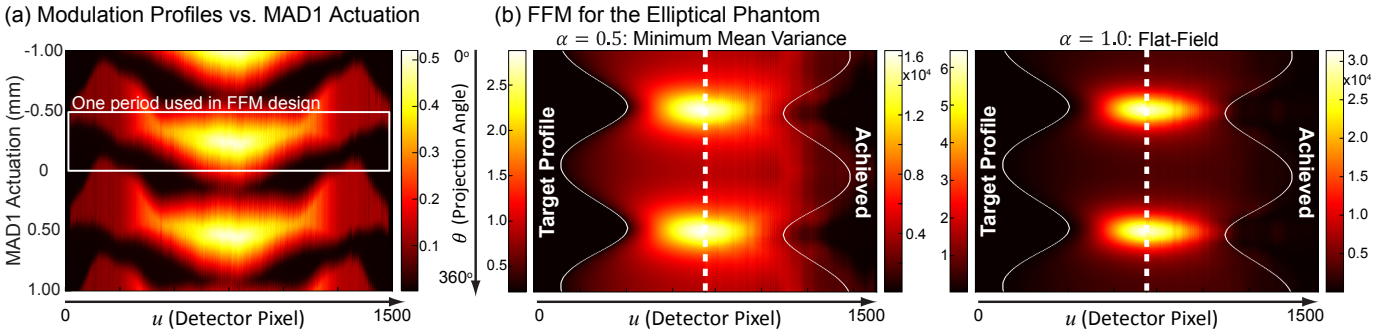


Fig. 2: (a) Modulation profiles (at a constant pulse width) achieved through actuation of MAD1 over a 2 mm range. The modulation exhibits periodic behavior. The FFM designs in this work were selected from one period from -0.5 to -0.06 mm as indicated. (b) Dynamic FFM designs with pulse with modulation for the ellipse phantom described in II-D1 following the $\alpha = 0.5$ (minimum mean variance) and $\alpha = 1.0$ (flat-field) design objectives. Target profiles (left) show good agreement with achieved (right). The boundary of the ellipse phantom in the sinogram domain is overlaid on the figures.

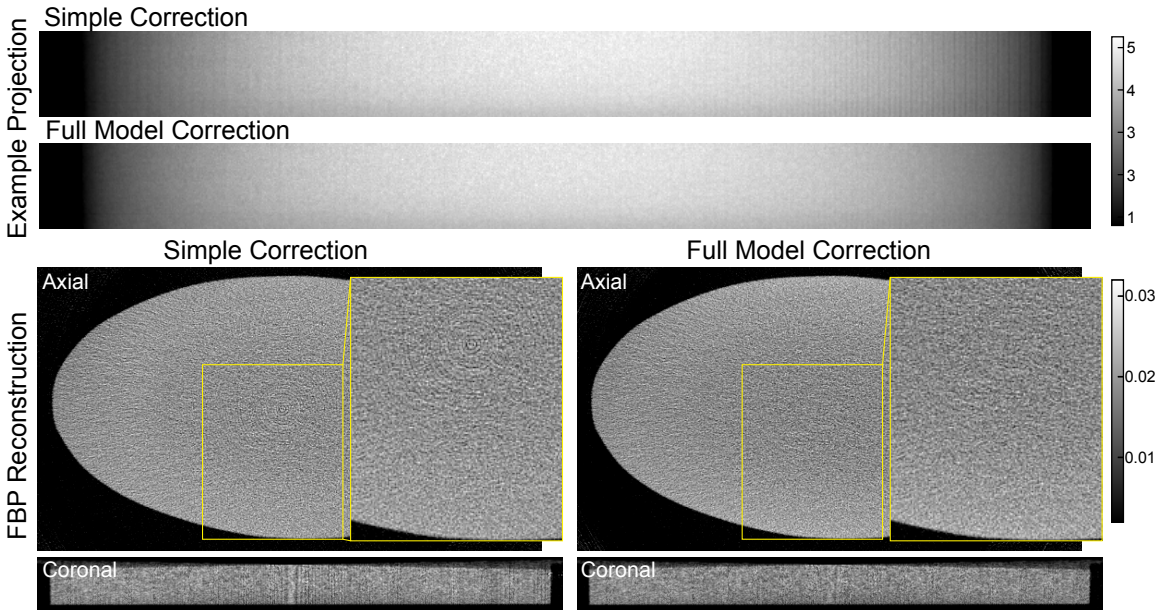


Fig. 3: Comparison of simple vs full model correction of the ellipse phantom. For the simple correction method, residual bars are apparent in the example projection as a result of imperfect gain correction. The reconstruction contains high frequency ring artifacts in the axial cut and high frequency bar patterns in coronal and sagittal (not shown) cuts. These artifacts are mostly removed by the full model correction method.

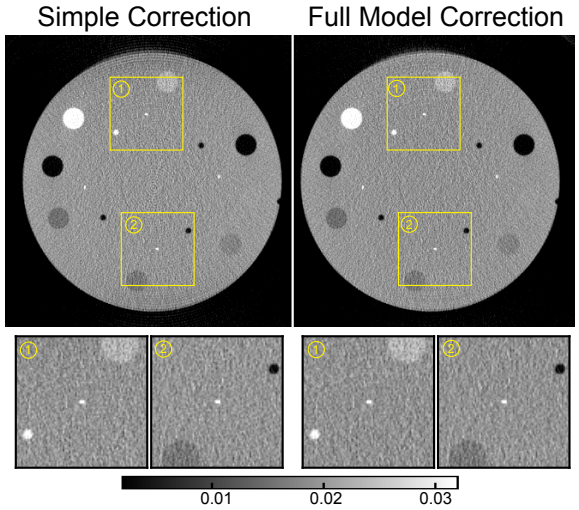


Fig. 4: Comparison of simple vs full model correction in a CatPhan module. The full model correction method effectively alleviates ring artifacts in the reconstruction. The wire and insert structures are unaffected.

$g_M(u,\Delta_\theta) \approx b(u) \circledast g^*_M(u,\Delta_\theta)$. The blur-corrected g_M can then be divided from y to minimize the effect of focal spot shift/blooming.

- Spectral calibration: Scans of PMMA blocks of four different thickness (constant thickness along u) with MADs at a range of actuation positions, 5 repeats at each position. At each thickness:

$$P = t_0 g_B(u,v) g_D(u,v) g^{**}_M(u,\Delta_\theta) e^{-\kappa(u,\Delta_\theta) l_P(u,v)} \quad (5)$$

To obtain κ, the blur estimation step was first performed between each P and the blur-corrected M in order to match the g_M in P and y, i.e., $g_M(u,\Delta_\theta) \approx b'(u) \circledast g^{**}_M(u,\Delta_\theta)$. Ideally, we can then calculate κ as the gradient of $-log(P/M)$ as a function of $l_P(u,v)$. While we could obtain $l_P(u,v)$ via a separate calibration scan of the PMMA blocks without MADs, we chose to compute the gradient as a function of the physical thickness of the blocks and find the proper scaling factor based on a flatness criterion in a portion of the corrected line integral data. The resulting $\kappa(u,\Delta_\theta)$ can then be divided from the exponent of y to arrive at line integrals of the phantom data, l.

D. Experimental Methods

1) Phantoms: Dynamic FFM designs and data acquisition were performed for an elliptical PMMA phantom (major axis=25.8 cm, minor axis=14.1 cm) following the $\alpha = 1.0$ and $\alpha = 0.5$ design objectives in Eq.1. To exercise the MAD correction algorithm on non-uniform phantoms, a CatPhan module (CTP401) containing four slanted wires and various contrast inserts was also imaged using the same FFM.

2) Image Reconstruction: Reconstructions were performed using the FBP algorithm to clearly illustrate the extent of ring artifacts from the MADs as well as noise characteristics as a result of fluence modulation. Images were reconstructed at an isotropic voxel size of 0.5 mm using a Hann filter with a cutoff frequency of 0.8 mm^{-1}.

3) Image Quality Analysis: To assess the effectiveness of the MAD correction algorithm, reconstructions following the simple correction and the full model correction methods (II-C) are presented for visual comparison. The effect of FFM on noise properties is illustrated in local noise power spectrum (NPS) measurements at multiple locations within the reconstructions of the ellipse for an unmodulated scan

(no MADs, constant pulse width), and acquisitions using the $\alpha = 0.5$ and $\alpha = 1.0$ FFM. The total bare-beam fluence for the three scans are matched approximately for comparison of noise magnitude. Due to limited vertical extent of the MADs and therefore the reconstructions, the NPS was measured in 2D from the differences of axial slices. Variance maps were computed within a 3×3 voxel neighbourhood at each voxel location on an axial slice through the reconstruction.

III. Results

A. Dynamic Fluence Field Design

The full range of FFM profiles achievable with MAD1 actuation positions are shown in Fig.2(a). Note that the modulation is periodic. For FFM design purposes (II-B), profiles were selected from a single period over the range of 0.44 mm to minimize actuator motion throughout the scan. The target (left) and achieved (right) FFM profiles for the ellipse phantom are shown in Fig.2(b). The pulse widths used in the $\alpha = 0.5$ FFM ranged from 2.6 to 6.5 ms; and in the $\alpha = 1.0$ FFM, 1.4 to 14.9 ms. The achieved FFM closely approximates the target profiles both in terms of beam widths and relative signal magnitude in each projection.

B. Reconstruction with MADs

Example projections and reconstructions following simple and full-model corrections are compared in Fig.3. Imperfect MAD gain correction by the simple method is apparent as residual bar patterns in the example projection shown. The reconstruction similarly contains high frequency rings in the axial slice and bar patterns in the coronal and sagittal (not shown) slices. Such artifacts are mostly suppressed in the full model correction method. An axial slice of the CatPhan reconstruction is shown in Fig.4 to examine the effect of the correction algorithm on non-uniform phantoms. The wires and contrast modules are preserved well in the full model correction method, while ring artifacts are mostly absent.

C. Effect of Dynamic FFM

Figure 5 shows the local NPS of the ellipse phantom at 7 locations for an unmodulated (no MADs, constant pulse width), an $\alpha = 0.5$, and an $\alpha = 1.0$ reconstruction. The NPS in the unmodulated reconstruction exhibits strongly anisotropic shapes associated with varying attenuation through the ellipse in each view. The magnitude of the NPS also varies across locations, with the central location being the noisiest. The $\alpha = 1.0$ case presents NPS that are much more isotropic and uniform in magnitude, indicating that the dynamic FFM through MADs closely homogenizes fluence across detector elements and projection angles. The variance maps of the three cases shows similar trends. The unmodulated scan shows the largest variation in magnitude within the phantom. The $\alpha = 1.0$ case is much flatter although not perfectly flat due to limitations in experimentally achievable modulation profiles. The $\alpha = 0.5$ case achieves the minimum mean variance over the extent of the phantom.

IV. Discussions and Conclusions

This work presents experimental realization of dynamic FFM using a combination of dual MADs and pulse width modulation on a CBCT bench. The MADs were capable of producing a wide range of beam modulation profiles using a limited range (0.44 mm) of relative motion. Example FFMs for an ellipse phantom are capable of achieving two design objectives: approximately homogeneous noise across the phantom, and reduced mean variance within the phantom. Methods developed in this work pave the way for dynamic FFM studies for a wide range of phantoms and design objectives. Ongoing work includes experimental evaluation

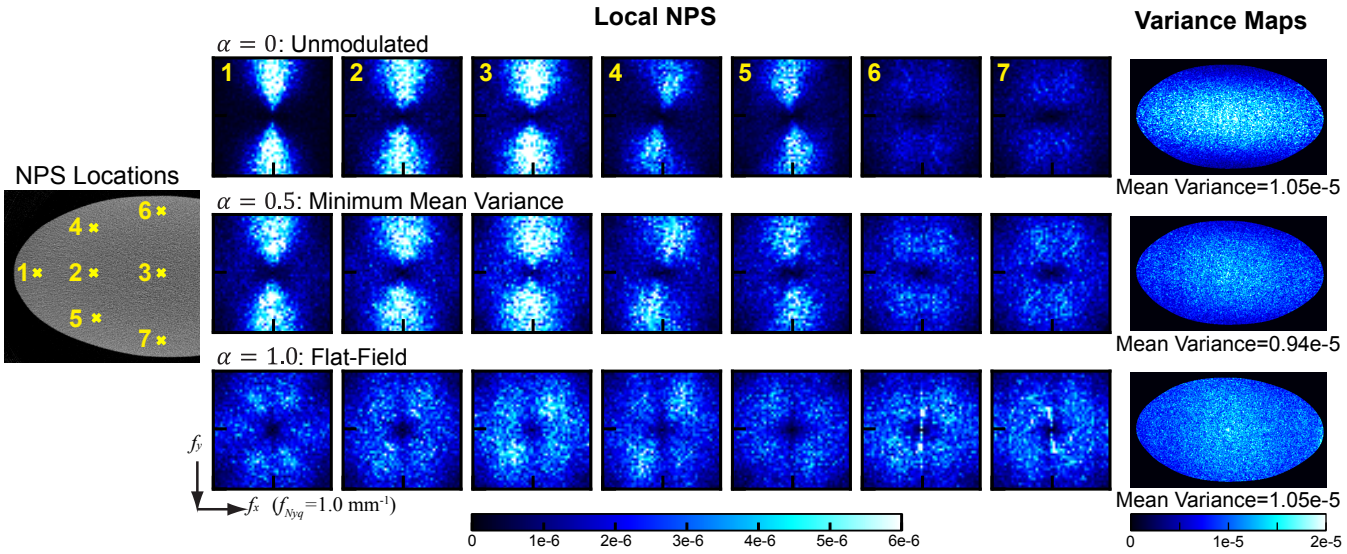


Fig. 5: The local NPS and variance maps of reconstructions for an unmodulated, $\alpha = 0.5$, and $\alpha = 1.0$ acquisition. The $\alpha = 1.0$ case exhibits close to isotropic NPS and a more homogeneous variance map compared to the unmodulated scan. The $\alpha = 0.5$ reconstruction has the minimum mean variance out of the three cases, in agreement with theoretical expectations.

of FFM designs that maximize task-based image quality [9], and integration of the dual-MAD system in a diagnostic CT gantry.

ACKNOWLEDGMENT

The authors would like to thank Dr. Reuven Levinson for helpful discussions regarding fluence field modulation.

REFERENCES

[1] S. Bartolac *et al.*, *Medical physics*, vol. 38 Suppl 1, no. S1, p. S2, jul 2011.

[2] T. L. Toth *et al.*, *Medical Physics*, vol. 34, no. 7, pp. 3093–3101, jun 2007.

[3] A. Mao *et al.*, in *SPIE Medical Imaging*, 2018.

[4] S. S. Hsieh and N. J. Pelc, *Medical physics*, vol. 40, no. 3, p. 31910, 2013.

[5] T. P. Szczykutowicz and C. A. Mistretta, *Physics in medicine and biology*, vol. 59, no. 5, p. 1305, mar 2014.

[6] J. W. Stayman *et al.*, in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2016, pp. 97 830X–97 830X.

[7] A. J. Mathews *et al.*, in *Conference proceedings. International Conference on Image Formation in X-Ray Computed Tomography*, vol. 2016. NIH Public Access, 2016, p. 29.

[8] M. D. Harpen, *Medical Physics*, vol. 26, no. 11, p. 2231, 1999.

[9] G. J. Gang *et al.*, *IEEE transactions on medical imaging*, vol. 36, no. 12, pp. 2424–2435, 2017.

First Multislit Collimator Prototype for SparseCT: Design, Manufacturing and Initial Validation

Baiyu Chen, Matthew J Muckley, Aaron Sodickson, Thomas O'Donnell, Matthias Berner, Thomas Allmendinger, Karl Stierstorfer, Thomas Flohr, Bernhard Schmidt, Daniel Sodickson, Ricardo Otazo[1]

***Abstract*— Compressed sensing represents a promising approach for CT dose reduction based on data undersampling. Alternative to reduced-view undersampling, which is difficult to implement, a within-view undersampling scheme named SparseCT has been proposed. SparseCT undersamples the beam within each view by interrupting the continuous beam with a multislit collimator (MSC). This work presents a SparseCT prototype, which to our knowledge is the first CS implementation on a clinical CT system. Details of the prototype design and manufacturing, data collection and processing steps, reconstruction steps unique to SparseCT, and initial phantom results are presented.**

***Index Terms*— CT, compressed sensing, SparseCT, prototype, undersampling, multi-slit collimator (MSC), penumbra**

I. Introduction

Compressed sensing (CS) is a promising technology for achieving order-of-magnitude CT dose reduction that uses undersampled projection data to reconstruct images [1]. The majority of CS studies have proposed a reduced-view undersampling scheme [1-3], which would require to pulse the x-ray tube to acquire reduced number of projections. However, this scheme is not feasible on a current clinical CT system, because the x-ray tube cannot be pulsed on the order of milliseconds due to the thermal inertia of the cathode.

We have developed an alternative to reduced-view undersampling, named SparseCT, which performs within-view undersampling [4, 5]. SparseCT does not pulse the x-ray beam, but instead interrupts the continuous beam with a multi-slit collimator (MSC). The MSC undersamples projection data within the view along the detector row direction (Figure 1). Furthermore, the MSC is jittered as the gantry rotates to change the undersampling pattern for each view to increase incoherence in the acquisition to improve the performance of compressed sensing.

Recently, we have built a SparseCT prototype, which to our knowledge is the first CS implementation on a clinical CT system. This work presents details of the prototype design and manufacturing, data collection and processing steps, a reconstruction algorithm tailored to SparseCT, and initial phantom results.

This work was supported by the National Institutes of Biomedical Imaging and Bioengineering under Grant U01 EB018760.

Baiyu Chen (e-mail: Baiyu.Chen@nyumc.org), Matthew J Muckley, Daniel Sodickson, and Ricardo Otazo (e-mail: Ricardo.Otazo @nyumc.org) are with the Department of Radiology, NYU Langone Health, New York, NY USA.

Thomas O'Donnell, Matthias Berner, Thomas Allmendinger, Karl Stierstorfer, Thomas Flohr, and Bernhard Schmidt are with Siemens Healthineers.

Aaron Sodickson is with Brigham and Women's Hospital, Boston, MA USA.

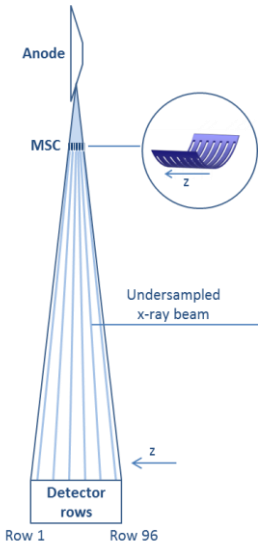


Figure 1: The multislit collimator (MSC) partially blocks the beam before it reaches the patient, such that undersampled projection data are acquired while reducing patient dose.

II. Methods

A. SparseCT prototype

The SparseCT prototype was designed and manufactured for the geometry of a Siemens SOMATOM Force scanner. The MSC was mounted onto the tube side of the scanner, 190 mm from the focal spot (Figure 2). The MSC is a tungsten plate curved along the fan angle and has laser-cut slits perpendicular to the z direction (patient bed direction). The plate is thick enough to block the beam so that the beam can only penetrate through the slits. The MSC can be moved along the z-direction with micrometer precision to undersample different detector rows.

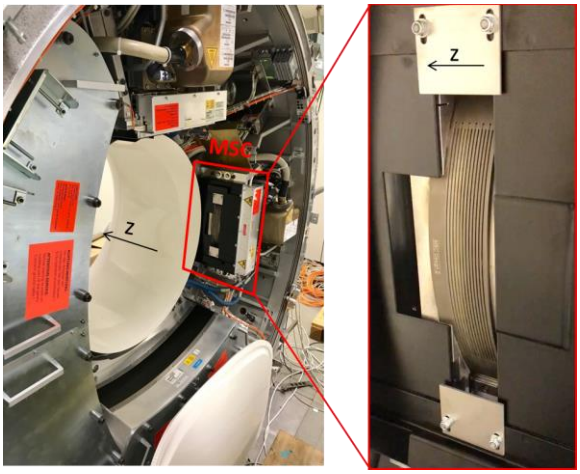


Figure 2: SparseCT prototype. The multislit collimator (MSC) is mounted to the tube side of the scanner and can be shifted along the z-direction.

Two MSC designs, named W4S12 and W4S16, were manufactured and tested on the prototype. "W" and "S" represent the slit width and the slit separation, respectively. For example, "W4S16" means that when projected to detector surface, each slit is 4-detector-row-wide and the separation between neighboring slits is 16-detector-row-wide. W4S12 and W4S16 have undersampling factors of 3 and 4, respectively.

The width of MSC slit was designed based on a previous study [5], which shows that the undersampled x-ray beam has penumbra regions on both sides of the beam, caused by the finite size of the focal spot (Figure **3**). The penumbra's width is independent of the slit width, about 3 rows on each side of the beam on detector surface. Consequently, as shown in Figure **4** (a simulation result from the previous study), the MSC slits need to be at least 3 to 4 rows wide (projected to detector surface) to ensure adequate separation between neighboring undersampled beam. Figure **4** also demonstrates that wider slits help reduce the amount of penumbra relative to umbra region, thus increasing dose efficiency. Based on aforementioned reasons, our prototype used a slit width of 4.

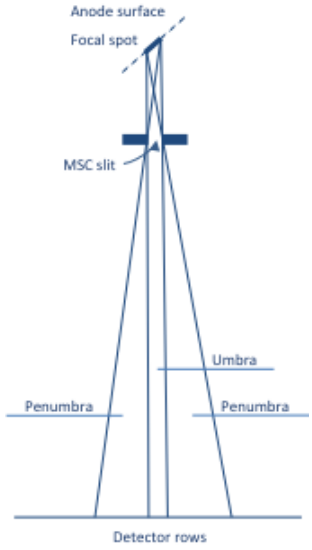


Figure 3: Due to the finite size of the focal spot, the beam through the MSC contains penumbra.

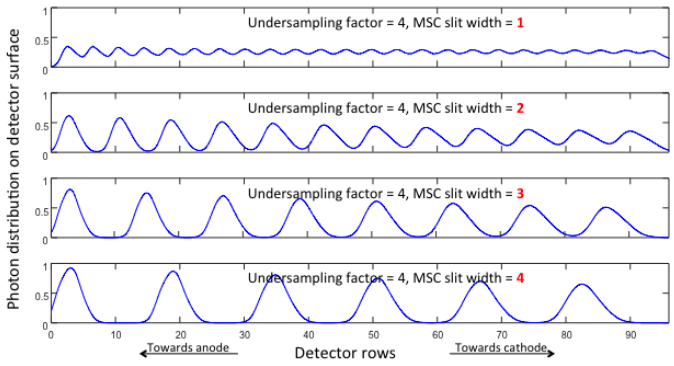


Figure 4: The photon distribution on the detector surface with different MSC designs, normalized by the photon distribution without MSC. All designs have the same undersampling factor of 4 but different slit width.

B. Collection of Projection Data

Both air scans and phantom scans were performed using the MSC prototype. For all scans, data were collected in a "step-and-shoot" fashion. The MSC was shifted along the z-direction to a prescribed location before an axial scan was taken. For W4S12 design, because the slit separation is 12 detector rows (projected to detector surface), we repeatedly scan at 12 different locations, each shifting the beam coverage by one row on the detector surface. For W4S16 design, scans were performed similarly at 16 different locations. All scans used a tube voltage of 120 kVp, a tube current of 800 mA, and a rotation time of 0.5 s.

C. Processing of Projection Data

In addition to the common preprocessing steps (e.g., air calibration and beam hardening correction), the projection data went through three steps unique to SparseCT.

First, a "MSC calibration" was performed to remove the attenuation of the MSC from the projection data of the phantom scans. This was done by subtracting the post-log projection data of the air scans from the post-log projection data of the phantom scans acquired at the same MSC locations.

Second, statistical weights were estimated based on the flux information (without MSC, provided by the vendor) and the phantom scans, which were inversely proportional to the variance of the post-log projection data and were used for reconstruction [6], as detailed in the next section.

Last, "dynamic MSC" scans were created by retrospectively drawing projections from scans of all MSC positions and stitching them together. The stitched scans simulate a linear MSC movement during the scan, shifting the beam by one row per projection.

D. Reconstruction of Projection Data

The processed projection data were reconstructed iteratively using a penalized weighted least squares cost function with OS-Momentum algorithm [7]. The reconstruction was tailored to SparseCT in two aspects.

First, to use the portion of projection data acquired in the penumbra region with reduced flux, we applied statistical weights to the reconstruction. The statistical weights accounted for the dramatic variation of exposure and quantum noise in the penumbra region by giving noisy data of less fidelity lower weights.

Second, to solve the "partial source obstruction" problem illustrated and explained in Figure 5, we analytically calculated the effective focal spot location corresponding to each detector row for each MSC design and MSC position, and passed on that information to the ray projector of the reconstruction algorithm.

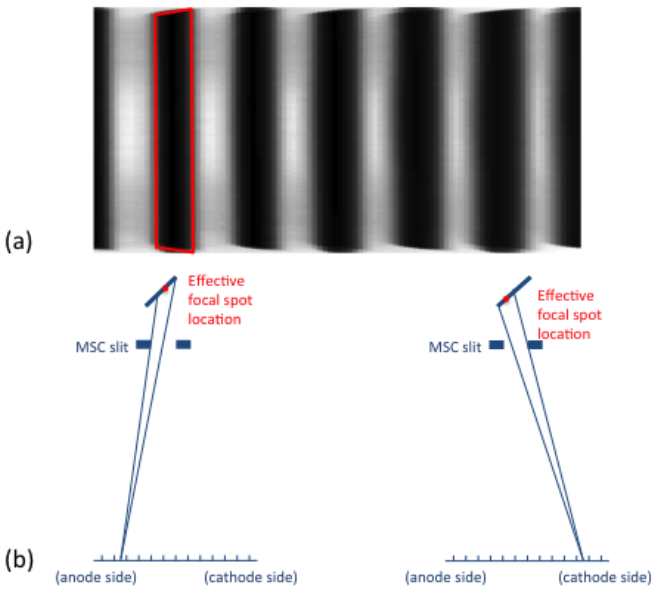


Figure 5: a) "Partial source obstruction" problem is when the projection of a SparseCT air scan captures each undersampled beam in trapezoidal shape instead of rectangular shape. b) This partial source obstruction is caused by the fact that the slit of the MSC is so narrow that part of the focal spot is obstructed by the slit. In other words, the detector rows in the beam can only "see" part of the focal spot. As a result, the effective locations of the focal spot (the centroid of the part of focal spot being seen) are different from row to row: The row on the anode side of the slit sees a focal spot further away from the MSC than the row on the cathode side of the slit, therefore having a smaller magnification ratio. Due to the difference in magnification ratio on the two sides of the beam, a rectangular beam was captured as a trapezoidal beam on the detector surface.

With these two aspects built into the reconstruction, the problem was formulated as the following optimization problem [8]:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_{\mathbf{W}}^2 + \beta \sum_{m=1}^{M} \psi\left([\mathbf{C}\mathbf{x}]_m\right)$$

where x is the image to be reconstructed, y is the projection data, A is the data acquisition operator (gantry geometry and undersampling pattern), β is the regularization parameter that weights the sparsity term relative to data consistency, C is a finite difference operator, and ψ is a hyperbola function. The subscript "W" indicates the inclusion of statistical weights.

III. Results

A. Projection data

Figure 6 shows the projection data of a SparseCT phantom scan with W4S16 MSC, the projection data of a SparseCT air scan with W4S16 MSC at the same location, and the process of MSC calibration.

Note that the penumbra effect can be observed in Figure 6 (b), where the beam (the black strips) and the MSC (the white strips) do not have clear boundaries. Also note that although ¾ of the beam is interrupted by the W4S16 MSC, about ½ of the detector rows are irradiated due to the penumbra effect.

B. Reconstruction

Figure 7 shows the reconstructed images of a liver phantom with two low contrast lesions. All images were reconstructed with 350 mm FOV and 3 mm slice thickness. Figure 7(a) is the reference image acquired with full dose (no MSC) and reconstructed with Siemens WFBP algorithm. Figure 7(b) is the SparseCT image acquired with 1/3 dose (W4S12 MSC) and reconstructed with SparseCT algorithm. Figure 7(c) is SparseCT image acquired with ¼ dose (W4S16 MSC) and reconstructed with SparseCT algorithm.

Under reduced dose, SparseCT images still show the two low contrast lesions. However, ring artifacts are visible in the images, possibly because detector was calibrated for relatively uniform flux distribution, but the flux of the undersampled beam varies drastically across detector rows, even within a single row. In addition, the W4S16 result is noisy, which might be improved by changing the beta or the motion pattern of the MSC (currently a linear motion pattern was simulated).

IV. Conclusion

A prototype of SparseCT has been built, which to our knowledge is the first CS CT system based on a clinical system. Phantom data have been collected from the prototype. Special processing and reconstruction procedures have been developed and tested on the SparseCT data.

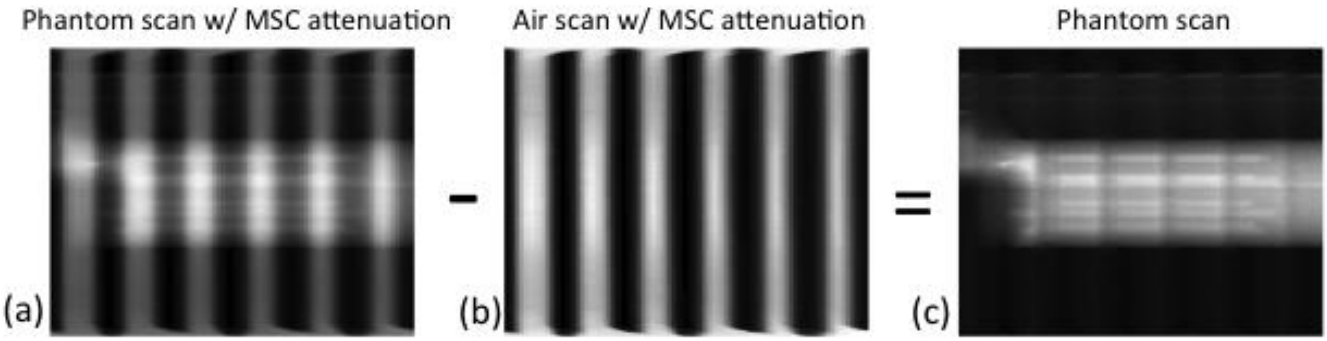


Figure 6: (a) A post-log projection of a SparseCT phantom scan with W4S16 MSC installed. The attenuation of the MSC is captured on top of the phantom attenuation in the projection. (b) A post-log projection of a SparseCT air scan with W4S16 MSC installed at the same location. (c) MSC calibration is performed by subtracting the aforementioned two, which removes the MSC attenuation from the projection.

ACKNOWLEDGMENT

This work was funded by National Institute of Biomedical Imaging and Bioengineering (U01 EB018760).

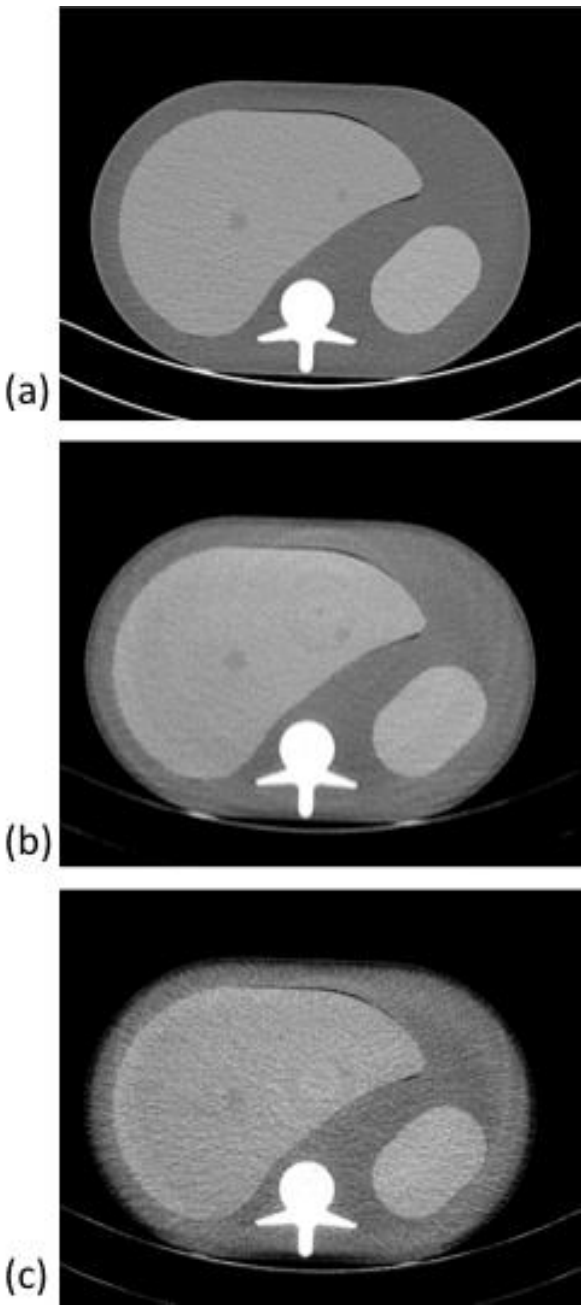


Figure 7: Reconstructed images of a liver phantom with two low contrast lesions: a) Full dose scan without MSC, WFBP reconstruction, b) 1/3 dose scan with W4S12 MSC, SparseCT reconstruction, and c) 1/4 dose scan with W4S16 MSC, SparseCT reconstruction.

REFERENCES

[1] G. H. Chen, J. Tang, and S. Leng, "Prior image constrained compressed sensing (PICCS): a method to accurately reconstruct dynamic CT images from highly undersampled projection data sets," *Med Phys*, vol. 35, pp. 660-3, Feb 2008.

[2] K. Choi, J. Wang, L. Zhu, T. S. Suh, S. Boyd, and L. Xing, "Compressed sensing based cone-beam computed tomography reconstruction with a first-order method," *Med Phys*, vol. 37, pp. 5113-25, Sep 2010.

[3] J. C. Ramirez-Giraldo, J. Trzasko, S. Leng, L. Yu, A. Manduca, and C. H. McCollough, "Nonconvex prior image constrained compressed sensing (NCPICCS): theory and simulations on perfusion CT," *Med Phys*, vol. 38, pp. 2157-67, Apr 2011.

[4] T. Koesters, F. Knoll, A. Sodickson, D. Sodickson, and R. Otazo, "SparseCT: Interrupted-beam acquisition and sparse reconstruction for radiation dose reduction," presented at the SPIE Medical Imaging, Orlando, FL, 2017.

[5] B. Chen, M. Muckley, T. O'Donnell, A. Sodickson, T. Flohr, K. Stierstorfer, *et al.*, "Realistic Undersampling Model for Compressed Sensing Using a Multi-Slit Collimator," presented at the The 14th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine, Xi'an, China, 2017.

[6] G. L. Zeng and W. Wang, "On Approximation of Compound Poisson by Poisson," presented at the The 4th International Conference on Image Formation in X-Ray Computed Tomography, 2016.

[7] D. Kim, S. Ramani, and J. A. Fessler, "Combining ordered subsets and momentum for accelerated X-ray CT image reconstruction," *IEEE Trans Med Imaging*, vol. 34, pp. 167-78, Jan 2015.

[8] M. Muckley, B. Chen, T. Vahle, A. Sodickson, F. Knoll, D. K. Sodickson, *et al.*, "Regularizer Performance for SparseCT Image Reconstruction With Practical Subsampling," presented at the The 14th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine, Xi'an, China, 2017.

Reconstruction of Reduced-Dose SparseCT Data Acquired With An Interrupted-Beam Prototype On A Clinical Scanner

Matthew J. Muckley, Baiyu Chen, Thomas O'Donell, Matthias Berner, Thomas Allmendinger, Karl Stierstorfer, Thomas Flohr, Bernhard Schmidt, Aaron D. Sodickson, Daniel K. Sodickson, and Ricardo Otazo

***Abstract*—Low-dose X-ray CT is a rapidly-advancing area in medical imaging. A number of paradigms for low-dose CT have emerged, some of which are based on sparse sampling. Sparse sampling approaches using view-based subsampling have been limited to simulations due to difficult hardware implementation, but recently a more practical within-view sparse sampling approach was proposed called SparseCT. SparseCT operates by interrupting the source beam with a multislit collimator before it reaches the patient, thus creating a row-subsampled projection view. By programming the multislit collimator to move throughout the scan, undersampling is possible in both the row and the projection angle dimensions. Here we develop reconstruction algorithms tailored to the interrupted-beam acquisition to compensate for partial source obstruction and nonideal statistical weighting. We present reconstruction of phantom data acquired with the first multislit collimator prototype on a clinical scanner using simulated motion patterns.**

I. Introduction

X-ray CT is currently one of the most important diagnostic tools available in the clinic. However, the ionizing radiation dose of CT remains a risk to patients, particularly pediatric patients and those with chronic illnesses. Thus, radiation dose reduction is an engaging topic of research in the CT field. A number of approaches have been proposed for reducing radiation dose in CT. The most common approach is to reduce the tube current. Iterative reconstruction techniques that compensate for the increased noise have been promising areas of research, with successful commercial examples such as ADMIRE [1] and research examples based on iterative algorithms with edge-preserving regularizers [2].

M. J. Muckley, B. Chen, and D. K. Sodickson are with the NYU School of Medicine, New York, NY.

T. O'Donnell, M. Berner, T. Allmendinger, K. Stierstorfer, T. Flohr, and B. Schmidt are with Siemens Healthineers.

A. D. Sodickson is with Harvard Medical School, Boston, MA.

R. Otazo is with the Memorial Sloan-Kettering Cancer Center, New York, NY.

An alternative approach to tube-current reduction is sparse sampling. These techniques show promise because they can be combined with standard compressed sensing theory [3]. This allows extremely high dose reductions. A practical way to achieve sparse sampling is to partially block the X-ray beam. This method, called SparseCT or interrupted-beam CT, has gained interest since it can achieve sparse sampling by using a multislit collimator (MSC) in replacement of the standard source collimator on current clinical systems [4]. Contrary to sparse-view sampling approaches that block entire views, SparseCT achieves sparse sampling within views. A number of aspects of SparseCT have already been studied in simulation. These discovered that penumbra effects can modify the noise level in the acquired sinogram [5]. The increased noise can be compensated for by using larger slits with larger spacings, and these have been observed to lead to only a modest drop in incoherence [6]. Nonetheless, none of these studies were done using real-world data.

In this work, we demonstrate the reconstruction of phantom data acquired using the first SparseCT MSC on a clinical scanner. The SparseCT collimator has four open rows out of every 16 (W4S16). We collect data while holding the MSC in 16 different static positions. We then retrospectively stitch the positions together to create a synthetic, linear motion pattern. We show new features that the proposed sampling scheme applies to the sinogram, including that of partial source obstruction. We outline a new reconstruction pipeline for the prototype, which includes modifications that take into account these new features. The images constitute the first demonstration of SparseCT using real-world data.

II. Methods

A. Prototype Design and Data Acquisition

SparseCT calls for an interrupted-beam scanner design. The ideal design is to apply a binary sampling

mask along the row dimension of the sinogram. During the scan, the MSC would be jittered along the z direction throughout a helical scan, giving incoherent θ-z undersampling. We desire incoherence in the acquisition to improve the performance of sparsity-promoting reconstruction as predicted by compessed sensing theory.

Such a design requires building a multislit collimator (MSC) to interrupt the X-ray beam prior to reaching the patient. Fig. 1 shows a schematic of such a design. The

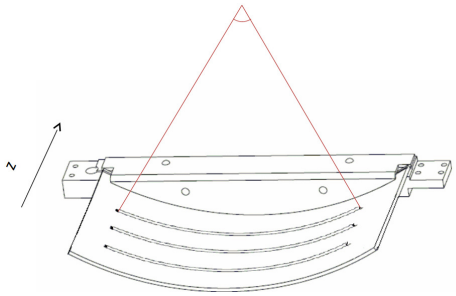


Fig. 1: Schematic of a multislit collimator (MSC) and source with periodic slit openings. In the W4S16 design, each slit illuminates four rows of the detector, and the repetition period is 16 detector rows.

design places the MSC into the gantry in place of the normal source collimator.

A number of parameters can govern the construction and operation of the MSC. Designing the MSC so that it has more, thinner slits produces more incoherence in the sampling pattern [6] at the cost of reduced dose efficiency due to penumbra effects [5]. To balance this tradeoff, we opted for a W4S16 scanner design, which exposes 4 detector rows out of every 16 in the ideal binary mask case. In reality, due to the penumbra, each slit illuminates more than 4 rows at a reduced intensity, but the W4S16 represents a balance point between these effects based on previous simulations [5], [6].

Data were acquired on a SOMATOM Force Scanner (Siemens Healthineers, Forccheim, Germany) with the MSC replacing the normal source collimator. The data were acquired with the MSC being located at 16 different static positions. We used a separate air scan to calibrate the incident flux into the bore of the scanner prior reaching the patient. After standard clinical preprocessing, we synthesized a data set with a moving MSC by retrospectively stitching together the 16 MSC positions. The final data set simulated a linear motion of the MSC from position 1 to position 16 over a single axial rotation. We designate the output of these preprocessing steps as $\mathbf{y}$, the data used for reconstruction.

B. Reconstruction

For reconstruction, we adopted an iterative approach that solves the following optimization problem:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \frac{1}{2} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2_{\mathbf{W}} + \beta \sum_{m=1}^{M} \psi\left([\mathbf{C}\mathbf{x}]_m\right), \quad (1)$$

where $\mathbf{x}$ is the image volume to be reconstructed, $\mathbf{y}$ is the preprocessed sinogram data, $\mathbf{A}$ is a Siddon line projector [7] incorporating sampling patterns and scanner geometry, $\mathbf{C}$ is a finite difference operator, and β is a regularization parameter. The quadratic term in this cost function enforces data consistency, and the regularizer promotes edge sparsity with β determining the level of regularization. $\psi(t)$ is a potential function that we define as the hyperbola function:

$$\psi(t) = \delta^2 \left[\sqrt{1 + (t/\delta)^2} - 1\right]. \quad (2)$$

This potential function makes the regularizer behave analogously to "corner-rounded" Total Variation.

$\mathbf{W}$ is a diagonal matrix constructed of elements w_n, $n \in [1, ..., N]$. We calculated the weights as $w_n = (b_n)^{0.5}$, where b_n is the nth prelog sinogram data point. This is a balanced method that preserves some of the noise mitigation benefits of reducing emphasis on low-signal projections [8] without giving too much weight to a small number of projections [9]. We solved (1) with the ordered subsets with momentum algorithm [2].

C. Partial Source Obstruction

Solving the problem in (1) requires building a routine that can apply $\mathbf{A}$ and $\mathbf{A}^T$, the forward and back-projection operations. We used the Siddon line integral method [7]. This method draws rays between the source location and each detector location. In normal settings, the source is approximated as a point source for such methods, but the insertion of the MSC partially obstructs the source. Thus, the assumed point source center no longer corresponds to the observed source center - effectively, a row-dependent shift in the source location. To precompensate for this effect, we simulated what portion of the source would be observable for each row of the detector. Then, we jitter the source location in the $\mathbf{A}$ operator during reconstruction. A schematic of this process is shown in Fig. 2. For each row of the detector, the assumed point source location is jittered based on the centroid of the observable portion of the detector. To do this, we approximated the source as a summation of Gaussian blobs of varying intensity, and then simulated how much of the source would be occluded in each row of our design geometry.

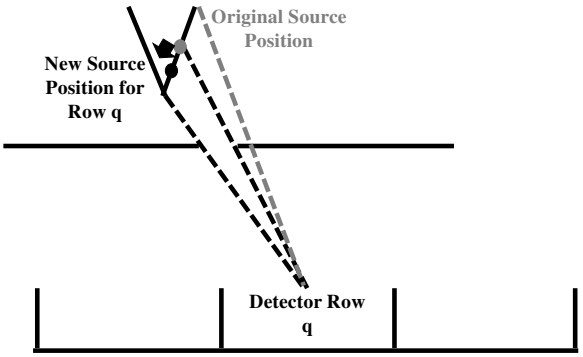


Fig. 2: Schematic of partial source obstruction. In the absence of the MSC, the assumed point source location for detector row q would correspond to the true source centroid. However, in the presence of the MSC, the observed centroid is shifted. This shift was applied to each of the ray casting steps in calculating $\mathbf{A}$.

III. Results

A. Partial Source Obstruction

Fig. 3 shows sinograms of an American College of Radiology (ACR) phantom acquired with and without the multislit collimator (MSC). The phantom has inserts that travel straight along the length of the phantom, as can be seen in the sinogram in Fig. 3a. When the MSC is inserted, the effects of partial source obstruction lead to a varying cone beam magnification in the row direction in Fig. 3b. The shift is most observable in slanted line patterns where the operating table is located in the sinogram.

To correct for partial source obstruction, we applied the technique outlined in Section II-C to a scan in which the MSC position was fixed. This means that parts of the object where the MSC was fully opened were fully sampled, and we should observe no subsampling artifacts. Prior to partial source obstruction correction, some areas are enlarged and others are diminished. Fig. 4 shows the effect of the partial source obstruction correction. The corrected version of the phantom is shown in Fig 4a, while the uncorrected version subtracted from the corrected version is shown in Fig. 4b. The dark outline around the edge of the ACR phantom implies that the uncorrected reconstruction was spatially diminished in this slice.

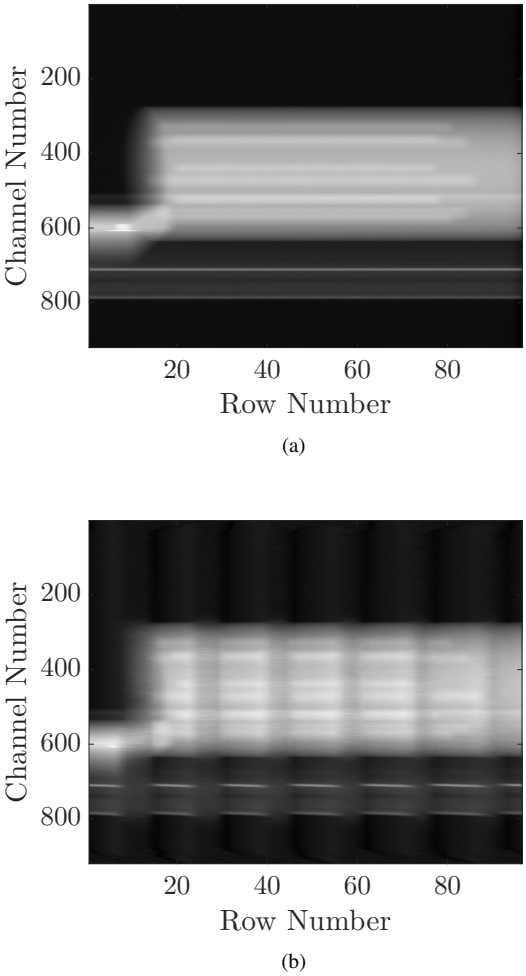


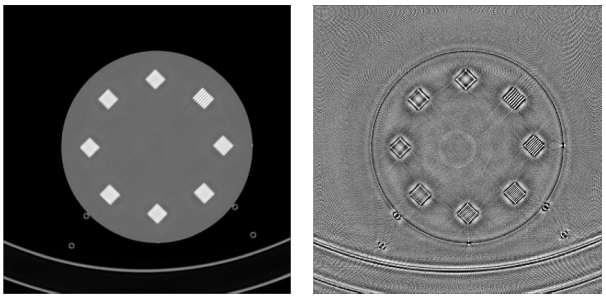
Fig. 3: Effects of partial source obstruction on one projection view. (a) shows the artifact-free ACR phantom with an open collimator, while (b) shows how a shift-variant magnification results with the MSC, most easily observed in the slanted lines of the operating table.

B. Linear Motion

Fig. 5 compares results with an open collimator vs. the MSC with linear motion by zooming in on the 7 lp/com insert of the ACR phantom. The open collimator result was obtained with the standard Siemens FBP pipeline. Some degradation occurs around the edge of the bar resolution pattern of the MSC image in Fig. 5d. Nonetheless, the reconstruction with the MSC while simulating linear motion is able to gain similar resolution to the fully-open, ideally sampled case. When no motion is applied, significant image degradation is observable in penumbra regions (not shown).

IV. Discussion

We have demonstrated reconstruction with an interrupted-beam CT prototype with simulated linear

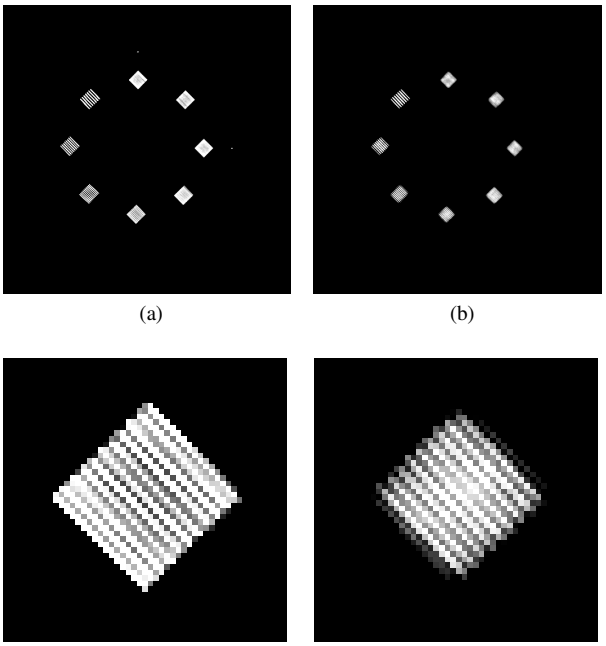


(a) (b)

Fig. 4: Effects of partial source obstruction. (a) shows the corrected image (-1000 HU to 1400 HU), while (b) shows the original uncorrected image subtracted from the corrected (-100 to 100 HU), demonstrating that the original image was minimized in this slice.

motion. We showed that the beam interruption can introduce new partial source obstruction effects and proposed a means to correct for them. We observed that this correction could change the size of the phantom in various slices. Subjectively, some blurring also occurred with the use of partial source obstruction correction - possibly due to incorrect assumptions on the source intensity distribution in our simulations. In the future, we aim to get accurate measurements of the source intensity distribution to improve this correction process.

Fig. 5 shows that SparseCT with a 4-fold dose reduction and linear motion can attain similar resolution to the standard full-dose case using the ACR phantom, which demonstrates practical sparsity-based dose reductions on a clinical scanner. Our results without any motion (not shown) indicated that moving the MSC throughout the acquisition is essential to preserving image quality. Without motion, significant image degradation is observable in the penumbra regions. In the future, we will perform reconstructions on cases involving real-world motion. This will require new advances in registering the air calibration to various motion states, as well as further developing the partial source obstruction model to handle motion states that are not explicitly simulated.



(a) (b)

(c) (d)

Fig. 5: Reconstructions of the ACR phantom with an open collimator vs. linear motion (900 HU to 1100 HU window). (a) reconstructed ACR slice with standard acquisition and reconstruction, with (c) showing a zoomed-in version of the 7 lp/cm insert. (b) reconstructed 7 lp/com insert with 4-fold reduction (W4S16 pattern) and linear motion, with (d) showing a zoomed-in version of the third insert. Both methods show similar resolution.

REFERENCES

[1] S. Gordic, F. Morsbach, B. Schmidt, T. Allmendinger, T. Flohr, D. Husarik, S. Baumueller, R. Raupach, P. Stolzmann, S. Leschka, T. Frauenfelder, and H. Alkadhi, "Ultralow-dose chest computed tomography for pulmonary nodule detection: first performance evaluation of single energy scanning with spectral shaping," *Investigative Radiology*, vol. 49, no. 7, pp. 465–73, 2014.

[2] D. Kim, S. Ramani, and J. A. Fessler, "Combining ordered subsets and momentum for accelerated X-ray CT image reconstruction," *IEEE Trans. Med. Imag.*, vol. 34, pp. 167–78, Jan. 2015.

[3] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," *IEEE Sig. Proc. Mag.*, vol. 25, pp. 21–30, Mar. 2008.

[4] T. Koesters, F. Knoll, A. Sodickson, D. Sodickson, and R. Otazo, "SparseCT: Interrupted-beam acquisition and sparse reconstruction for radiation dose reduction," in *SPIE Med. Im.*, 2017.

[5] B. Chen, M. J. Muckley, T. O'Donnell, A. Sodickson, T. Flohr, K. Stierstorfer, B. Schmidt, F. Knoll, A. Primak, D. Faul, D. Sodickson, and R. Otazo, "Realistic undersampling model for compressed sensing using a multi-slit collimator," in *Proc. Intl. Mtg. on Fully 3D Image Recon. in Rad. and Nuc. Med.*, 2017.

[6] M. J. Muckley, B. Chen, T. Vahle, F. Knoll, A. Sodickson, D. K. Sodickson, and R. Otazo, "Regularizer performance for SparseCT image reconstruction with practical subsampling," in *Proc. Intl. Mtg. on Fully 3D Image Recon. in Rad. and Nuc. Med.*, 2017.

[7] R. L. Siddon, "Fast calculation of the exact radiological path for a three-dimensional CT array," *Med. Phys.*, vol. 12, pp. 252–5, Mar. 1985.

[8] J.-B. Thibault, K. Sauer, C. Bouman, and J. Hsieh, "A three-dimensional statistical approach to improved image quality for multi-slice helical CT," *Med. Phys.*, vol. 34, pp. 4526–44, Nov. 2007.

[9] G. L. Zeng and W. Wang, "Does noise weighting matter in CT iterative reconstruction?," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 1, no. 1, pp. 68–75, 2017.

Stack Transition Artifact Removal for Cardiac CT using a Symmetric Demons Algorithm

Sergej Lebedev, Eric Fournie, Karl Stierstorfer, and Marc Kachelrieß

***Abstract*—Cardiac CT can be achieved by performing short scans with prospective gating. As the collimation of multi–slice CT scanners generally does not allow for a coverage of the entire heart, sequence scans, or step-and-shoot, can be used, where irradiation is performed multiple times for varying positions. Each of these short scans yields data, generally with a longitudinal overlap, that can be reconstructed into a sub-volume, or stack. The latter ideally corresponds to the same phase. The issue addressed in this work is irregular motion, such as irregular heart motion. It leads to stacks that do not represent exactly the same volume, resulting in discontinuities at stack transitions when assembling the complete CT volume. We propose a stack transition artifact removal method including a symmetric deformable registration approach, a symmetric demons algorithm. The deformation vector fields are extended from the overlapping regions in order to maintain smooth and anatomically meaningful images. We validated the method using clinical data sets. By applying a symmetric registration method to cardiac data, we show that the stack transition artifacts can be addressed in this fashion. The method was able to considerably improve image quality and its performance and stability can be further improved in the future.**

I. Introduction

Cardiac CT can be achieved by performing short scans with prospective ECG–gating and reconstructing one phase. Reconstructions of cardiac data can include motion artifacts due to heart motion. The issue to be addressed here is irregular motion. It can occur in every patient and might originate from irregular heart and respiratory motion or basically any movement of the patient during the scan. As the collimation of multi–slice CT scanners generally does not allow for a coverage of the entire heart, irradiation has to be performed multiple times for varying positions. Such a sequence scan, also known as step–and–shoot, can be utilized for multiple cardiac imaging tasks [1]–[3]. Each of these short scans yields data that can be reconstructed into a sub-volume, or stack. The stacks are used to assemble the complete CT volume, where each one corresponds to a different time during the scan, but ideally to the same phase. Such stacks can also be reconstructed using data from a retrospectively ECG–gated spiral scan, if the pitch is chosen correctly with respect to the heart rate [4]–[7]. In this case, data from other phases would be available as well.

Sergej Lebedev is with the German Cancer Research Center (Dkfz), Heidelberg, Germany, with Siemens Healthineers, Forchheim, Germany and with the University of Heidelberg, Germany. Email: sergej.lebedev@dkfz.de. Eric Fournie and Karl Stierstorfer are with Siemens Healthineers, Forchheim, Germany. Marc Kacherlrieß is with the German Cancer Reasearch Center (Dkfz), Heidelberg, Germany and with the University of Heidelberg, Germany.

Irregular motion leads to stacks that do not represent exactly the same volume and as a consequence discontinuities at stack transitions arise when stitching the stacks together to yield the complete CT volume. Such discontinuities can be very apparent depending on the severity of the irregular motion encountered. To remove artifacts and discontinuities from irregular motion we propose a stack transition artifact removal method. The method in this work makes use of symmetric image registration, in particular a symmetric demons algorithm. Herein, symmetric means that deformations in one volume are symmetrically applied in the opposite direction in the other volume. Therefore, intermediate volumes are computed, as opposed to a more classical approach with one volume being registered to another, static volume. Neighboring stacks will generally have a longitudinal overlap that can be used to compute a transformation that, applied to both stacks, yields modified stacks ("intermediate stacks") that are mostly identical within the overlapping region. A series of registrations in the overlapping regions using the demons algorithm will compute a series of deformation vector fields (DVF) that can be extended on to the non–overlapping parts via linear interpolation between the overlapping regions. This framework enables continuous and smooth DVFs covering all stacks entirely and is illustrated in figure I.

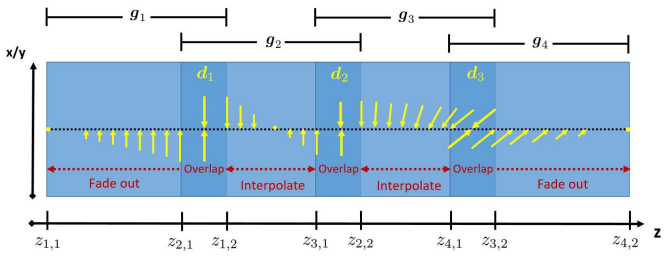


Fig. 1. Illustration of a volume consisting of four overlapping stacks g with DVFs $\boldsymbol{d}$ and interpolation. The DVFs are illustrated by one constant vector, each. Stacks not to scale.

II. Methods and Materials

Given a set of longitudinally overlapping stacks, the goal is to compute DVFs that are applied to the stacks yielding similar, intermediate volumes. Given two volumes f_1, f_2, the aim is to compute a DVF $\boldsymbol{d}(\boldsymbol{r})$ that will match the two. While a more common approach assumes a static volume and a moving volume that is registered onto the static one, the herein implemented method is symmetric in terms of the deformations that are applied to both volumes so that the transformed volumes $T_{+\boldsymbol{d}}f_1(\boldsymbol{r})$ and $T_{-\boldsymbol{d}}f_2(\boldsymbol{r})$ match:

$$T_{+\boldsymbol{d}} f_1(\boldsymbol{r}) = f_1(\boldsymbol{r} + \boldsymbol{d}(\boldsymbol{r}))$$
$$T_{-\boldsymbol{d}} f_2(\boldsymbol{r}) = f_2(\boldsymbol{r} - \boldsymbol{d}(\boldsymbol{r})). \quad (1)$$

Given the deformation model, an additive, symmetric demons algorithm is derived similar to Thirion's demons algorithm [8], [9]. In accordance with past works in the field, a cost function C is defined, consisting of a similarity measure, also known as image distance, a regularization applied to the DVF $\boldsymbol{d}$ and an auxiliary term named correspondence including the difference between two non identical vector fields $\boldsymbol{d}$ and $\boldsymbol{e} = \boldsymbol{d} + \Delta\boldsymbol{d}$ that enable a decoupling of the optimization of the similarity and the regularization in two separate steps for computational performance. Given N voxels at positions $\boldsymbol{r}_n$, the DVF is stored as N vectors $\boldsymbol{d}(\boldsymbol{r}_n)$. Assuming a vector notation with $f_1, f_2 \in \mathbb{R}^N$ one can write:

$$C(\boldsymbol{d}) = \| \frac{1}{\sigma_{\mathrm{i}}(\boldsymbol{r}_n)} (T_{+e} f_1 - T_{-e} f_2)(\boldsymbol{r}_n) \|_2^2 + \frac{1}{\sigma_{\mathrm{x}}^2} \|\boldsymbol{d} - \boldsymbol{e}\|_2^2 + \frac{1}{\sigma_{\mathrm{T}}^2} \|\nabla \boldsymbol{d}\|_2^2.$$

The cost function term weights σ_{i} ($[\sigma_{\mathrm{i}}]$ = HU) and σ_{T}^{-2} ($[\sigma_{\mathrm{T}}]$ = mm) account for the image noise and the regularization strength, while σ_{x}^{-2} ($[\sigma_{\mathrm{x}}]$ = mm) weights the correspondence and therefore affects the spatial uncertainty between $\boldsymbol{d}$ and $\boldsymbol{e}$ [9]. The choice of these parameters will be discussed in a later paragraph. The optimization for one iteration is performed in two steps separating the cost function:

$$\text{Step 1: } \Delta\hat{\boldsymbol{d}} = \arg\min_{\Delta\boldsymbol{d}} \| \frac{1}{\sigma_{\mathrm{i}}(\boldsymbol{r}_n)} \big(T_{+d+\Delta d} f_1 - T_{-d-\Delta d} f_2\big)(\boldsymbol{r}_n) \|_2^2 + \frac{1}{\sigma_{\mathrm{x}}^2} \|\Delta\boldsymbol{d}\|_2^2$$

$$\text{Step 2: } \hat{\boldsymbol{d}} = \arg\min_{\boldsymbol{d}} \frac{1}{\sigma_{\mathrm{x}}^2} \|\boldsymbol{d} - \boldsymbol{e}\|_2^2 + \frac{1}{\sigma_{\mathrm{T}}^2} \|\nabla \boldsymbol{d}\|_2^2$$

In the first step an optimization with respect to $\Delta\boldsymbol{d}$ is performed, given a DVF $\boldsymbol{d}$ that originates from the previous iteration or in case of the first iteration from the initialization that can be zero. Given the thus obtained $\Delta\hat{\boldsymbol{d}}$, the new DVF $\boldsymbol{e} = \boldsymbol{d} + \Delta\hat{\boldsymbol{d}}$ is computed and used in the optimization in the second step to acquire a new DVF $\hat{\boldsymbol{d}}$ that is then carried over to the next iteration. Similar to Thirion's demons algorithm the similarity term can be linearized via a first order approximation. Using the latter and evaluating the cost function at each voxel individually yields an analytic formula for $\Delta\hat{\boldsymbol{d}}(\boldsymbol{r}_n)$ in the first optimization step that can be applied for each voxel independently:

$$\Delta\hat{\boldsymbol{d}}(\boldsymbol{r}_n) = \frac{(T_{+d} f_1 - T_{-d} f_2)(\boldsymbol{r}_n)(T_{+d} \nabla f_1 + T_{-d} \nabla f_2)(\boldsymbol{r}_n)}{\frac{\sigma_{\mathrm{i}}^2}{\sigma_{\mathrm{x}}^2} + (T_{+d} \nabla f_1 - T_{-d} \nabla f_2)(\boldsymbol{r}_n)^2}.$$

As for the second step, the cost function can once again be evaluated at each voxel individually. Using differential calculus and applying the Fourier transform, forth and back, while using one of it's unique characteristics, $\mathcal{F}\left(\frac{d^2 f(x)}{dx^2}\right) = -(2\pi u)^2 F(u)$, yields a convolution. We use a Gaussian kernel for the convolution. Finally, the symmetric demons algorithm yields the DVF for the iteration:

$$\hat{\boldsymbol{d}}(\boldsymbol{r}) = \boldsymbol{e}(\boldsymbol{r}) * \frac{1}{\sqrt{2\pi}(\sigma_{\mathrm{x}}/\sigma_{\mathrm{T}})} \exp\left(-\frac{\boldsymbol{r}^2}{2(\sigma_{\mathrm{x}}/\sigma_{\mathrm{T}})^2}\right).$$

The volume is divided in S overlapping stacks g_s, so that $S-1$ registrations between neighboring stack pairs have to be performed. One registration includes the upper overlapping region of a stack s and the lower overlapping region of the subsequent stack $s+1$, which correspond to f_1, f_2 in eq. (1). Each stack g_s covers a z-axis range $z_{s,1}$ to $z_{s,2}$. An illustration of the setup is shown in figure I. All registrations yield DVFs $\boldsymbol{d}_s$ that are only valid in the upper overlapping region of their respective stack ($z_{s+1,1} < z < z_{s,2}$). In order to compute smooth DVFs that are valid for an entire stack each, linear interpolations between the deformation vectors from the lower and upper overlapping regions are performed. Based on eq. (1), the DVF for the lower overlapping region of a stack g_s is $-\boldsymbol{d}_{s-1}(\boldsymbol{r})$. The transformed stacks $Tg_s(\boldsymbol{r})$ can be computed as

$$Tg_s(\boldsymbol{r}) = g_s\Big(\boldsymbol{r} - \left(\frac{z_{s+1,1} - z}{z_{s+1,1} - z_{s-1,2}}\right) \boldsymbol{d}_{s-1}(\boldsymbol{r}_{x,y}, z_{s-1,2}) + \left(\frac{z - z_{s-1,2}}{z_{s+1,1} - z_{s-1,2}}\right) \boldsymbol{d}_s(\boldsymbol{r}_{x,y}, z_{s+1,1}) \Big), \quad (2)$$

where $\boldsymbol{d}_s(\boldsymbol{r}_{x,y}, z)$ represents deformation vectors within a plane at a fixed z-coordinate. Furthermore, let $\boldsymbol{d}_0, \boldsymbol{d}_{S+1} = 0$ as well as $z_{0,2} = z_{1,1}$ and $z_{S+1,1} = z_{S,2}$ in order for eq. (2) to be valid for the outer stacks ($s = 1; S;$) so that the DVFs is faded out towards their respective outer boundaries. The interpolation is illustrated in figure I.

Given the update equations, one has to choose the parameters σ_{i}, σ_{x}, σ_{T}. The voxel dependent parameter accounting for the image noise $\sigma_{\mathrm{i}}(\boldsymbol{r}_n)$ is set to a local estimation of the image noise. This further allows the use of σ_{x} to directly limit the update length $\sqrt{\|\Delta\boldsymbol{d}_n\|_2^2} \leq \sigma_{\mathrm{x}}/2$ [9]. Finally one can define task-specific, meaningful parameters to control the algorithm: Δ is the maximum step length for the first optimization step and σ defines the kernel width for the second optimization step, where $\sigma_{\mathrm{x}} = 2\Delta$ and $\sigma_{\mathrm{T}} = 2\Delta/\sigma$. Displacements in the range of 10 mm are expected and one can set Δ to similar values, e.g. $\Delta = 2$ mm. In order to be able to deform relatively small objects, while maintaining a sufficiently smooth DVF, we set $\sigma = 1$ mm.

The registration is stopped by a convergence criterion or after reaching a maximal amount of iterations. The convergence criterion checks if the slope of a similarity measure, here the mean squared difference between the volumes, as a function of iteration number, drops to approximately zero. A maximum iteration number of 200 was found sufficient to ensure convergence, also with larger displacements between the volumes.

III. Results

The original and corrected volumes are assembled from the individual stacks, where one stack at a time is used and the stack transition is in the middle of each overlapping region.

Figure 2 shows the results for the stack transition artifact removal, along with the original data (case 1). Grid like patterns that illustrate the deformation via the DVFs are included as well. This particular case includes severe stack transition artifacts. The method is able to remove most artifacts within the image. Feature (A), as marked in figure 2, could not be corrected entirely. Figure 3 shows another data set (case 2) that features smaller displacements and part of a coronary artery.

In this case, the method is able to remove almost all artifacts. All DVFs appear to be smooth and continuous within their respective stacks. Figure 4 shows aquivalent axial slices at a stack transition (at (A), see figure 2) from the lower and upper stack for case 1. The upper stack slices include an overlay that indicates the (absolute) differences to the lower stack. Appart from noise, the differences are considerably reduced by the registration, while maintaining a natural anatomy after the deformation. Some local bluring of the noise occurs.

Discussion and Conclusion

By using a symmetrical demons algorithm, the stack transition artifact removal method is able to considerably improve image quality. Almost all stack transition artifacts could be removed. However, in case of large displacements in the range of multiple mm, some artifacts may remain. While the demons algorithm is generally able to achieve precise registrations, the herein implemented, symmetrical version may fail to compute large displacements. To overcome this problem, another method that is more robust towards large displacements, but does not have to be as precise, might be used as initialization. For example, our previous work includes an approach that uses patch-based similarities, i.e. the comparison of displaced sub volumes, to compute an approximate DVF [10]. This method qualifies as potential initializer for the symmetric demons method.

Some local bluring may occur due to interpolation when applying the DVF. Finally, it is also of interest how well a stack transition artifact removal method performs with motion compensated data. Due to the reduction of motion artifacts, that potentially obscure stack transition artifacts, it is expected that the latter are more apparent in motion-compensated data.

Acknowledgment

Parts of the reconstruction software were provided by RayConStruct® GmbH, Nürnberg, Germany.

References

[1] H. Scheffel, H. Alkadhi, S. Leschka, A. Plass, L. Desbiolles, I. Guber, T. Krauss, J. Gruenenfelder, M. Genoni, T. F. Luescher, B. Marincek, and P. Stolzmann, "Low-dose CT coronary angiography in the step-and-shoot mode: Diagnostic performance," *Heart*, vol. 94, no. 9, pp. 1132–1137, 2008.

[2] H. Alkadhi, P. Stolzmann, L. Desbiolles, S. Baumueller, R. Goetti, A. Plass, H. Scheffel, G. Feuchtner, V. Falk, B. Marincek, and S. Leschka, "Low-dose, 128-slice, dual-source CT coronary angiography: Accuracy and radiation dose of the high-pitch and the step-and-shoot mode," *Heart*, vol. 96, no. 12, pp. 933–938, 2010.

[3] J.-F. Paul, A. Rohnean, E. Elfassy, and A. Sigal-Cinqualbre, "Radiation dose for thoracic and coronary step-and-shoot CT using a 128-slice dual-source machine in infants and small children with congenital heart disease," *Pediatric Radiology*, vol. 41, no. 2, pp. 244–249, 2011.

[4] M. Kachelrieß and W. A. Kalender, "Electrocardiogram–correlated image reconstruction from subsecond spiral computed tomography scans of the heart," *Med. Phys.*, vol. 25, no. 12, pp. 2417–2431, Dec. 1998.

[5] M. Kachelrieß, S. Ulzheimer, and W. A. Kalender, "ECG–correlated image reconstruction from subsecond multi–slice spiral CT scans of the heart," *Med. Phys.*, vol. 27, no. 8, pp. 1881–1902, Aug. 2000.

[6] ——, "ECG–correlated imaging of the heart with subsecond multislice CT," *IEEE Transactions on Medical Imaging*, vol. 19, no. 9, pp. 888–901, Sep. 2000.

[7] B. M. Ohnesorge, C. R. Becker, T. Flohr, A. Knez, and M. Reiser, *Multi-slice and Dual-source CT in Cardiac Imaging*, second edition ed. Springer Berlin Heidelberg, 2007.

[8] J.-P. Thirion, "Image matching as a diffusion process: an analogy with Maxwell's demons," *Medical Image Analysis*, vol. 2, no. 3, pp. 243–260, 1998.

[9] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, "Diffeomorphic demons: Efficient non–parametric image registration," *NeuroImage*, vol. 45, pp. S61–S72, 2009.

[10] S. Lebedev, E. Fournie, K. Stierstorfer, and M. Kachelrieß, "Stack transition artifact removal for cardiac CT using patch-based similarities," in *Medical Imaging: Physics of Medical Imaging*, ser. Proc. SPIE, 2018.

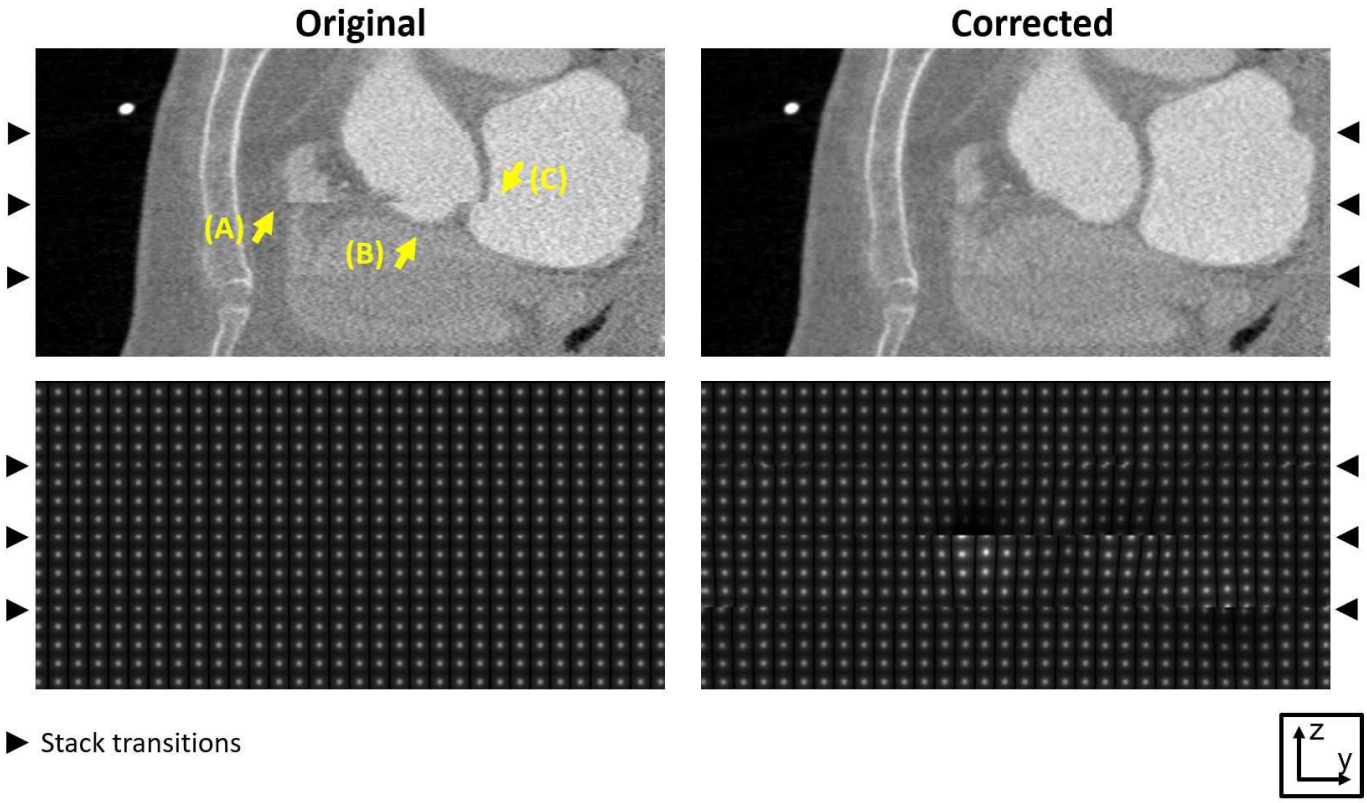


Fig. 2. Sagittal view of an original CT volume (case 1) and the volume after processing with stack transition artifact removal (top row). Grid like patterns, original and deformed using the demons algorithm computed DVFs, assembled with stacks just like the CT volume (bottom row). Some dominant stack transition artifacts are marked with (A),(B),(C). The stack transitions are marked with arrows. The patterns illustrates the DVF. The distance between its dots is 3.2 mm and 4.8 mm in y and z–direction. Brighter and darker dots indicate deformation out of and in to the image plane (White $\hat{=}$ -1.6 mm, black $\hat{=}$ +1.6 mm). CT images: $C = 0$ HU, $W = 2000$ HU.

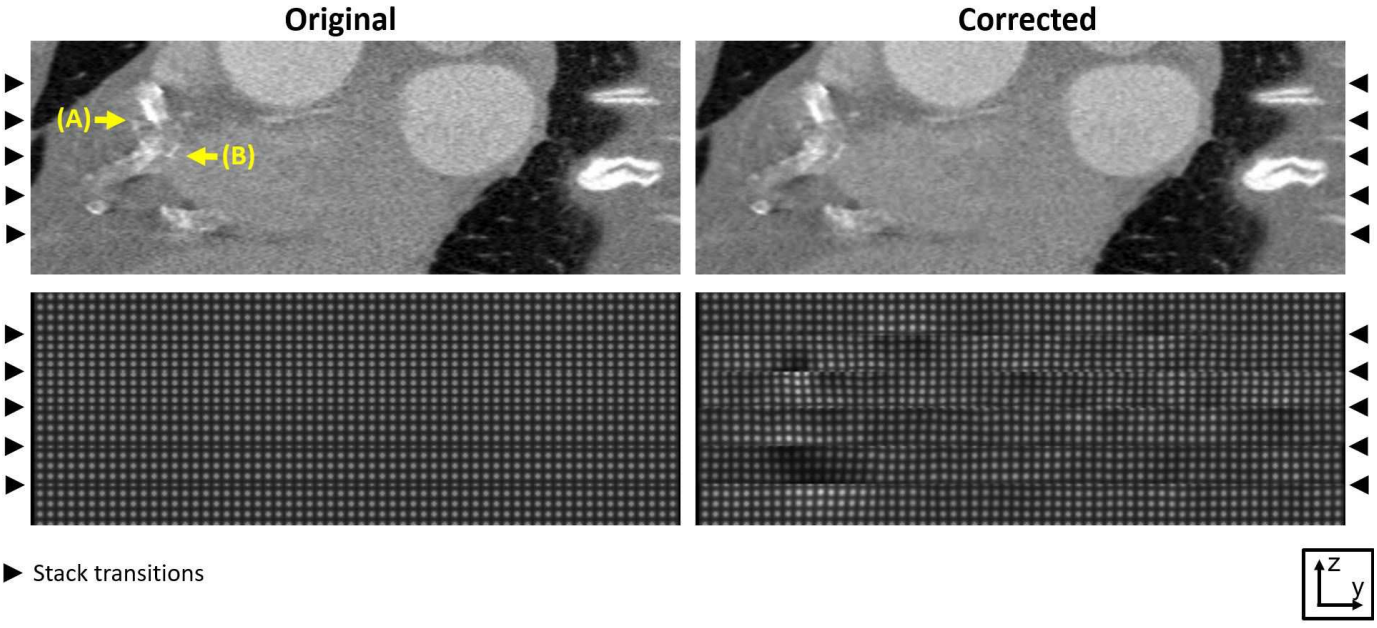


Fig. 3. Sagittal view of an original CT volume (case 2) and the volume after processing with stack transition artifact removal (top row). Grid like patterns, original and deformed using the demons algorithm computed DVFs, assembled with stacks just like the CT volume (bottom row). Some dominant stack transition artifacts are marked with (A),(B). The stack transitions are marked with arrows. The patterns illustrates the DVF. The distance between its dots is 2.4 mm and 3.6 mm in y and z–direction. Brighter and darker dots indicate deformation out of and in to the image plane (White $\hat{=}$ -1.2 mm, black $\hat{=}$ +1.2 mm). CT images: $C = 0$ HU, $W = 2000$ HU.

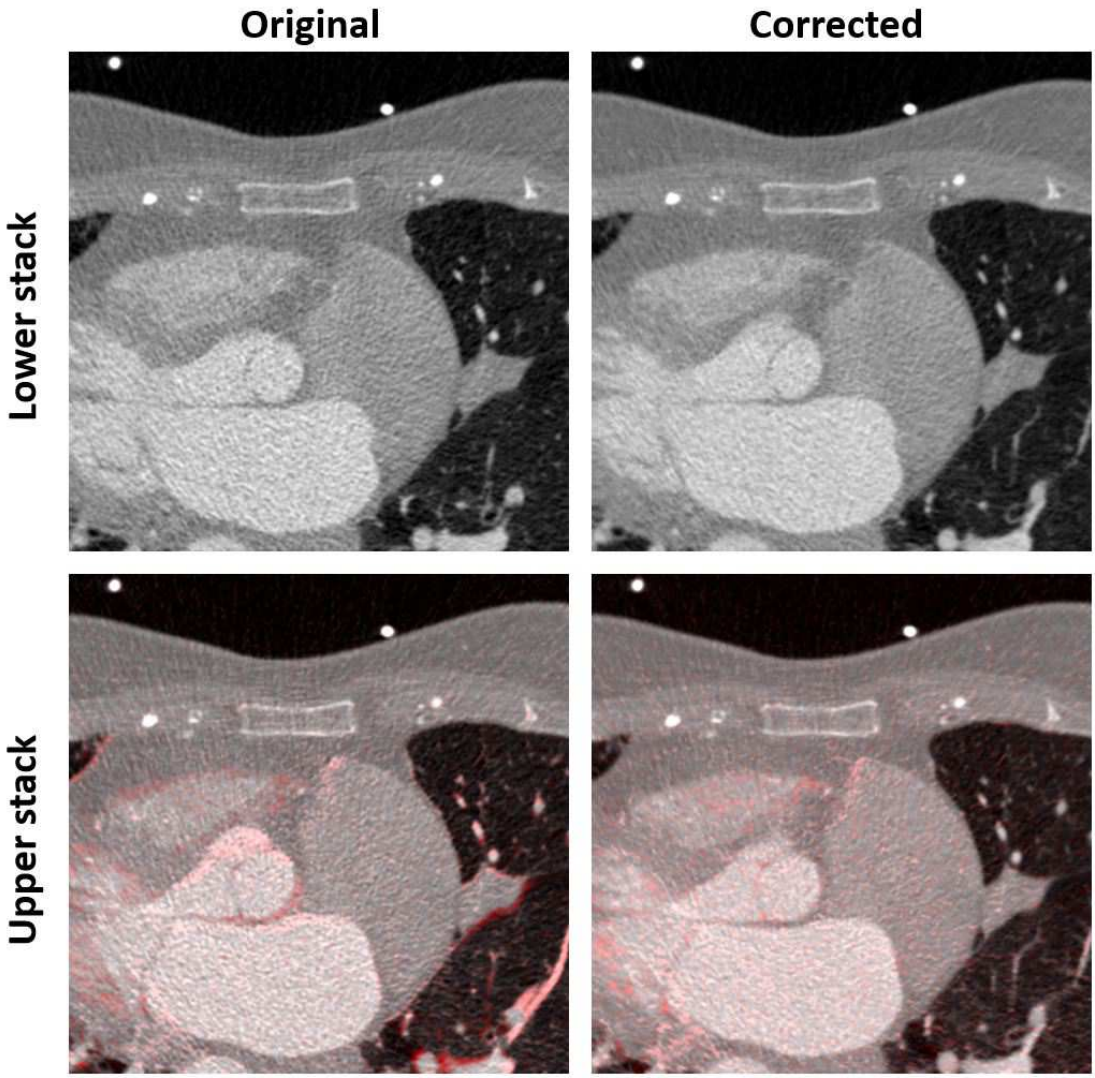


Fig. 4. Aquivalent axial slices at a stack transition, from a lower and upper stack, case 1 (see figure 2, transition at (A)). Original and corrected data. The upper stack images include a color overlay indicating (absolute) differences to the lower stack slice. Darker red indicates strong gray value differences between lung and soft tissue. Lighter red indicates differences between soft tissues and from noise. $C = 0$ HU, $W = 2000$ HU.

Noise Reduction via Filtering Temporal Differences

Brian E. Nett, Ph.D., Chuang Miao, Ph.D. and Jed D. Pack, Ph.D

***Abstract*—In cardiac CT the temporal resolution and the image noise are both important parameters. In state-of-the-art CT scanners it is possible to acquire the complete heart from an axial scan trajectory. This acquisition methodology inherently provides freedom in selection of cardiac phases, i.e. the user may select a narrow or wider range of cardiac phases. In the case that more data than the minimal amount is acquired it is useful to incorporate the additional information into a single image volume. In this work, an adaptive method is presented for incorporating the additional data in the static regions of the image volume. The method is also demonstrated on non-cardiac data as well.**

Keywords—cardiac CT

INTRODUCTION

In cardiac computed tomography (CT) there are several clinical tasks of interest including coronary angiography of iodinated vessels, calcium scoring, myocardial assessment from static or dynamic perfusion protocols, and functional imaging of the cardiac muscle movement throughout the cardiac cycle. In each of these exams it is important to reduce motion artifacts and reduce image noise. A number of methods have been proposed which apply multidimensional adaptive smoothing operations for noise reduction with emphasis on cardiac anatomy [1-2]. In this work we propose an adaptive filtration method which preserves the noise texture of direct Filtered BackProjection (FBP) reconstruction by applying the filtration operation on a difference image volume between the minimum motion artifacts image and the expanded time range image.

In generating high quality cardiac images of rapidly moving objects it may be beneficial to apply motion compensated image reconstruction [3]. The aim of motion compensated reconstruction methods is to provide an image with minimal motion artifacts, and thus the data used during backprojection corresponds to the minimal trajectory, i.e. the short scan angular range [4]. However, it is often the case that more data is available. In theory, it would also be possible to apply motion compensated reconstruction to utilize a wider range of projection data to achieve lower noise in the reconstructed image. Such a method would provide for optimal use of projection data and reduced image noise. However, this technique would involve a significant computational burden as robust motion vector fields for the entire image volume would be required to utilize the additional projection data. In this work, we present a more pragmatic approach wherein the motion compensated reconstruction is used to generate the minimal motion artifacts image and an additional image is generated from a wider scan range with reduced temporal resolution, but also reduced image noise.

Brian E. Nett, Ph.D. is with GE Healthcare, 3000 N. Grandview Blvd., Waukesha, WI, USA brian.nett@ge.com

Chuang Miao, Ph.D. wass formerly with GE Healthcare, 3000 N. Grandview Blvd., Waukesha, WI, USA

Jed D. Pack, Ph.D. is with GE Global Research Center One Research Circle, Niskayuna, NY, USA

METHODS

A) Temporal Difference Filtering

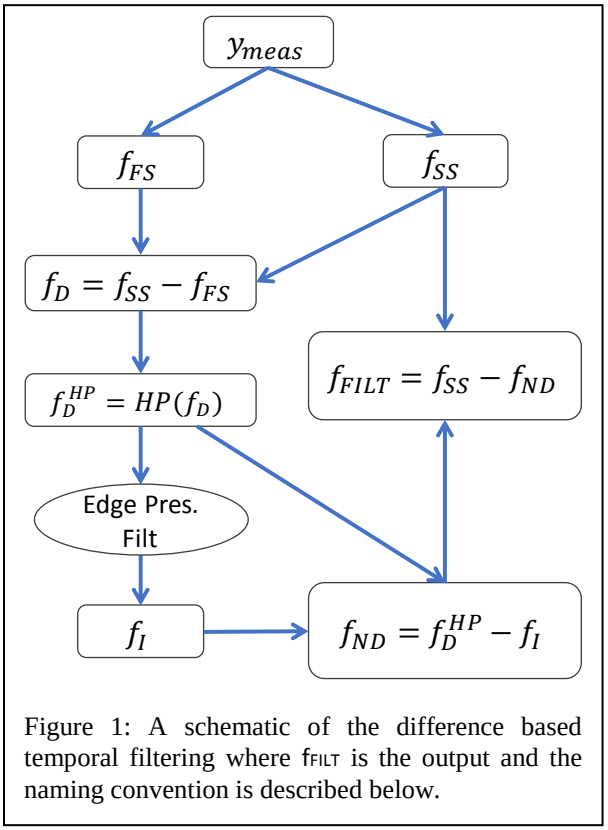


Figure 1: A schematic of the difference based temporal filtering where f_{FILT} is the output and the naming convention is described below.

Definition of the terminology used in describing the algorithm:

- y_{meas} -the measured projection data
- f_{FS} -the reconstructed image with a larger temporal footprint such as a Full Scan
- f_{SS} -the reconstructed image with a smaller temporal footprint such as a Short Scan, in the case when motion compensation is used the motion compensation image is used as this input image
- f_D - difference between the full scan and short scan image data

- f_D' - filtered version where very low frequencies are attenuated
- f_I - inconsistency between the Full Scan and the Short Scan where the quantum noise has been suppressed
- f_{ND} - image volume representing the noise difference
- f_{FILT} - the output filtered image which aims to improve the noise in the stationary portions of the image

The algorithmic flow is laid out in Figure 1. The input to the algorithm is a full scan (f_{FS}) and a short scan (f_{SS}) image reconstruction which have already been corrected for cone-beam artifacts in the case of large cone-angle acquisitions. The first step is to take a simple difference operation between the two volumes yielding the difference image volume (f_D). This volume contains all the difference information between the full and short scans. There will be some low frequency differences due to the difference in the effective weighting function. These very low frequencies do not contain significant quantum noise fluctuations as these frequencies are attenuated by the ramp filter. A frequency filter is applied to the difference image where the low frequencies are removed for the further processing. This image is referred to as (f_D'), and it is a high pass filtered version of the original. This image which contains both noise and differences between the full and short scan due to temporal inconsistency are then filtered using an edge preserving filter such as: Total Variation (TV) minimization, bilateral filtering, or non-local means [5-6]. The output after the filtering operation is that the image contains very little quantum noise and the remaining image content is attributable to the inconsistency between f_{FS} and f_{SS}, therefore this image volume will be referred to as f_I . Finally, to isolate just the noise difference between f_{FS} and f_{SS} the inconsistency image f_I is subtracted from the high pass filtered difference image (f_D'). The resulting noise difference image is referred to as f_{ND}. After the noise difference has been isolated from the temporal inconsistencies this noise difference may be subtracted from the short scan image such that the final image, (f_{FILT}) will preserve the temporal resolution of the short scan image but have noise properties very similar to the full scan image.

B) Edge Preserving Filters

Many potential methods exist to perform the edge preserving image filtration step. In implementing this method we compared the bilateral filter [5] and non-local means filter [6] and for each filter tested both 2D and 3D versions. The bilateral filter was selected for its speed and straightforward manner of parameterization. In the results shown here a fast approximation of the bilateral filter was used where a single convolution is performed in a higher dimensional space [7]. When comparing the 2D and 3D bilateral filter only marginal gain was achieved in the noise reduction with the 3D filter and therefore the 2D bilateral filter was used for the results shown here. In the standard bilateral filter there is a distance and a range component to the filter and they are both filtered with Gaussian kernels parameterized by σ_d. and σ_r. In order to adapt the filtration to the noise content of each image volume the value of σ_r was calculated for each image volume. In this study

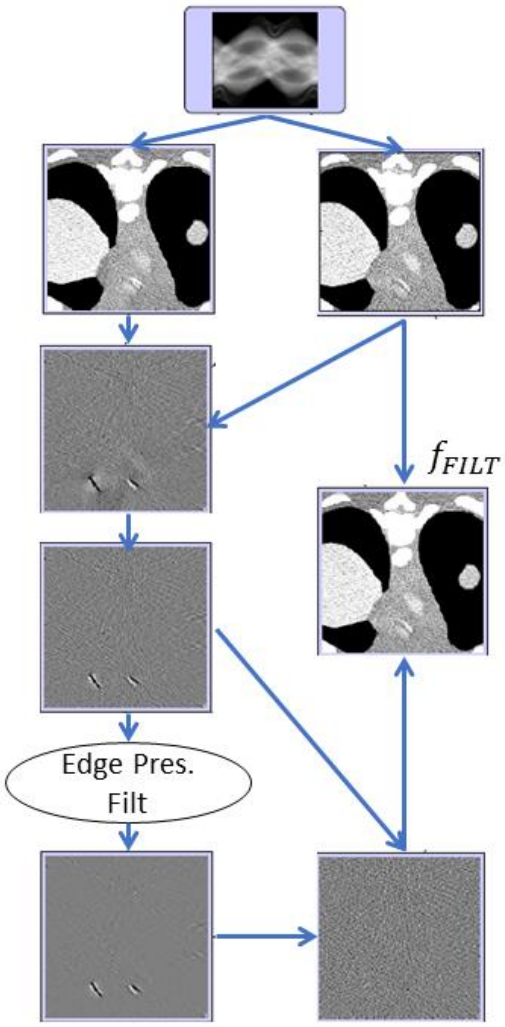


Figure 2 A schematic figure matching the same flow as Figure 1 where sample images at each point in the image chain demonstrate the algorithm. The filtering operation is applied in a more sparse domain than the original image space. The coronary vessel which is split in the fullscan reconstruction is preserved in the filtered output and the noise is suppressed compared with the shortscan reconstruction.

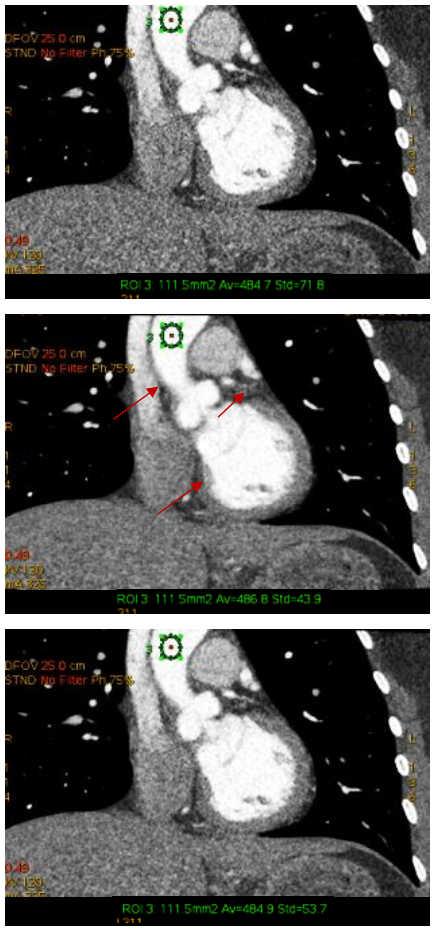


Figure 3: A representative cardiac case where the images from top to bottom are: short scan reconstruction, full scan reconstruction, and filtered output. The red arrows point to vessels and boundaries which are blurred in the fullscan but crisp in the other two images. The green ROIs are reported below the figure and report the std of the noise.

axial image volumes were used and the noise was estimated based on the central slice of the image volume. A noise mask was calculated based on the soft tissue and fat pixels by first performing a hard threshold operation selecting pixels in the range [-126 44] HU. The noise mask was applied to the difference image f_D' and standard deviation within the mask was calculated to achieve a patient specific noise level (σ_{PS}). The selection of the range filter parameter was then performed relative to σ_{PS}, i.e. $\sigma_r = \alpha \cdot \sigma_{ps}$. Given the differences in the expected level of motion the parameter α was selected separately for cardiac cases and non-cardiac cases. Image reviews were conducted with application specialists in order to determine the parameter level such that no edge degradation was perceived. In the cardiac specific case image reviews with application specialists yielded $\alpha_{cardiac}$ =1.5 and in the non-cardiac case $\alpha_{non\text{-}cardiac}$ =2.0. These parameters were used for all of the results shown here as well as: DFOV 250mm, matrix size 512x512, σ_S =16 pixels, and the first high pass filter

Figure 4: A representative non-cardiac case where the images from top to bottom are: short scan reconstruction, full scan reconstruction, and filtered output. The red arrow point to an artifact introduced by motion in the fullscan image which is not present in the other two images. The green ROIs are reported below the figure and report the std of the noise.

is implemented in Fourier space with a Gaussian parameterized by σ =64.

C) Evaluation

Clinical data and anthropomorphic phantom data was evaluated from the Revolution-CT system. Noise measurements were made from manually placing 2D regions of interest (ROIs) on uniform soft tissue regions in axial, sagittal and coronal slices. The results reported below represent the average of these measurements. The noise reduction is calculated as,

$$NI = \frac{\sigma_{SS} - \sigma_{meas}}{\sigma_{SS}}$$

where σ_{meas} will be either σ_{FS} or σ_{FILT}. Additionally, the ratio of the noise reduction achieved with the Filtered results will be compared with the noise reduction of a Full Scan as this is the upper limit on the noise reduction expected.

RESULTS

Representative data is presented comparing the standard short scan reconstruction, a full scan reconstruction and the output of the temporal difference filtration [Figures 3-4]. These figures demonstrate the ability of the method to reduce image noise without introducing motion related artifacts, that would be more likely when using a full scan image reconstruction. As noted above the parameterization of the method was chosen separately for cardiac and non-cardiac data. The parameters where chosen such that in both cases there was no perceivable degradation in resolution and boundary delineation. The algorithm was applied to 5 cardiac and 7 non-cardiac datasets and the results of the noise reduction are given in Table 1 and Table 2 respectively. The tables present the percentage of noise reduction in comparison with that enabled by full scan image reconstruction. In the cardiac parameterization the filtration was able to achieve about 2/3 the noise reduction of the full scan and in the non-cardiac case the filtration was able to achieve over 90% of the full scan result.

Cardiac Clinical Cases	**(NI Filt/ NI Full)*100**
Clinical case 1 (no SSF)	76.5
Clinical case 2 (SSF)	75.6
Clinical case 3 (no SSF)	62.3
Clinical case 4 (no SSF)	66.5
Clinical case 5 (SSF)	57
Mean	**67.6**

Table 1: Summary of the noise reduction measurements for five clinical cardiac cases. Overall, the noise improvement achieved ~2/3 the results of the fullscan. The fullscan provided a mean noise reduction of 26%, thus the filtering **reduced noise by ~18% on average**.

Non-Cardiac Cases	**(NI Filt/ NI Full)*100**
Pediatric Motion Phantom 1	94.7
Pediatric Motion Phantom 2	94.1
Pediatric Motion Phantom 3	96
Pediatric Motion Phantom 4	93.9
Clinical Abdomen 1	99.0
Clinical Abdomen 2	83.8
Clinical Abdomen 3	95.7
Mean	**93.9**

Table 2: Summary of the noise reduction measurements for seven non-cardiac cases. Overall, the noise improvement achieved ~94% the results of the fullscan. The fullscan provided a mean noise reduction of 28%, thus the filtering **reduced noise by ~26% on average**.

CONCLUSIONS

A method has been presented for explicitly filtering the temporal differences in between images which use different amounts of projection data. One specific instance of this method was tested in which axial data is the input and one reconstruction is from a fullscan while another reconstruction is from a shortscan. In this case the initial measurements estimate that the noise can be reduced by ~18% in cardiac images and ~26% in non-cardiac images without noticeable difference in spatial resolution. The larger improvements in non-cardiac imaging was attributable to less severe motion during the scanning. The method presented can work in conjunction with motion compensation algorithms. The method as presented does not modify the images where significant motion is occurring and thus it is not recommended to use very different inputs for the fullscan and shortscan.

REFERENCES

[1] H. Bruder, et al, "Spatio-temporal filtration of dynamic CT data using diffusion filters.", SPIE Medical Imaging, 2009

[2] H. Bruder, et al, "4D iterative reconstruction in cardiac CT", SPIE Medical Imaging, 2012

[3] B. Nett, et al, "Task based assessment of a motion compensation algorithm via simulation of a moving …," SPIE Medical Imaging, 2013.

[4] D. L. Parker, "Optimal short-scan convolution reconstruction for fan beam CT," Med. Phys., vol. 9, pp. 254–257, 1982.

[5] C. Tomasi and R. Manduchi, Bilateral Filtering for Gray and Color Images, IEEE CCV 1998.

[6] A. Buades, et.al., A non-local algorithm for image denoising, IEEE CVPR 2005.

[7] S. Paris and F. Durand, et. al, *A Fast Approximation of the Bilateral Filter using a Signal Processing Approach*, ECCV, 2006.

Projective Invariants for Geometric Calibration in Flat-Panel Computed Tomography

André Aichert, Bastian Bier, Leonhard Rist and Andreas K. Maier

***Abstract*—We present a new phantom design for geometric calibration of flat-panel detector CT systems. This work does not address a specific phantom but rather a toolbox for building very different phantoms for special trajectories and a general software to detect them and determine the geometric parameters of the scanner. Flexible robot trajectories and almost arbitrary distribution of metal beads in space are supported. A calibration algorithm is devised, which exploits a projectively invariant descriptor of four collinear points to solve the correspondence problem and determine the projection matrix for each projection. A proof-of-concept numerical study is presented with a randomly generated example phantom. We present a comparison to the frequently used PDS2 phantom.**

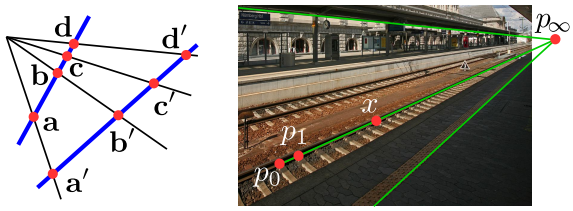


Figure 1. Example 1 (left): The corss-ratio as a projectively invariant property $\mathrm{cr}(\mathbf{a},\mathbf{b};\mathbf{c},\mathbf{d}) = \mathrm{cr}(\mathbf{a}',\mathbf{b}';\mathbf{c}',\mathbf{d}')$. Example 2 (right): Projective scales. Three collinear points p_0, p_1 and the horizon p_∞ define a projective scale which allows us to take distance measurements directly in a photograph. Here, $\mathrm{cr}(p_\infty, p_0; p_1\, x) = 7$ is the distance of x to p_0 in units of the distance between p_0 and p_1.

I. Introduction

Robotic C-arms with a flat-panel detector are becoming increasingly flexible and support circle-line, saddle, rectangular and other trajectories. Using non-circular trajectories may be beneficial in some applications for more complete data, less redundancy or simply working with hardware and space constraints. However, the geometric calibration of a flat-panel CT system based on robot odometry alone is difficult, since small variations in joint angles have a large effect in terms of detector pixels. Combined approaches are feasible [8]. Image-based approaches are therefore preferable for reproducible trajectories and it remains common practice to calibrate scanner geometry using an X-ray compatible phantom prior to data acquisition. These calibration phantoms are typically comprised of radiopaque spherical markers, which are manufactured at high accuracy, although other forms exist [3]. In addition, not all phantoms work for general projections due to their shape [7]. See Mennessier et al. for some design considerations [6].

Most calibration algorithms determine only few parameters of the (perfectly elliptical, helical etc.) trajectory, instead of the projection for individual projections, e.g. [1], [9]. Other approaches, including this work, understand the trajectory more generally as a set of independent projections, in no particular order [7], [6], [3], which makes them generally applicable. In this context, calibration with a flat-panel detector is a standard computer vision problem, merely with X-ray images instead of visible light photography. It is well-known, that the determination of an 11 DOF projection matrix requires six images of known 3D points, no more than four of which may be coplanar. If we assume a decently manufactured X-ray detector has square pixels, only 9 DOF remain [2]. The reason why established methods for pose estimation, factorization and auto-calibration are not directly applicable, is that they all rely on descriptors for matching corresponding points. Descriptors are usually based on texture, color or local gradients, none of which are salient for X-ray projections of a bead-phantom.

The contribution of this paper is a flexible and general phantom design comprised of metal beads for the determination of geometric parameters of an X-ray source and detector from its projections images.We suggest a descriptor based on four collinear metal beads, among them exactly one larger bead, in the following referred to as “pins”. A phantom may be comprised of five to about thirty of such pins, arbitrarily arranged in space. Their detection is robust, since it is unambiguous and independent. We are able to solve the matching problem with as little as three correctly detected pins. This paper presents the underlying theory and a proof-of-concept based on numerical simulation, as well as comparison to the established PDS2 phantom [7]. We show that our phantom works equally well when projected from arbitrary angles and therefore supports not just circular, but also unusual (arbitrary) trajectories. Manufacturing accuracy does not limit calibration accuracy. We suggest cheap manufacturing using 3D-printing, since our solution of the matching problem allows standard Computer Vision algorithms to replace accurate manufacturing of phantom hardware with an accurate measurement process of both structure and motion.

II. Estimation of Projective Transformations

A. Objective

In the following, we outline the established gold-standard algorithm in computer vision for the estimation of the projection matrix from detected 2D points and a known 3D geometry. We understand geometric calibration in the sense that a single 3×4 projection matrix $\mathbf{P}$ must be estimated for each projection

All authors are with the Pattern Recognition Lab, Friedrich-Alexander Universität Erlangen-Nürnberg, Germany.

image, using a pre-defined 3D phantom. The phantom consists of metal beads, which are easy to detect in 2D projection images (see Section IV-A). The problem is then to find the linear transformation between two unordered sets of points. The algorithms for both geometric calibration (i.e. estimation of a projection matrix) and registration (i.e. estimation of a linear 2D or 3D transformation) or fiducial-based rigid body tracking (i.e. rigid 3D pose) from points, are very similar and can be addressed by the algorithms in this work. For a concise mathematical notation, we will restrict ourselves to the calibration problem and present an overview of the process in this section.

Given detected 2D image points $\mathbf{x}_i \in \mathbb{P}^2$ in the projective plane (image) and known 3D points $\mathbf{y}_j \in \mathbb{P}^3$ in projective three-space (world), we seek to estimate a projection matrix $\mathbf{P}$, which minimizes the reprojection error

$$\underset{\mathbf{P}}{\operatorname{argmin}} \frac{1}{|\mathcal{M}|} \sum_{(i,j)\in\mathcal{M}} \mathrm{d}\left(\mathbf{x}_i,\, \mathbf{P}\cdot\mathbf{y}_j\right), \tag{1}$$

where $\mathcal{M} \subset \mathbb{N}_0^2$ is a set of index matches between the detected 2D points and the known 3D points and $\mathrm{d}(\cdot,\,\cdot)$ is the euclidian distance.

Given a set of at least 6 point matches one may obtain an algebraic estimate of $\mathbf{P}$ using the Direct Linear Transformation (DLT) [2, Ch. 7]. In order to be robust against outliers however, RANdom SAmpling Consensus (RANSAC) applies DLT to six randomly selected matches many times and determines the quality of the current estimate. A good measure of quality is the proportion of detected 2D points, for which a projected 3D point is close-by. The algorithm terminates when that proportion reaches a certain upper threshold, or, after a fixed number of iterations. Once RANSAC has produced a stable algebraic estimate of $\mathbf{P}$, bundle adjustment [2, Ch. 18] can be used to refine the solution by minimizing the non-linear geometric error according to Eq. 1.

B. Matching and Invariant Descriptors

The performance of RANSAC strongly depends on the frequency of outliers in the data, i.e. the chance that a random candidate match is incorrect. Unfortunately, there exist $\frac{N!}{(N-6)!}$ possibilities for a random point match with N points, so a better heuristic for establishing candidate matches than random guessing is mandatory. The common approach is to find a local descriptor of the points, typically based either on neighboring image data or local structure of the data. A good descriptor is (1) invariant to the observed transformations, so it will be possible to identify it across two images (2) local, so that occlusion or overlap in one part of the object does not affect the matching in other regions (3) salient, so that no two different points shall have a similar descriptor.

We can specifically design a calibration phantom, so prior knowledge reduces the number of possible point matches. We demand that any metal bead in the phantom is collinear with three other beads and that exactly one of these beads has a significantly larger radius. This enables us to determine a sequence of the points. We refer to a a set of four beads in this configuration as a "pin". Given four collinear points, the cross-ratio can be used as an invariant descriptor. By combining prior knowledge and a descriptor, we will show that the matching problem is effectively solved for a wide range of possible bead configurations in space. The main novelty of the paper lies in an elegant solution to the matching problem.

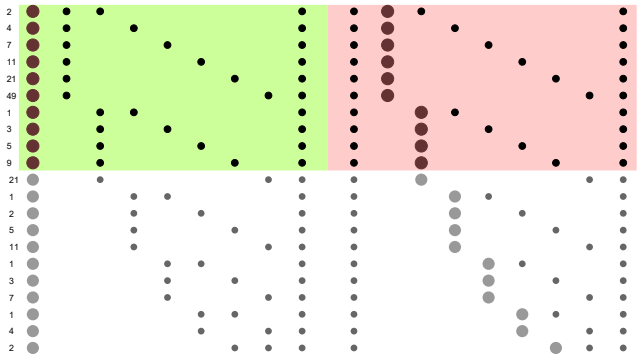


Figure 2. Example pin configurations. The left pins have a positive descriptor (green background) while the right pins have negative descriptor (red backgr.) due to the position of the big bead. Note that the 11 rows at the bottom show pin configurations with identical descriptor to one of the top ten rows.

III. A Projective Invariant for Matching Points

A. The Cross-Ratio

Many familiar properties of an object change under projective transformation, notably, length, area, angles and ratio of lengths are not generally preserved under X-ray projection. However, some invariant properties prevail, notably incidence relations, such as two lines meet in a point and the cross-ratio. The cross-ratio bears its name since it is also the ratio of ratios of the distances between four points on a line. Find an illustrative example of its application to take distance measurements directly in photographs in Figure 1. We are interested in this quantity, because it remains constant under projective transformation, including translation, rotation, scaling and especially projective distortion. Using coordinates $a, \ldots, d$ on the line, the cross-ratio is defined

$$\lambda = \mathrm{cr}\,(a, b; c, d) \overset{\text{def}}{=} \frac{(a-c)\cdot(b-d)}{(a-d)\cdot(b-c)} \in \mathbb{R}. \tag{2}$$

B. Practical Implementation in Arbitrary Dimension

In order to work with measured 2D (or, analogously 3D) points, we take the practical approach of projecting a set of approximately collinear points in arbitrary dimension $\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$, $\mathbf{a} \in \mathbb{R}^n$ to the real line $\mathbb{R}$ with the scalar product and use Equation 2 to compute the cross-ratio. W.l.o.g. let the euclidean points $\mathbf{c}$, $\mathbf{d}$ lie between $\mathbf{a}$ and $\mathbf{b}$. The scalar product with the correctly scaled vector defines a coordinate frame

$$\mathbf{i} = \frac{\mathbf{a}-\mathbf{b}}{\|\mathbf{a}-\mathbf{b}\|^2}, \tag{3}$$

between $b = 0$ and $a = 1$ on the line with

$$c = \mathbf{i}^\top(\mathbf{c}-\mathbf{b}) \;<\; d = \mathbf{i}^\top(\mathbf{d}-\mathbf{b}) \;\in\;]\,0, 1\,[\,. \tag{4}$$

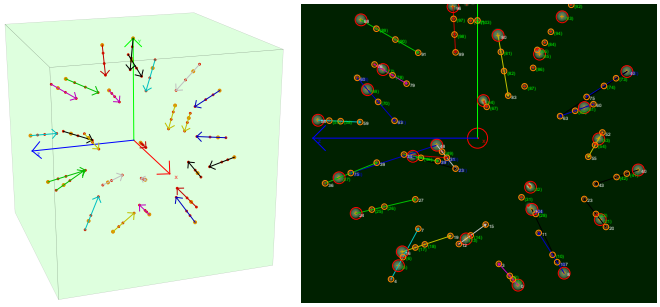


Figure 3. Left: Example of a randomly generated phantom containing several randomly distributed pens (sets of four collinear beads each). Lines with different cross-ratios are color-encoded by their index. Right: Imperfect detection of the same cross-ratios in a simulated projection. In this case there were 30 correct and 8 incorrect initial point matches in $\mathcal{M}_\star$. After RANSAC, all 96 detected beads were correctly matched, while 6 beads were truncated at the top and 6 were not detected due to an overlap.

For points not exactly the measured cross-ratio is

$$\tilde{\mathrm{cr}}\,(\mathbf{a},\,\mathbf{b},\,\mathbf{c},\,\mathbf{d}) \overset{\text{def}}{=} \mathrm{cr}\,(a,b,c,d) = \frac{d-cd}{c-cd} > 0. \qquad (5)$$

Note that $\mathbf{i}^\top \cdot$ is itself a linear projection, to which the cross-ratio is invariant, so this is consistent with Equation 2 provided that the points are exactly collinear. We observe a hyperbolic growth from $\lambda = 0$ at $d = 0$ towards infinity at $d = 1$.

C. Phantom Design

We suggest a phantom design based on several sets of four collinear metal beads. We allow two different radii of beads $r_l > r_s$. The difference between the radii should be sufficient to make classification into large and small beads easy, even under projective distortion (size is *not* a projective invariant). Each collinear four-set shall contain exactly one large bead, because the cross-ratio is symmetric with respect to the order of points $\mathrm{cr}\,(a,b;c,d) = \mathrm{cr}\,(d,c;b,a)$. We define that the large bead shall be named either $\mathbf{b}$ or $\mathbf{c}$, thus resolving the ambiguity. In applications where differently sized beads are not an option, $p2$ invariants [5] provide an alternative.

IV. Calibration Algorithm

A. Bead Detection and Approximate Collinearity

We employ the fast radial symmetry transform (FRST) [4] for bead detection, parametrized by two sets of radii in image pixels which correspond to the sizes of projected large and small beads. In experiments we found that running the algorithm twice sometimes produces spurious detections of small beads. We therefore remove detected small beads, when their center is less than one radius away from the center of a detected large bead. Selection of parameters for the FRST is done manually within this work. Next, we extract sets of collinear points. We loop over all four-sets of detected points which contain a large bead and test for approximate collinearity with a simple distance-threshold, called candidate pins. We suggest two other heuristic priors for better stability. First, we ignore all large beads which form part of more than three candidate pins. If two coplanar 3D pins project to the same 2D line, this produces $2 \cdot \mathrm{binomial}(3,6) = 40$ candidates of which only two are correct detections, so it is safer to just ignore all of them. Second, we assume that beads are spread out along the pin. If the length of a pin is larger than 20 times the shortest distance between any two of its beads, it is an unlikely candidate and we also ignore it.

B. Descriptor and Initial Matching

Since we require $b < c < d < a$, we always obtain cross-ratios $\mathrm{cr}\,(a,b,c,d) > 0$. This will be identical whether $\mathbf{b}$ or $\mathbf{c}$ is the large bead in the configuration. To differentiate between those two cases, we use λ as a descriptor if $\mathbf{b}$ is a large bead, and $-\lambda$ otherwise. The initial matching is defined by assignment of the best match of all detected sets of four collinear points, based on the descriptor. The initial matching is given by $\mathcal{M}_\star \subset \mathbb{N}_0^2$, which contains indices of two points which are candidate matches. However, $\mathcal{M}_\star$ may still contain a relatively large number of outliers. An example for this case is visualized in Figure 3, right.

C. Robust Estimation of Projection Geometry

A algebraic estimate of the projection matrix $\mathbf{P}$ is unreliable in the presence of outliers. We employ RANSAC to find, with high probability, a set $m \subset \mathcal{M}_\star$ with $|m| = 6$, whose estimated projection matrix $\mathbf{P}_m$ explains most detected points given the known 3D bead locations. By assignment of these beads to the closest projection of the known 3D beads, we can establish an improved and relatively complete set $\mathcal{M} \subset \mathbb{N}_0^2$, which likely contains many more point matches than $\mathcal{M}_\star$ but no outliers. The final step is a re-estimation of $\mathbf{P}$ using all points contained in $\mathcal{M}$. This work is restricted to a straightforward algebraic estimation using DLT. A slightly better solution may be found using non-linear optimization of the geometric error and bundle-adjustment.

V. Validation and Experiments

A. Validation with Source Positions on a Sphere

For validation we created a digital phantom based on an approximately equal distribution of points on a sphere surface using the Fibonacci series and the golden angle. It is randomized by misaligning collinear points from the ray though to the center of the sphere with $m = 27$ lines and $n = 108$ beads at a size of 3.2 mm for large beads and 1.6 mm for small beads (same bead sizes and number of points as the PDS2 phantom). An instance of such a phantom is shown in Figure 3. We present 3456 noise-free projections of 1240 px $\times$ 960 px of a phantom . We presents a validation with with a spacing of $0.308\,\frac{\mathrm{mm}}{\mathrm{px}}$ from all directions (sampled by equal longitudinal and latitudinal angles) with a source-to-detector-distance of 325 mm and a source-isocenter-distance of 200 mm, where the source positions are distributed on a sphere instead of a circle. We simulate the phantom instance from Figure 3, to compute projection matrix $\tilde{\mathbf{P}}$ with the proposed algorithm (without non-linear refinement) and compare to the ground truth projection $\mathbf{P}$. The target projection

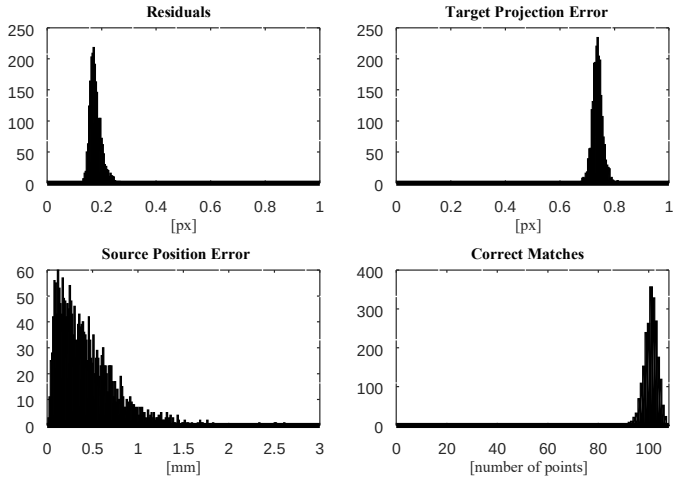


Figure 4. Distribution of errors and number of points used in 3456 calibrations from Figure 5, top. There were no outliers.

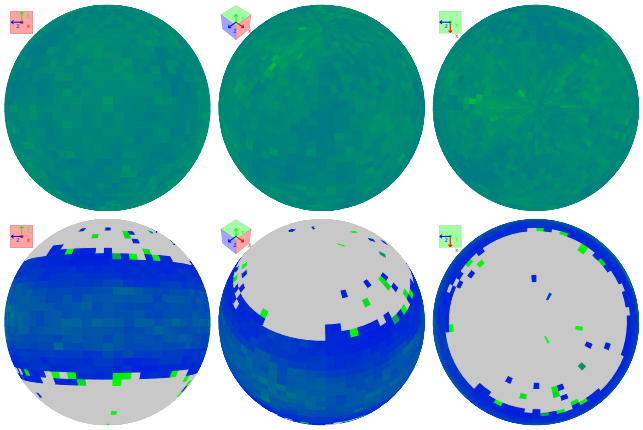


Figure 5. Evaluation of matching algorithm. Color encodes the number of points correctly detected and used for estimation of the projection matrix. The displayed range is from 54 (deep blue) to 108 (bright green). Top row: Results for an evaluation with projection direction vectors on a sphere. See also Figure 4, right. Bottom row: Results for the same evaluation using a naive algorithm for the PDS2 phantom.

error (TPE) is evaluated on 500 random points within the field of view of ~ 20 cm radius according to

$$\text{TPE} = \sum_i \mathrm{d}\left(\tilde{\mathbf{P}} \cdot \mathbf{y}_i, \mathbf{P} \cdot \mathbf{y}_i\right). \tag{6}$$

The results are themselves projected to a sphere for visualization, see Figure 5. A distribution of errors, residuals and number of points can be found in Figure 4. For all 3456 projections more than 54 points were correctly matched, with an average of 100.1 out of 108 points detected. A mean error of 0.73 px was achieved at a mean residual of just 0.177 px. The mean error of the estimated source position was 0.43 mm.

We repeated the experiment for projections of the PDS2 phantom. The PDS2 phantom is designed for circular trajectories so its detection for steep angles is more difficult and a comparison is fair only close to the equator. There were 1243 usable projections with more than 54 points correctly matched, compare Figures 5, bottom. The results for those views are comparable to the suggested phantom with a mean target projection error of 0.79 px. The residual of 0.21 px. The mean error of the estimated source position was 0.63 mm.

VI. Discussion and Conclusion

We present a new phantom design for geometric calibration of FD-CT systems. Users may quickly build phantoms of arbitrary size and shape, all of which can be analyzed with the same software. The phantoms are comprised of short pens that contain four metal beads each. The detection and matching is based on the cross-ratio, which allows robust detection, even when parts of the phantom are truncated. We present a proof of concept with a randomized numerical phantom. The suggested phantoms can work equally well if viewed from all spacial directions and support a more equal distribution of beads in space. Manufacturing accuracy does not limit the calibration accuracy, since the descriptor is the missing link to apply standard Computer Vision algorithms such as factorization and bundle-adjustment to recover both geometry of the phantom and parameters of the projection. The long-term goal of the project is to 3D print low-cost elements (pens), which the user can freely distribute in space. Application with unusual trajectories or varying size of the scanned objects are ideal for an application in material testing, for instance. The algorithm may also be useful to research as a tool for marker based tracking and fiducial-based registration.

Acknowledgments: We are supported by the German Research Foundation; DFG MA 4898/3-1 “Consistency Conditions for Artifact Reduction in Cone-beam CT”.

References

[1] M. J. Daly, J. H. Siewerdsen, Y. B. Cho, D. A. Jaffray, and J. C. Irish. Geometric calibration of a mobile c-arm for intraoperative cone-beam ct. *Medical physics*, 35(5):2124–2136, 2008.

[2] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521623049, 2000.

[3] M. W. Jacobson, M. D. Ketcha, Sa. Capostagno, A. Martin, A. Uneri, J. Goerres, T. De Silva, S. Reaungamornrat, R. Han, A. Manbachi, J. W. Stayman, S. Vogt, G. Kleinszig, and J. H. Siewerdsen. A line fiducial method for geometric calibration of cone-beam ct systems with diverse scan trajectories. *Physics in Medicine and Biology*, 2017.

[4] G. Loy and A. Zelinsky. Fast radial symmetry for detecting points of interest. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(8):959–973, Aug 2003.

[5] Peter Meer, Reiner Lenz, and Sudhir Ramakrishna. Efficient invariant representations. *International Journal of Computer Vision*, 26(2):137–152, 1998.

[6] C. Mennessier, R. Clackdoyle, and F. Noo. Direct determination of geometric alignment parameters for cone-beam scanners. *Physics in medicine and biology*, 54(6):1633, 2009.

[7] N. K. Strobel, B. Heigl, T. M. Brunner, O. Schuetz, M. M. Mitschke, K. Wiesent, and T. Mertelmeier. Improving 3 d image quality of x-ray c-arm imaging systems by using properly designed pose determination systems for calibrating the projection geometry. In *Proceedings of SPIE*, volume 5030, pages 943–954, 2003.

[8] S. Thürauf, O. Hornung, M. Körner, F. Vogt, M. A. Nasseri, and A. Knoll. Evaluation of a 9d-position measurement method of a c-arm based on x-ray projections. In *Proc. of MICCAI: 1st Interventional Microscopy Workshop. Munich, Germany*, pages 9–16, 2015.

[9] L. Von Smekal, M. Kachelrieß, E. Stepina, and W. A. Kalender. Geometric misalignment and calibration in cone-beam tomography. *Medical physics*, 31(12):3242–3266, 2004.

Multi-Motion Compensation for High-Quality Cone-Beam CT of the Head

A. Sisniega, W. Zbijewski, P. Wu, J. W. Stayman, N. Aygun, R. Stevens, X. Wang, D. H. Foos, J. H. Siewerdsen

***Abstract*—Recently developed CBCT systems offer the potential for compact, portable imaging of the brain with image quality suitable for prompt evaluation of intracranial hemorrhage – a potentially major advantage to imaging at the point of care in the critical care setting. However, slow image acquisition speed often results in artifacts due to motion, since such patients are frequently unable to follow commands and remain still, and effective immobilization is limited in the presence of life-support systems. Moreover, the motion patterns encountered in critical care brain imaging are uniquely challenging because movements of the head tend to be rapid, aperiodic, and large amplitude (~1 - 3 cm). Previously developed autofocus motion compensation methods fail to recover from such large motion amplitude, with structured similarity index (SSIM, computed relative to a static object) reduced from ~0.95 for 10 mm motion to <0.75 for 30 mm motion.**

We propose a new framework for motion compensation that expands the autofocus method in two important aspects. First is a pre-conditioning stage in which a rough estimate of rapid, high amplitude components of motion is obtained using 3D-2D registration between the motion-contaminated CBCT and the projection data. Simulation studies show that this approach yields ~25% improvement in SSIM for large, impulse-like motions of the head compared to conventional autofocus. Second is the ability to estimate motion separately in multiple regions of the image – e.g., in one field-of-view (FOV) corresponding to the central region of the image within the cranium and a second FOV corresponding to the head support. The method allows the two FOVs to move independently (with the FOV of the head holder assumed to be static in work described below), and a novel forward projection model was developed to enable high-quality model-based image reconstruction (MBIR) via penalized weighted least squares (PWLS). Compared to a conventional approach that applies the motion pattern of the head to the entire FOV (and thus erroneously assumes motion of the head holder), the proposed multi-motion PWLS method resulted in a 35% reduction of residual artifact magnitude (RMSD). Together, the motion compensation methods were shown to recover image quality at a level sufficient for low-contrast (~30-60 HU) visualization of fresh intracranial hemorrhage for realistic motion patterns.

***Index Terms*—CBCT, motion compensation, model-based iterative reconstruction, autofocus, intracranial hemorrhage, traumatic brain injury.**

I. Introduction

Recent developments in dedicated cone-beam CT (CBCT) demonstrate the feasibility for soft-tissue imaging of the brain with compact CBCT systems through careful system design [1], advanced artifact correction [2], and model-based iterative image reconstruction [3]. The resulting portable systems for point of care imaging could enable prompt evaluation of intracranial hemorrhage in traumatic brain injury or stroke.

One of the remaining challenges in brain CBCT is the relatively long image acquisition time (typically 15 – 60 s) that results in susceptibility to artifacts due to involuntary patient motion. Head motion often presents relatively sudden, large amplitude (> 10 mm) trajectories [4] and results in streak artifacts, blurring, and double contours. Motion compensation methods for such motion patterns have been proposed using fiducial markers or external trackers for motion estimation [4, 5]. However, the use of markers or trackers leads to undesirable disruption of the workflow in point-of-care brain imaging.

An image-based autofocus method has shown successful rigid motion estimation and compensation in extremities CBCT [6]. In such approach, a motion trajectory is estimated by iteratively updating a set of candidate motions that are used to generate a population of reconstructed volumes. The updates are guided by minimization of a cost function with an image sharpness term computed on the reconstructed volumes and a regularization term that penalizes motion roughness. The method was successfully applied to brain CBCT for moderate motion amplitudes of ~10 mm [7]. However, the movements of the head encountered in critical care imaging often involve even larger amplitude and rapid displacement, which challenges the autofocus optimization and requires improved preconditioning to converge upon an accurate motion estimate.

An additional challenge in clinical application is the presence of elements in the FOV that do not follow the same motion trajectory as the head (e.g. the head holder or "cradle," which remains relatively static during acquisition). This results in artifacts arising from application of the motion trajectory derived from moving elements (i.e., the head) to static elements (i.e., the head holder), even after perfect motion estimation. The artifacts are especially conspicuous when applying model-based iterative image reconstruction approaches, which involve a data-fidelity term that requires an accurate forward projection of the entire FOV. Such algorithms (e.g. PWLS) are important for high-quality soft-tissue imaging, motivating development of modified reconstruction methods to handle a multi-motion FOV.

We tackle these challenges to motion compensation through two important generalizations of our previously developed autofocus framework [6]. A multi-stage approach is incorporated that includes a pre-conditioning stage in

This work was supported by NIH Grant R01-EB018896 and industry-academy collaboration with Carestream Health.
A. Sisniega, W. Zbijewski, P. Wu, J. W. Stayman, and J. H. Siewerdsen are with the Dept. of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, USA.
N. Aygun is with the Dept. of Radiology, Johns Hopkins University, Baltimore, MD, USA.
R. Stevens is with the Dept. of Neurology, Johns Hopkins University, Baltimore, MD, USA.
X. Wang and D.H. Foos are with Carestream Health, Rochester, NY, USA.

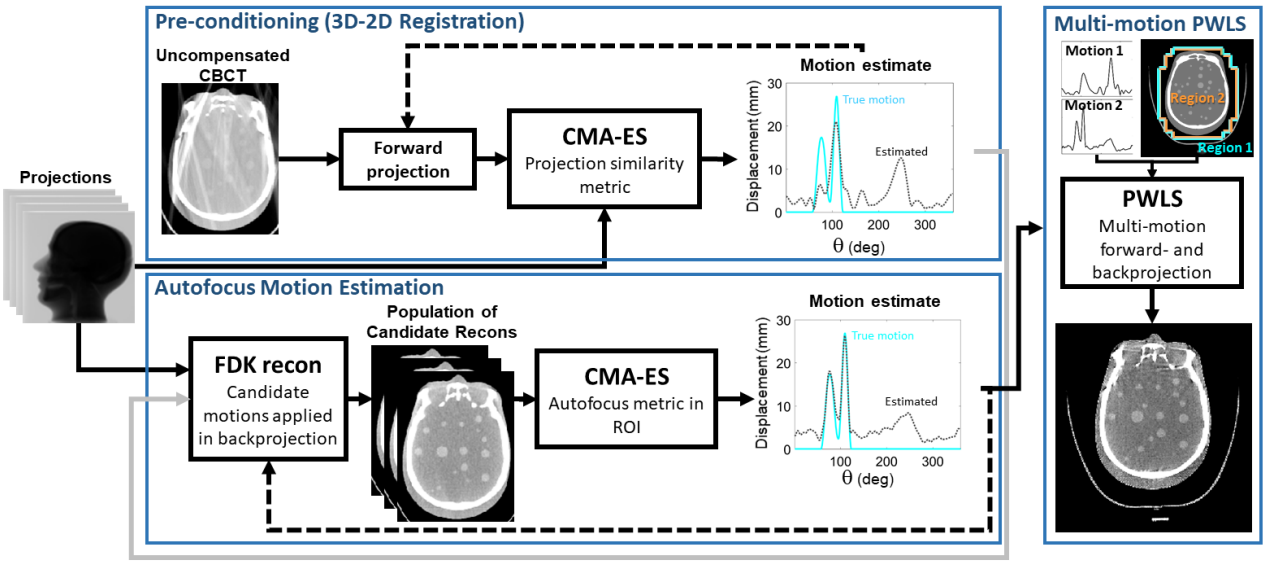


Figure 1. Workflow of the motion estimation and multi-motion reconstruction framework. The pre-conditioning stage estimates large amplitude, sudden motion components via 3D-2D registration. The second stage uses a (conventional) autofocus approach to fine-tune the initial motion trajectory by maximizing image sharpness (viz., variance of gradients). The resulting estimate feeds a multi-motion iterative reconstruction method that applies a separate motion trajectory for different regions of the image – e.g., the outer region (head holder / cradle) and the inner region (head).

which an initial estimation of the motion trajectory is obtained with 3D-2D registration using the motion-contaminated CBCT and projection data. The approach is similar to that in [8], but in that case the authors used a motion-free prior CBCT scan, which is usually not available in the clinical scenario considered here. However, the motion-contaminated CBCT provides sufficient detail to estimate large amplitude, rapid components of head motion using 3D-2D registration. This initial estimate is then refined using the image-based autofocus algorithm.

Moreover, the framework includes a modified PWLS reconstruction method accommodating multiple FOVs in motion estimation – in this case, splitting the volume in two regions with different motion trajectory. The methods are tested in simulations over a broad range of realistic, known motion trajectories and translated to first clinical studies using a prototype CBCT system developed for imaging in the neuro critical care environment.

II. Methods

This section presents the proposed motion compensation framework for brain CBCT, involving (i) improved pre-conditioning of the motion estimation algorithm (subsection A), and (ii) an iterative reconstruction approach that includes motion compensation for multiple, independently moving objects in the field of view ("multi-motion," subsection B).

A. Motion Estimation Framework

A flowchart of the proposed motion estimation algorithm is shown in Fig 1. Analogous to our previous approach [6], the head motion trajectory (T) consisted of a sequence of rigid (6 degrees of freedom) displacements (one for each projection angle θ) modelled using cubic b-splines (B):

$$T(\theta, j) = \sum_{i=0}^{N} c_{ij} B(\theta - \theta_i) \tag{1}$$

where j is the degree of freedom ($j = 1, \ldots, 6$), N is the number of spline knots, θ_i is the projection angle corresponding to spline knot i, and c_{ij} are the b-spline coefficients to be estimated.

The preconditioning stage (top box in Fig. 1) obtained the initial motion estimate T_0 by minimizing the following 3D-2D registration cost function:

$$\widehat{T_0} = \arg\min_T -\mathrm{GC}(l, \mathbf{A}(T)\mu) + \beta_m R_m(T) \tag{2}$$

where l are the log-corrected projections, $\mathbf{A}(T)$ is the forward projection operator for motion trajectory T, GC is the similarity metric (in this work gradient correlation, as defined below), and $R_m(T)$ is a regularization term that penalizes abrupt motion, weighted by the scalar β_m.

The gradient correlation (GC) similarity term was given by:

$$\mathrm{GC}(f, m) = \frac{1}{2}\{\mathrm{NCC}(\nabla_u f, \nabla_u m) + \mathrm{NCC}(\nabla_v f, \nabla_v m)\} \tag{3}$$

where ∇_u and ∇_v are gradient operators along the u and v directions of the projection domain, and:

$$\mathrm{NCC}(f, m) = \frac{\sum_i (f_i - \bar{f})(m_i - \bar{m})}{\sqrt{\sum_i (f_i - \bar{f})^2}\sqrt{\sum_i (m_i - \bar{m})^2}} \tag{4}$$

where f is the fixed image (projection data l), and m is the moving image (forward projection $\mathbf{A}(T)\mu$).

Analogous to [6], the term $R_m(T)$ penalizes the first order difference of the position of the moving patient volume in subsequent projections:

$$R_m(T) = \sum_{k=1}^{8} \sum_{q=2}^{N_\theta} \left\| \mathbf{r}_{k,q} - \mathbf{r}_{k,q-1} \right\| \tag{5}$$

where $\mathbf{r}_{k,q}$ are the vector coordinates of the k-th corner of the volume after the application of motion T for projection q.

The 3D-2D solution T_0 is used as an initial guess for the autofocus motion compensation (bottom box in Fig. 1). Analogously to [6], the following cost function is used:

$$\hat{T} = \arg\min_T -S(T,\mu) + \beta_m R_m(T), \text{ with } \hat{T}^{(0)} = \widehat{T_0} \qquad (6)$$

where $S(T,\mu)$ is a metric of image sharpness, in this case the variance of the spatial gradient [6] computed on volume μ after application of candidate motion trajectory T.

The cost functions for both stages (Eqs. 1 and 6) are not convex and exhibit local minima that challenge gradient-based methods. The minimization was performed with the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) [9].

B. *Motion Compensated Iterative Reconstruction*

Even when a perfect estimate of head motion is available, the presence of elements that remain static or follow different motion trajectories than the head introduces artifacts in the reconstructed volume if the reconstruction algorithm assumes that the entire FOV follows the same motion pattern. We propose a "multi-motion" Penalized Weighted Least Squares (PWLS) approach in which the volume is split into a number of regions (in this application, two regions) that follow different motion trajectories.

Conventional PWLS uses the following cost function:

$$\hat{\mu} = \arg_\mu \min \|\mathbf{A}\mu - l\|^2_{\mathbf{W}} + \beta R(\mu) \qquad (7)$$

where $\hat{\mu}$ is the image estimate, $\mathbf{A}$ is the forward projector operator, l are the log-corrected projections, $\mathbf{W}$ is the matrix of weighting terms (set to the inverse of the variance of the projection data, which are approximated by the raw projection measurements y, assuming a Poisson distribution), and R is an image roughness penalty in the form of a Huber function.

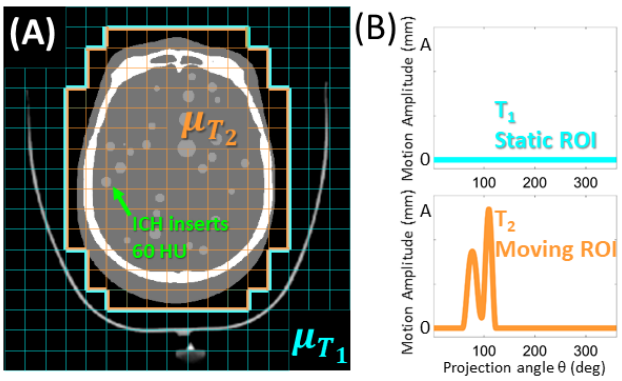


Figure 2. (A) Partition of the image volume into regions with different motion trajectory μ_{T1} (blue) and μ_{T2} (orange). The digital head phantom (following T_2) presents a natural cranium and spherical intracranial hemorrhages (ICH). The head holder follows a different trajectory (T_1) from that of the head, modeling the realistic scenario in which the head rocks or twitches within the cradle. (B) Motion trajectories corresponding to the two subvolumes in (A), μ_{T1} containing the head holder (assumed static) and μ_{T2}, containing the moving head.

In our approach, $\mathbf{A}$ is decomposed into two forward projection operators, following a derivation similar to the one used for multi-resolution PWLS, as described in [10, 11]. Matrices $\mathbf{A}_{T_1}$ and $\mathbf{A}_{T_2}$ represent the forward projection operators for regions affected by motion T_1 and T_2, respectively. The forward model can therefore be written as:

$$\bar{y} = \mathbf{D}(g)\exp\left(-\tilde{\mathbf{A}}\mu\right) = \mathbf{D}(g)\exp\left(\left[\mathbf{A}_{T_1}\ \mathbf{A}_{T_2}\right]\begin{bmatrix}\mu_{T_1}\\ \mu_{T_2}\end{bmatrix}\right) \qquad (8)$$

where $\mathbf{D}(g)$ is a diagonal matrix containing the detector gain, and μ_{T_1} and μ_{T_2} are the subvolumes following motion trajectories T_1 and T_2, respectively, as illustrated in Fig. 2. In the current work, the transition between regions was assumed discontinuous – a reasonable choice since the intermediate region is "air" (a pillow). Ongoing work investigates smooth transition between regions.

PWLS optimization was performed using the separable quadratic surrogate method with ordered subsets (OS-SQS) [12] with 20 subsets and 50 iterations.

III. Experimental Evaluation

Performance of motion estimation was evaluated in simulations using a digital head phantom (see Fig. 2) including a heterogeneous brain background with spherical inserts simulating ICH with diameter ranging from 4 to 12 mm, and 60 HU contrast. The brain background is combined with a skull with variable density extracted from an MDCT scan of a real skull.

Motion was simulated with amplitudes ranging from A= 5 mm to A = 30 mm, following the trajectory depicted in Fig. 2B. The trajectory represents a realistic motion pattern that was estimated from clinical CBCT data using the prior-image-based method in [8] (the clinical study protocol involves CBCT and MDCT imaging of the same patient [7], the MDCT was used as a prior). Projection data were obtained with and without presence of a (static) head holder, which was modelled from a segmented MDCT scan of an actual head holder used in clinical routine.

Motion estimation was performed using the two-stage approach, with 1000 iterations of the 3D-2D registration in the pre-conditioning stage and 10,000 iterations of the "multi-motion" autofocus method, with a population of 10 volumes in CMA-ES optimization, β_m = 0.05, and N = 60. Performance of motion compensation was quantified in terms of structural similarity index (SSIM) between single-motion PWLS reconstructions obtained using estimated and known motion trajectories.

Multi-motion PWLS was applied with the ROIs μ_{T_1} and μ_{T_2} delineated manually in a motion-contaminated reconstruction. Residual artifacts were visually assessed and quantified in terms of RMSD from reconstructions of simulated projections that involved the same motion trajectory, but did not include the head holder in the FOV.

IV. Results

The performance of the proposed motion compensation method is evaluated in Fig. 3 and Fig. 4. Uncompensated datasets (Fig. 3A, 3F) show severe motion-induced artifacts. For moderate motion (10 mm), motion estimation via conventional autofocus (without preconditioning) was sufficient to reduce the artifacts and achieve image quality

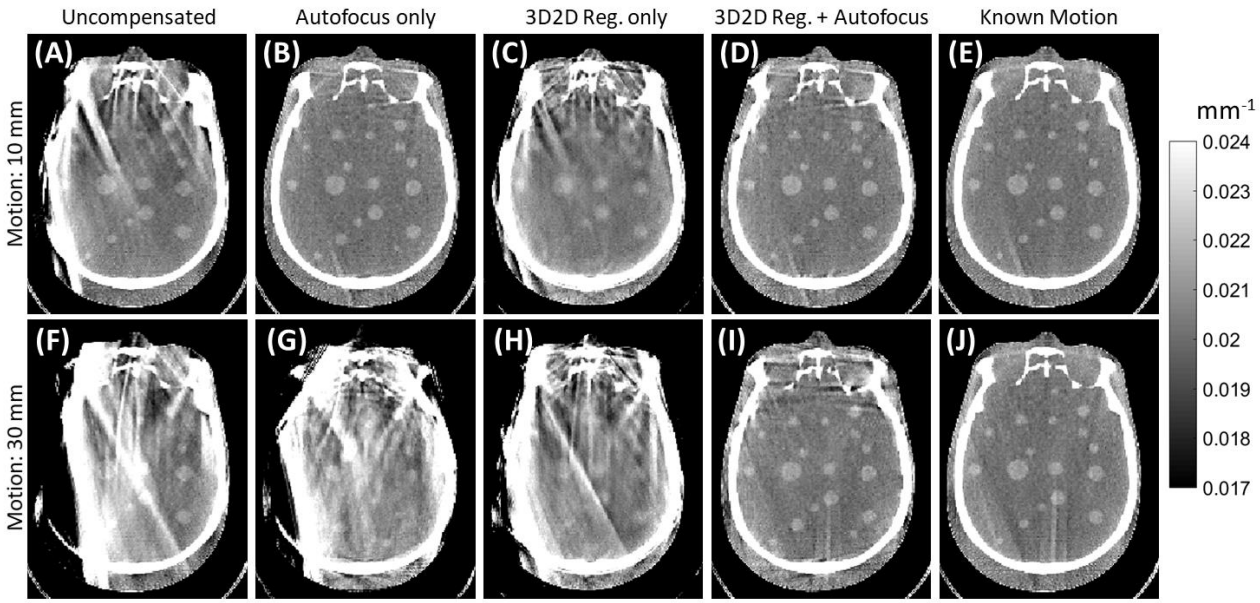


Figure 3. Motion compensation in simulated data, for moderate motion amplitude of 10 mm (A, B, C, D, E), and large motion amplitude of 30 mm (F, G, H, I, J).

comparable to reconstruction with exactly known motion (Fig. 4B, SSIM = 0.96). For a large motion of 30 mm, however, conventional autofocus did not yield an accurate motion estimate (Fig. 3G), yielding residual artifacts (SSIM = 0.75). The estimation using only 3D-2D registration was also insufficient to completely suppress motion artifacts (Fig. 3H), though it captured the major large-amplitude components of the motion trajectory (Fig. 4A). When this initialization was applied in the autofocus stage, improved reduction of motion artifacts (Fig. 3I) and a motion trajectory closer to true motion (Fig. 4A) were achieved. The resulting SSIM = 0.94 indicates good agreement with know-motion reconstruction. Final SSIM after motion compensation (Fig. 4B) shows significantly improved image quality for the pre-conditioned approach for amplitudes larger than 10 mm (minimum SSIM of 0.91).

Computational burden for autofocus motion compensation in the non-optimized version used here was previously explored in [6, 7]. In cases involving moderate motion amplitudes, solvable with only autofocus (e.g. 10 mm motion), the better initialization provided by the 3D-2D registration stage resulted in ~30% reduction in the number of iterations required for convergence (2900 iterations with autofocus only vs 2000 iterations for 3D-2D initialized autofocus, for 10 mm motion), yielding equivalent reduction in running time for the autofocus stage. However, the computational burden added by the 3D-2D registration stage resulted in similar running time (~3000 s) for both cases.

Residual artifacts were evident even when using the known motion trajectory. The artifacts were introduced by the assumption that the static head holder followed the same motion as the head. The performance of the proposed multi-motion PWLS method in reducing those artifacts is shown in Fig. 5. Even for moderate motion amplitude (10 mm), conventional PWLS resulted in distortion of the reconstructed head holder and streak artifacts at the back of the head (Fig. 5B), with RMSD = 3.2×10^{-4} mm^{-1}.

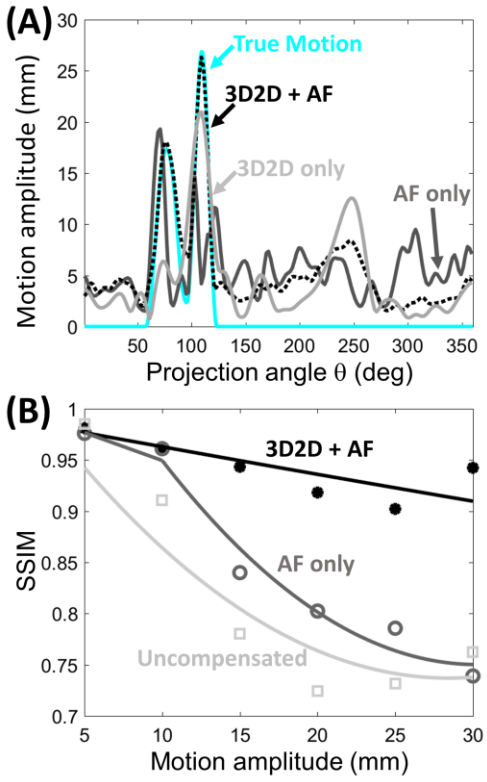


Figure 4. (A) Motion trajectories obtained with the autofocus motion estimation with no initialization (dark grey curve), with only the 3D2D initialization stage (light grey) and with the two-stage approach (black), compared to the true motion (blue). (B) SSIM between the current image estimate and that for the known trajectory plotted as a function of iteration number for the autofocus (AF) motion estimation with conventional initialization and with the two-stage approach (3D2D+AF). (C) SSIM for the resulting motion-compensated images show better performance for the two-stage approach, especially for large amplitude motions.

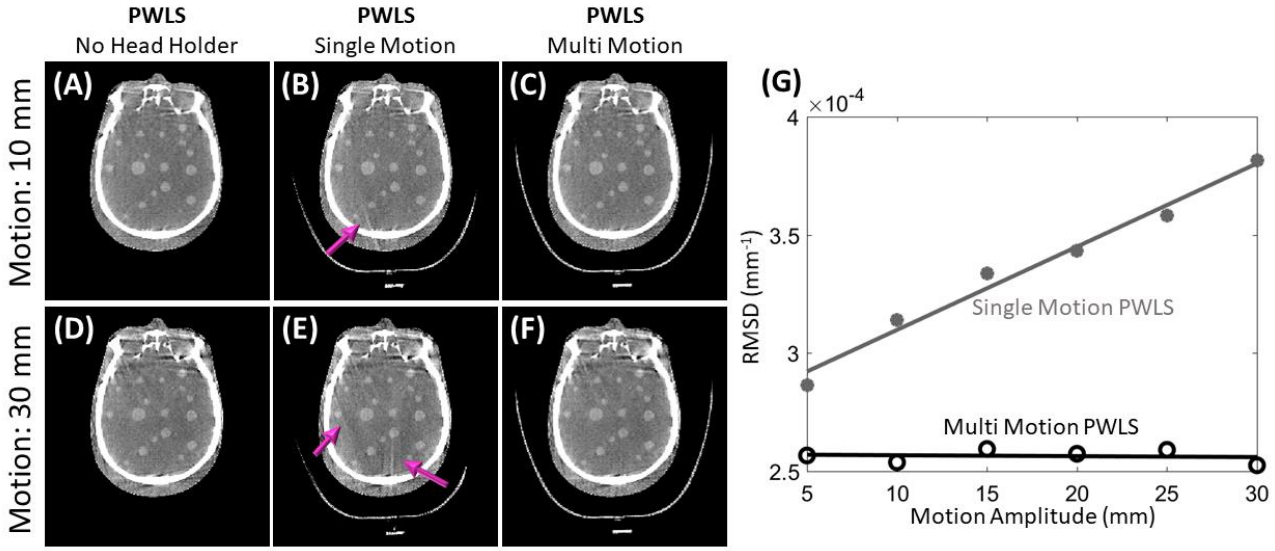


Figure 5. Comparison between single-motion and multi-motion PWLS reconstruction for moderate motion (10 mm amplitude in A, B, C) and large motion (30 mm amplitude in D, E, F). (G) RMSD (relative to a simulation without the head holder) for single- and multi-motion PWLS reconstructions.

Those artifacts were mitigated when using the multi-motion PWLS (Fig. 5C), with image quality comparable to the reconstruction with no head holder shown in Fig. 5A (RMSD = 2.56×10^{-4} mm^{-1}). The benefits of multi-motion PWLS increased with motion amplitude (Fig. 5D-F): using the conventional approach, the RMSD increased to 3.81×10^{-4} mm^{-1} for 30 mm motion (Fig. 5G). The multi-motion PWLS, on the other hand, yielded RMSD that was invariant with motion amplitude (RMSD = 2.52×10^{-4} mm^{-1} for 30 mm motion).

VI. Conclusion

The combination of 3D-2D registration-based preconditioning (using the motion-contaminated reconstruction), autofocus multi-motion estimation, and multi-motion PWLS reconstruction demonstrates the capability to perform purely image-based motion estimation (i.e., without external trackers) in the challenging scenario of high-quality CBCT of the brain. Ongoing application to clinical data of a prototype head CBCT scanner shows promising results for motion compensation in clinical environments.

References

[1] Xu J, et al. 2016. Phys. Med. Biol. 61 3180.
[2] Sisniega A, et al. 2015. Phys. Med. Biol. 60 1415.
[3] Dang H, et al. 2015. Phys. Med. Biol. 60 6153.
[4] Jang S, et al. 2017. Med. Phys. 12705.
[5] Jacobson M W and Stayman J W, 2008. IEEE Nucl. Sci. Symp. Conf. Rec., pp. 5240.
[6] Sisniega A, et al, 2017. Phys. Med. Biol. 62 3712.
[7] Sisniega A, et al, 2017. Proc. SPIE Medical Imaging, 9783, 97830K.
[8] Ouadah S, et al, 2017. Phys. Med. Biol. 62 8813.
[9] N. Hansen and S. Kern, 2004. Proc. 8th Int. Conf. Parallel Probl. Solving from Nat. - PPSN VIII, vol. 3242/2004, pp. 282–291.
[10] Cao Q, et al, 2016. Phys. Med. Biol. 61, 7263.
[11] Dang H, et al, 2016. Phys. Med. Biol. 62, 539.
[12] Erdogan H and Fessler J A, 1999. Phys. Med. Biol. 44 2835.

Estimation of the Source-Detector Alignment of Cone-Beam X-ray Systems using Collimator Edge Tracking

Christoph Luckner, Thomas Mertelmeier, Andreas Maier, and Ludwig Ritschl

***Abstract*—Upcoming applications for clinically well-established digital X-ray systems, like workflow automation, high-quality free exposures with mobile detectors, free tomosynthesis or weight-bearing full-body acquisitions with dynamic wireless detectors require a precise and reproducible method to determine the source-detector alignment. This alignment is usually obtained once in a calibration step through the imaging of a phantom with known marker geometry or via online-calibration methods, which are currently still subject to research. The former approach usually suffers from a degrading image quality over time and the latter complicates the clinical workflow due to the cumbersome and prone-to-error positioning of the patient along the additional hardware. The proposed three-step method, in contrast, only uses the existing collimator of the X-ray system and its projection and does not require any additional hardware. We assume that the extrinsic projection parameters and the orientation of the source are already known considerably well by the system, while the intrinsic projection parameters still have to be estimated individually for each scan. For evaluation, we compared the result of the proposed method to the parameters obtained through the imaging of a calibration phantom. It could be shown that the proposed method is able to achieve a high accuracy for the estimation of the intrinsic projection parameters, i. e. focal length and principal point, with a mean relative error lower than 0.5 %.**

I. Introduction

In the clinical field, X-ray imaging and X-ray computed tomography (CT) is widely used for visualizing the inside of the human body. With upcoming applications for those clinically well-established digital X-ray systems, like workflow automation, high-quality free exposures with mobile detectors, or free tomosynthesis acquisitions with dynamic wireless detectors a precise and reproducible method to determine the source-detector alignment is required. Moreover, novel X-ray systems with independently movable source and detector like the Multitom Rax (Siemens Healthcare) which can be seen in Figure 1a might benefit from the presented method. Usually, the source-detector alignment is determined once in a calibration step through the imaging of a phantom with known marker geometry [1], [2]. However, one drawback of such a one-time calibration is that the behavior of the system changes over time, which leads to a degrading image quality. There also exist several online calibration methods which rely on either phantoms or fiducial markers [3], [4], [5], other calibration objects [6] which have to be visible in the acquired image or are purely image-based [7]. However, the positioning of the patient along the additional hardware is a cumbersome and prone-to-error process and complicates the clinical workflow.

The proposed method, in contrast, uses only the already existing collimator of the X-ray system and its projection to estimate the source-detector alignment and does not require any additional hardware.

Christoph Luckner and Andreas Maier are with the Pattern Recognition Lab, Friedrich-Alexander University Erlangen-Nürnberg, Germany Email: christoph.luckner@fau.de
Christoph Luckner, Thomas Mertelmeier, and Ludwig Ritschl are with Siemens Healthcare GmbH, Forchheim, Germany

II. Materials and Methods

Any arbitrary 3-D point p^{3D} in a world coordinate system can be mapped onto the 2-D detector plane of an X-ray imaging system using a projection matrix $\mathbf{P}$

$$p^{2D} = \mathbf{P} \cdot p^{3D} = \mathbf{K}\,[\mathbf{R}\ \mathbf{t}] \cdot p^{3D}, \tag{1}$$

where $\mathbf{R}$ and $\mathbf{t}$ denote the extrinsic parameters of the imaging system, i. e. rotation and translation form the world coordinate system to the detector coordinate system, and $\mathbf{K}$ the intrinsic parameters, i. e. focal length and the principal point, as illustrated in Figure 1b.

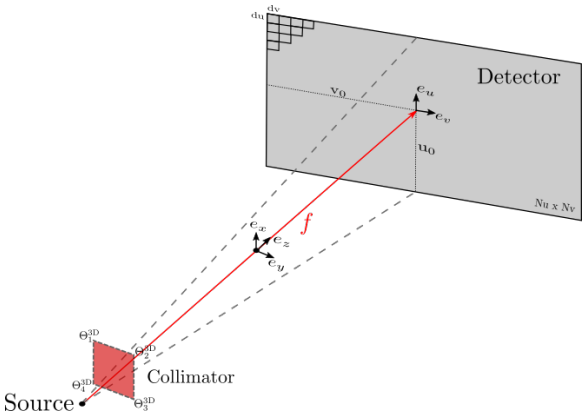
We assume, that the position of the imaging system relative to a point in the exam room is known. Hence, rotation and translation are considered to be known, whereas the intrinsic parameters still have to be estimated.

A. Algorithm

In the following a three-step algorithm to estimate the source-detector alignment using only the collimator of the imaging system will be presented.

1) Corner detection:
   In the first step the corners of the collimator c^{2D} in the X-ray image I have to be detected. This can either be done by corner detection algorithms like the Harris corner detector [8] or by an intersection of the collimator edges.
   Due to the fact that not it might happen, that not all corners of the collimator might be visible in the image, we decided to use the second approach. Therefore, we computed the Hough-transform $H(I)$ of the input image I to detect the outline of the collimator, indicated as colored lines in Figure 2. Afterwards, the intersection

(a) Twin-robotic X-ray system Multitom Rax (Siemens Healthcare) with two independently movable ceiling-mounted robotic arms for source and detector.



(b) Schematic illustration of acquisition geometry. The focal length f is indicated as red arrow, the point where the principal ray hits the detector perpendicularly is called principal point (u_0, v_0). The collimator of the X-ray system is indicated as semi-transparent red square.

Fig. 1. The used imaging system Multitom Rax (Siemens Healthcare) and a schematic illustration of the acquisition geometry.

of all lines with each other was computed, which results in the four corners of the collimator $c^{2\mathrm{D}}$.

2) Initial estimate:
The detected corners $c^{2\mathrm{D}}$ were then used as input to a least square approximation in order to compute an initial estimate for the focal length $\dot{f}$ and the principal point $(\dot{u}_0, \dot{v}_0)$

$$\dot{f}, \dot{u}_0, \dot{v}_0 = \underset{f,u_0,v_0}{\arg\min} \sum_{i=1}^{n} |\mathbf{P}(f, u_0, v_0) \cdot \Theta_i^{3\mathrm{D}} - c_i^{2\mathrm{D}}|_2^2, \tag{2}$$

where $\Theta_i^{3\mathrm{D}}$ denotes the i-th corner of the collimator in the world coordinate system.
This step was performed using a grid search approach, which exhaustively considered all possible parameter combinations.

3) Refinement step:
Since the corner detection itself might be prone to errors, e. g. due to not clearly visible collimator edges, we propose to use a refinement step which does not rely on the detected corners but uses the gradient information in the image. First, the gradient magnitude image G of the input image I has to be computed

$$G(u, v) = \sqrt{g_u^2 + g_v^2}, \tag{3}$$

with

$$\nabla I = \begin{bmatrix} g_u \\ g_v \end{bmatrix} = \begin{bmatrix} \frac{\partial I}{\partial u} \\ \frac{\partial I}{\partial v} \end{bmatrix}. \tag{4}$$

Then, based on the set of initial parameters $(\dot{f}, \dot{u}_0, \dot{v}_0)$ from the previous step, we project two adjacent collimator points $\Theta_i^{3\mathrm{D}}$ onto the detector and evaluate the line integral E between those two points utilizing the fact that there are only vertical and horizontal edges

$$E(p_i^{2\mathrm{D}}, p_{i+1}^{2\mathrm{D}}) = \int_{p_i}^{p_{i+1}} G(u, v) \, \mathrm{d}u \, \mathrm{d}v \tag{5}$$

and

$$p_i^{2\mathrm{D}} = \mathbf{P}(f, u_0, v_0) \cdot \Theta_i^{3\mathrm{D}}. \tag{6}$$

Since the gradient magnitude has its maximum at the exact location of the edge (see Figure 3), we aim to maximize the sum over all line integrals, in order to obtain a refined estimate $(\ddot{f}, \ddot{u}_0, \ddot{v}_0)$ of the intrinsic parameters

$$\ddot{f}, \ddot{u}_0, \ddot{v}_0 = \underset{f,u_0,v_0}{\arg\max} \sum_{i=1}^{n} E(p_i^{2\mathrm{D}}, p_{i+1}^{2\mathrm{D}}). \tag{7}$$

B. Experiment: Proof of Concept

In order to evaluate the proposed method, we conducted the following experiment. We acquired 3-D scans with the twin robotic X-ray system Multitom Rax (Siemens Healthcare) consisting of 152 projections along a circular trajectory.

Each scan was performed twice: first with calibration phantom (PDS-2 phantom [1]) which was later used to determine the ground truth and a second scan without phantom. We used a Matlab implementation of the proposed method to estimate the intrinsic parameters and evaluated the relative error $\varepsilon(p)$ for each intrinsic parameter $p \in \{f, u_0, v_0\}$ separately

$$\varepsilon(p) = \frac{p_{\mathrm{GT}} - \ddot{p}}{p_{\mathrm{GT}}}. \tag{8}$$

For an additional visual inspection, we performed a forward projection of the collimator edges with both the ground truth projection matrix and the reassembled projection matrix with the estimated intrinsic parameters and compared the results to each other.

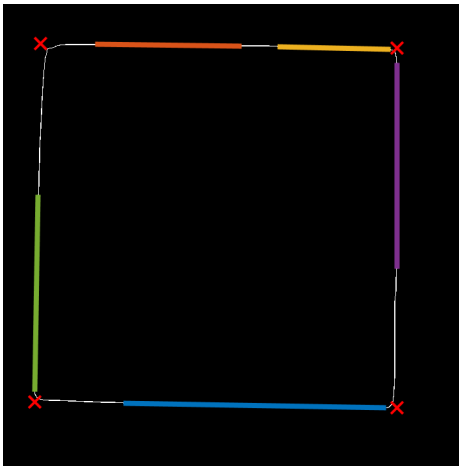


Fig. 2. Hough transformed image $H(I)$ with detected edges (colored lines) and the computed corners (red crosses).

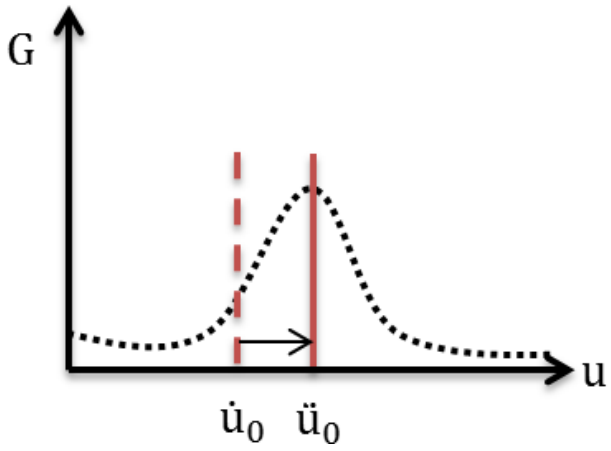


Fig. 3. Schematic drawing of the refinement step. The initial estimate $\dot{u}_0$ is moved towards the maximum of the gradient magnitude G, which yields that $\ddot{u}_0$ finally ends up at the exact location of the edge.

III. Results

Figure 4 shows the relative errors ε for each intrinsic parameter. An overall high accuracy with a mean relative error of $\bar{\varepsilon}(f) = 0.29\,\%$, $\bar{\varepsilon}(u_0) = 0.06\,\%$, and $\bar{\varepsilon}(v_0) = 0.43\,\%$ could be achieved. The computation time for each projection was about 4 seconds in a non-optimized CPU implementation. Furthermore, an exemplary result after each step as described in Section II-A can be seen in Figure 5.

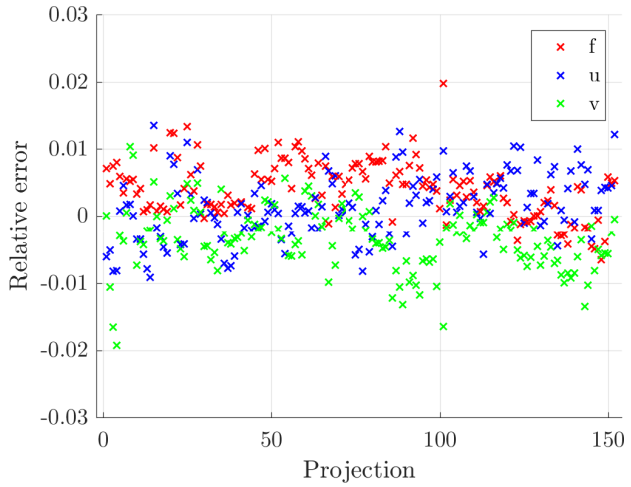


Fig. 4. Plot of the relativ error for each intrinsic variable, $\varepsilon(f)$ in red, $\varepsilon(u_0)$ in blue, and $\varepsilon(v_0)$ in green

IV. Conclusion

The presented algorithm is capable of estimating the source-detector alignment of cone-beam X-ray systems, utilizing only already existing information of the X-ray system. The proposed method might open up the possibility of further workflow automation and image quality improvement in well-established digital X-ray systems. In terms of online calibration, this method enables free tomosynthesis acquisitions in case of an exactly known detector position and orientation. Moreover, especially purely line-based trajectories where source and detector move simultaneously in parallel planes [9] can benefit from such a method since the orientation of the source as well as the extrinsic parameters of the system remain nearly constant during the entire scan. Additionally, since the method is purely image-based we do not introduce any complications in clinical workflow since no additional hardware, such as calibration phantoms or markers are required. As the initial estimation of the source-detector alignment relies on the detected corners of the collimator, a more sophisticated algorithm which is already used for auto-cropping of X-ray images [10] might lead to further improvements in the estimates.

As a topic for future research, instead of using the grid search approach in the second stage of the algorithm, a solution using, for instance, an SVD approach, could speed up the algorithm.

Disclaimer

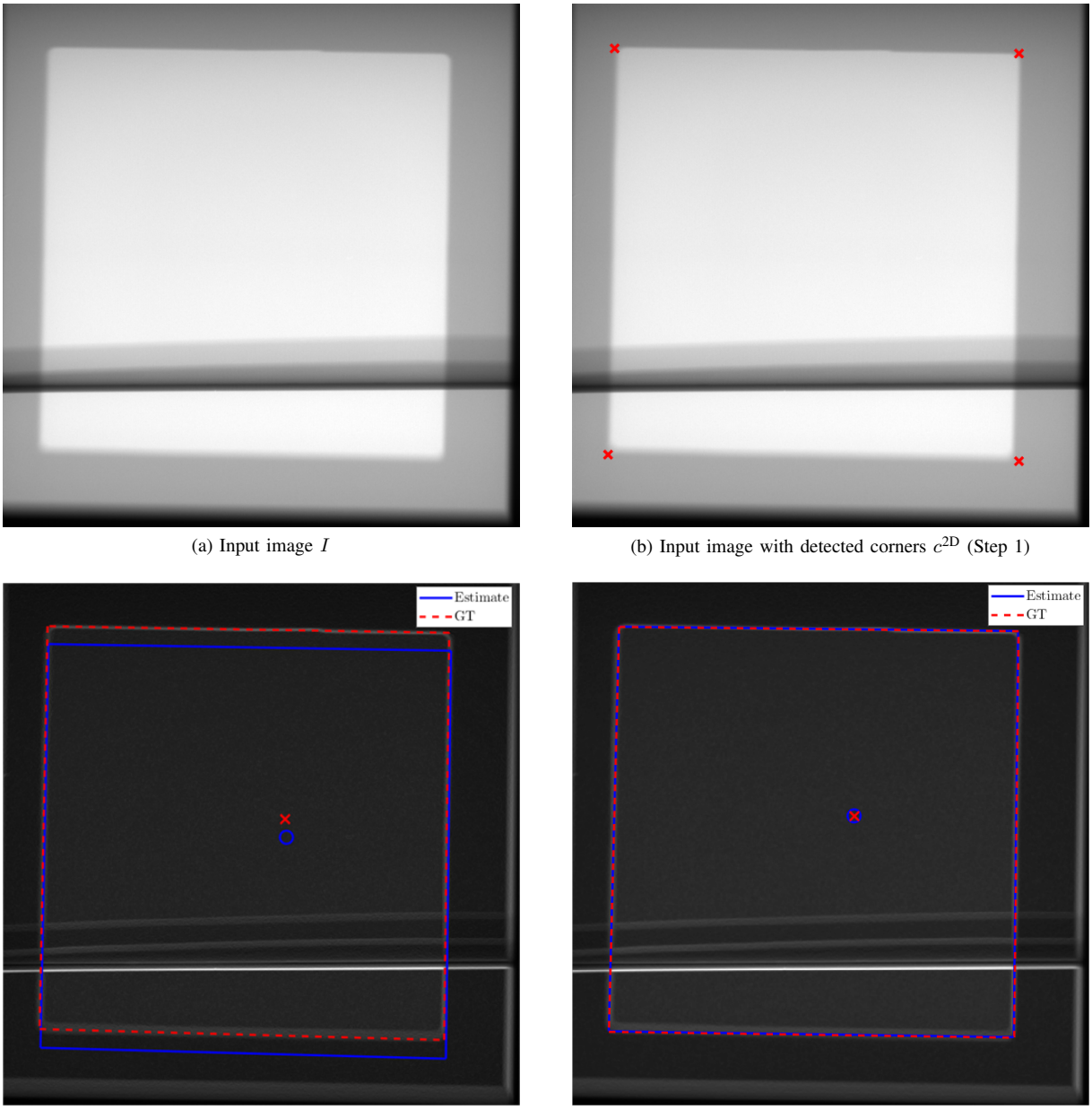
The presented method is commercially not available, its availability cannot be guaranteed. The Siemens Healthineers Multitom Rax is not available in all countries, its availability cannot be guaranteed.

References

[1] N. K. Strobel, B. Heigl, T. M. Brunner, O. Schuetz, M. M. Mitschke, K. Wiesent, and T. Mertelmeier, "Improving 3 d image quality of x-ray c-arm imaging systems by using properly designed pose determination systems for calibrating the projection geometry," in *Proceedings of SPIE*, vol. 5030, pp. 943–954, 2003.

[2] A. Maier, J. H. Choi, A. Keil, C. Niebler, M. Sarmiento, A. Fieselmann, G. Gold, S. Delp, and R. Fahrig, "Analysis of Vertical and Horizontal Circular C-Arm Trajectories," in *Proc. SPIE Vol. 7961* (SPIE, ed.), pp. 7961231–7961238, 2011.

[3] J.-H. Choi, A. Maier, A. Keil, S. Pal, E. J. McWalter, G. S. Beaupré, G. E. Gold, and R. Fahrig, "Fiducial marker-based correction for involuntary motion in weight-bearing c-arm ct scanning of knees. ii. experiment," *Medical physics*, vol. 41, no. 6, 2014.



(a) Input image I

(b) Input image with detected corners c^{2D} (Step 1)

(c) Initial estimate (Step 2)

(d) Refined estimate (Step 3)

Fig. 5. Input and output images. The forward projection of the collimator using the ground truth projection matrix is indicated as red dashed line, the estimated one as solid blue line.

[4] C. Syben, B. Bier, M. Berger, A. Aichert, R. Fahrig, G. Gold, M. Levenston, and A. Maier, "Self-Calibration and Simultaneous Motion Estimation for C-arm CT using Fiducial Markers," in *Bildverarbeitung für die Medizin 2017: Algorithmen - Systeme - Anwendungen.* (K. Maier-Hein, T. Deserno, H. Handels, and T. Tolxdorff, eds.), (Springer), pp. 56–61, 2017.

[5] C. Syben, B. Bier, M. Berger, A. Aichert, R. Fahrig, G. Gold, M. Levenston, and A. Maier, "Joint Calibration and Motion Estimation in Weight-Bearing Cone-Beam CT of the Knee Joint using Fiducial Markers," in *Proceedings of the 2017 IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (G. Egan and O. Salvado, eds.), pp. 494–497, 2017.

[6] K. Sato, T. Ohnishi, M. Sekine, and H. Haneishi, "Geometry calibration between x-ray source and detector for tomosynthesis with a portable x-ray system," *International Journal of Computer Assisted Radiology and Surgery*, vol. 12, no. 5, pp. 707–717, 2017.

[7] Y. Meng, H. Gong, and X. Yang, "Online geometric calibration of cone-beam computed tomography for arbitrary imaging objects," *IEEE transactions on medical imaging*, vol. 32, no. 2, pp. 278–288, 2013.

[8] C. Harris and M. Stephens, "A combined corner and edge detector.," in *Alvey vision conference*, vol. 15, pp. 10–5244, Manchester, UK, 1988.

[9] C. Luckner, S. Sesselmann, T. Mertelmeier, A. Maier, and L. Ritschl, "Parallel-shift tomosynthesis for orthopedic applications," in *Medical Imaging 2018: Physics of Medical Imaging*, vol. 10573, p. 105730G, International Society for Optics and Photonics, 2018.

[10] L. Zhao, Z. Peng, and X. S. Zhou, "Automated region of interest detection using machine learning and extended hough transform," Mar. 14 2016. US Patent App. 15/068,896.

Evaluation of Optimization-based Reduction of Truncation Artifacts in C-arm CBCT

Dan Xia, Yu-Bing Chang, Joe Manak, Zheng Zhang, Buxin Chen, Emil Y. Sidky, and Xiaochuan Pan

***Abstract*—C-arm cone-beam computed tomography (CBCT) has increasingly been used as an imaging tool for providing 3D anatomical information about the subjects in surgical and interventional procedures. However, measured data are often truncated due to the limited field-of-view (FOV) of the C-arm CBCT, resulting in image artifacts that can obscure low-contrast between soft tissue and tumor. Optimization-based reconstruction with an additional data-derivative fidelity term can effectively suppress the truncation artifacts. In this work, in an attempt to evaluate the performance of the optimization-based reconstruction in truncation-artifact reduction, we propose two quantitative metrics to characterize the recovery of the low-contrast objects and the reduction of streak artifacts. Quantitative investigation of the reconstruction performance demonstrates that the appropriately designed optimization-based reconstruction method can significantly reduce streak artifacts, leading to improved visualization of low-contrast structures in the reconstruction relative to clinical FDK reconstruction, in C-arm CBCT.**

I. Introduction

C-arm cone-beam computed tomography (CBCT) is used frequently as an tomographic imaging tool for providing 3D anatomical information about the subjects in surgical and interventional procedures. While it is of utility for revealing high contrast anatomic structures, there exist increased needs for localizing soft-tissue tumors in images of C-arm CBCT procedure guidance and possible complication detection. In a clinical C-arm CBCT system, the flat-panel detector forms a limited field of view (FOV), thus may resulting in considerable angularly-varying-data truncation when the scanned object is larger than the FOV. Such problem becomes increasingly serious as the surgical and interventional procedures often involve devices and tubes that are placed outside the FOV of C-arm CBCT.

Our previous work shows that optimization-based reconstruction may be exploited for reducing image artifacts caused by data truncation observed in the reconstructions obtained with the current, standard algorithm [1], [2]. In the optimization-based reconstruction method, the reconstruction problem is formulated as a constrained optimization program in which a ℓ_2-norm of data-derivative fidelity is included for effectively suppressing image artifacts caused by the data truncation, and the generic Chambolle-Pock (CP) algorithm is tailored to solve the optimization program [3], [4]. The results of visual inspection suggest that the optimization-based program with a data-derivative fidelity and the corresponding CP algorithm can yield images with significantly reduced artifacts. In this work, in addition to investigate the optimization-based reconstruction method for image-artifact reduction caused by angularly-varying-data truncation, we design and use two quantitative metrics of practical interest to evaluate the effectiveness of the optimization-based reconstruction method in terms of the recovery of low-contrast structures and the level of streak-artifact reduction.

D. Xia, Z. Zhang, B. Chen, E. Y. Sidky, and X. Pan are with the Department of Radiology, The University of Chicago, Chicago, IL 60637, USA. Y.-B. Chang and J. Manak are with Canon Medical Research USA, Inc., Vernon Hills, IL 60061, USA.

II. Data Acquisition

In this work, we collected projection data by using a Canon C-arm system in which the X-ray source has distances of 700 mm and 1100 mm to the rotation center and to the detector. The flat-panel detector consists of a 1024x1024 array of detector bins of size 0.29 x 0.29 mm^2, which is rebinned often to an 512 x 512 array with a bin size of 0.58 x 0.58 mm^2, and forms an FOV of diameter 18.7 cm. Such an FOV may be insufficient to cover the scanned subject, especially with some interventional devices presented during the intra-operative imaging process. With this system, C-arm CBCT data were collected from a physical head phantom at 247 views over a short-scan angular range of $\sim 194°$. Therefore, the data-array size for the phantom study is 512 x 512 x 247. The standard calibration scheme of the clinical C-arm system was performed on data prior to the image reconstruction.

III. Optimization-based Reconstruction

The problem of image reconstruction is formulated as a constrained optimization program [2], [5] below

$$\mathbf{f}^\star = \underset{\mathbf{f}}{\operatorname{argmin}} D_W(\mathbf{g}_m, \mathcal{H}\mathbf{f}) \text{ s.t. } ||\mathbf{f}||_{TV} \leq t_1 \text{ and } f_j \geq 0, \quad (1)$$

where $\mathbf{f}$ and $\mathbf{g}_m$ denote image vector of size N and measured data of size M, respectively, $\mathcal{H}$ the system matrix of size $M \times N$, $||\mathbf{f}||_{TV}$ the image total-variation (TV), $t_1 > 0$ a constraint parameter on the image TV, and f_j entry j in image vector $\mathbf{f}$, and $D_W(\mathbf{g}_m, \mathcal{H}\mathbf{f})$ the data divergence. In an attempt to reduce the truncation artifacts, we consider [1]

$$D_W(\mathbf{g}_m, \mathcal{H}\mathbf{f}) = \|[c\mathcal{I} + (1-c)\mathcal{D}_u](\mathcal{H}\mathbf{f} - \mathbf{g}_m)\|_2^2, \quad (2)$$

where $\mathcal{I}$ denotes the identity matrix of size $M \times M$, and matrix $\mathcal{D}_u$ of size $M \times M$ represents a two-point-difference operation (i.e., an approximation of the derivative) carried out along the horizontal axis of the detector, and c a parameter that weights the contributions of $(\mathcal{H}\mathbf{f} - \mathbf{g}_m)$ and its derivative to the data divergence. It has been shown [1], [2] that the appropriate incorporation of data derivative $\mathcal{D}_u$ can suppress effectively truncation artifacts in reconstructed images. In the work, we

tailor the Chambolle-Pock (CP) algorithm [4] to solve the program because the CP algorithm is capable of solving the non-smooth convex program in Eq. (1).

While the form of the optimization program is given by Eq. (1), the complete specification of the program depends also upon the selection of program parameters: $\mathcal{H}$, c, and t_1. Clearly, program parameters can significantly impact the program and thus its solution (i.e., the reconstruction.) We will report the strategy for the selection of these program parameters at the conference.

IV. Analysis of Images Reconstructed

As a leading interest of the clinical application is in the detection of low-contrast structures in reconstructed images, we define a contrast-resolution metric below to characterize the recovery of low-contrast anatomies:

$$\mathrm{CNR} = \frac{|\mathbf{f}_s - \mathbf{f}_b|}{\sqrt{\sigma_s^2 + \sigma_b^2}}, \tag{3}$$

where $\mathbf{f}_s$ and σ_s denote the mean and standard deviation within a region of interest (ROI), and $\mathbf{f}_b$ and σ_b the mean and standard deviation within a selected background region.

In an attempt to quantitatively characterize the streak artifacts, we use the Hough transform to detect straight lines within a two-dimensional image [6]. The Hough transform is a 2D array specified by a pair of parameters, and each element of the Hough transform has a value that is the summation of the pixel values on a line specified by the pair of parameters. The elements with the value exceeding a threshold suggest potential straight lines in the input image (i.e., the image containing streak artifacts.) Therefore, the number of the elements with values exceeding a selected threshold provides an estimate of the number of streaks and the severity of streak artifacts. Our study suggests that the Hough transform appears to be robust to gaps in straight lines and image noise.

V. Results

The scanning geometry information was incorporated into system matrix $\mathcal{H}$ in which an element was defined as the intersection length of X-ray measured and a voxel in the image array. The clinical FDK reconstructions is used as a reference in the study. We use the same image voxel sizes as those in the clinical reconstruction. The image arrays of sizes $N = 700 \times 560 \times 512$ were used for the physical head phantom, in which the voxel size is $0.47 \times 0.47 \times 0.47$ mm^3. The program parameters c and t_1 are chosen to be 0.05 and 70000 in this study.

1) Visual inspection of images reconstructed: In Figs. 1 and 2, we display the images within three transverse and sagittal slices of the phantom reconstructed, with a narrow display window for revealing details of both low-contrast structures and truncation artifacts. As a result of the metal object placed outside the FOV of the system, significant streak artifacts can be observed in the reconstructions obtained with the FDK algorithm. Conversely, the artifacts are suppressed in the reconstructions of the CP algorithm, thus some of the low-contrast structures can be discerned more readily than those obscured in the FDK reconstructions. Moreover, it is interesting to observe that some of the cone-beam artifacts near the jaw region, as pointed by the arrows in the FDK reconstruction, appear to be reduced considerably in the reconstruction of the CP algorithm in Fig. 2.

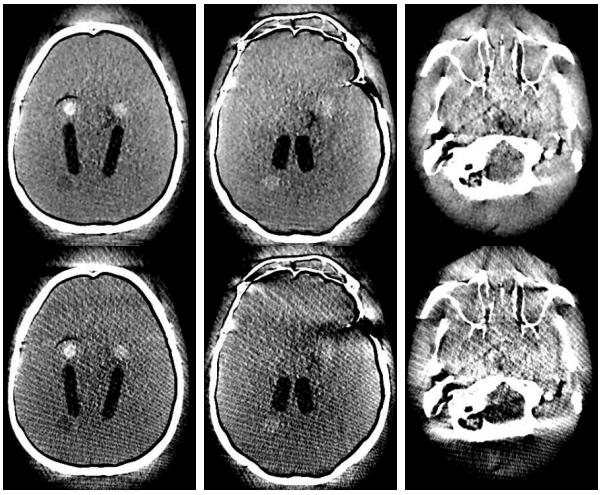


Figure 1. Reconstructed (row 1) and reference (row 2) images of the head phantom within transverse slices specified by $z = 37.6$ mm (a), 12.2 mm (b), and -41.8 mm (c), respectively. Display window: [0, 200] HU.

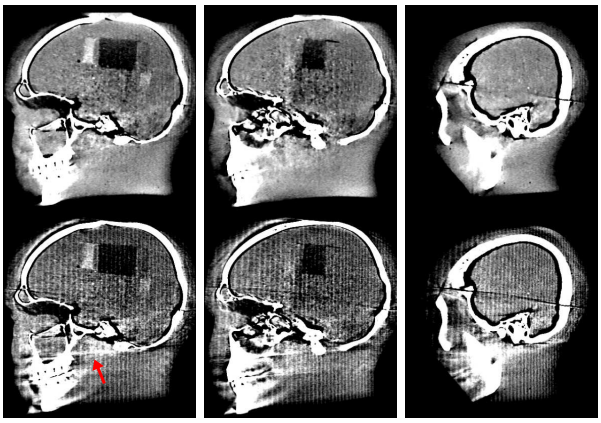


Figure 2. Reconstructed (row 1) and reference (row 2) images of the head phantom within sagittal slices specified by $y = -24.9$ mm (a), 12.2 mm (b), and -44.2 mm (c), respectively. Display window: [0, 200] HU. The arrow in the reference image indicates the cone-beam artifacts.

2) Quantitative analysis of images reconstructed: We compute the CNR metrics defined in Eq. (3) within ROIs selected in reconstructed images. As an example, three ROIs containing the cylindrical inserts at contrast level around 20 HU were selected, as enclosed by the circles in the left panel of Fig. 3. We show in the right panel of Fig. 3 the CNRs calculated within these ROIs, along with the CNRs computed from the corresponding reference image. It can be observed that the optimization-based reconstruction yields CNR higher than that obtained with the reference image, consistent with the visual inspection of Fig. 1.

Additionally, we display in Fig. 4 the Hough transforms

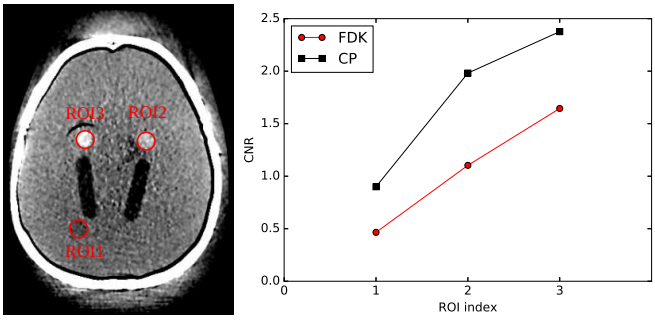


Figure 3. Reconstructed image of the head phantom within a transverse slice within three ROIs enclosed by the circles (left) and CNRs within the ROIs computed from the reconstructed (square) and reference (circle) images.

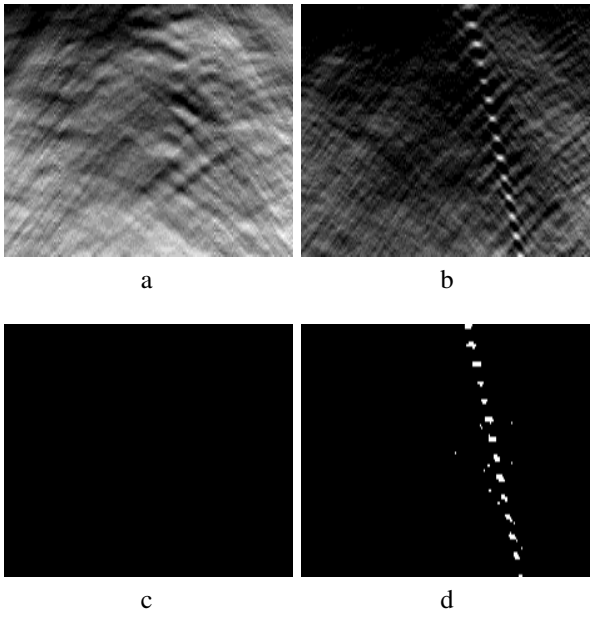


Figure 4. The Hough transforms of the reconstructed (a) and reference (c) images within a transverse slice of the head phantom, and the same Hough transforms in which the values below thresholds 120 (b) and 190 (d), respectively, are set to zero. The number, N_{streaks}, of non-zero values in (b) or (d) provides an estimate of the number of leading streak artifacts in (a) or (c).

of reconstructed and reference images within a sagittal slice, and a set of elements within the Hough transform of the reference image can be observed. Without loss of generality, we threshold the Hough transforms in Fig. 4 with a thresholding value selected and then estimate the number of the non-zero elements within each of the thresholded images. This number, denoted as N_{streaks}, is used for quantitatively characterizing the streak artifacts in the reconstructed images. We estimate N_{streaks} for all of the transverse and sagittal slices within a reconstructed image, and show in Fig. 5 the result estimated from the slices shown in Figs. 1 and 2. It can be observed that $N_{\text{streaks}} \approx 0$ for the CP images, suggesting that the optimization-based reconstruction reduces effectively the leading streak artifacts observed in the reference images with significantly non-zero N_{streaks}s.

The convergent reconstructions shown above were obtained

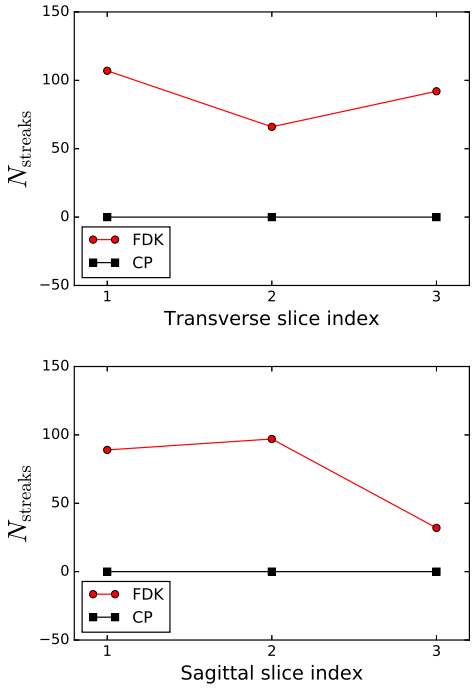


Figure 5. N_{streaks} estimated from reconstructed (square) and reference (circle) images within transverse (top) and sagittal (bottom) slices shown in Figs. 1 and 2, respectively.

with about 900 iterations. The number of iterations is neither a parameter nor a concern, as our goal is to achieve the convergent reconstruction thus avoiding the possible ambiguity in results stemmed from the iteration number. However, it is often of practical interest to inspect the reconstruction evolution as a function of iteration numbers. We display in Fig. 6 the reconstructions within a transverse slice at iteration numbers 100, 200, and 400, along with the convergent reconstruction (i.e., at $n = 900$), and in Fig. 7 CNR computed and N_{streaks} estimated within the transverse slice as a function of iteration numbers. It can be observed that the reconstruction at earlier iteration (i.e., at $n = 400$) begins to resemble the convergent reconstruction in terms of visualization and metrics CNR and N_{streaks}.

VI. Conclusions

In this work, we have carried out an optimization-based reconstruction in C-arm CBCT with an aim to reduce image artifacts caused by the angularly-varying-data truncation caused by the high contrast objects outside FOV of C-arm CBCT system. The reconstruction results demonstrate that our optimization-based reconstruction method can significantly reduce streak artifacts and improve the visualization of low-contrast structures in comparison to the clinical FDK reconstruction. In order to quantitatively evaluate the performance of the optimization-based reconstruction, two quantitative metrics, CNR and N_{streak} were introduced for characterizing the the recovery of the low-contrast objects and the reduction of streak artifacts. Results indicate that appropriately designed optimization-based reconstruction may improve image quality over the clinical FDK reconstruction.

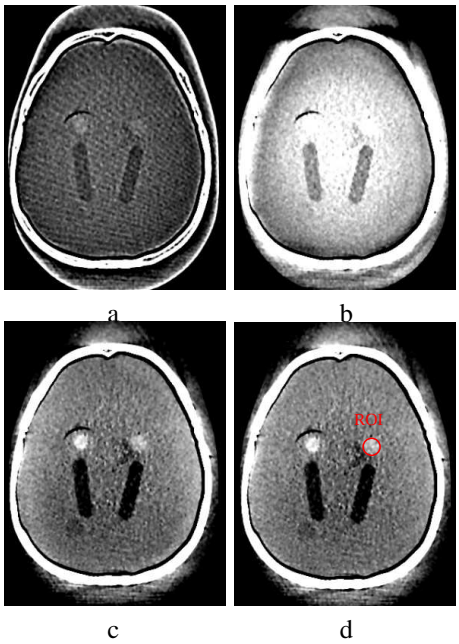


Figure 6. Reconstructions of the head phantom within a transverse slice specified by $z = 37.6$ mm at iterations n =50 (a), 200 (b), and 400 (c), respectively, along with the convergent reconstruction at n = 900 (d). Different display windows: $[-300, -100]$, $[-300, -100]$, $[0, 200]$, and $[0, 200]$ are used for revealing reconstruction details at different iterations.

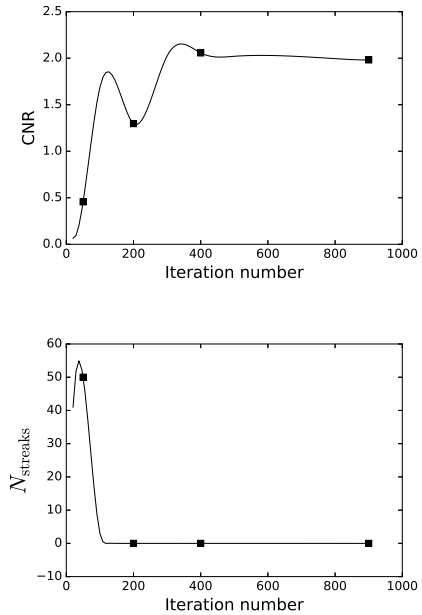


Figure 7. CNRs (left) in the ROI enclosed by the circle in, and N_{streaks} (right) estimated from, reconstructions within the transverse slice shown in Fig. 6 as functions of iterations.

ACKNOWLEDGMENT

This work was supported in part by NIH R01 Grants Nos. CA182264 and EB018102. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official NIH views.

REFERENCES

[1] E. Y. Sidky, D. N. Kraemer, E. G. Roth, C. Ullberg, I. S. Reiser, and X. Pan, "Analysis of iterative region-of-interest image reconstruction for x-ray computed tomography," *J. Med. Imaging*, vol. 1, pp. 031 007–031 007, 2014.

[2] D. Xia, D. A. Langan, S. B. Solomon, Z. Zhang, B. Chen, H. Lai, E. Y. Sidky, and X. Pan, "Optimization-based image reconstruction with artifact reduction in C-arm CBCT," *Phys. Med. Biol.*, vol. 61, pp. 7300–7333, 2016.

[3] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Math. Imag. Vis.*, vol. 40, pp. 120–145, 2011.

[4] E. Y. Sidky, J. H. Jørgensen, and X. Pan, "Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm," *Phys. Med. Biol.*, vol. 57, pp. 3065–3091, 2012, PMCID:PMC3370658.

[5] Z. Zhang, X. Han, E. Pearson, C. Pelizzari, E. Y. Sidky, and X. Pan, "Artifact reduction in short-scan CBCT by use of optimization-based reconstruction," *Phys. Med. Biol.*, vol. 61, pp. 3387–3406, 2016.

[6] R. O. Duda and P. E. Hart, "Use of the hough transformation to detect lines and curves in pictures," *Communications of the ACM*, vol. 15, pp. 11–15, 1972.

“X-map 2.0” and “iBHC2” for consistent Edema Signal Enhancement for Acute Ischemic Stroke using Non-Contrast-Enhanced Dual-Energy CT

Katsuyuki Taguchi, Toshihide Itoh, Matthew K. Fuld, Eric Fournie, Okkyun Lee, and Kyo Noguchi

***Abstract*—In this study, we propose a novel way to use material or spectral information for improved diagnoses and characterization: to suppress “noise” that makes diagnosis challenging and enhance “signals” to detect. An imaging technique (“X-map”) has recently been developed to identify acute ischemic lesions for stroke patients using non-contrast-enhanced dual-energy computed tomography (NE-DE-CT). Using the 3-material decomposition technique, the original X-map (“X-map 1.0”) eliminates fat and bone from the images, suppresses the gray matter (GM)–white matter (WM) tissue contrast, and makes signals of edema induced by severe ischemia easier to detect. There were the following two problems with the X-map 1.0: (1) biases near the skull of NE-DE-CT images; and (2) large intra- and inter-patient variations in X-map 1.0 values. In this study, we aim at addressing these problems by improving both an iterative beam hardening correction method (iBHC) and the X-map algorithm. The new iBHC (“iBHC2”) modeled x-ray physics more accurately. The new X-map (“X-map 2.0”) estimated patient-specific, local GM values—thus, maximizing the ability to suppress the GM–WM contrast, make edema signals quantitative, and enhance the edema signals that denote an increased water density for each pixel.**

***Index Terms*—CT, dual-energy, material decomposition, stroke**

I. Purposes

STROKE is the 5th leading cause of death in the United States and 87% of those events are acute ischemic stroke [1, 2] (Fig 1D) while 10% are intracranial hemorrhage stroke (Fig 1C). Early treatment (within 4–6 hours from symptom onset) can significantly improve outcomes for acute ischemic stroke patients [3, 4]; however, performing treatments such as mechanical thrombectomy to every patient could have adverse effects. It has suggested that aggressive intra-arterial mechanical thrombectomy increases the risk of hemorrhagic transformation, which can be fatal, when recanalization is successful—particularly when the infarct core is large [5]. When ischemic lesions involve more than one third of the middle cerebral artery territory, it is a strong contraindication to treatment [3, 5]. Thus, while it is critical to minimize the “door-to-needle time” in order to obtain the best outcome of the therapy, it is also essential to perform non-invasive imaging and to evaluate the presence and the extent (size) of infarction in acute ischemic stroke patients.

Non-contrast-enhanced x-ray computed tomography (NE-CT) is fast, easy to perform, available at many emergency departments, and very sensitive for detecting intracranial hemorrhage [6] (Fig 1C). The problem is that NE-CT indicates only subtle signs of ischemic stroke within the early phase of acute stroke (i.e., < 2–3 hours of symptom onset) (Fig 1D), which are mainly a loss of gray matter (GM)–white matter (WM) differentiation and hypodensity in parenchyma. At an acute ischemic site, cytotoxic edema evolves over time and decreases pixel values of the ischemic lesion due to increased water densities [7, 8]. This edema signal in acute settings, however, is difficult to detect consistently because the signal is as small as the anatomical signal difference between GM and WM in the healthy brain, which is typically 5–10 HU.

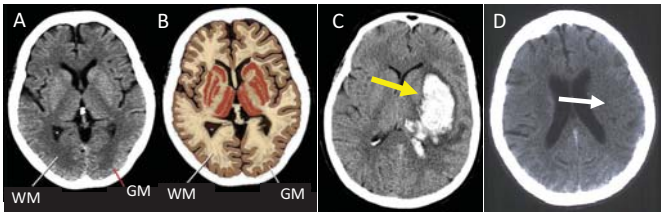


Figure 1. A NE-CT image (A) delineates the brain anatomy (B) accurately, including gray matter (GM)–white matter (WM) differentiation with 5–10 HU. WM has a greater fat content than GM, which makes it appear darker than GM on NE-CT. NE-CT clearly indicates a hyper-dense intracranial hemorrhage lesion (C, arrow), while NE-CT shows only a subtle hypo-dense signal at an acute ischemic lesion within a few hours of onset (D, arrow). Images are obtained from (A–B) http://www.strokecenter.org/professionals/brain-anatomy/cellular-injury-during-ischemia/edema-formation/, (C) https://www.researchgate.net/figure/a-Ischemic-strokeshown-as-dark-black-area-b-Hemorrhagic-strokeshown-as-bright_fig1_40452206, and (D) http://slideplayer.com/slide/3883198/.

Often a loss of GM–WM differentiation (or a regional decrease of GW–WM contrast) is a critical sign of ischemic lesion.

A novel imaging method, X-map, has been developed that uses non-contrast-enhanced dual-energy compute tomography (NE-DE-CT) [9] with the dose and time duration comparable to the conventional NE-CT. X-map uses the dual-energy material decomposition technique [10, 11] (Fig. 2a) and projects all of the pixel values along the direction parallel to the fat slope onto the GM line. The projection process essentially eliminates the anatomical GM–WM contrast, making edema lesions easier to detect. X-map values denote relative positions of the projected point on the GM slope—thus, relative edema levels—100 for tissues without edema and <100 with edema at different levels. While X-map algorithm is technically a derivative of “3-material decomposition method” [10] and “virtual non-contrast-enhanced imaging,” [11] the concept of X-map is a diagnostic task-specific material decomposition, which is novel and deserve recognition.

There have been two independent studies, which demonstrated the potential viability of X-map. One study compared X-map with diffusion weighted imaging (DWI)

K. Taguchi and O. Lee are with Johns Hopkins University School of Medicine. T. Itoh, M. K. Fuld, and E. Fournie are with Siemens Healthineers. K. Noguchi is with University of Toyama.

acquired 30–60 min after NE-DE-CT, and concluded that there was a very good agreement in acute infarctions between X-map and DWI, including ischemic regions that recovered after treatments [9]. Another study assessed X-map with 24-hour follow-up NE-CT. They concluded that X-map can detect edema and quantify the core infarct volume more accurately than the conventional NE-CT [12]. The sensitivity, specificity, positive predictive value, and negative predictive value of X-map were 93.3%, 100%, 100%, and 91.7%, respectively, while those of the conventional NE-CT were 80%, 72.7%, 80%, and 72.7%, respectively.

It has been found, however, that the original X-map (which we call "X-map 1.0" in this paper) has the following two problems despite these promising results. First, even when an iterative beam-hardening correction ("iBHC1") is employed, NE-DE-CT images have smearing artifacts near the brain–skull boundaries. Second, X-map values have large inter-patient and intra-patient variations. We propose to address these problems by improving NE-DE-CT images and developing a new version of X-map (which we call "X-map 2.0") and a new scheme to enhance edema signals (ESE for edema signal enhancement) in this study. Quantitative assessments will be performed on the accuracy and variations of NE-DE-CT images and X-map images in this study. A part of the results has been presented at RSNA 2017 as a 7-min presentation.

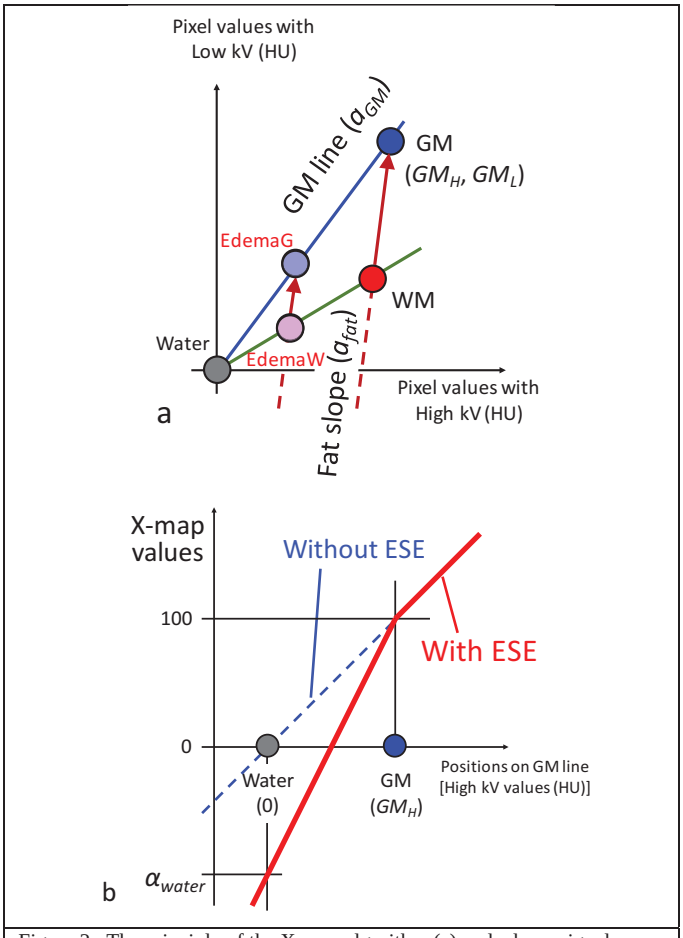


Figure 2. The principle of the X-map algorithm (a) and edema signal enhancement (ESE) (b). (a) WM contains more fat than GM, and thus, the WM point is located on the fat slope through the GM point. With edema evolving due to the acute ischemia, both GM and WM points move toward water and get to EdemaG and EdemaW, respectively. X-map projects all of the pixel values onto the GM line along the direction parallel to the fat slope. WM and EdemaW are projected onto GM and EdemaG, respectively. This projection process eliminates the presence of fat, hence, GM–WM contrast. (b) An X-map value denotes a relative position of the projected point on the GM line, a value of 100 corresponds to the GM point. The values below the GM are scaled by $100/GM_H$ without ESE and by $(100+\alpha_{water})/GM_H$ with ESE. (GM_H, GM_L) are the pixel values of GM at (high, low) kV images.

II. Methods

We briefly describe (A) the patients, (B) CT acquisition and image reconstruction with iBHC, (C) X-map, (D) Image and statistical analyses.

A. *Patients*

Eleven patients were included in this study, among which 6 patients had DWI-confirmed acute ischemic stroke and received DE-NE-CT within 3 hours from the symptom onset and 5 patients with negative CT findings.

B. *CT Acquisition, image reconstruction with iBHC*

A NE-DE-CT scan was performed with two X-ray tubes being operated at 80 kV for 800 mAs and 150 kV with a tin filter for 533 mAs (SOMATOM Force, Siemens Healthineers, Forchheim, Germany). The other scan parameters were: 64×0.6 mm; 1.0 s/rot; pitch, 0.7; CTDIvol, 80.08 mGy, which was the standard care protocol for the institution.

Four sets of NE-DE-CT images—high and low kV images with iBHC1 for skull and another pair of images with the new iBHC ("iBHC2")—were reconstructed using an iterative reconstruction algorithm ADMIRE (Siemens Healthineers, Forchheim, Germany) with a quantitative intermediately sharp soft tissue kernel (Qr40). Both of the iBHC methods were essentially a generic 2-pass beam-hardening correction discussed by Hsieh [13] and others; iBHC2 modeled X-ray physics more accurately than iBHC1 did.

C. *X-map with or without ESE*

A non-local mean filter was applied to all of the four sets of images independently, from which the following four X-map images were computed: X-map 1.0 without ESE (i.e., the original X-map algorithm), X-map 1.0 with ESE, X-map 2.0 without ESE, and X-map 2.0 with ESE.

X-map 1.0 projected all of the pixel values along the direction parallel to the fat slope onto the GM line (Fig 2a) and eliminated GM–WM contrast. The relative positions of the projected positions on the GM line were then converted to X-map values either with or without ESE (Fig 2c). The X-map value of 100 referred to brain tissues with no edema and a decrease of X-map values indicated edema at different levels. Parameters for the fat slope and GM values (see Fig 1a) were fixed for all of the patients with X-map 1.0. The slope of the GM line (a_{GM}), calculated as $a_{GM}=GM_L / GM_H$, was also fixed implicitly.

X-map 2.0 estimated and used the local GM values to compute X-map values using the following 5 steps: **Step 1.** Perform image segmentation on low kV images and segment the brain into GM, WM, and other low-density materials. **Step**

2. Estimate *local* GM values for high kV and low kV independently by using a 3-D box kernel and averaging pixel values within the kernel that were labeled as GM in Step 1. **Step 3.** Calculate the slope of the *local* GM line (a_{GM}). **Step 4.** Project the pixel value onto the *local* GM line adaptively, either along the fat slope as X-map 1.0 did if the pixel value was below the GM line (Fig 2a) or to the GM point if the pixel value was above the GM line. The former process eliminated the presence of fat and GM–WM contrast on a local basis while the latter process decreased the source of potential false positives by eliminating unnaturally small X-map values. **Step 5.** Convert the relative position of the projected point on the local GM line to the X-map value of the pixel, either with or without ESE (Fig 1c). Both versions of the X-map were implemented using Matlab (Mathworks, Inc.; Natick, MA).

D. *Image analyses*

Bias in CT pixel values: The pixel value was measured at GM, 6–10 locations near the skull and 2–3 locations near the center of the brain. The mean of the peripheral region-of-interest (ROI) values and the mean of the central ROI values were calculated. The difference between the central GM value and the peripheral GM value was considered a bias due to beam hardening artifacts.

Inter- and intra-patient variation of X-map values: X-map values were measured at 5–9 locations per patient on GM without infarct, resulting in 76 locations for 11 patients. The patient-mean X-map value was calculated by averaging multiple ROI values per patient. A zero-mean X-map value was calculated for each patient by subtracting the corresponding patient-mean X-map value from ROI X-map values. The standard deviation of the patient-mean X-map values over multiple patients denotes the inter-patient variation; the standard deviation over multiple zero-mean X-map values denotes the intra-patient variation.

Correlation with DWI: The ROI values at and near ischemic lesions were measured at 7–14 locations per patient (thus, in total of 45 samples) with X-map images, CT images, and DWI images using five patients who had DWI within 2 hours of NE-DE-CT. The DWI values were then converted to a relative increase of DWI signals from the background as d-DWI = (DWI_{lesion}–$DWI_{background}$) / $DWI_{background}$ ×100 (%). Correlation coefficients with d-DWI were calculated.

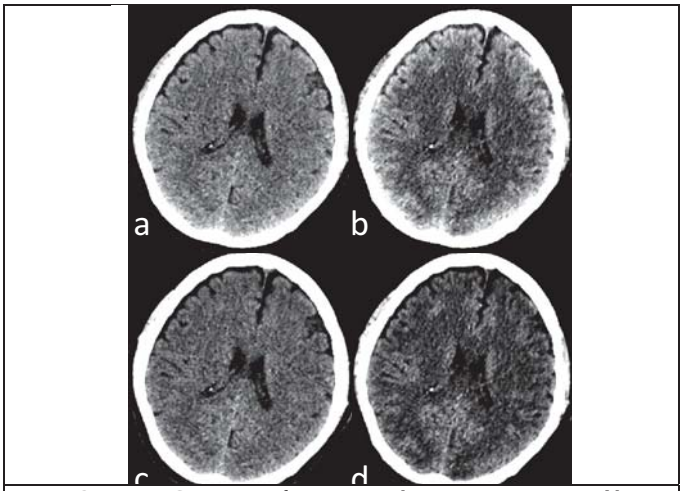


Figure 3. NE-DE-CT images of an acute stroke patient, reconstructed by the old iBHC (iBHC1) (A,B) and the new iBHC (iBHC2) (C,D). (A,C) High kV images, (B,D) low kV images. Biases (or whitening cupping artifacts) near the skull, observed with iBHC1 (A–B), were significantly improved by iBHC2. Window width is 50 HU.

III. Results

Figure 3 shows typical CT images reconstructed with iBHC1 and iBHC2. The biases or whitening/cupping artifacts near the skull were improved substantially by iBHC2 compared with iBHC1 and a Wilcoxon signed-rank test indicated the changes were statistically significant: 2.5±2.0 HU for iBHC2 versus 6.9±2.3 HU for iBHC1 with high kV images, P<0.01; 1.5±3.6 HU versus 12.8±3.3 HU with low kV images, P<0.01.

Figure 4 presents images of the acute stroke patient with NE-DE-CT performed 40 min after the symptom onset and DWI obtained 11.5 hours after the NE-DE-CT. The DWI image (Fig. 4A) showed an infarct in the left hemisphere; neither the high kV image (Fig. 4B) nor the low kV image (Fig. 4C), both after non-local mean filter being applied,

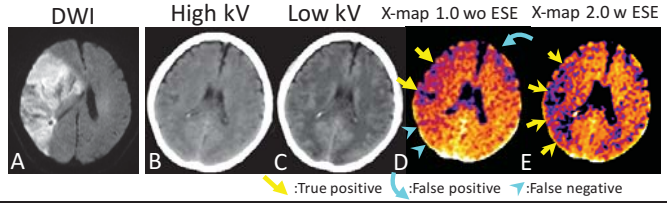


Figure 4. NE-DE-CT was taken 40 min after the symptom onset; DWI 11.5 hours after CT. The color images were generated using "fire" in ImageJ.

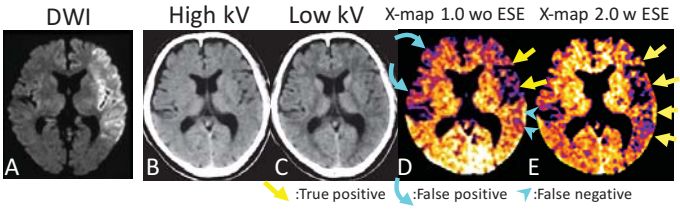


Figure 5. NE-DE-CT was taken 53 min after the symptom onset; DWI 40 min after CT. The color images were generated using "fire" in ImageJ.

showed any change in GM–WM contrast. The X-map 1.0 image without ESE presented using ImageJ's color scheme "fire" (Fig. 4D) had larger pixel values toward the posterior direction, presenting intra-patient variations of pixel values. A part of the infarct was presented clearly (arrows), while missing the rest of the infarct region (arrow-heads) due to larger values toward the posterior side and showing a potential false positive in the frontal lobe (curved arrows) due to smaller values toward the anterior side. In contrast, the X-map 2.0 image with ESE had consistent pixel values throughout the brain region (Fig. 4E) and showed the infarct/ischemic lesion with an improved contrast thanks to ESE (arrows). The extent of the infarct lesion delineated by X-map 2.0 with ESE agreed quite well with the lesion DWI presented (Figs. 4E and 4A).

Figure 5 shows images of another patient with NE-DE-CT performed 53 min after the symptom onset and DWI obtained 40 min after the NE-DE-CT. The DWI image (Fig. 5A) showed an infarct in the left insula, frontal lobe, and temporal lobe; neither the high kV image (Fig. 5B) nor the low kV image (Fig. 5C) showed noticeable changes in GM–WM

contrast. The X-map 1.0 image without ESE had larger pixel values toward the posterior direction, due to a large intra-patient variation of pixel values; the window center was set at X-map value of 105 to make the appearance of images similar to Fig. 4, in order to compensate for a constant bias of X-map value of +5 added to this patient data. A part of lesion was clearly seen with both of the X-map 1.0 images (Fig. 5D arrows), although a potential false positive (curved arrows) and false negative (arrow-heads) were observed. The X-map 2.0 image with ESE had consistent pixel values for the entire brain and showed decreased values at the infarct lesion clearly (Fig. 5E, arrows), while the pixel values of the opposite hemisphere was maintained at ~100. The extent of the infarct lesion was consistent with the finding in DWI (Fig. 5A).

The inter-patient variation of X-map values was smaller with X-map 2.0, either with or without ESE, than with the corresponding X-map 1.0. Levene's *F*-test was performed to evaluate the statistical significance in a difference between algorithms in variances: The standard deviation values were 4.3 for X-map 2.0 versus 19.0 for X-map 1.0, both with ESE, $P<0.01$; 3.0 versus 12.0 without ESE, $P<0.01$. The intra-patient variation was smaller with X-map 2.0 than with X-map 1.0: 6.2 versus 8.5 with ESE, $P<0.01$; 4.1 versus 6.3 without ESE, $P<0.01$.

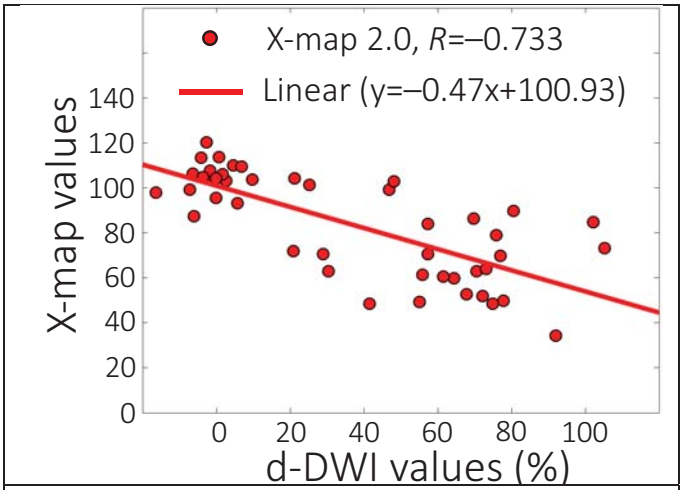


Figure 6. X-map 2.0 values with ESE correlates very well with DWI signals.

Figure 6 shows a scatter plot of X-map 2.0 with ESE against d-DWI values. The correlation coefficient was the strongest between these values and it was –0.733, 95% confidence interval, –0.845 to –0.560; $P<0.001$. The fitted line went through 100.9 when d-DWI was 0 (healthy). The other images had weaker correlation coefficients and they were in the order of –0.642 for high kV CT, –0.631 for X-map 1.0 with ESE, –0.609 for X-map 1.0 without ESE, etc.

It took as little as 7.1 seconds per patient to process Step 1 through Step 5 of the X-map 2.0 algorithm using a laptop PC without any code optimization for efficiency.

IV. Conclusions

We have proposed a novel way to use material or spectral information: to suppress "noise" for a given diagnostic task and enhance "signals" for lesion detection and characterization. To our best knowledge, this is the first approach that can be generalized for and applied to other diagnostic tasks. We have improved both iBHC and X-map algorithm for acute ischemic stroke imaging. The iBHC2 improved the bias and visibility of GM–WM contrast throughout the brain space. The combination of iBHC2 and X-map 2.0 with ESE decreased both intra- and inter-patient variations significantly, improved the quantitative accuracy of X-map values, and provided a strong correlation with DWI signals.

Acknowledgement—We thank Karl Stierstorfer, Ph.D. of Siemens Healthcare for his help with iBHC2 and valuable discussion, and Jeffrey H. Siewerdsen, Ph.D., Nafi Aygun, M.D., and Steven R. Zeiler, M.D., Ph.D. of the Johns Hopkins University for their feedback and discussion.

References

[1] S. L. Murphy, K. D. Kochanek, J. Xu, and E. Arias, "Mortality in the United States, 2014. National Center for Health Statistics, Centers for Disease Control and Prevention (CDC), Department of Health and Human Services," vol. NCHS Data Brief, No. 229 (December 2015), ed. (Hyattsville, MD), 2015.

[2] A. S. Go, D. Mozaffarian, V. L. Roger, E. J. Benjamin, J. D. Berry, M. J. Blaha, *et al.*, "Heart Disease and Stroke Statistics—2014 Update: A Report From the American Heart Association," *Circulation*, vol. 129, pp. e28-e292, 1/20 2014.

[3] E. C. Jauch, J. L. Saver, H. P. Adams, A. Bruno, J. J. B. Connors, B. M. Demaerschalk, *et al.*, "Guidelines for the Early Management of Patients With Acute Ischemic Stroke: A Guideline for Healthcare Professionals From the American Heart Association (AHA) and American Stroke Association (ASA)," *Stroke*, vol. 44, pp. 870-947, February 25, 2013 2013.

[4] J. Saver, M. Goyal, A. van der Lugt, B. Menon, C. Majoie, D. Dippel, *et al.*, "Time to treatment with endovascular thrombectomy and outcomes from ischemic stroke: A meta-analysis," *JAMA: The Journal of the American Medical Association*, vol. 316, pp. 1279–1289, 2016.

[5] R. E. Latchaw, M. J. Alberts, M. H. Lev, J. J. Connors, R. E. Harbaugh, R. T. Higashida, *et al.*, "Recommendations for Imaging of Acute Ischemic Stroke," *A Scientific Statement From the American Heart Association*, vol. 40, pp. 3646-3678, 2009.

[6] C. P. Chao, A. L. Kotsenas, and D. F. Broderick, "Cerebral Amyloid Angiopathy: CT and MR Imaging Findings," *RadioGraphics*, vol. 26, pp. 1517-1531, 2006.

[7] J. Bardutzky and S. Schwab, "Antiedema Therapy in Ischemic Stroke," *Stroke*, vol. 38, pp. 3084-3094, 2007.

[8] G. A. Rosenberg, "Ischemic brain edema," *Progress in Cardiovascular Diseases*, vol. 42, pp. 209-216, 1999.

[9] K. Noguchi, T. Itoh, N. Naruto, S. Takashima, K. Tanaka, and S. Kuroda, "A Novel Imaging Technique (X-Map) to Identify Acute Ischemic Lesions Using Noncontrast Dual-Energy Computed Tomography," *Journal of Stroke and Cerebrovascular Diseases*, vol. 26, pp. 34-41, 2017.

[10] X. Liu, L. Yu, A. N. Primak, and C. H. McCollough, "Quantitative imaging of element composition and mass fraction using dual-energy CT: Three-material decomposition," *Medical Physics*, vol. 36, pp. 1602-1609, 2009.

[11] M. Lundin, M. Liden, A. Magnuson, A. A. Mohammed, H. Geijer, T. Andersson, *et al.*, "Virtual non-contrast dual-energy CT compared to single-energy CT of the urinary tract: a prospective study," *Acta Radiol*, vol. 53, pp. 689-94, Jul 2012.

[12] M. F. Mohammed, O. Marais, A. Min, D. Ferguson, S. Jalakhan, F. Khosa, *et al.*, "Unenhanced Dual-Energy Computed Tomography: Visualization of Brain Edema," *Invest Radiol*, Sep 14 2017 (Epub ahead of print).

[13] J. Hsieh, R. C. Molthen, C. A. Dawson, and R. H. Johnson, "An iterative approach to the beam hardening correction in cone beam CT," *Medical Physics*, vol. 27, pp. 23-29, 2000.

Eigentissue Decomposition for Multi-Energy CT Reconstruction

Mikaël Simard[1], Arthur Lalonde[1] and Hugo Bouchard[1,2]

***Abstract*— A novel material decomposition framework is proposed in the context of multi-energy CT. Photon attenuation coefficients are parametrized as a linear combination of eigentissue cross sections weighted by partial electron densities. The concept of eigentissue was previously proposed with the intent of optimally representing human tissue information. By defining virtual materials which variations of density-weighted elemental fractions are orthogonal, the method allows for direct reconstruction of electron density and elemental fractions from multi-energy CT raw data using less information than the amount of elements found in human tissues. In this work, we compare eigentissue decomposition with other formalisms established for dual-energy CT and we further show the generalizability of the approach for more than two energies. The main impact of this method is to provide accurate characterization of human tissues in the context of radiotherapy dose calculation without the need to parametrize tissues using unphysical parameters, such as the energy-dependent effective atomic number.**

I. Introduction

Tissue characterization using computed tomography (CT) allows the quantification of physical properties, such as electron density (ED), from measurements of X-ray effective attenuation coefficients. One main field of application of tissue characterization is radiotherapy. During radiotherapy treatment planning, single-energy CT (SECT) is used to provide patient-specific ED maps as an input for calculating radiation dose. For treatment planning systems based on Monte Carlo radiation transport simulations, the elemental composition of tissues is additionally required. While these physical parameters can be obtained for SECT using the gold standard stoichiometric calibration of Schneider *et al.* [1], it has recently been shown that dual-energy CT (DECT) has the potential to provide more accurate radiotherapy-related quantities for protontherapy [2] and low energy brachytherapy [3] than SECT.

These studies have been performed using post-reconstruction formalisms adapted to DECT [4], [5] or more general models compatible with multi-energy CT (MECT) [6]. While accurate, such post-reconstruction approaches may suffer from the presence of systematic errors such as beam hardening artefacts in reconstructed attenuation coefficient maps. As such, the accuracy of post-reconstruction methods is limited by systematic errors introduced in the reconstruction process, as well as by the inherent accuracy of semi-empirical beam hardening correction methods, which justifies the need of pre-reconstruction approaches for tissue characterization [2], [7]. In this work, we extend the post-reconstruction tissue characterization method of Lalonde & Bouchard [8] to an image reconstruction framework. The method, called eigentissue decomposition (ETD), is a material decomposition approach that uses virtual tissues, named eigentissues, constructed in such a way that they optimally describe human tissues in terms of partial elemental electron density. ETD allows for the direct reconstruction of the elemental composition and ED and can therefore be used to directly provide inputs for treatment planning systems using Monte Carlo simulations.

The adaptation of ETD to a pre-reconstruction scheme is the main goal of the present work and is described at first. The accuracy of the parametrization in evaluating various radiotherapy-related quantities is then evaluated and compared with other well-established parametrization schemes. Finally, the generalizability of the method to MECT is shown.

II. Methods

A. Eigentissue decomposition

The mass attenuation coefficient $\frac{\mu_{\text{med}}(\mathbf{r},E)}{\rho}$ of a medium can be expressed as the weighted sum of the mass attenuation coefficients of the M elements that constitute the medium:

$$\frac{\mu_{\text{med}}(\mathbf{r},E)}{\rho} = \sum_{m=1}^{M} w_m \frac{\mu_m(\mathbf{r},E)}{\rho} \tag{1}$$

where w_m is the mass fraction of element m. An alternative expression for the linear attenuation coefficient of the medium consists in using partial electronic densities $x_m = \rho_e \lambda_m$, where the fraction of electrons of the m^{th} element in the medium is $\lambda_m = \frac{w_m (Z/A)_m}{(Z/A)_{\text{med}}}$ and ρ_e is the ED of the medium:

$$\mu(\boldsymbol{r},E) = \sum_{m=1}^{M} x_m(\mathbf{r}) \sigma_{e,m}(E) \tag{2}$$

where $\sigma_{e,m}(E)$ is an elemental electronic cross section and the x_m sum up to the electronic density, i.e., $\rho_e = \sum_m x_m$.

For Monte Carlo dose calculation, one is interested in extracting the elemental compositions w_m and ED. Including trace elements, human tissues can be characterized fully with $M = 13$ elements [9], leaving the number of unknowns to 13, including ED. In the case of MECT, where the number of energies acquired is $K \geq 2$, tissue characterization is generally

[1] Université de Montréal, Département de Physique. 2900 Boul. Édouard Montpetit, Montréal, Québec H3T 1J4.
[2] Centre de recherche du Centre hospitalier de l'Université de Montréal, Montréal, Québec H2X 3H8

an underdetermined problem as the number of unknowns is greater than the number of measurements per voxel.

The elemental composition of human tissues is often redundant between tissues, or in some cases, such as carbon and oxygen in soft tissues or phosphorus and calcium in bones, some elements are strongly correlated [1], [8]. Thus, the elemental basis used to represent $\mu(E)$ (equation 2) does not form an orthogonal basis and is sub-optimal - a similar amount of information on human tissues may be extracted by expressing tissues with an orthogonal basis and a reduced set of variables. In order to be mapped back to elemental composition, such a new basis should be obtained through a base transformation of the elemental composition. To perform this transformation, Lalonde & Bouchard [8] applied principal component analysis (PCA) on a reference dataset of human tissues of known elemental composition [9]. This allows reducing the dimensionality of the problem by extracting new materials, called eigentissues, which are composed of elements in such a way that the most variability throughout human tissues information is contained in the first few eigentissues.

While M eigentissues are extracted when performing PCA, the highest variance in chemical composition is enclosed within the first few eigentissues - using K=2 eigentissues explains respectively 96 and >99% of the variation of the elemental composition of soft tissues and bones [8]. As such, one can represent human tissues as a linear combination of the $K<M$ first eigentissues, and assume that the remaining $(M\text{-}K)$ eigentissues are constant for all media. This results in an additional term, the residual eigentissue, corresponding to the average of the remaining eigentissues over all tissues from the original database. This makes the approach compatible with MECT, as one extracts the partial electron density y_k of the K first eigentissues from K measurements obtained from MECT, while assuming that the remaining eigentissues, which explains only a limited amount of the variation in chemical composition, are constant.

The partial electron density of an eigentissue, y_k, is related to that of the elements, x_m, by a transformation matrix $\mathbf{P}$ obtained from PCA [8]. As such, extracting the y_k using an approach similar to material decomposition leads to the knowledge of the x_m, which maps back to the elemental fractions w_m through λ_m. The error introduced in the estimation of w_m is related to the truncation of the remaining $M\text{-}K$ eigentissues, and might be reduced as the number of energies K increases. In the eigentissue basis, the attenuation coefficient is then expressed as:

$$\mu(\mathbf{r},E)=\sum_{k=1}^{M} y_k(\mathbf{r})\sigma_{e,k}^{\mathrm{ET}}(E)\approx\sum_{k=1}^{K} y_k(\mathbf{r})\sigma_{e,k}^{\mathrm{ET}}(E)+y_0\sigma_{e,0}^{\mathrm{ET}}(E) \tag{3}$$

Equation 3 corresponds to the eigentissue decomposition (ETD) approach. Here, $\sigma_{e,k}^{\mathrm{ET}}(E)$ represents the electronic cross section of the k^{th} eigentissue, calculated from its elemental composition extracted from PCA. y_0, $\sigma_{e,0}^{\mathrm{ET}}(E)$ respectively represent the residual partial electron density and cross section of the $(M-K)$ remaining eigentissues. The estimator of the ED is then given by $\hat{\rho}_e \approx y_0+\sum_{k=1}^{K} y_k$.

B. Adaptation to sinogram space

In CT, one is interested in the transmission $t_i(\xi,\theta)$ produced by a ray of path $\ell(\xi,\theta)$ defined with its projection angle θ and detector location ξ. $t_i(\xi,\theta)$ represents the detected X-ray intensity normalized by a blank scan, using a normalized X-ray spectrum $\psi_i(E)$ (which includes contributions of the source and detector response). A simple model to evaluate the transmission is given by

$$t_i(\xi,\theta)=\int \psi_i(E)\exp\left(-\int_{\ell(\xi,\theta)}\mu(\mathbf{r},E)\mathrm{d}\ell\right)\mathrm{d}E \tag{4}$$

with the sinogram being equal to $-\ln t_i(\xi,\theta)$. Combining ETD (equation 3) with equation 4 and separating the energy and position dependent terms, an estimator of the transmission in the ETD formalism is given by

$$\hat{t}_i(\xi,\theta)=\int \psi_i(E)\exp\left(-\sum_{k=1}^{K} Y_k(\xi,\theta)\sigma_{e,k}^{\mathrm{ET}}(E)-G_0(E)\right)\mathrm{d}E \tag{5}$$

Where $Y_k(\xi,\theta)=\int_{\ell(\xi,\theta)} y_k(\mathbf{r})\mathrm{d}\ell$ is the 2D Radon transform of the partial electron density of the k^{th} eigentissue and $G_0(E)=\int_{\ell(\xi,\theta)} y_0\sigma_{e,0}^{ET}(E)d\ell$ is the 2D Radon transform of the residual eigentissue.

Following the methodology introduced by Alvarez & Macovski [10], one can perform a non-linear least squares fit at each sinogram voxel (ξ,θ) to solve for the K values of $Y_k(\xi,\theta)$ given K transmission values obtained using different spectra. Partial electron density maps $y_k(\boldsymbol{r})$ can then be reconstructed using a filtered back-projection. Maps of ED and elemental fractions w_m can be obtained from the $y_k(\boldsymbol{r})$ according to the methods described in section II-A.

Practically, as there are notable differences between the elemental composition of soft tissues and bones, PCA is performed independently on both tissue groups (labelled $\boldsymbol{s}$ and $\boldsymbol{b}$) in the reference database. This results in two distinct sets of eigentissues, which are represented by different values of $\sigma_{e,k}^{ET}(E)$ and y_0 in equation 3. In order to assign the most likely decomposition basis $\boldsymbol{i}$ (where $\boldsymbol{i}\in\{\boldsymbol{s},\boldsymbol{b}\}$) for each image voxel, partial electron density maps are obtained by solving equation 5 separately using the two sets of eigentissues. Overall, this yields, for each image voxel, two $K\times 1$ vectors of partial electron densities: $\tilde{\boldsymbol{y}}_{\boldsymbol{s}}$ and $\tilde{\boldsymbol{y}}_{\boldsymbol{b}}$, respectively obtained using soft tissue and bone eigentissues.

The set of partial electron densities of the soft tissues and bones in the reference database, labelled $\{\boldsymbol{y}_{\boldsymbol{s}}^{(p)}\}$ and $\{\boldsymbol{y}_{\boldsymbol{b}}^{(p)}\}$ (where p denotes a specific tissue) are considerably different. As such, in order to assign the correct set of partial

electron densities on a voxel-wise basis, one can evaluate the similarity between each $\tilde{\boldsymbol{y}}_{\boldsymbol{i}}$ and the corresponding distribution $\{\boldsymbol{y}_{\boldsymbol{i}}^{(p)}\}$, and assign the most likely class $\boldsymbol{i}$. Formally, we use a kernel density estimator (KDE) [6] to provide an estimation of the probability density function for the partial electronic density of soft tissues and bones. The probability $p_i(\tilde{\boldsymbol{y}}_{\boldsymbol{i}})$ that a set of partial electron densities $\tilde{\boldsymbol{y}}_{\boldsymbol{i}}$ at a given voxel belongs to class $\boldsymbol{i}$ is defined as (with $\mathcal{K}$ the kernel)

$$p_i(\tilde{\boldsymbol{y}}_{\boldsymbol{i}}) = \frac{1}{P}\sum_{p=1}^{P}\mathcal{K}(\boldsymbol{y}_{\boldsymbol{i}}^{(p)};\tilde{\boldsymbol{y}}_{\boldsymbol{i}}) \tag{6}$$

the class is assigned such that the probability is maximized. In this work, we use a multivariate gaussian kernel:

$$\mathcal{K}(\boldsymbol{y}_{\boldsymbol{i}}^{(p)};\tilde{\boldsymbol{y}}_{\boldsymbol{i}}) = \frac{1}{\sqrt{(2\pi)^K|\boldsymbol{\Sigma}_{\boldsymbol{i}}|}}\mathrm{e}^{-\frac{1}{2}\left(\tilde{\boldsymbol{y}}_{\boldsymbol{i}}-\boldsymbol{y}_{\boldsymbol{i}}^{(p)}\right)^{\mathrm{T}}\boldsymbol{\Sigma}_{\boldsymbol{i}}^{-1}\left(\tilde{\boldsymbol{y}}_{\boldsymbol{i}}-\boldsymbol{y}_{\boldsymbol{i}}^{(p)}\right)} \tag{7}$$

The probability density function is then modelled as a sum of gaussians each centered around a tissue of the database, with $\boldsymbol{\Sigma}_{\boldsymbol{i}}$ being the diagonal covariance matrix of the eigentissues. This classification approach is particularly well suited for ETD, as prior knowledge on the expected distribution of the partial electron densities for human tissues arises from PCA.

C. Simulation framework

A simulation framework is developed to evaluate the performance of ETD for MECT in extracting physical parameters of tissues and quantities relevant to radiotherapy in two specific situations. First, ETD is compared against other reference methods for DECT ($K = 2$). Second, we demonstrate the potential gains and generalizability of ETD when increasing the number of energies to $K = 3$, $i.e.$ using 3 eigentissues instead of 2 in equation 3.

Polyenergetic sinograms (equation 4) are modelled as a weighted sum of monoenergetic projections at 1 keV intervals [11], and projections are performed using a simple ray-tracing algorithm [12]. For DECT, the 80 and 140/Sn kVp spectra of the Siemens SOMATOM Definition Flash provided by the manufacturer (Siemens Healthcare, Forcheim, Germany) are used. For MECT ($K = 3$), the 80 kVp spectrum is divided into bins of equals fluence to reproduce idealized measurements from a photon counting system. Sinograms are generated for various noise levels (0, 5, 8 and 15 HU).

For DECT, ETD is compared with two other formalisms: the physical interactions decomposition of Alvarez and Macovski (AM) [10] optimized for the reference database tissues [11] and a two-material decomposition (2MD) [13] using cortical bone and water. Elemental photon cross sections are obtained from the XCOM database [14]. For AM and 2MD, elemental fractions are evaluated indirectly using the parametrization of Hünemohr *et al.* [5].

Formalisms are compared on their ability to accurately reconstruct physical parameters useful for various radiotherapy modalities, which directly translates the accuracy of the reconstructed ED and elemental fractions. Specifically, for hadron therapy, we estimate the stopping power relative to water (SPR) of 200 MeV protons. For brachytherapy, we evaluate the photon energy attenuation coefficient (EAC) of 21 KeV photons from ^{103}Pd seeds. The SPR is estimated from the Bethe-Bloch equation using the Bragg additivity rule for the estimation of the I-value, while the EAC is obtained using equation 1.

The simulated phantom contains 16 plugs with the same geometry as in Tremblay *et al.* [11], each containing a human tissue randomly selected from the 71 tissues in White & Woodard's database of human tissues [9]. The RMS error and overall bias (mean error) on the quantities of interest, averaged over all plugs, is compared across formalisms. Plugs containing soft tissues and bone are analyzed separately. To obtain reliable statistics, the randomized phantom is generated and analyzed 100 times.

III. Results

Figures 1 and 2 show the mean error (bias) and RMS error for SPR and EAC for the 3 formalisms applied to DECT, respectively for soft tissues and bones. Without noise, ETD outperforms all other formalisms in providing lower RMS errors and a negligible bias on SPR and EAC, for soft tissues and bone. As the noise increases, the RMS error for ETD stays smaller or equal to 2MD, while the bias always remains negligible with respect to the other formalisms.

Figure 3 shows the theoretical performance gains of ETD in estimating elemental fractions when going from 2 to 3 energies. The error is largely reduced for carbon and oxygen fractions, while it increases slightly for remaining elements. Overall, this leads to a reduction in the RMS error on the SPR from 0.25 to 0.22% when increasing the number of energies.

IV. Conclusion

This work shows the potential of using ETD as a parametrization for the attenuation coefficient to accurately and precisely describe human tissues. Concerning DECT, for all noise levels considered, ETD provides more accurate quantities related to dose calculation (SPR, EAC), which suggests that ETD may be less biased in estimating radiotherapy-related quantities for human tissues, as other parametrizations evaluated can produce biased estimates of ED and effective atomic number [11]. The method is also at least as precise as 2MD with increasing noise. For radiotherapy purposes, the advantage of ETD resides in the fact that it allows a direct extraction of ED and elemental composition instead of the effective atomic number, which is an energy dependent quantity.

ETD has the advantage of being generalizable to more than two energies, as shown in figure 3. This indicates that ETD is potentially well suited for tissue characterization using spectral CT, and is not limited by a two-parameter parametrization.

Future work include developing a regularization scheme on partial electronic densities [6] to increase the robustness of the method with respect to noise and weak conditioning in MECT.

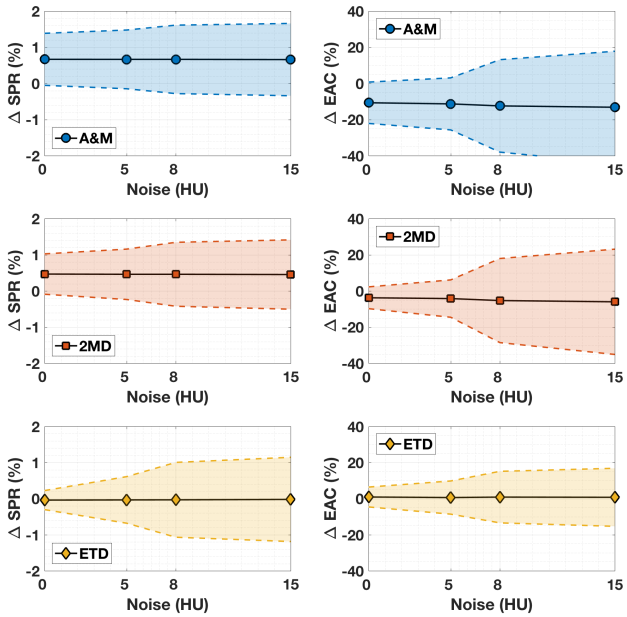


Fig. 1. Mean error (full lines) and RMS errors (coloured bands) on SPR (left) and EAC (right) averaged over all soft tissue plugs using various parametrizations of the attenuation coefficient (top to bottom).

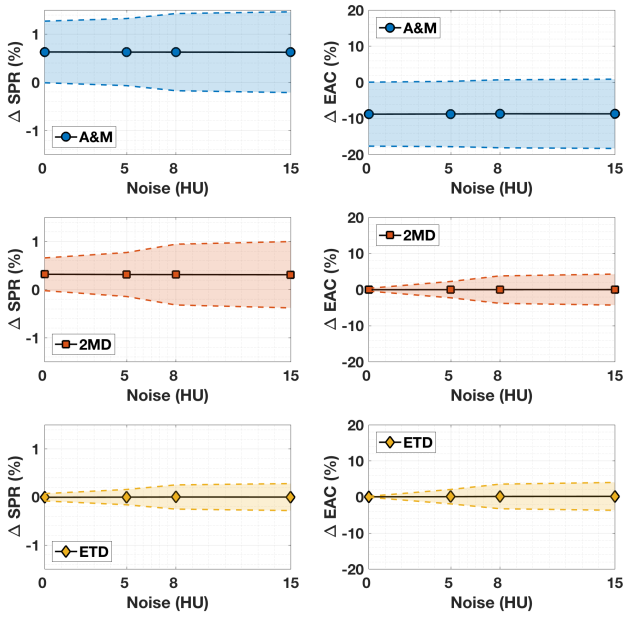


Fig. 2. Mean error (full lines) and RMS errors (coloured bands) on SPR (left) and EAC (right) averaged over all bone plugs using various parametrizations of the attenuation coefficient (top to bottom).

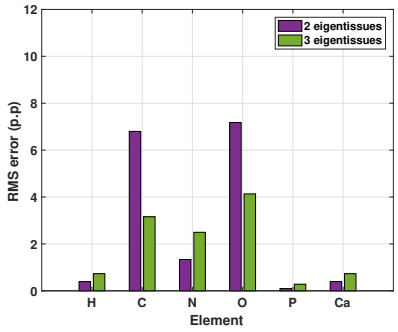


Fig. 3. RMS errors (in percentage points) on the main elemental fractions of human tissues as a function of the number of energies used, for soft tissues.

References

[1] W. Schneider, T. Bortfeld, and W. Schlegel, "Correlation between ct numbers and tissue parameters needed for monte carlo simulations of clinical dose distributions," *Physics in medicine and biology*, vol. 45, no. 2, p. 459, 2000.

[2] E. Bär, A. Lalonde, R. Zhang, K.-W. Jee, K. Yang, G. Sharp, B. Liu, G. Royle, H. Bouchard, and H.-M. Lu, "Experimental validation of two dual-energy ct methods for proton therapy using heterogeneous tissue samples," *Medical physics*, 2017.

[3] C. Remy, A. Lalonde, D. Béliveau-Nadeau, J.-F. Carrier, and H. Bouchard, "Dosimetric impact of dual-energy ct tissue segmentation for low-energy prostate brachytherapy: a monte carlo study," *Physics in medicine and biology*, 2017.

[4] G. Landry, K. Parodi, J. E. Wildberger, and F. Verhaegen, "Deriving concentrations of oxygen and carbon in human tissues using single-and dual-energy ct for ion therapy applications," *Physics in medicine and biology*, vol. 58, no. 15, p. 5029, 2013.

[5] N. Hünemohr, H. Paganetti, S. Greilich, O. Jäkel, and J. Seco, "Tissue decomposition from dual energy ct data for mc based dose calculation in particle therapy," *Medical physics*, vol. 41, no. 6, 2014.

[6] A. Lalonde, E. Bär, and H. Bouchard, "A bayesian approach to solve proton stopping powers from noisy multi-energy ct data," *Medical physics*, 2017.

[7] E. Bär, A. Lalonde, G. Royle, H.-M. Lu, and H. Bouchard, "The potential of dual-energy ct to reduce proton beam range uncertainties," *Medical Physics*, 2017.

[8] A. Lalonde and H. Bouchard, "A general method to derive tissue parameters for monte carlo dose calculation with multi-energy ct," *Physics in medicine and biology*, vol. 61, no. 22, p. 8044, 2016.

[9] H. Woodard and D. White, "The composition of body tissues," *The British journal of radiology*, vol. 59, no. 708, pp. 1209–1218, 1986.

[10] R. E. Alvarez and A. Macovski, "Energy-selective reconstructions in x-ray computerised tomography," *Physics in medicine and biology*, vol. 21, no. 5, p. 733, 1976.

[11] J.-É. Tremblay, S. Bedwani, and H. Bouchard, "A theoretical comparison of tissue parameter extraction methods for dual energy computed tomography," *Medical physics*, vol. 41, no. 8, 2014.

[12] R. L. Siddon, "Fast calculation of the exact radiological path for a three-dimensional ct array," *Medical physics*, vol. 12, no. 2, pp. 252–255, 1985.

[13] L. Lehmann, R. Alvarez, A. Macovski, W. Brody, N. Pelc, S. J. Riederer, and A. Hall, "Generalized image combinations in dual kvp digital radiography," *Medical physics*, vol. 8, no. 5, pp. 659–667, 1981.

[14] M. J. Berger and J. Hubbell, "Xcom: Photon cross sections on a personal computer," tech. rep., National Bureau of Standards, Washington, DC (USA). Center for Radiation Research, 1987.

Optimization of Dose Distribution and Filter Thickness in Energy-integrating-detector-based Multi-energy CT

Liqiang Ren, Cynthia H. McCollough, Lifeng Yu

***Abstract*— Dose distribution and filter thickness were optimized for a novel energy-integrating-detector (EID) based multi-energy CT (MECT) platform through theoretical analysis and simulation. The optimization was based on a material-specific imaging task that involves multiple contrast agents: iodine, gadolinium, and water. A generic image-based method without/with prior information (volume conservation) was used for material decomposition. The dose distribution among triple or quadruple beams was optimized to minimize the noise in each basis material map. Three sets of split filters (Sn/Au) with thicknesses of 0.3/0.09, 0.4/0.12, 0.5/0.14 mm for high energy beam were employed to investigate the impact of split-filter thickness on noise in each basis material map. Theoretical analysis demonstrated that, when the prior information was incorporated into the material decomposition process, the noise levels in iodine, gadolinium, and water maps were almost minimized simultaneously by distributing the dose equally among all the beams. Although better spectral separation was achieved by utilizing thicker filters, simulation results showed that filter thickness has little impact on image noise of basis materials when prior information is incorporated in the material decomposition.**

***Index Terms*—Multi-energy CT (MECT), basis material decomposition**

I. Introduction

MULITI-ENERGY CT (MECT) enabled by energy-resolved photon-counting computed tomography (PCCT) is promising for material-specific imaging with multiple contrast agents. However, non-idealities of the photon-counting detector (PCD) such as pulse pileup, K-edge escape, and charge sharing may degrade the spectral performance [1-8]. Because of these non-idealities, the dose efficiency of PCCT did not show obvious advantages over existing energy-integrating-detector (EID) based CT scanners for typical dual-energy applications [9]. To perform material-specific imaging with multiple contrast agents, EID-based MECT approaches can still be employed by acquiring measurements at three or more different x-ray spectra. One such approach was to acquire triple or quadruple x-ray beam measurements on a dual-source CT scanner using the split filter technique [10]. Computer simulations showed that this EID-based MECT method appears to have at least similar or even better spectral separation and noise performance as in a PCCT scanner [10]. The main advantages of this EID-based MECT approach include (1) easy to implement; (2) similar or better spectrum separation than current PCCT scanners; (3) flexible to adjust dose distribution among different spectra.

In this work, we will perform a theoretical analysis and computer simulation study to optimize two of the important technical parameters in the new EID-based MECT platform: dose distribution and filter thickness. In the analysis, we will focus on a specific MECT imaging task, multi-phase liver imaging in one single MECT scan [6, 7, 11], which involves three materials, iodine, gadolinium, and water.

This work was supported by a NIH grant R21 EB024071.

Liqiang Ren, Cynthia H. McCollough, and Lifeng Yu are with the Department of Radiology, Mayo Clinic, Rochester, MN, 55905 USA. (e-mail: Ren.Liqiang@mayo.edu, McCollough.Cynthia@mayo.edu, Yu.Lifeng@mayo.edu).

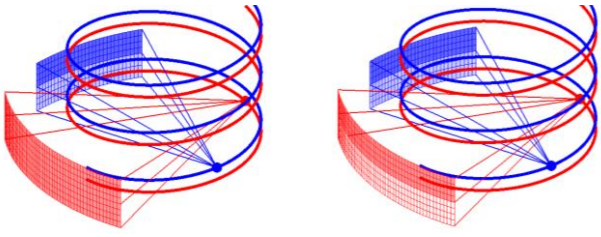
II. Methods

A. An EID-based MECT method

There are many possible ways to acquire more than two x-ray beam measurements using EID-based CT platform. We recently proposed a scanner configuration that can acquire triple- or quadruple-beam measurements nearly simultaneously [10]. This scanner configuration originates from the "Twin Beam" design proposed by Siemens on a single-source CT scanner to enable DECT capability. In this design, the same x-ray beam is pre-filtered by two different materials, e.g., gold (Au) and tin (Sn), each of which covers half of the detector rows along the longitudinal direction. The x-ray beam filtered by Sn is hardened to form the "high-kV" and the x-ray beam filtered by Au is softened because of the K-edge to form the "low-kV". This split filter design can be readily extended to a dual-source CT scanner by operating one (triple-beam configuration, Fig. 1a) or both (quadruple-beam configuration, Fig. 1b) sources in the "Twin Beam" mode to acquire three or four distinct x-ray beam measurements.

B. Material Decomposition Method for EID-based MECT

A generic image-based material decomposition method without any noise reduction was applied on the images reconstructed with a filtered-back-projection (FBP) method, to generate three mass density maps for iodine, gadolinium, and water. This material decomposition method for the EID-based MECT with triple-beam configuration was formulated in (1), where $\mu(E)$ is the effective linear attenuation



(a) (b)

Fig. 1. (a) Triple-beam configuration on a dual-source CT with a split filter added to one of the sources; (b) Quadruple-beam configuration with split filters added to both sources.

coefficient from measurement at energy E, $\mu_{m,I}/\mu_{m,Gd}/\mu_{m,W}$ and $\rho_I/\rho_{Gd}/\rho_W$ are the mass attenuation coefficients and mass density for basis materials iodine/gadolinium/water, and the last row represents the volume conservation (prior information) that is optionally incorporated into the material decomposition [12-14].

$$\begin{bmatrix} \mu(E_L) \\ \mu(E_{H1}) \\ \mu(E_{H2}) \\ 1 \end{bmatrix} = \begin{bmatrix} \mu_{m,I}(E_L) & \mu_{m,Gd}(E_L) & \mu_{m,W}(E_L) \\ \mu_{m,I}(E_{H1}) & \mu_{m,Gd}(E_L) & \mu_{m,W}(E_L) \\ \mu_{m,I}(E_{H2}) & \mu_{m,Gd}(E_{H2}) & \mu_{m,W}(E_{H2}) \\ 1/\rho_{I0} & 1/\rho_{Gd0} & 1/\rho_{W0} \end{bmatrix} \begin{bmatrix} \rho_I \\ \rho_{Gd} \\ \rho_W \end{bmatrix} \quad (1)$$

For quadruple-beam configuration, the two measurements from low energy x-ray beam are described as $\mu(E_{L1})$ and $\mu(E_{L2})$, respectively, with totally four x-ray beam measurements.

C. Dose Distribution for Triple-Beam Configuration

The total dose was distributed to minimize the noise in each basis image. In general, the radiation dose, $D(E)$ is inversely proportional to the variance of the measurement $Var[\mu(E)]$, as described in (2), where k is the dose factor that could be determined empirically [9].

$$D(E) = \frac{k(E)}{Var[\mu(E)]} \quad (2)$$

In the EID-based MECT with triple beams, the total dose equals to the dose from all of the three beams, as in (3).

$$\begin{aligned} D_{tot} &= D(E_L) + D(E_{H1}) + D(E_{H2}) \\ &= \frac{k(E_L)}{Var[\mu(E_L)]} + \frac{k(E_{H1})}{Var[\mu(E_{H1})]} + \frac{k(E_{H2})}{Var[\mu(E_{H2})]} \end{aligned} \quad (3)$$

A dose distribution coefficient α, the ratio between the dose delivered to the low energy beam and the total dose is defined. The doses delivered to those two high energy beams are assumed to be approximately the same.

$$D(E_L) = \alpha D_{tot} \quad (4)$$

$$D(E_{H1}) = D(E_{H2}) = \frac{(1-\alpha)D_{tot}}{2} \quad (5)$$

Based on the noise propagation analysis from (1), the variance in the mass density map $Var(\rho_i)$ for material i can be expressed as the weighted sum of the variances from all of the three measurements, as in (6), where w_{iL}^2, w_{iH1}^2, and w_{iH2}^2 are the weighted coefficients.

$$\begin{aligned} Var(\rho_i) &= w_{iL}^2 Var[\mu(E_L)] + w_{iH1}^2 Var[\mu(E_{H1})] + w_{iH2}^2 Var[\mu(E_{H2})] \\ &= w_{iL}^2 \frac{k(E_L)}{D(E_L)} + w_{iH1}^2 \frac{k(E_{H1})}{D(E_{H1})} + w_{iH2}^2 \frac{k(E_{H2})}{D(E_{H2})} \\ &= w_{iL}^2 \frac{k(E_L)}{\alpha D_{tot}} + w_{iH1}^2 \frac{k(E_{H1})}{\frac{(1-\alpha)D_{tot}}{2}} + w_{iH2}^2 \frac{k(E_{H2})}{\frac{(1-\alpha)D_{tot}}{2}} \\ &= \frac{k(E_L)w_{iL}^2}{\alpha D_{tot}} + \frac{2[k(E_{H1})w_{iH1}^2 + k(E_{H2})w_{iH2}^2]}{(1-\alpha)D_{tot}} \end{aligned} \quad (6)$$

Differentiate $Var(\rho_i)$ with respect to α, and set the result to 0.

$$\frac{\partial Var(\rho_i)}{\partial \alpha} = \frac{-k(E_L)w_{iL}^2}{\alpha^2 D_{tot}} + \frac{2[k(E_{H1})w_{iH1}^2 + k(E_{H2})w_{iH2}^2]}{(1-\alpha)^2 D_{tot}} = 0$$

$$\alpha = \left\{ 1 \pm \sqrt{\frac{2[k(E_{H1})w_{iH1}^2 + k(E_{H2})w_{iH2}^2]}{k(E_L)w_{iL}^2}} \right\}^{-1} \quad (7)$$

D. Dose Management for Quadruple-Beam Configuration

In the EID-based MECT with quadruple beams, the total dose equals to the dose from all of the four beams, as in (8).

$$\begin{aligned} D_{tot} &= D(E_{L1}) + D(E_{L2}) + D(E_{H1}) + D(E_{H2}) \\ &= \frac{k(E_{L1})}{Var[\mu(E_{L1})]} + \frac{k(E_{L2})}{Var[\mu(E_{L2})]} + \frac{k(E_{H1})}{Var[\mu(E_{H1})]} + \frac{k(E_{H2})}{Var[\mu(E_{H2})]} \end{aligned} \quad (8)$$

Similarly, a dose distribution coefficient α, the ratio between the dose delivered to one of the low energy beams and the total dose is defined. In addition to the assumption of equivalent dose for two high energy beams as in triple beams, the doses delivered to those two low energy beams are also approximately the same. Thus,

$$D(E_{L1}) = D(E_{L2}) = \alpha D_{tot} \quad (9)$$

$$D(E_{H1}) = D(E_{H2}) = \frac{(1-2\alpha)D_{tot}}{2} \quad (10)$$

The dose distribution coefficient for quadruple-beam configuration is determined as in (11).

$$\alpha = \left\{ 2 \left[1 \pm \sqrt{\frac{k(E_{H1})w_{iH1}^2 + k(E_{H2})w_{iH2}^2}{k(E_{L1})w_{iL1}^2 + k(E_{L2})w_{iL2}^2}} \right] \right\}^{-1} \quad (11)$$

E. Computer Simulation for Filter Thickness Optimization

The optimization of filter thickness was performed using a simulation tool (DRASIM, Siemens Healthcare) [15]. A cylindrical water phantom in a diameter of 20cm containing 9 solutions was used in the simulation. One sample CT image indicating the solution types and concentrations was shown in Fig. 2.

In this preliminary study, only the optimization of thickness for split filter associated with the high energy beams with 140kV was investigated. The low energy beam with 80kV in EID-based MECT with triple beams has no additional

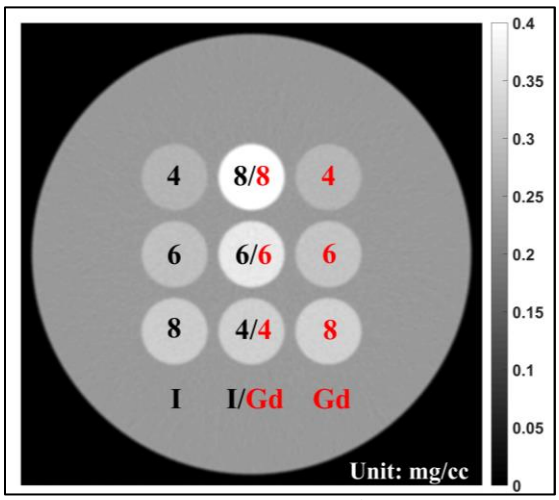


Fig. 2. The linear attenuation map of the cylindrical water phantom acquired with 80kV with all the material type and concentrations of the inserted solutions labelled.

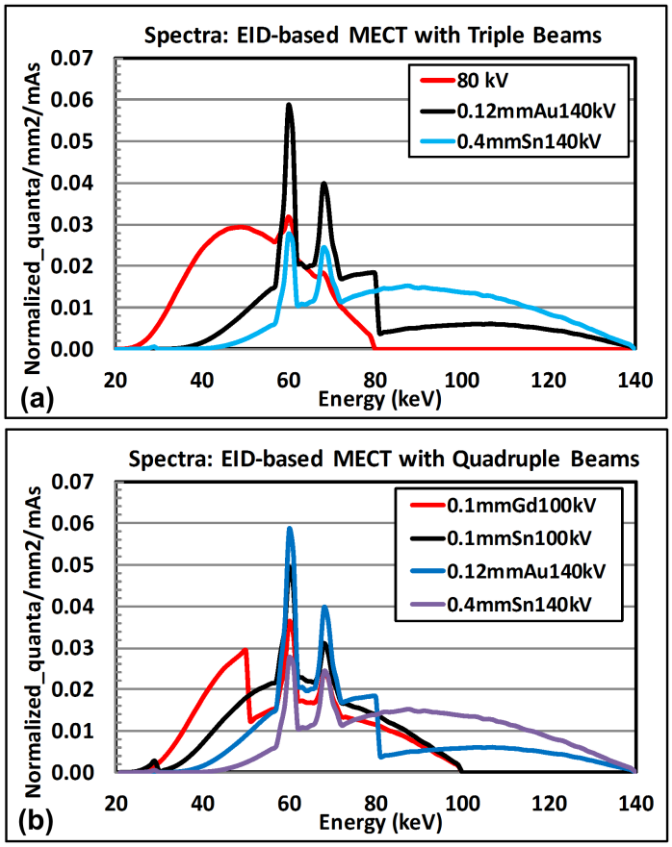


Fig. 3. Examples of spectra used in EID-based MECT with (a) triple beams, and (b) quadruple beams.

filtration, whereas the low energy beam with 100kV in EID-based MECT with quadruple beams uses a split filter of Sn/GOS (gadolinium oxysulfide) with fixed thickness of 0.1/0.1mm. The high energy beam with 140kV was pre-filtered by a split filter made by Au/Sn with three different thickness levels: 0.09/0.3, 0.12/0.4, 0.14/0.5mm. Examples of x-ray spectra for triple- (80/Au140/Sn140kV) and quadruple-(GOS100/Sn100/Au140/Sn140kV) beam configurations were plotted in Fig. 3. The tube current-time products (mAs) were determined for each beam based on the optimization results from Sections II. C and D. The total radiation dose output (CTDIvol) for all the scan configurations was determined as 16.4mGy.

After performing the material decomposition, we obtained three material maps for iodine, gadolinium, and water for each filtration configuration, and then the noise levels were determined and compared.

III. Results

A. Dose Distribution with Triple-Beam Configuration

Without loss of generality, Eq. (7) was used to determine the dose distribution among triple beams with beam spectral configurations of 80kV for the low energy beam, whereas 140kV for the high energy beam with split filter of 0.12mm Au and 0.4mm Sn. The determined dose distribution coefficients were 3.2%, 13.8%, and 9.4% for water, iodine, and gadolinium, respectively, when the volume conservation (prior information) was excluded from the material decomposition process. In contrast, when the volume conservation was included, the dose distribution coefficients are 38.3%, 34.5%, and 28.0%, with an average of 33.6%, indicating an approximately equivalent dose distribution among three beams.

B. Dose Distribution with Quadruple-Beam Configuration

Eq. (11) was used to determine the dose distribution among four beams with beam spectral configurations of 100kV for the low energy beam with split filter of: 0.1mm GOS and 0.1 mm Sn, whereas 140kV for the high energy beam with split filter of 0.12mm Au and 0.4mm Sn. The determined dose distribution coefficients were 26.3%, 37.9%, and 36.9% for water, iodine, and gadolinium, respectively, when the volume conservation (prior information) was not used. In contrast, when the volume conservation was included into the material decomposition process, the dose distribution coefficients were 23.1%, 27.9%, and 26.4%, with an average of 25.8%, also indicating an approximately equivalent dose distribution among four beams.

C. Impact of Filter Thickness

The impact of filter thickness on noise levels in each basis material map was summarized in Fig. 4. As the filter thickness increases, the noise levels in all basis material maps were reduced, which was attributable to better spectral separation among triple/quadruple beams. Though noise reductions were observed in both material decompositions without/with the volume conservation (prior information), the variation of the filter thickness has little impact on basis image noise after material decomposition with volume conservation.

IV. Discussions and Conclusions

For an EID-based MECT platform, the optimal dose distribution among triple/quadruple beams was determined to minimize the noise of basis materials for a specific clinical application that involves multiple contrast agents. The impact of filter thickness for the high energy beam was also

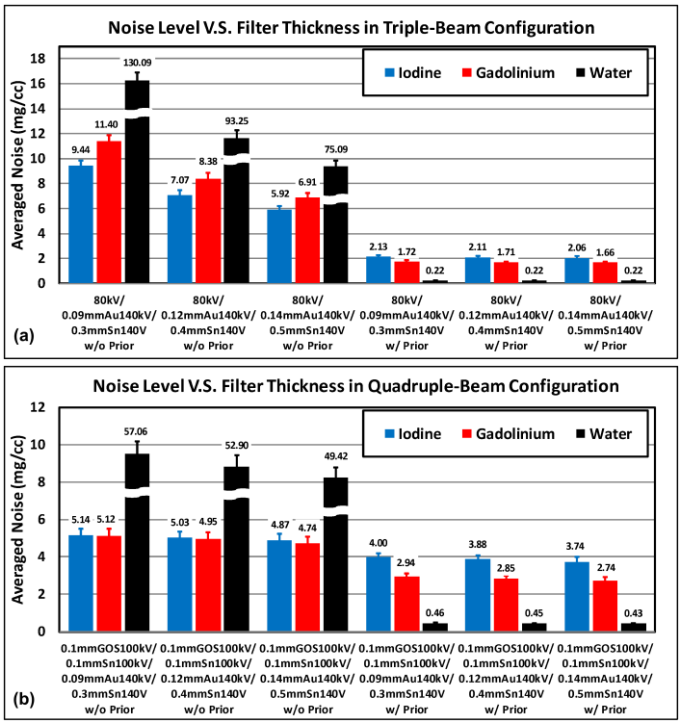


Fig. 4. Impact of split filter thickness for high energy x-ray beams on noise levels in basis material maps in EID-based MECT with (a) triple beams; (b) quadruple beams.

evaluated. By incorporating the volume conservation as an additional physical constraint, all the noise levels in three material maps may be minimized simultaneously by distributing the total dose approximately equally among triple/quadruple beams. Overall noise reduction was observed with increased thickness of the filter for high energy beams. However, the selection of filter thickness will also need to take into account other practical factors, such as maximum x-ray tube output capacity.

ACKNOWLEDGMENT

Research reported in this publication was supported by the National Institute of Biomedical Imaging and Bioengineering of the National Institutes of Health under Award Number EB024071. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES

[1]. Wang, A.S., et al., *Pulse pileup statistics for energy discriminating photon counting x-ray detectors.* Medical Physics, 2011. **38**(7): p. 4265-4275.

[2]. Wang, X., et al., *Material separation in x-ray CT with energy resolved photon-counting detectors.* Medical Physics, 2011. **38**(3): p. 1534-1546.

[3]. Taguchi, K. and J.S. Iwanczyk, *Vision 20/20: Single photon counting x-ray detectors in medical imaging.* Medical Physics, 2013. **40**(10): p. 100901-n/a.

[4]. Kappler, S., et al. *Photon counting CT at elevated X-ray tube currents: contrast stability, image noise and multi-energy performance*. in *SPIE Medical Imaging*. 2014. SPIE.

[5]. Yu, Z., et al. *Initial results from a prototype whole-body photon-counting computed tomography system*. in *SPIE Medical Imaging*. 2015. SPIE.

[6]. Muenzel, D., et al., *Spectral Photon-counting CT: Initial Experience with Dual–Contrast Agent K-Edge Colonography.* Radiology, 2017. **283**(3): p. 723-728.

[7]. Symons, R., et al., *Photon-counting CT for simultaneous imaging of multiple contrast agents in the abdomen: An in vivo study.* Medical Physics, 2017. **44**(10): p. 5120-5127.

[8]. Ren, L., B. Zheng, and H. Liu, *Tutorial on X-ray photon counting detector characterization.* Journal of X-ray science and technology, 2017.

[9]. Faby, S., et al., *Performance of today's dual energy CT and future multi energy CT in virtual non-contrast imaging and in iodine quantification: A simulation study.* Medical Physics, 2015. **42**(7): p. 4349-4366.

[10]. Yu, L., et al. *Dual-source multi-energy CT with triple or quadruple x-ray beams*. 2016.

[11]. Roessl, E. and R. Proksa, *K-edge imaging in x-ray computed tomography using multi-bin photon counting detectors.* Physics in Medicine & Biology, 2007. **52**(15): p. 4679.

[12]. Kelcz, F., P.M. Joseph, and S.K. Hilal, *Noise considerations in dual energy CT scanning.* Medical Physics, 1979. **6**(5): p. 418-425.

[13]. Mendonça, P.R.S., P. Lamb, and D.V. Sahani, *A Flexible Method for Multi-Material Decomposition of Dual-Energy CT Images.* IEEE Transactions on Medical Imaging, 2014. **33**(1): p. 99-116.

[14]. Xue, Y., et al., *Statistical image-domain multimaterial decomposition for dual-energy CT.* Medical Physics, 2017. **44**(3): p. 886-901.

[15]. Fung, G.S.K., et al. *XCAT/DRASIM: a realistic CT/human-model simulation package*. in *SPIE Medical Imaging*. 2011. SPIE.

Generalized least squares for spectral and dual energy CT: a simulation study

Cyril Mory, Bernhard Brendel, Klaus Erhard, Simon Rit

***Abstract*—In the presence of noise, decomposing spectral CT projections into materials generates anti-correlated noise. Estimating the covariance of this noise and taking it into account in the reconstruction process, by minimizing a GLS data-attachment term, is expected to lower the impact of the noise on the reconstruction. GLS has already been used in dual energy or spectral computed tomography, but always coupled with a regularization term, which raises the question of the relative impact of regularization and GLS. To our knowledge, a fair comparison between plain OLS and plain GLS is still missing. We provide one in this paper, with OLS and GLS reconstruction results from simulated projections, and discuss the relevance of using GLS. Pixels of the projection data are assumed to be independant, neglecting spatially correlated noise, and focusing on the inter-material noise correlation only. With these hypotheses, GLS brings little reduction of the noise level, while significantly increasing algorithmic complexity, slowing down convergence and requiring increased numerical precision with respect to OLS. Furthermore, in real situations, the covariance matrix has to be estimated, which adds another level of complexity and a potential source of inaccuracies.**

I. Introduction

In the presence of noise, decomposing spectral CT photon counts projections into material projections generates anti-correlated noise, since an over-estimation of the length traversed through one material must be compensated by an under-estimation of the length traversed through another to match the total attenuation. Estimating the covariance of this noise and taking it into account in the reconstruction process, by minimizing a Generalized Least Squares data-attachment term, theoretically guarantees that the estimate (the reconstructed volume) has the smallest possible variance, i.e. the least noise. GLS data-attachment terms taking into account the inter-materials correlation have already been used in dual energy [1] or spectral computed tomography [2] from material-decomposed projections, but since these studies use regularization, the impact of GLS alone cannot be evaluated from them. Throughout the paper, in order to avoid confusions, we use the term 'pixel' to describe an element of the projection data, and the term 'voxel' to describe an element of the reconstructed volume, even though the numerical experiments are actually 2-dimensional.

A. Principles of Ordinary and Generalized Least Squares

Ordinary least squares (OLS) and Generalized Least Squares (GLS) are two ways to design the cost function in an inverse problem. They do not dictate which algorithm should be used to perform minimization. When trying to retrieve a vector of parameters f from measurements p through a system of matrix R, OLS consists in minimizing $\|Rf - p\|_2^2$. If the errors on the measurements p are uncorrelated and of identical variance, minimizing the OLS cost function yields the Best Linear Unbiased Estimator (BLUE). If the errors on p are of different variances and/or correlated, the BLUE is obtained by minimizing the GLS cost function, i.e. $(Rf-p)^T C^{-1}(Rf-p)$, where C is the covariance matrix of p. Choosing $C = I$ yields the OLS cost function. OLS is therefore a specific case of GLS. In our case, the matrix R is the forward projection matrix, and f and p are column vectors representing 4D datasets: f is the set of 3D volumes to reconstruct, one per material, and p is the set of decomposed 2D material projections, one per X-ray source position and material. Throughout this document, we use the following dimension orders, from fastest to slowest:

- for f: material, then x, y, z axes of the volume
- for p: material, then u,v axes of the detector, then source position

B. Application to spectral CT

Since the decomposition method we used [3] processes each pixel of the projection data separately, it cannot provide spatial covariance information. The covariance matrix C is therefore restricted to inter-material covariance. With the dimension order specified in section I-A, it is block-diagonal, each block representing the $m \times m$ inter-material covariance matrix for a given pixel, where m is the number of materials. The forward projection matrix R is designed accordingly, and since all materials are projected the same way, it is made of blocks of m identical rows.

II. Material and methods

A. Minimization algorithm

Both OLS and GLS cost functions can be minimized by the linear conjugate gradient (CG) algorithm. CG solves problems of the type $Af = b$, where A is a symmetric positive definite matrix, f is the unknown vector and b is a known vector. Let us derive the expressions of A and b in the OLS and GLS cases:

- OLS: the minimum of the cost function is reached when its gradient is null, i.e. when

$$R^T(Rf - p) = 0$$
$$R^T Rf = R^T p$$

  Identifying the terms, $A = R^T R$ and $b = R^T p$. Note that if $R^T R$ can be inverted in a reasonable time, the Moore-Penrose pseudo inverse reconstruction $\widetilde{f} = (R^T R)^{-1} R^T b$ yields the same solution as a converged conjugate gradient initialized from zero.

C. Mory & S. Rit are with the Université de Lyon, CREATIS ; CNRS UMR5220 ; Inserm U1044 ; INSA-Lyon ; Université Lyon 1 ; Centre Léon Bérard, France.

B. Brendel & Klaus Erhard are with Philips GmbH Innovative Technologies, Research Laboratories, Hamburg, Germany

- GLS: let $U^TU = C^{-1}$ be the Cholesky decomposition of the inverse covariance matrix, with U an upper-triangular matrix. Then the GLS cost function can be rewritten as $\|U(Rf - p)\|_2^2$, and its gradient is null when

$$\begin{aligned} R^TU^TU(Rf - p) &= 0 \\ R^TU^TURf &= R^TU^TUp \\ R^TC^{-1}Rf &= R^TC^{-1}p \end{aligned}$$

  Identifying the terms, $A = R^TC^{-1}R$ and $b = R^TC^{-1}p$. The Moore-Penrose pseudo inverse reconstruction is $\widetilde{f} = (R^TC^{-1}R)^{-1}R^TC^{-1}b$.

The conjugate gradient algorithm is guaranteed to have converged after n iterations, where n is the number of elements in f. In practice, however, iterations are usually stopped far before that.

B. Simulation study

Simulations were performed with RTK [4]. We used a 2-dimensional, 3-materials (iodine, gadolinium and water) phantom, in which each material volume is made of 350×350 voxels. The phantom is composed of a small cylinder of iodine at concentration $1\,\text{mg.ml}^{-1}$, a small gadolinium cylinder at concentration $1\,\text{mg.ml}^{-1}$, both inside a large water cylinder at concentration $1\,\text{g.ml}^{-1}$, as shown in Fig. 1. 2400 projections of 643×1 pixels were computed analytically, using the geometry of the spectral CT scanner prototype (Philips Healthcare, Israël) installed at the CERMEP, Lyon, France. They were then converted to photon counts with 5 energy bins using the scanner's spectrum, detector response, and X-ray attenuation functions of each material. Poisson noise was added to the photon counts, then the noisy photon counts were decomposed into material projections using the RTK implementation of the method described in [3]. Since with noisy input photon count data, this method generates aberrant results for some pixels, we applied a Hampel filter [5] to remove the outliers. For each pixel, in each material, a $3 \times 3 \times 3$ neighborhood centered on the pixel is considered, and the median and standard deviation in this neighborhood are computed. If the pixel's value is off the median by more than 2σ, it is replaced by the median, otherwise it is left unchanged. In order to obtain a reliable estimate of the covariance of the noise on the decomposed projections, we repeated the last three steps (Poisson noise, decomposition and denoising) a hundred times, each time with a different realization of Poisson noise, and for each pixel, we computed the 3×3 covariance matrix on these hundred (3-materials) values. For the reconstructions, only one set of material projections was required, so we kept only the one from the first simulation.

Fig. 1: The three channels of the phantom used for simulations on RTK, shown side-by-side. From left to right: iodine, gadolinium, water

C. Evaluation of inter-material correlation coefficients

On the RTK simulated data, during the computation of the covariance matrix, we also computed the standard deviation of the noise in each material and each pixel, on the hundred realizations. Combining the covariance matrix and the standard deviation values, we were able to calculate correlations coefficients. Obviously, like the covariances, these vary from one pixel to the other, but the mean and standard deviation over all pixels is nevertheless interesting. The correlation coefficients we measured were the following:

- Iodine-gadolinium correlation's mean = -0.3963, std = 0.1185,
- Iodine-water correlation's mean = -0.1571, std = 0.1135,
- Gadolinium-water correlation's mean = -0.5489, std = 0.07953.

As expected, the noise is anti-correlated between materials.

D. Convergence issues and workarounds

While the mathematics of GLS are appealingly simple, we encountered a number of practical difficulties during our investigation of this topic. These difficulties are listed in the present section.

1) First CG iterates: It is well known that the first CG iterates of an OLS reconstruction are (roughly) low-pass approximations of the result CG yields at convergence, and high frequencies build up over the course of the iterations. In a GLS reconstruction, the first CG iterates are also low frequency, but they display a lot of cross-talk between materials (see Fig. 3). This cross-talk disappears over the course of the iterations. Stopping the iterations early is therefore much more problematic in GLS than in OLS. This problem can be alleviated by initializing the GLS conjugate gradient reconstruction with an early iterate of the corresponding OLS conjugate gradient reconstruction: with this approach, early stopping the GLS part no longer yields a result with a lot of cross-talk. However, it does not seem to reduce the number of iterations required to attain convergence. Therefore all results presented in this paper were obtained from a zero initialization.

2) Relative convergence speed: On the experiments we have carried out, GLS required many more iterations than OLS to reach convergence. Having to perform a lot of iterations for the GLS case brought in additional problems: after a certain point, each additional iteration of CG increased the noise in the reconstructions. This was due to the forward and a back projector not being the adjoint of one another with enough precision.

3) Adjoint operators: Having unmatched forward and back projectors is usually not a problem if the number of iterations remains low (typically below 100), but we had to do more iterations than that. We adopted the method described in [6], i.e. we computed the ratio between $< Rf, p >$ and $< f, R^Tp >$, where f and p are a random volume and a random set of projections, respectively, and $< ., . >$ denotes the dot product. The precision with which this ratio matches 1 is a measure of the 'adjointness' of R and R^T. In our case, using RTK's matched implementations of the Joseph forward and back projector [7], we obtained $1 - ratio \approx 10^{-5}$, which was insufficient. Switching from single precision (32-bits floats) to double precision

(64-bits) yielded $1 - ratio \approx 10^{-13}$, and got rid of the noise divergence problem. With our implementations, having to use matched forward and back projectors implies that at least one is sub-optimally implemented (either the ray-based back projector, like in our case, or a voxel-based forward projector, which is not available in RTK), which slows down calculations. In addition, having to use double precision both increases the reconstruction time again and doubles the memory requirements.

4) Scaling as preconditioning: Preconditioning consists in inserting a matrix D, and solving by CG the following problems instead of the ones described in section II-A:

- OLS:
$$DR^T RD^T \hat{f} = DR^T p$$
Identifying the terms, this means one must run the conjugate gradient algorithm with $A = DR^T RD^T$ and $b = DR^T p$, and obtain $\hat{f}$. Then $f = D^T \hat{f}$.
- GLS:
$$DR^T C^{-1} RD^T f = DR^T C^{-1} p$$
Identifying the terms, this means one must run the conjugate gradient algorithm with $A = DR^T C^{-1} RD^T$ and $b = DR^T C^{-1} p$, and obtain $\hat{f}$. Then, again, $f = D^T \hat{f}$.

Any matrix D can be used, but the goal of preconditioning is to obtain a new matrix A with a condition number closer to 1 than the original matrix A, because this condition number has a large impact on convergence speed (the closer it is to 1, the faster CG converges). In case of a diagonal preconditioner, $D = D^T$, and multiplying by D consists in a voxel-wise multiplication. We used

$$D = I_N \otimes \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

i.e. each material was simply scaled by a fixed factor λ_i. In our specific case, scaling can be performed either before or after forward projection, therefore $RD = D'R$ with D' a matrix similar to D but scaling projections instead of volumes. Defining $C' = D'C^{-1}D'$, the matrix A becomes $A = R^T C' R$ and the vector b becomes $b = R^T C' D'^{-1} p$. Preconditioning can therefore be obtained without increasing the amount of calculations at each iteration, simply by dividing the material projections by the $\lambda_i, i \in \{1..m\}$, performing a GLS reconstruction with a scaled covariance matrix C', and then multiplying the obtained material volumes by λ_i. We tried several choices for the λ_i: setting them all to 1, i.e. no preconditioning; setting them to the mean of the ground truth projection of each material; and $\lambda_i = \underset{E}{\text{mean}} \left(\frac{\mu_{iodine}(E)}{\mu_i(E)} \right)$, where $\mu_i(E)$ is the mass-attenuation of material i at energy E. The third method, although it does not depend of the object, and is therefore probably suboptimal in some cases, lead to the fastest convergence in our experiments.

5) Fair comparison: Because OLS and GLS converge at different speeds, choosing one iterate of each to perform a fair comparison between them is a delicate problem. The only possibility seems to be to wait until both have converged. After 1000 iterations, the GLS cost function reduction between successive iterates had dropped to 10^{-6}, which was deemed sufficient, while OLS had reached convergence after 150 iterations. Fig. 3 shows the convergence curves in logarithmic scale.

	OLS	GLS	Gain
MSE iodine	8.79e-07	7.85e-07	-10.7%
MSE gadolinium	4.62e-07	4.19e-07	-9.2%
MSE water	0.0352	0.0327	-7.0%
SNR iodine	0.478	0.486	+1.6%
SNR gadolinium	0.606	0.631	+4.1%
SNR water	2.497	2.572	+3.0%

TABLE I: Results on RTK simulations: compared MSEs and SNRs between OLS and GLS

III. Results

A. Photon counting experiments

Fig. 2 shows the results obtained with RTK, with one realization of the Poisson noise on photon counts. The GLS results are slightly less noisy than the OLS ones, but the improvement is only noticeable in the numerical analysis, and not visually. The MSE with respect to the ground truth and the SNR, computed as $\text{mean}(C)/\text{std}(C)$, where C is the set of voxels inside the cylinder of the material considered, are displayed in Table I.

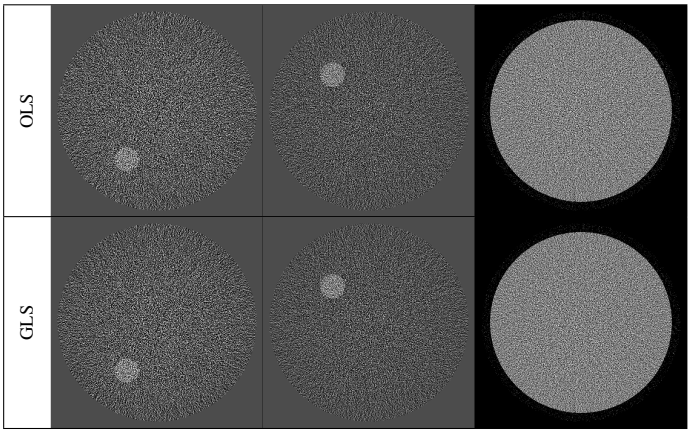


Fig. 2: RTK reconstructions with converged OLS and GLS. The GLS reconstructions are slightly less noisy, as shown in Table I, but it is not obvious on any of the reconstructions

B. Dual energy experiments

The small impact of GLS in the spectral CT experiments raises a question: could GLS be more effective in reducing the noise in a dual energy case, given that the anti-correlation between materials is much stronger in dual energy ? We ran the same RTK simulations as for spectral CT, only with a 2-materials phantom (water and iodine). The products of the detector responses by the incident spectrum were obtained by digitizing the green and pink curves in Fig. 1 of [8]. With this setup, we measured a mean correlation coefficient of -0.899. A Hampel filtering was applied on the material-decomposed projections, just like in spectral CT, and all subsequent results (covariance matrix, OLS and GLS reconstructions) were computed from the denoised projections. Despite the preconditioning, it took about 50 iterations for the OLS reconstruction to converge, and about 3000 iterations for the GLS reconstruction. Since, just like in the spectral case, the reconstructions OLS and GLS results are visually identical, they are not shown here, but image quality metrics are provided in Table II.

IV. Discussion

In an attempt to isolate how GLS improves over OLS from other contributions to noise reduction, we have chosen to perform unregularized reconstructions only, and have noticed

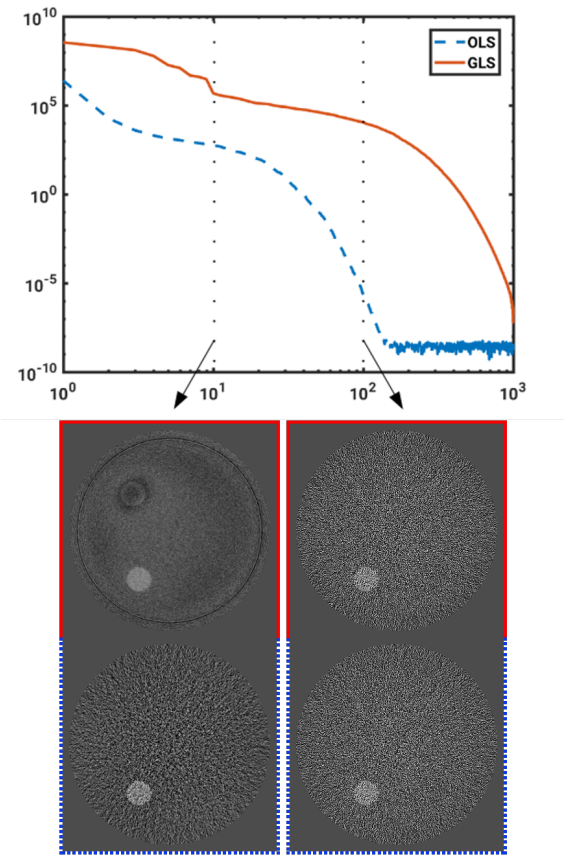


Fig. 3: Cost function after each iteration minus cost function after iteration 1000, in logarithmic scale. Iodine volumes at iterates 10 and 100 are shown below the graph for GLS and OLS.

	OLS	GLS	Gain
MSE iodine	9.18e-7	8.12e-7	-11.6%
MSE water	0.407	0.405	-0.57%
SNR iodine	0.442	0.470	+6.27%
SNR water	3.067	3.197	+4.25%

TABLE II: Results on RTK dual energy simulations: compared MSEs and SNRs between OLS and GLS

that GLS brought a very limited benefit. Whether this conclusion also holds for regularized reconstructions, and with which regularization method and parameters, remains an open question. Taking into account spatial correlations by accurately modeling advanced effects like scattering or charge sharing should increase the benefit of using GLS. However, it would also make implementation more complex, since C^{-1} would no longer be block-diagonal, therefore no longer "pixel-separable". In this work, the covariance matrix is estimated from a large number of repetitions of the same simulation (here 100), where only the Poisson noise realization changes. This empirical estimation is undoubtedly a source of inaccuracies, which could undermine the efficiency of GLS. To evaluate the importance of this error in our simulations, we ran a side experiment, adding noise with exactly the right covariance matrix (the one estimated by the simulations) to the noiseless projections, and reconstructing from those. This had no noticeable effect on the image quality of the GLS results, leading us to think that the limited impact of GLS is not caused by inaccuracies in the covariance matrix estimation. In addition, in a real situation, the covariance matrix estimation would most likely be less accurate than in the presented simulations.

V. Conclusion

Even with a quite precise knowledge of the covariance matrix, GLS only brings a moderate improvement of SNR in the RTK simulated case we have studied. The inevitable inaccuracies in estimating the covariance matrix in real situations are likely to reduce this improvement even more. Overall, the gain in SNR GLS can provide over OLS does not seem to be worth the efforts it requires in implementation, the risk that an incorrect covariance matrix might degrade the reconstruction, and the drop in performance and increase in memory footprint implied by matched projectors with double precision. Note that these conclusions hold for unregularized OLS and GLS: adding a regularizer may lead to different results.

Acknowledgements

This project was done in the framework of the EU's H2020 research and innovation program under the grant agreement No. 633937. This work was partially supported by grant ANR-17-CE19-0006-02 (ROIdoré project) from the French National Research Agency (ANR). This work was performed within the framework of the SIRIC LYric Grant INCa-DGOS-4664 and the LABEX PRIMES (ANR-11-LABX-0063) of Université de Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR).

References

[1] Y. Liu, Z. Yu, and Y. Zou, "Impact of covariance modeling in dual-energy spectral CT image reconstruction," p. 94123Q, Mar. 2015.

[2] Q. Xu, A. Sawatzky, M. A. Anastasio, and C. O. Schirra, "Sparsity-regularized image reconstruction of decomposed K-edge data in spectral CT," *Physics in Medicine and Biology*, vol. 59, pp. N65–N79, May 2014.

[3] J. P. Schlomka, E. Roessl, R. Dorscheid, S. Dill, G. Martens, T. Istel, C. Bäumer, C. Herrmann, R. Steadman, G Zeitler, A. Livne, and R. Proksa, "Experimental feasibility of multi-energy photon-counting K-edge imaging in pre-clinical computed tomography," *Physics in Medicine and Biology*, vol. 53, no. 15, p. 4031, 2008.

[4] S. Rit, M. Vila Oliva, S. Brousmiche, R. Labarbe, D. Sarrut, and G. C. Sharp, "The Reconstruction Toolkit (RTK), an open-source cone-beam CT reconstruction toolkit based on the Insight Toolkit (ITK)," in *Proceedings of the International Conference on the Use of Computers in Radiation Therapy (ICCR)*, 2013.

[5] R. K. Pearson, Y. Neuvo, J. Astola, and M. Gabbouj, "Generalized Hampel Filters," *EURASIP Journal on Advances in Signal Processing*, vol. 2016, p. 87, Dec. 2016.

[6] F. Arcadu, M. Stampanoni, and F. Marone, "On the crucial impact of the coupling projector-backprojector in iterative tomographic reconstruction," *arXiv:1612.05515 [cs]*, Dec. 2016. arXiv: 1612.05515.

[7] P. M. Joseph, "An Improved Algorithm for Reprojecting Rays through Pixel Images," *Medical Imaging, IEEE Transactions on*, no. 3, pp. 192 – 196, 1982.

[8] K. Brown, S. Zabic, and G. Shechter, "Impact of Spectral Separation in Dual-Energy CT with Anti-Correlated Statistical Reconstruction," pp. 491–494, 2015.

Model-based Multi-material Decomposition using Spatial-Spectral CT Filters

J. Webster Stayman and Steven Tilley II*

***Abstract*—Spectral CT with multiple contrast agents has been enabled by energy-discriminating detectors with multiple spectral channels. We propose a new approach that uses spatial-spectral filters to provide multiple beamlets with different incident spectra for spectral channels based on "source-side" control. Since these spatial-spectral filters yield spectral channels that are sparse, we adopt model-based material decomposition to directly reconstruct material densities from projection data. Simulation studies in three- and four-material decomposition experiments show the underlying feasibility of the spatial-spectral filtering technique. This methodology has the potential to facilitate imaging of multiple contrast agents simultaneously with relatively simple hardware, or to improve spectral CT performance via combination with other established spectral CT methods for additional control and flexibility.**

I. Introduction

Spectral CT is an emerging technology that permits decomposition and density estimation for multiple material components within an image volume. In particular, spectral CT has enabled simultaneous imaging of multiple contrast agents for applications including simultaneous iodine-bismuth imaging for angiography/lung imaging [1] and three-agent iodine-gadolinium-bismuth imaging [2] for multiphasic renal [3] studies. A number of new contrast agents are also in development including *gold nanoparticles* for angiography [4], mammography [5], and targeted imaging of HeLa cells [6], lung adenocarcinoma [7], and colorectal liver metastasis [8]; *xenon* for lung ventilation [9]; *bismuth sulphide nanoparticles* for lymph nodes [10]; and *tantalum oxide nanoparticles* for cartilage [11]. However, most spectral CT has focused on single contrast agent imaging (e.g. iodine) – inevitably leading to systems that are optimized for that agent. With the emergence of new contrast agents and simultaneous imaging of multiple agents, there is need for sufficient flexibility in data acquisition to acquire high-quality spectral data for many contrast agents.

A number of different strategies have been investigated to enable spectral CT. Methods include the use of dual-sources [12], kV-switching [13], split-filters [14], dual-layer detectors [15], and energy-discriminating photon-counting detectors [16]. Many photon-counting detectors have the flexibility to provide several energy bins for spectral discrimination which enables multi-material decomposition. However, individually, most of the other methods typically only easily allow two different spectral channels limiting their use in multiple contrast agent studies. For example, "source-side" spectral variation is often limited to two x-ray sources, alternation between two energies in kV-switching, or two filters in a split-filter design where spectral filters cover exactly one-half of the fan-beam x-ray. With only two spectral channels, only two (or, three using a constraint) different materials may be estimated as part of a decomposition. Thus, a strategy for more control over the number and form of x-ray spectra available for data acquisition has the potential to enable and improve multi-material decomposition in CT.

In this work, we introduce a new concept wherein the x-ray beam is spectrally modulated across the face of the detector using a repeating pattern of filter materials. Such *spatial-spectral* filters allow for collection of many different spectral channels using "source-side" control. However, in contrast to other spectral techniques that provide mathematically complete projection data, spatial-spectral filtered data is sparse in each spectral channel – making traditional projection-domain [17] or image-domain [18] material decomposition difficult to apply. Thus, we adopt direct model-based material decomposition [19]–[21] which combines reconstruction and multi-material decomposition, and permits arbitrary spectral, spatial, and angular sampling patterns.

Spatial-spectral filters can be interpreted as an extension of split-filters methods [14] that divide the beam into two different spectra for each half-fan – instead, dividing the x-ray beam into several different beamlets. The proposed spatial-spectral approach also includes mechanical translation of the filter to vary spectral sampling patterns. Such an approach shares similarities with other recent "source-side" filtering innovations in CT acquisition including beam-shaping approaches using multiple aperture devices [22], interrupted-beam acquisitions for sparse data [23], and grating-oriented line-wise filtration for dual-energy CT [24].

In this paper we introduce the concept of spatial-spectral filtering for multi-material CT decomposition. Simulation experiments are conducted demonstrating the basic feasibility of the approach. Various spatial-spectral filter designs are explored and applied to multi-contrast imaging studies in simple digital phantom studies.

II. Methods

A. Spatial-Spectral Filtering

An overview of the spatial-spectral filtering approach is shown in Figure 1. The main idea is to place a tiled filter with different materials in front of the x-ray source. For example, a repeating pattern of several material types can be used to shape the spectrum of a number of beamlets across the face of the detector. This kind of spectral encoding through spatial filtering is widely applied in

*Biomedical Engineering, Johns Hopkins University, Baltimore, MD. email: web.stayman@jhu.edu

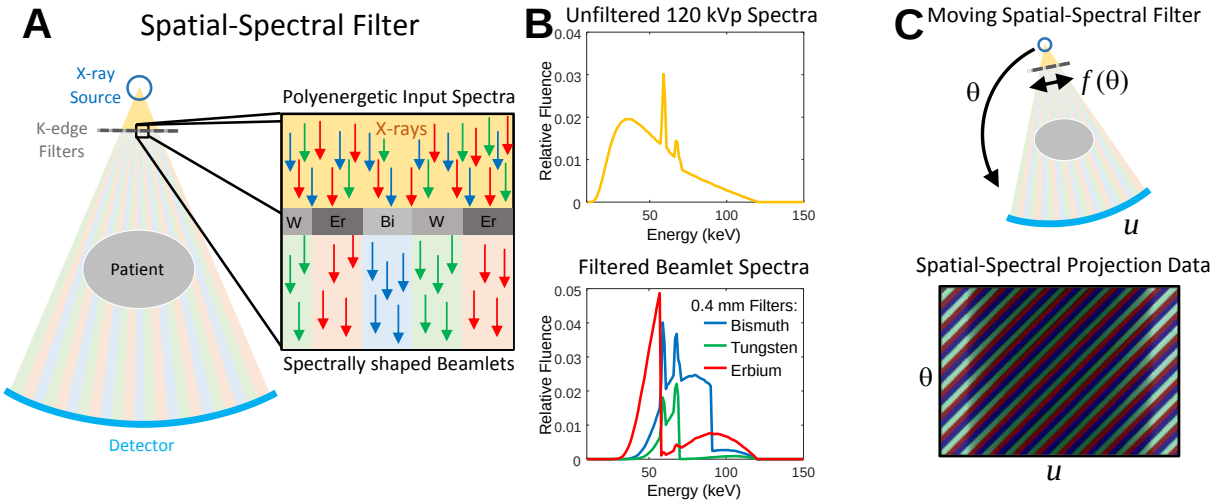


Fig. 1. Illustration of the spatial-spectral filtering concept. A) The spatial-spectral filter is composed of a repeating pattern of attenuating materials used to shape the x-ray beam into beamlets with different spectra. Materials with k-edges in the diagnostic energy range are potentially of greatest utility for spectral shaping. B) In a sample spatial-spectral filter application, a 120 kVp spectrum emitted from the x-ray tube is modified with a repeating pattern of bismuth, tungsten, and erbium filters (0.4 mm thickness) to achieve three different spectral channels. The total fluence over all energies is normalized to the integral of the bismuth-filtered spectrum. C) Adding the ability to translate the spatial-spectral filter permits variation of the sampling pattern. A simple constant velocity linear translation as a function of rotation angle is shown.

optical imaging – e.g. Bayer filters [25] used in visible light detectors for color imaging. For x-ray imaging, practical filters may be constructed from materials with k-edges in the diagnostic range. This includes elements from Z = 50 (tin, k-edge at 29.2 keV) to Z = 85 (bismuth, k-edge at 90.5 keV). A sample repeating filter that uses a tiled pattern of bismuth, tungsten, and erbium filters is shown in Figure 1. The spectra produced by a 0.4 mm thick filter and a typical 120 kVp x-ray spectrum are shown in Figure 1B. Note the distinct edges in the spectra that coincide with the k-edges of those material filters. The exact number and combination of materials for filter construction can be optimized for particular contrast agents; however, it is straightforward to include several different filters for a diversity of spectral channels.

While one could potentially implement spectral CT with a static spatial-spectral filter, this has the potential to have poor overall sampling. For example, for a static filter, the center of the object being scanning may only be probed with a single spectrum (based on the filter placed at the central ray). To provide increased flexibility in the spatial-spectral sampling pattern, we propose to translate the filter with rotation. Even a simple constant velocity linear translation will improve the sampling homogeneity across spectral channels. Such a linear translation and an illustration of the resulting projection data is shown in Figure 1C. Note that each spectral channel is sparse but all channels can be collected in a standard acquisition.

B. *Model-based Material Decomposition*

In order to reconstruct such sparse data, we will adopt a direct model-based material decomposition (MBMD) approach. The authors have previously developed a MBMD algorithm [26] based on related development in advanced high-resolution CT reconstruction [27]. This MBMD approach uses a forward model for projection data where mean measurements are

$$y_i = \sum_e s_{e,i} \exp\left(-\sum_m p_{e,m} [\mathbf{A} x_m]_i\right)$$

where the i^{th} measurement is formed by projection (via $\mathbf{A}$) of m material density maps, x_m, and scaling by the energy-dependent mass attenuation coefficients, $p_{e,m}$. Each measurement has an energy-dependent factor, $s_{e,i}$, which incorporates the local incident spectrum induced by the moving spectral-spatial filter as well as the detector sensitivity (e.g., energy-dependent scintillator stopping power, secondary quanta proportional to energy, etc.).

The objective function for the MBMD approach uses a Gaussian log-likelihood function and standard roughness penalty regularization for each material basis. The nonlinear objective function is solved using an optimization transfer approach based on separable surrogate functions. Details of the algorithm can be found in [26] and [27]. For all reconstructions/material decompositions in this work, 500 iterations of the algorithm are applied initializing the water material basis with a thresholded, binarized FBP image with a density of 1000 mg/ml, and all other material bases set to zero. All reconstructions used a 256 x 256 volume of 0.5 mm voxels.

C. *Simulation Experiments*

Two preliminary investigations of spatial-spectral filters are explored in this work. First, we consider a three-material decomposition example using a digital phantom composed of water, iodine, and gold – emulating a multiple contrast agent study with both iodinated contrast agent and gold nanoparticles (an emerging contrast agent). The digital phantom is illustrated in Figure 3A and includes a 10 cm cylinder of water and various concentrations of iodine (2.5-15 mg/ml) and gold (1-4 mg/ml) in 2 cm diameter cylinders. We simulated a CT system with 120 cm source-to-detector distance and 60 cm source-to-axis distance with an indirect energy integrating detector (with 600 μm thick CsI scintillator) with 512-0.5 mm flat-panel pixels, and

	Spatial-Spectral Filter	Iodine RMSE (mg/ml)	Iodine Rank	Gold RMSE (mg/ml)	Gold Rank
4-Element Filters	Bi-Au-W-Lu	0.2974	8	0.1852	5
	Bi-Au-W-Er	0.2980	9	0.1838	3
	Bi-Au-Lu-Er	0.2642	2	0.1833	2
	Bi-W-Lu-Er	0.2646	3	0.2120	8
	Au-W-Lu-Er	0.4945	13	0.3624	13
3-Element Filters	Bi-Au-W	0.4434	11	0.1870	6
	Bi-Au-Lu	0.2706	5	0.1711	1
	Bi-Au-Er	0.2804	7	0.1846	4
	Bi-W-Lu	0.2705	4	0.2287	9
	Bi-W-Er	0.2708	6	0.1964	7
	Bi-Lu-Er	0.2619	1	0.2756	10
	Au-W-Lu	0.6998	15	0.5995	15
	Au-W-Er	0.5436	14	0.3826	14
	Au-Lu-Er	0.4705	12	0.3556	12
	W-Lu-Er	0.4368	10	0.2824	11

<0.15 <0.20, >0.15 <0.25, >0.20 <0.30, >0.25 <0.40, >0.30 >0.40

Fig. 2. A summary of all spatial-spectral filter performance for the water-iodine-gold material decompositions. The RMSE for iodine and gold contrast agents are reported and ranked.

rotated 360° in 360-1° increments. We used a 120 kVp source spectrum (prior to spatial-spectral filtration but with typical inherent filtration due to the housing, and shown in Figure 1B).

Various combinations of filtering materials were explored to determine their relative performance for the three-material water-iodine-gold decomposition. To keep the search space reasonable, we focused on five metals with k-edges in the diagnostic range, that are readily available, machinable, and with few physical hazards. These were bismuth (k-edge at 90.5 keV), gold (80.7 keV), tungsten (69.5 keV), lutetium (63.3 keV), and erbium (57.5 keV). For all investigations, a 0.25 mm thick filter was simulated passing 17.5%, 6.5%, 6.5%, 17.7%, and 8.3% of the incident 120 kVp spectra for each material, respectively. All data was normalized so that the bare-beam fluence for the bismuth-filtered beamlet was $5x10^5$ photons/pixel. Filters were designed so that each beamlet was 8 detector pixels wide and the spatial-spectral filter was translated 1 pixel per rotation angle. For these preliminary studies, perfect alignment of beamlets with pixel boundaries was presumed. All 3- and 4-element spatial-spectral filters possible with the above five materials were investigated. Root mean squared error (RMSE) was computed for material decomposition density estimates. Regularization parameters for material bases were chosen to be 3 x 10^8, 5 x 10^{11}, and 1 x 10^{12} for the water, iodine, and gold bases, respectively; and the Huber penalty delta was set to 10^{-3}.

A second study considered a four-material decomposition for a digital phantom with water, iodine (2-8 mg/ml), gold (1-4 mg/ml), and gadolinium (1-4 mg/ml), which is illustrated in Figure 4B. The same CT system geometry, source, detector, and filter parameters were used. In this case all possible 4-element filters and the one 5-element filter from the five aforementioned filter materials were explored. RMSE was computed for each material decomposition. Regularization parameters for material bases were chosen to be 3 x 10^8, 1 x 10^{11}, 1 x 10^{12} and 1 x 10^{12} for the water, iodine, gold, and gadolinium bases, respectively; and the Huber penalty delta was set to 10^{-3}.

III. Results/Discussion

The results of the three-material decomposition with spatial-spectral filtering are shown in Figure 2. RMSE for both the iodine and gold contrast agents are reported and ranked. The best filter for estimating iodine density is the 3-element Bi-Lu-Er filter; whereas the best filter for estimating gold density is the Bi-Au-Lu filter. The 4-element Bi-Au-Lu-Er filter achieves the 2nd best performance for both iodine and gold – suggesting that the 4-element filter provides some level of compromise, and

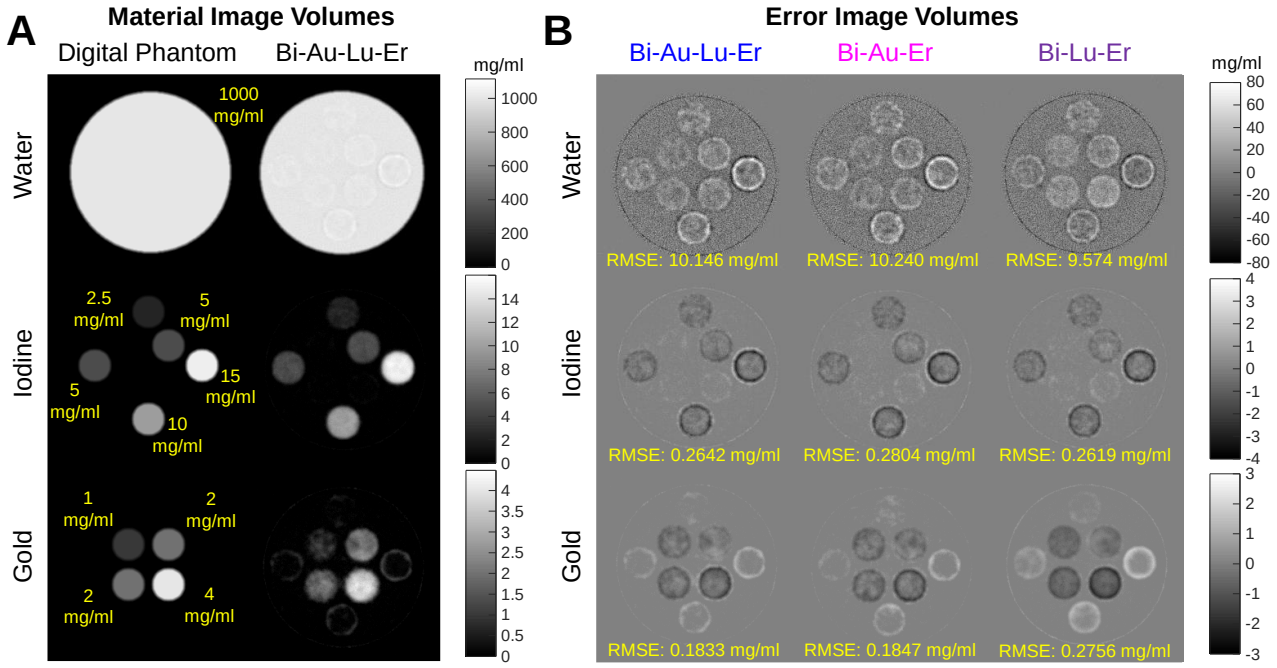


Fig. 3. Three-material decompositions using spatial-spectral filters. A) The ground-truth digital phantom and a representative reconstruction/material decomposition for the Bi-Au-Lu-Er spatial-spectral filter. B) Error image volumes for each material basis for three different spatial-spectral filters: Bi-Au-Lu-Er (best 4-element filter); Bi-Au-Er (best 3-element filter for gold contrast); and Bi-Lu-Er (best 3-element filter for iodine contrast).

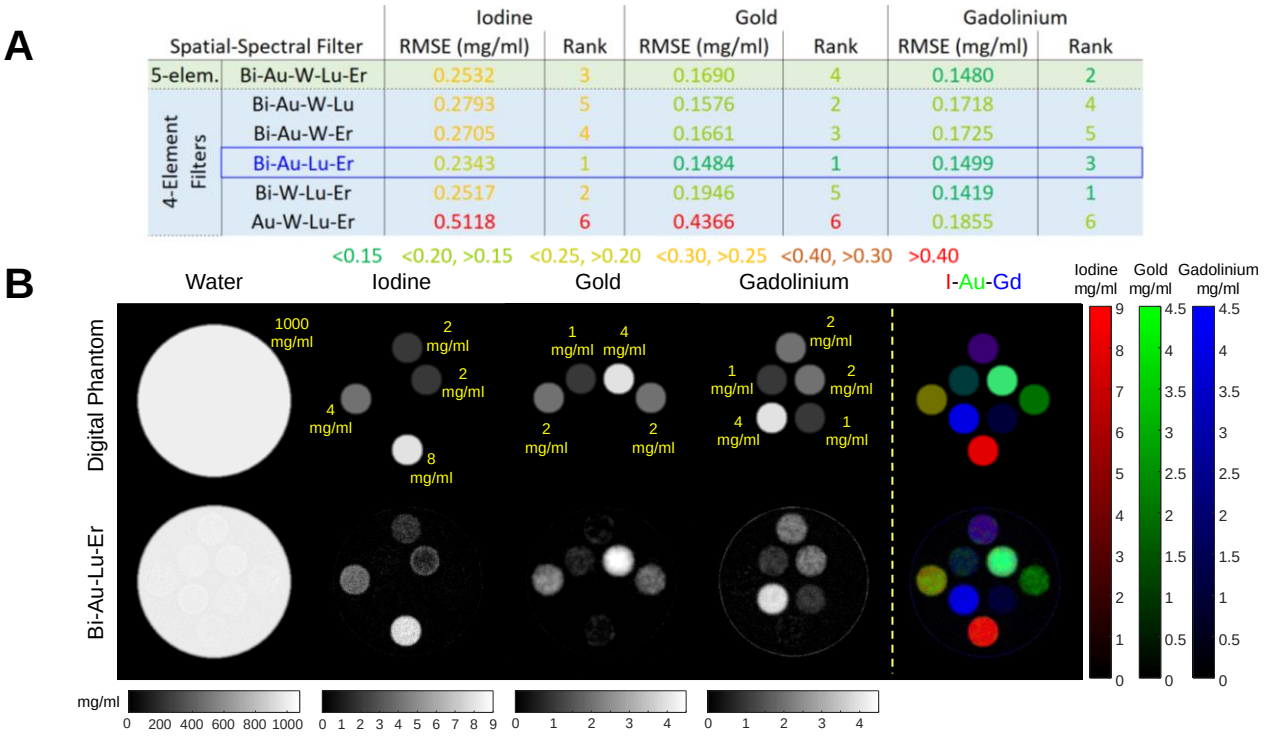


Fig. 4. Four-material decompositions using spatial-spectral filters. A) Summary of the performance of all 4-element filters and the 5-element filter. B) Ground truth and estimated densities using the Bi-Ai-Lu-Er filter for the four-material decomposition experiment.

gets the best of both 3-element filters. Note that the bismuth filter appears important for good performance. A reconstruction and error images associated with these three spatial-spectral designs is shown in Figure 3.

A summary of the four-material decomposition results are shown in Figure 4A. The best overall filter was the 4-element Bi-Au-Lu-Er filter (although the Bi-W-Lu-Er had very slightly better gadolinium estimates). Decomposition images for this filter are shown in Figure 4B including a colorized iodine-gold-gadolinium image. Again, the bismuth filter appears critical in delivering a good decomposition.

These preliminary studies suggest the feasibility of spatial-spectral filters to provide spectral CT using relatively simple hardware. The filters permit a larger number of spectral channels that traditional "source-side" modifications which facilitates material decompositions with multiple contrast agents. Moreover, spatial-spectral filters may be combined with other spectral CT approaches for additional advantages. For example, combinations may help to improve the relationship between density estimates and radiation dose, or to help improve the low concentration density estimation limits (especially important for highly specific and targeted contrast agents).

While these preliminary results are encouraging a number of ongoing studies are required to fully realize the spatial-spectral filter concept in a physical system. In particular, due to finite x-ray focal spot size, misalignment of beamlets with pixel boundaries, etc., more complete modeling of the transition regions between beamlets is required. Similarly, kV optimization and an exploration of tube efficiency and tube loading must be undertaken. Optimal sampling patterns driving filter actuation and the effects of regularization also need to be fully explored.

Work was supported, in part, by NIH grant U01EB018758.

References

[1] N. G. Anderson *et al.*, *Eur. Radiol.*, 20(9), pp. 2126–2134, Sep. 2010.
[2] R. Symons *et al.*, *Med. Phys.*, 44(10), pp. 5120–5127, Oct. 2017.
[3] M. Kekelidze *et al.*, *Radiology*, 255(2), pp. 508–516, May 2010.
[4] C. Alric *et al.*, *J. Am. Chem. Soc.*, 130(18), pp. 5908–5915, May 2008.
[5] L. E. Cole, *et al.*, *ACS Nano*, 9(9), pp. 8923–8932, Sep. 2015.
[6] C. Peng *et al.*, *Polym. Chem.*, 4(16), p. 4412, Jul. 2013.
[7] H. Wang *et al.*, *Biomaterials*, 34(2), pp. 470–480, Jan. 2013.
[8] I.-C. Sun *et al.*, *Pharm. Res.*, 31(6), pp. 1418–1425, Jun. 2014.
[9] E. J. Chae *et al.*, *Radiology*, 248(2), pp. 615–624, Aug. 2008.
[10] O. Rabin *et al.*, *Nat. Mater.*, 5(2), pp. 118–122, Feb. 2006.
[11] J. D. Freedman *et al.*, *Angew. Chem. Int. Ed. Engl.*, 53(32), pp. 8406–10, Aug. 2014.
[12] T. G. Flohr *et al.*, *Eur. Radiol.*, 16(2), pp. 256–268, Feb. 2006.
[13] D. Xu *et al.*, *SPIE Medical Imaging*, 2009, p. 72583T.
[14] B. Rutt, A. Fenster, *J. Comput. Assist. Tomogr.*, 4(4), pp. 501–9, Aug. 1980.
[15] R. Carmi et al., *IEEE Nuclear Science Symposium Conference Record*, 2005, vol. 4, pp. 1876–1878.
[16] J.-P. Schlomka *et al.*, *Phys. Med. Biol.*, 53(15), pp. 4031–47, Aug. 2008.
[17] R. E. Alvarez, A. Macovski, *Phys. Med. Biol.*, 21(5), p. 2, Sep. 1976.
[18] X. Liu *et al.*, *Med. Phys.*, 36(5), pp. 1602–9, May 2009.
[19] Y. Long, J. A. Fessler, *IEEE Trans. Med. Im.*, 33(8), pp. 1614–26, Aug. 2014.
[20] R. Foygel Barber *et al.*, *Inst. Phys. Eng. Med. Print. UK Phys. Med. Biol*, 61, pp. 3784–3818, 2016.
[21] R. Zhang *et al.*, *IEEE Trans. Med. Im.*, 33(1), pp. 117–34, Jan. 2014.
[22] J. W. Stayman *et al.*, *SPIE Medical Imaging*, 2016, p. 97830X.
[23] T. Koesters *et al.*, *SPIE Medical Imaging*, 2017, p. 101320Q.
[24] Y. Xi *et al.*, *Sens Imaging*, 18, 2017.
[25] P. M. Hubel *et al.*, *SPIE Electronic Imaging*, 2004, p. 402.
[26] S. Tilley II *et al.*, *SPIE Medical Imaging*, 2018, accepted.
[27] S. Tilley II *et al.*, *IEEE Trans. Med. Imaging*, 2018, accepted.

Prior-Based Multi Material Decomposition (PBMMD) for Dual Energy CT

Sabrina Dorn, Shuqing Chen, Stefan Sawall, Joscha Maier, Michael Knaup, Andreas Maier, Michael Lell, and Marc Kachelrieß

***Abstract*—With the advent of dual energy CT it is possible to obtain a quantitative analysis of the material composition of the anatomical structures in the human body. However, the task of material decomposition is still challenging, as organs might be a composite of numerous materials: e.g. blood, fat, water, air, bone and contrast media. Although multi energy CT achieves promising material decompositions into two, three, or more materials, this technology is not yet prevalent in the clinical routine. In this work, we present a prior–based multi material decomposition (PBMMD) for dual energy computed tomography (DECT) data. The method consists of multiple organ–dependent three material decompositions exploiting prior anatomical information. Using a cascaded 3D fully convolutional neural network, the CT data set is segmented and classified into different organs. Depending on the segmented structure, the basis material triplet is selected organ–dependently leading to overlapping triangles in the dual energy space. The PBMMD approach is evaluated using patient data in the arterial and portal venous phase acquired with a dual source CT system. The materials are quantified according to the anatomical structure they belong to. We conclude that the proposed method provides physically plausible volume fractions that bear the potential to improve the material quantification for diagnosis, particularly to evaluate the vascularization of tumors.**

I. Introduction

Dual energy computed tomography (DECT) acquisition samples the spatial distribution of the energy–dependent linear attenuation coefficient at two different x–ray spectra. There are two main physical mechanisms dominating the attenuation of biological materials in the clinically relevant energy range of 30 to 150 keV: the photoelectric effect that is dominated by the atomic number of the composite materials and the Compton effect that depends on the irradiated beam energy. Hence, dual energy (DE) allows the differentiation between two materials with high atomic numbers, e.g. iodine, calcium or gadolinium, or in clinical practice, two other basis materials, e.g. water and bone, if no K–edge is present in the energy range. However, there is a clinical interest in decomposing the DE data into more than three materials. For example, the quantification of the fat content in the liver requires the decomposition into four basis components: liver tissue, fat, blood and contrast media (CM) [7], [8], [11]. By making an additional assumption on the mixing behavior, it is possible to decompose DE data into three material components [6]. However, there are two problems regarding the three material decomposition: the basis materials at the vertices of the triangle may not represent the components of the organ of interest. For instance, bone can mathematically be represented by fat, blood and iodine, but this representation may lack a physical interpretation. The volume fractions may not result in meaningful values if the voxel is outside the triangle where at least one basis material coefficient becomes negative. One possible solution is the tessellation of the DE diagram resulting in more than three basis materials [8] (see Fig. 2 lower left). The three basis materials for the decomposition are determined by the location in the DE diagram of the voxel that should be decomposed. Due to noise, the voxel could fall into an incorrect triangle leading to ambiguities in the decomposition. Each point that is located in the convex hull of the tessellation is decomposed into its three basis materials. Even if the resulting volume fractions are non–negative and sum up to one, wrong vertices might be used that do not always result in a physiologically meaningful decomposition. The major disadvantage of this method is the isolated treatment of each voxel. Neither the information from adjacent voxels nor the location of the voxel within the volume is used. In our preliminary work, a segmentation–assisted material quantification method couples the material decomposition with a simple threshold–based segmentation [5]. In this approach, the segmentation is very limited because no assignment to an anatomical structure is possible. Moreover, we have demonstrated the benefit of more sophisticated prior anatomical knowledge for a context–

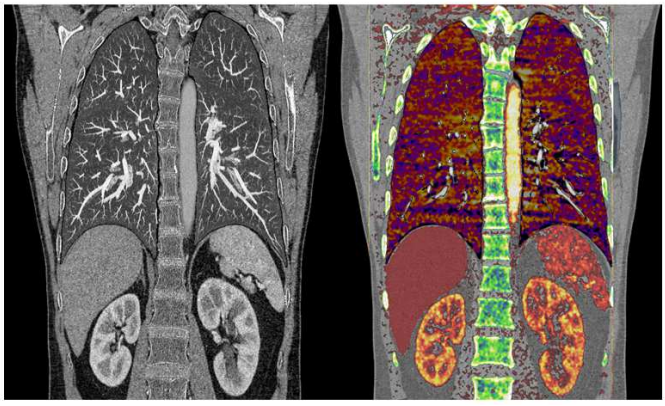


Fig. 1. Left: context–sensitive resolution and display. Right: context–sensitive dual energy evaluation using conventional DE applications [3], [4].

S. Dorn, J. Maier, Dr. S. Sawall and Prof. Dr. M. Kachelrieß are with the German Cancer Research Center (DKFZ), Heidelberg, Germany and the Ruprecht–Karls–University Heidelberg, Heidelberg, Germany. Corresponding author e-mail: sabrina.dorn@dkfz.de.

Dr. M. Knaup is with the German Cancer Research Center (DKFZ), Heidelberg, Germany.

S. Chen and Prof. Dr. A. Maier are with the Pattern Recognition Lab, Friedrich–Alexander University Erlangen–Nuremberg, Erlangen.

Prof. Dr. M. Lell is with the Klinikum Nuremberg and with the Paracelsus Medical University.

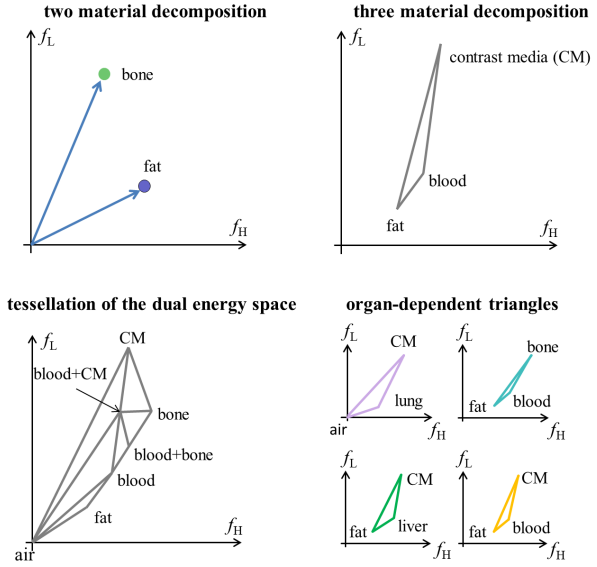


Fig. 2. Schematic representation of the dual energy diagram, where the CT–values of the high energy data are plotted against the CT–values of the low energy data (upper left). Each voxel falling into one triangle can be decomposed into its thee basis material fractions (upper right). By using a tessellation of triangles that is constructed by a set of basis materials, it is possible to decompose the voxel into more than three basis material fractions (lower left) [8]. However, this does not always lead to meaningful material fractions. An organ–dependent selection of the basis materials leads to overlapping triangles. Prior anatomical information is used to resolve the resulting ambiguities in an organ–specific manner (lower right).

sensitive CT image reconstruction, display and analysis [3], [4]. Depending on the a–priori information, the individual settings for each organ are adjusted separately. In this work, we are aiming at transferring the principle of exploiting prior anatomical knowledge to DECT material decomposition. We propose a prior–based multi material decomposition (PBMMD) that consists of multiple organ-specific three material decompositions. We perform an organ–dependent selection of the basis materials leading to overlapping triangles in the DE diagram. Prior anatomical knowledge is obtained by means of an automatic neural network multi–organ segmentation.

II. Materials and Methods

TABLE I
TRIPLET MATERIAL LIBRARY

organ	Basis material library		
	material 1	***material 2***	***material 3***
lung	air	blood	contrast media
liver	fat	liver tissue	contrast media
bone	fat	contrast media	cortical bone
muscles	fat	soft tissue	contrast media
fatty tissue	fat	blood	contrast media
aorta,vasculature	blood	contrast media	cortical bone
spleen, kidneys	fat	blood	contrast media
…	…	…	…

For each DE image pair, there exist two basis material images α_1 and α_2. The decomposition in image domain is given by the linear system of equations

$$\begin{pmatrix} f_{\mathrm{L},1} & f_{\mathrm{L},2} \\ f_{\mathrm{H},1} & f_{\mathrm{H},2} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \begin{pmatrix} f_{\mathrm{L}} \\ f_{\mathrm{H}} \end{pmatrix}.$$

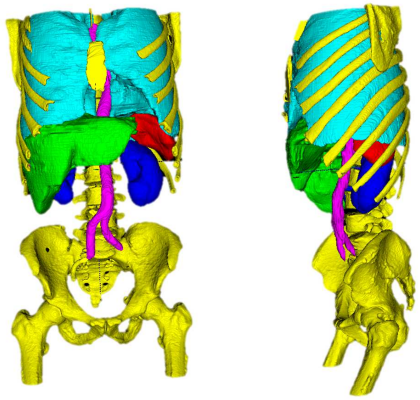


Fig. 3. 3D representation of the automatic segmentation.

In order to decompose the DE data into three basis material components, additional constraints on volume or mass preservation are necessary. We assume that each voxel consists of at most three material constituents and is therefore a compound of three basis materials that are known a priori. We assume that the mixture is volume preserving, meaning that the material composition consisting of M basis material volume fractions α_i satisfies the ideal solution constraint

$$\sum_{i=1}^{M} \alpha_i = 1.$$

Related work [8] pointed out that each reasonable decomposition into more than two basis materials implicitly makes an assumption on mass conservation. Therefore, the assumption results in a realistic approximation of the mixing behavior in the human body. Three material decomposition assuming volume preservation is given as follows:

$$\begin{pmatrix} f_{\mathrm{L},1} & f_{\mathrm{L},2} & f_{\mathrm{L},3} \\ f_{\mathrm{H},1} & f_{\mathrm{H},2} & f_{\mathrm{H},3} \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} f_{\mathrm{L}} \\ f_{\mathrm{H}} \\ 1 \end{pmatrix},$$

subject to $\alpha_i \geq 0,\ i = 1,2,3$. This positivity constraint must be fulfilled for all volume fractions to provide physically meaningful results. As shown in Fig. 2 (upper right), the three material decomposition can be interpreted as a triangle in the DE diagram. Each point that is located inside one specific triangle can be decomposed into its basis material fractions. In order to decompose the DE data into more than three basis material components, we propose a prior–based multi material decomposition (PBMMD). Instead of using a complete tessellation [8] of the DE space (see Fig. 2 lower left), the basis materials are selected based on prior anatomical knowledge. The prior knowledge is gained from an automatic multi–organ segmentation using an cascaded fully convolutional neural network [1] consisting of two U–Nets [10]. The first U–Net is trained to detect the abdominal cavity in order to reduce the search space for the automatic segmentation. The second U–Net is then applied to segment the organs liver, kidneys, spleen, aorta, lung and bone. The remaining yet unlabeled voxels in the background of the organs are further classified into the four tissue types muscle, fat, vasculature and air inclusions

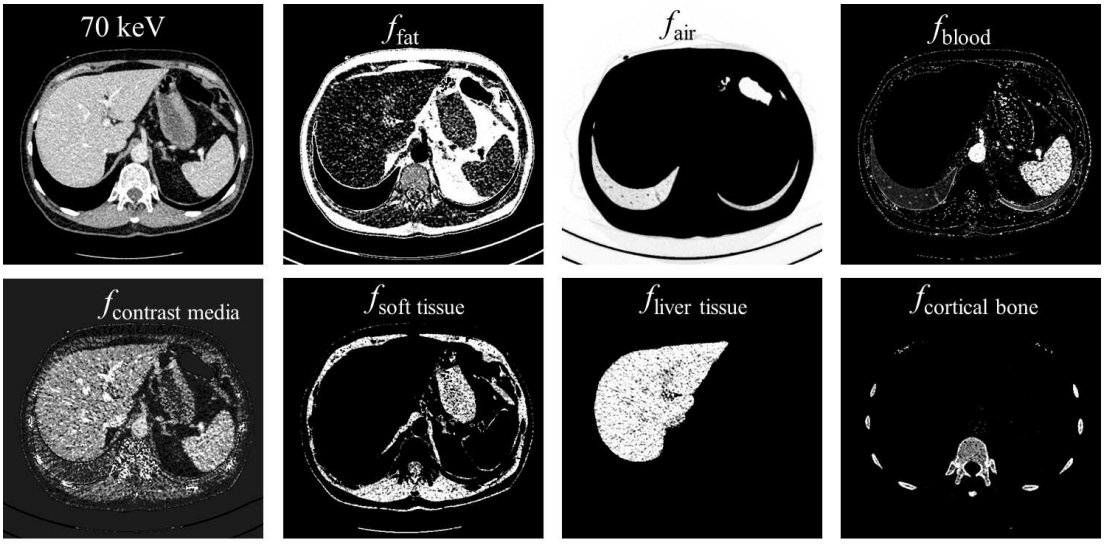


Fig. 4. Decomposition of an abdominal DECT data set into the seven constituent materials. From left to right (first row): monoenergetic reconstruction at 70 keV, volume fraction fat, air, blood. From left to right (second row): volume fraction contrast media (CM), soft tissue, liver tissue and cortical bone. Except for CM the volume fractions are displayed in the range between 0 and 1. In order to visualize the iodine distribution, a thinner range from 0 to 0.02 is chosen.

using a simple thresholding approach. Depending on the organ or tissue assignment, the basis material triplet is selected for the decomposition (see Fig. 2 lower right). Possible basis material triplets are listed in TABLE I. Hence, we assign one basis material triplet to each voxel. The organ–dependent assignment allows for potentially overlapping triangles in the DE space. For instance, the triangle that is assigned to bone overlaps with the triangle that is assigned to the liver. Each voxel that is assigned to one triangle has to be located inside the triangle in order to yield meaningful volume fractions. However, due to the local noise distribution, some voxels could be outside the triangle. In order to cope with these voxels, they are projected onto the closest edge of the triangle [9]. Instead of only considering the measured CT–values at two different energies, we incorporate prior anatomical knowledge to guide the decomposition process. Besides presenting more and more material images to the radiologist, the volume fractions of the material compound can be rescaled to organ–dependent material scores (fat content of liver, iodine content of liver, soft tissue content of liver, absolute or relative) that can be visualized via a reasonable visualization, e.g. pie charts, and later correlated with different diagnoses.

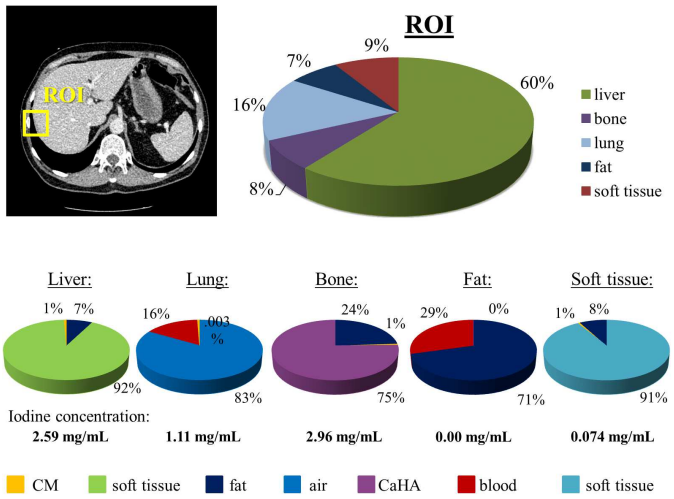


Fig. 5. Pie chart visualization of the decomposition results. Instead of presenting evermore material images to the radiologist, the material composition is visualized as pie chart. An arbitrarily shaped ROI is evaluated and decomposed into its material components. The exact delineation of one specific anatomical structure is not necessary.

III. Results and Discussion

Computed tomography patient data of the thorax and abdomen were acquired using a dual source CT system (Definition Force, Siemens Healthineers, Forchheim, Germany) in the dual energy mode. The x–ray tube voltages were set to 70 and 150 kV, respectively. Ultravist 370 was administered as contrast agent with body weight adapted volumes. The study was performed for contrast–enhanced DECT patient data sets in the arterial and in the portal venous phase. The final segmentation result is shown in Fig. 3. Since the segmentation accuracy is not yet completely sufficient in some cases, we need to manually refine the organ boundaries. The used basis materials are blood, contrast media (Ultravist 370 – iodine contrast agent containing 370 mg iodine/mL), cortical bone (CaHA with 250 mg/mL), fat, soft tissue, liver tissue and air as tabulated in the literature [2]. The basis materials liver tissue, soft tissue and blood differ by their relative density during the simulation using the Evaluated Photon Data Library (EPDL97) [2]. We also allow for mixtures of cortical bone with contrast media and contrast media with fat, since in a clinical case the contrast media is also diluted into these anatomical structures. The triplet material library used in our experiments (shown in TABLE I) consists of seven varying basis material combinations. Fig. 4 shows the monoenergetic reconstruction and decomposition of an abdominal DECT data set into its seven basis material volume fractions using the proposed PBMMD. Since the basis material CM is a dilution

consisting of blood and iodine, the volume fraction of contrast media is much smaller. Hence, we present the volume fraction within another window. However, the volume fractions of

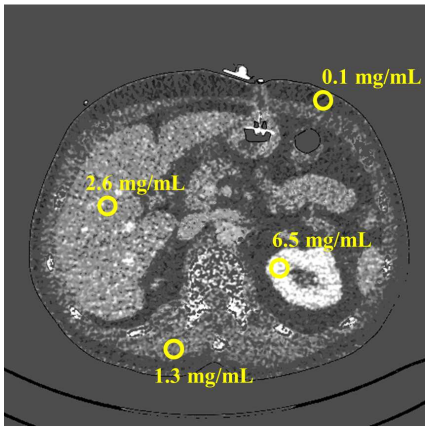


Fig. 6. Quantification of iodine concentration ($C = 5$ mg/mL, $W = 10$ mg/mL).

the contrast media can easily be converted to an iodine concentration. The location–dependent mass concentration of any material is given by

$$\gamma(\boldsymbol{x}) = \frac{m}{V} = f_{\text{basis}} \cdot \rho_{\text{basis}}.$$

The mass concentration of iodine in each voxel is therefore $\gamma_{\text{iodine}} = f_{\text{CM}} \cdot 370$ mg/mL. Fig. 6 illustrates the distribution of the mass concentration of iodine in each voxel. By evaluating small regions of interest, we obtain a quantitative measure of the average values of absolute iodine concentrations. The value of the bone mineral density, i.e. the concentration of pure CaHA cannot be converted as easily, since the real concentration in the human body is not known exactly. In our experiments, we calibrated the value with an arbitrarily chosen value. Besides the quantification of iodine content (fat content etc.) the proposed method is also able to detect calcifications in the vascular system (see Fig. 7).

IV. Conclusion and Outlook

In this work, we have demonstrated that PBMMD is able to decompose patient data into more than three basis materials. Compared to the tessellation approach, where the assignment of one DE data point to one triangle is restricted to the location of the voxel in the DE space leading to ambiguities in the decomposition, our novel approach shows robust decomposition results. The number of possible material triplets is not restricted to a specific number and can be extended on demand. Exploiting the prior anatomical knowledge, each organ or tissue type is exactly assigned to one decomposition triangle. The materials are therefore decomposed into its material compounds according to the anatomical structure they belong to. Furthermore, PBMMD enables the quantification of the iodine content in each voxel and facilitates the detection of small calcifications in the vascular system. While we showed that the decomposition into more than three materials is possible, a quantitative evaluation of the material compound, particularly for the quantification of the bone mineral density (BMD),

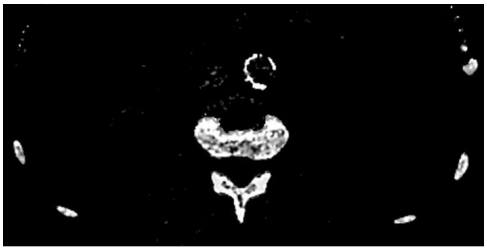


Fig. 7. Calcifications of the aorta are also apparent in the volume fraction of the cortical bone.

is still needed. Future work includes to perform a patient-specific calibration of the basis materials by means of the automatic segmentation. The method yields volume fractions that bear the potential to improve the material quantification for diagnoses, particularly to evaluate the vascularization of tumors.

Acknowledgment

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) under grant KA 1678/20-1, LE 2763/2-1 and MA 4898/5-1. Parts of the reconstruction software were provided by RayConStruct® GmbH, Nürnberg, Germany.

References

[1] S. Chen, H. Roth, S. Dorn, M. May, A. Cavallaro, M. Lell, M. Kachelrieß, H. Oda, K. Mori, and A. Maier. Towards automatic abdominal multi-organ segmentation in dual energy CT using cascaded 3D fully convolutional network. *CoRR*, 2017.

[2] D. Cullen, S. Perkins, and J. Rathkopf. The 1989 livermore evaluated photon data library (EPDL). Technical report, March 1990.

[3] S. Dorn, S. Chen, F. Pisana, J. Maier, M. Knaup, S. Sawall, A. Maier, M. Lell, and M. Kachelrieß. Organ–specific context–sensitive single and dual energy CT (DECT) image reconstruction, display and analysis. In *103rd Scientific Assembly and Annual Meeting of the Radiological Society of North America (RSNA), November 2017*, 2017.

[4] S. Dorn, S. Chen, S. Sawall, D. Simons, M. May, J. Maier, M. Knaup, H.-P. Schlemmer, A. Maier, M. Lell, and M. Kachelrieß. Organ–specific context–sensitive CT image reconstruction and display. In *Proc. SPIE 10573, Medical Imaging 2018: Physics of Medical Imaging*, 2018.

[5] S. Kuchenbecker, S. Faby, D. Simons, M. Knaup, H.-P. Schlemmer, and M. Kachelrieß. Segmentation-assisted material decomposition in dual energy computed tomography (DECT). In *101st Scientific Assembly and Annual Meeting of the Radiological Society of North America (RSNA), November 2015*, 2015.

[6] X. Liu, L. Yu, A. N. Primak, and C. H. McCollough. Quantitative imaging of element composition and mass fraction using dual-energy CT: Three-material decomposition. *Med. Phys.*, 36(5):1602–1609, 2009.

[7] Y. Long and A. Fessler, Jeffrey. Multi-material decomposition using statistical image reconstruction for spectral CT. *IEEE Transactions on Medical Imaging*, 33(8):1614–1626, 2014.

[8] P. R. S. Mendonça, P. Lamb, and D. V. Sahani. A flexible method for multi-material decomposition of dual-energy CT images. *IEEE Transactions on Medical Imaging*, 33(1):99–116, 2014.

[9] T. P. O'Donnell, D. P. Cormode, A. Halaweish, R. Cheheltani, Z. Fayad, and V. Mani. Material decomposition in an arbitrary number of dimensions using noise compensating projection. *Biomedical Physics and Engineering Express*, 2017.

[10] O. Ronneberger, P. Fischer, and T. Brox. U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer–Assisted Interventions (MICCAI)*, volume abs/1505.04597 of *LNCS*, pages 234–241. Springer, 2015. available on arXiv:1505.04597 [cs.CV].

[11] R. Zhang, J.-B. Thibault, C. A. Bouman, and K. D. Sauer. Soft classification with gaussian mixture model for clinical dual–energy CT reconstructions. In *Proceedings of the 12th International Meeting on Fully 3D Image Reconstruction in Radiology and Nuclear Medicine*, pages 408–411, 2013.

Divergent-beam backprojection-filtration formula with applications to region-of-interest imaging

Aymeric Reshef, Cyril Riddell, Yves Trousset, Saïd Ladjal, and Isabelle Bloch

***Abstract*—We propose a new backprojection-filtration (BPF) method for cone-beam computed tomography (CBCT) with flat-panel detectors over circular orbits. The method is exact in the fan-beam geometry and provides an approximate CBCT reconstruction that is different from the standard Feldkamp-Davis-Kress (FDK) method. More interestingly, it can be used for region-of-interest (ROI) reconstruction by complementing a truncated low-noise acquisition with dense angular sampling by additional non-truncated views that are either high-noise or angularly undersampled.**

I. Introduction

Filtered backprojection (FBP) performs poorly when projections are truncated, unless data extrapolation is performed prior to filtering. Alternative direct reconstruction methods were derived to address the issue of ROI reconstruction [1]. However, when not in parallel-beam geometry, they require a dense angular sampling of the projections.

In the case of interior tomography, it was shown that no unique solution could be obtained from the truncated projections only. However, a tiny additional information, such as a prior knowledge on the image itself [2] or a few additional untruncated measurements [3] is enough to stabilize the problem. Unfortunately, no closed-form analytical solution exists and iterative reconstruction has been used instead.

In this work, we propose a new backprojection-filtration (BPF) formula in cone-beam geometries with flat-panel detectors and circular orbits, which is used to design a direct reconstruction method for ROI imaging. The formula is derived in Section II. The application of the method to ROI imaging is described in Section III. Experiments are presented in Section IV and results are shown in Section V.

II. Method

A. Cone-beam geometry

Let $\Theta = [0, 2\pi]$. For $\theta \in \Theta$, we write $\boldsymbol{\theta} = (\cos\theta, \sin\theta, 0)^T$ and $\boldsymbol{\theta}^\perp = (-\sin\theta, \cos\theta, 0)^T$. The X-ray source is located at point $\boldsymbol{\xi}_\theta = -d\boldsymbol{\theta}$, where d is the source-to-rotation-axis distance. The detector is located at a distance D from the X-ray source. It is a plane orthogonal to the line passing through the center of rotation and the source point $\boldsymbol{\xi}_\theta$. A 3D point $\boldsymbol{x} \in \mathbb{R}^3$ projects onto the detector plane at coordinates $(u_\theta(\boldsymbol{x}), v_\theta(\boldsymbol{x}))$; without loss of generality, we write, for $\alpha \in \Theta$:

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x} \cdot \boldsymbol{\alpha} \\ \boldsymbol{x}_{\alpha^\perp} \end{pmatrix}, \tag{1}$$

where $\boldsymbol{x}_{\alpha^\perp} \in \mathbb{R}^2$ consists of the coordinates of $\boldsymbol{x}$ in the plane of equation $\boldsymbol{x} \cdot \boldsymbol{\alpha} = 0$. When looking only at points $\boldsymbol{x}$ belonging to a plane of equation: $\boldsymbol{x} \cdot \boldsymbol{\alpha} = x_\alpha$, where $x_\alpha \in \mathbb{R}$, the relationship between $(u_\theta(\boldsymbol{x}), v_\theta(\boldsymbol{x}))$ and $\boldsymbol{x}$ is given by:

$$\begin{pmatrix} s_\theta(\boldsymbol{x}) u_\theta(\boldsymbol{x}) \\ s_\theta(\boldsymbol{x}) v_\theta(\boldsymbol{x}) \\ s_\theta(\boldsymbol{x}) \end{pmatrix} = H_\theta^\alpha(x_\alpha) \begin{pmatrix} \boldsymbol{x}_{\alpha^\perp} \\ 1 \end{pmatrix}, \tag{2}$$

where the matrix $H_\theta^\alpha(x_\alpha) \in \mathbb{R}^{3\times 3}$ is a homography matrix.

The cone-beam projection of an image f at angle θ is denoted p_θ. It is defined at each detector coordinate (u, v) as the integral of f along the line joining $\boldsymbol{\xi}_\theta$ to (u, v). The full-scan, cone-beam tomographic acquisition over a circular orbit is the collection $p = \{p_\theta \mid \theta \in \Theta\}$. We define the backprojection from angle θ of a single projection p_θ as: $\mathcal{B}_\theta[p_\theta] = p_\theta(u_\theta, v_\theta)$.

B. Feldkamp-Davis-Kress reconstruction

The Tuy conditions [4] are not satisfied in the cone-beam geometry with a circular orbit. Hence, only approximate direct reconstruction methods exist, such as the Feldkamp-Davis-Kress method [5], which is a direct extension of the fan-beam FBP to cone-beam data. Given a full-scan tomographic acquisition p, FDK reconstructs an image f_{FDK} as:

$$f_{\mathrm{FDK}} = \int_0^{2\pi} \frac{D^2}{s_\theta^2} \mathcal{B}_\theta \mathcal{D}[\widetilde{p}_\theta] \mathrm{d}\theta, \tag{3}$$

where $\mathcal{D}$ is the ramp filter, and:

$$\widetilde{p}_\theta(u, v) = \frac{1}{2} \cdot \frac{d}{D} \cdot \frac{D}{\sqrt{D^2 + u^2 + v^2}} \cdot p_\theta(u, v). \tag{4}$$

By design, FDK is equal to FBP when $z = 0$, yielding an exact reconstruction in the midplane. If the true image f is invariant along the z-axis, the reconstructed image f_{FDK} is exact [5]; otherwise, it deviates from f as the cone angle increases, yielding cone-beam artifacts.

In the following, we define $\{\Theta_k\}_{k=1\cdots K}$ as a subdivision of Θ. Then by linearity of the integral:

$$f_{\mathrm{FDK}} = \sum_{k=1}^{K} g_{\Theta_k}, \text{ where } g_{\Theta_k} = \int_{\Theta_k} \frac{D^2}{s_\theta^2} \mathcal{B}_\theta \mathcal{D}[\widetilde{p}_\theta] \mathrm{d}\theta. \tag{5}$$

A. Reshef (corresponding author: aymeric.reshef@ge.com) is with GE Healthcare, Buc, France, and LTCI, Télécom ParisTech, Université Paris-Saclay, Paris, France. C. Riddell and Y. Trousset are with GE Healthcare, Buc, France. S. Ladjal and I. Bloch are with LTCI, Télécom ParisTech, Université Paris-Saclay, France. This work was supported by the CIFRE grant No. 873/2014 from the French Association Nationale de la Recherche et de la Technologie (ANRT).

$\phi(t)$	$\mathcal{H}[\phi](t)$
$\mathcal{H}[\psi](t)$	$-\psi(t)$
$t\psi(t)$	$t\mathcal{H}[\psi](t) - \frac{1}{\pi}\int_{-\infty}^{+\infty}\psi(t')\mathrm{d}t'$
$\delta(t)$	$\frac{1}{\pi t}$
$\frac{1}{h_3 t+h_4}\psi\left(\frac{h_1 t+h_2}{h_3 t+h_4}\right)$, where $h_1h_4 - h_2h_3 = \pm 1$	$\frac{\mathrm{sgn}(h_1h_4-h_2h_3)}{h_3t+h_4}\mathcal{H}[\psi]\left(\frac{h_1t+h_2}{h_3t+h_4}\right)$

Table I: Useful Hilbert transforms to prove Eq. (7).

C. Proposed backprojection-filtration method

The proposed backprojection-filtration (BPF) formula is first derived in the fan-beam geometry. We rely on the decomposition of the ramp filter $\mathcal{D}$ into a spatial derivative operator and a Hilbert transform operator $\mathcal{H}$:

$$\mathcal{D}[\tilde{p}_\theta] = \frac{1}{2\pi}\mathcal{H}\left[\frac{\partial \tilde{p}_\theta}{\partial u}\right]. \tag{6}$$

When $\Theta_k = \{\theta\}$, we write $g_\theta = g_{\{\theta\}}$. Let $\alpha \in \Theta$ such that for any point $\boldsymbol{x}$ in the field of view, $\det\left(H_\theta^\alpha(\boldsymbol{x}\cdot\boldsymbol{\alpha})\right) \neq 0$. We define $\sigma_\theta^\alpha(\boldsymbol{x}) = \mathrm{sgn}\left(\det\left(H_\theta^\alpha(\boldsymbol{x}\cdot\boldsymbol{\alpha})\right)\right)$. The following holds:

$$g_\theta = \frac{\sigma_\theta^\alpha}{2\pi}\mathcal{H}_\alpha[b_\theta], \text{ where } b_\theta = \frac{D^2}{s_\theta^2}\mathcal{B}_\theta\left[\frac{\partial\tilde{p}_\theta}{\partial u}\right], \tag{7}$$

and $\mathcal{H}_\alpha$ applies the one-dimensional Hilbert transform to each line of the 2D plane that is colinear to $\boldsymbol{\alpha}^\perp$. We sketch the proof of Eq. (7). The homography matrix $H_\theta^\alpha(\boldsymbol{x}\cdot\boldsymbol{\alpha})$ is such that $\det\left(H_\theta^\alpha(\boldsymbol{x}\cdot\boldsymbol{\alpha})\right) \neq 0$, hence the determinant of the matrix $\left|\det H_\theta^\alpha(\boldsymbol{x}\cdot\boldsymbol{\alpha})\right|^{-1/2} H_\theta^\alpha(\boldsymbol{x}\cdot\boldsymbol{\alpha})$ is equal to σ_θ^α. We use the last row of Table I to obtain the intermediate result: $s_\theta g_\theta = \frac{\sigma_\theta^\alpha}{2\pi}\mathcal{H}_\alpha\left[\frac{D^2}{s_\theta}\mathcal{B}_\theta\left[\frac{\partial\tilde{p}_\theta}{\partial u}\right]\right]$. We then apply again $\mathcal{H}_\alpha$ to each side of this equality and observe that s_θ is an affine function of $\boldsymbol{x}$. Using the other properties of the Hilbert transform recalled in Table I, we obtain Eq. (7).

If $\Theta_k = [\theta_{k-1}, \theta_k]$ and $|\theta_k - \theta_{k-1}| < \pi - \gamma$, where γ is the fan angle, we can find a common admissible α value, denoted α_k, such that Eq. (7) holds for all $\theta \in \Theta_k$; hence:

$$g_{\Theta_k} = \frac{1}{2\pi}\int_{\Theta_k}\mathcal{H}_{\alpha_k}[b_\theta]\,\mathrm{d}\theta = \frac{1}{2\pi}\mathcal{H}_{\alpha_k}[b_{\Theta_k}], \tag{8}$$

where $b_{\Theta_k} = \int_{\Theta_k} b_\theta \mathrm{d}\theta$. We propose to reconstruct an image f_{BPF} as:

$$f_{\mathrm{BPF}} = \frac{1}{2\pi}\sum_{k=1}^{K}\mathcal{H}_{\alpha_k}[b_{\Theta_k}]. \tag{9}$$

As in FDK, the reconstruction formula from Eq. (9) is always exact in the midplane; it is exact everywhere when the true image f is invariant along the z-axis [5].

D. Implementation

The support of the backprojection over each angular subset Θ_k depends both on Θ_k and on the value of α_k. Provided that this support is large enough to compute all the backprojected lines colinear to $\boldsymbol{\alpha}_k^\perp$ and intersecting the image field of view (Fig. 1), the Hilbert transform can be computed directly in the Fourier domain using zero-padding and periodicity.

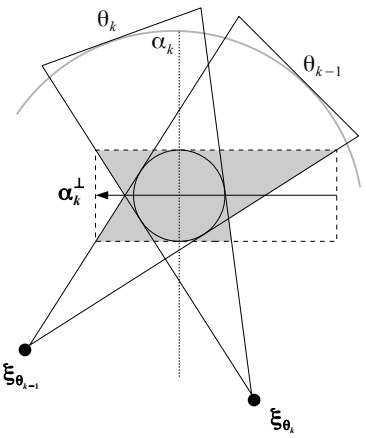


Fig. 1: Extended support (dashed rectangle) for filtering b_{Θ_k}. Lines colinear to $\boldsymbol{\alpha}_k^\perp$ (arrow) crossing the circular field of view (black circle) are all compactly supported (shaded area).

We divide Θ into frontal and lateral views. Frontal views correspond to $\Theta_{\mathrm{FRT}} = \left[\frac{\pi}{4}, \frac{3\pi}{4}\right] \cup \left[\frac{5\pi}{4}, \frac{7\pi}{4}\right]$. With $\alpha_{\mathrm{FRT}} = \frac{\pi}{2}$, image $b_{\Theta_{\mathrm{FRT}}}$ is filtered horizontally. Lateral views correspond to $\Theta_{\mathrm{LAT}} = \Theta \backslash \Theta_{\mathrm{FRT}}$. With $\alpha_{\mathrm{LAT}} = 0$, image $b_{\Theta_{\mathrm{LAT}}}$ is filtered vertically. Thus:

$$f_{\mathrm{BPF}} = \frac{1}{2\pi}\left(\mathcal{H}_{\frac{\pi}{2}}[b_{\Theta_{\mathrm{FRT}}}] + \mathcal{H}_0[b_{\Theta_{\mathrm{LAT}}}]\right). \tag{10}$$

III. Application to region-of-interest imaging

Following [3], we define a dual-rotation acquisition as follows. Let p_{T} be a set of truncated projections that finely sample the source-detector orbit. Projections p_{T} are complemented by additional un-truncated projections p_{F}. We do not intend to bring too much additional dose to the patient with the un-truncated projections, by either lowering the dose level per view or by reducing the number of views in p_{F}.

Since local operations are applied to projections p_{T} prior to backprojection, they correctly sample the unfiltered backprojected image b_{T} within the ROI, denoted Ω'. Outside the ROI, however, each backprojected point is observed over a limited angular range, which differs from one point to the other. We thus merge the ROI of b_{T} with the unfiltered backprojected image b_{F} obtained from p_{F}, yielding image $\mathcal{M}(b_{\mathrm{F}}, b_{\mathrm{T}})$ such that:

$$\mathcal{M}(b_{\mathrm{F}}, b_{\mathrm{T}}) = \begin{cases} \eta\cdot b_{\mathrm{F}} + (1-\eta)\cdot b_{\mathrm{T}} & \text{inside } \Omega'; \\ b_{\mathrm{F}} & \text{outside } \Omega'. \end{cases} \tag{11}$$

The function $\eta : \mathbb{R}^3 \to [0,1]$ is a continuous function that ensures a smooth transition from b_{T} to b_{F} at the boundaries of the ROI. The Fourier-based filtering step is then performed on the hybrid image $\mathcal{M}(b_{\mathrm{F}}, b_{\mathrm{T}})$. Using our BPF method, the merging step is performed separately for backprojections of the frontal views and of the lateral views.

IV. Simulations

All images were reconstructed on a $256 \times 256 \times 256$ grid with isotropic voxels of size 1.17 mm^3.

A. Full-volume reconstruction

A diagnostic CT scan of a head was forward-projected over an ideal circular orbit using $D = d = 1180$ mm. A total of 1440 projections sampling Θ was generated. The projections were reconstructed using FDK (yielding image f_{FDK}) and Eq. (10) (yielding image f_{BPF}). We computed the mean relative error (MRE) over a mask Ω_0, denoted $\Delta_{\Omega_0}(f_{\mathrm{BPF}}, f_{\mathrm{FDK}})$, using the formula:

$$\Delta_{\Omega}(f, f^*) = \frac{1}{\mathrm{Card}(\Omega)} \sum_{\boldsymbol{x}\in\Omega} \frac{|f(\boldsymbol{x}) - f^*(\boldsymbol{x})|}{|f^*(\boldsymbol{x})|} \qquad (12)$$

Mask Ω_0 was defined in order to keep only the voxels higher than -250 HU.

We repeated the experiment using modified projection data corresponding to $1.6 \cdot 10^6$ photons per ray emitted from the X-ray source, in order to check the stability of the method with respect to noise, yielding images $f_{\mathrm{FDK}}^{\mathrm{noisy}}$ and $f_{\mathrm{BPF}}^{\mathrm{noisy}}$.

B. Region-of-interest reconstruction

The truncated projections p_{T} were simulated by applying a digital transaxial truncation to the previous set of 1440 noisy projections, corresponding to a cylindrical, centered field of view Ω' whose edges cross the head skull. It is thus expected that empirical projection extrapolation methods would not perform as well. Such a reconstruction was computed using [6], yielding image $f_{\mathrm{FDK}}^{\mathrm{ROI}}$. For the un-truncated projections p_{F}, we simulated two configurations. In the first configuration, we simulated an acquisition of 1440 projections corresponding to 10^5 photons per ray emitted from the X-ray source, yielding image $f_{\mathrm{BPF}}^{\mathrm{ROI}(1)}$. In the second configuration, we simulated an acquisition of 90 projections corresponding to $1.6 \cdot 10^6$ photons per ray, yielding image $f_{\mathrm{BPF}}^{\mathrm{ROI}(2)}$. In both cases, the dose ratio between the un-truncated and the truncated acquisitions is fixed to 1/16. The merging step was performed using the following weighting function:

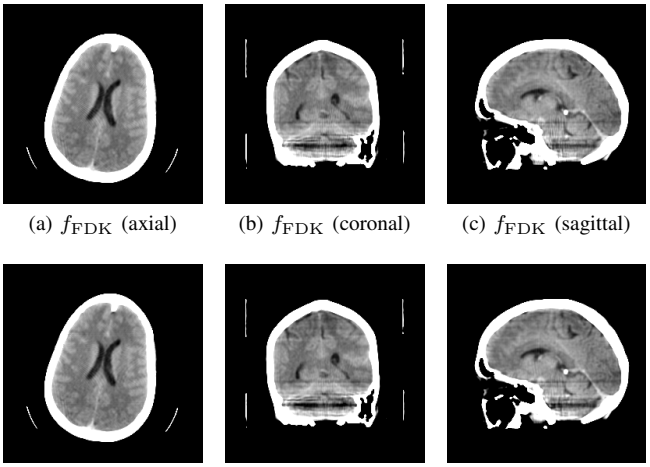
$$\eta(\boldsymbol{x}) = \frac{1}{2}\left(1 - \cos\left(\pi \cdot \frac{|\boldsymbol{x}| - r_{\Omega'}}{\Delta r}\right)\right), \qquad (13)$$

where $r_{\Omega'}$ denotes the radius of the cylindrical ROI Ω', and Δr is the transition zone radial width. In the following, Δr was arbitrarily set to 5 voxels. The MRE over the intersection set $\Omega = \Omega' \cap \Omega_0$ was computed with respect to the un-truncated FDK reconstruction $f_{\mathrm{FDK}}^{\mathrm{noisy}}$.

V. Results

A. Full-volume reconstruction

Noise-free reconstructed images are shown in Fig. 2. The images f_{FDK} and f_{BPF} are visually very similar. Both reconstructions are exact and identical in the fan-beam geometry of the midplane. However, f_{BPF} is more sensitive to the cone-beam incomplete sampling over a circular orbit (see the dark streaks near the temporal bones in the coronal and sagittal slices). Similar noise behavior occurs for both methods when reconstructing from noisy projections (images not shown). On average, the MRE inside Ω_0 is equal to 0.42% in the noise-free case and to 0.43% in the noisy case (Table II), the higher errors being located towards points with high cone angles.



(a) f_{FDK} (axial) (b) f_{FDK} (coronal) (c) f_{FDK} (sagittal)

(d) f_{BPF} (axial) (e) f_{BPF} (coronal) (f) f_{BPF} (sagittal)

Fig. 2: Noise-free, full-volume reconstructions. Display window: [10 HU, 60 HU].

	Full-volume		ROI	
Ω	Ω_0		$\Omega' \cap \Omega_0$	
f	f_{BPF}	$f_{\mathrm{BPF}}^{\mathrm{noisy}}$	$f_{\mathrm{BPF}}^{\mathrm{ROI}(1)}$	$f_{\mathrm{BPF}}^{\mathrm{ROI}(2)}$
f^*	f_{FDK}	$f_{\mathrm{FDK}}^{\mathrm{noisy}}$	$f_{\mathrm{FDK}}^{\mathrm{noisy}}$	$f_{\mathrm{FDK}}^{\mathrm{noisy}}$
$\Delta_{\Omega}(f, f^*)$	0.42%	0.43%	0.44%	0.50%

Table II: Mean relative errors in region Ω.

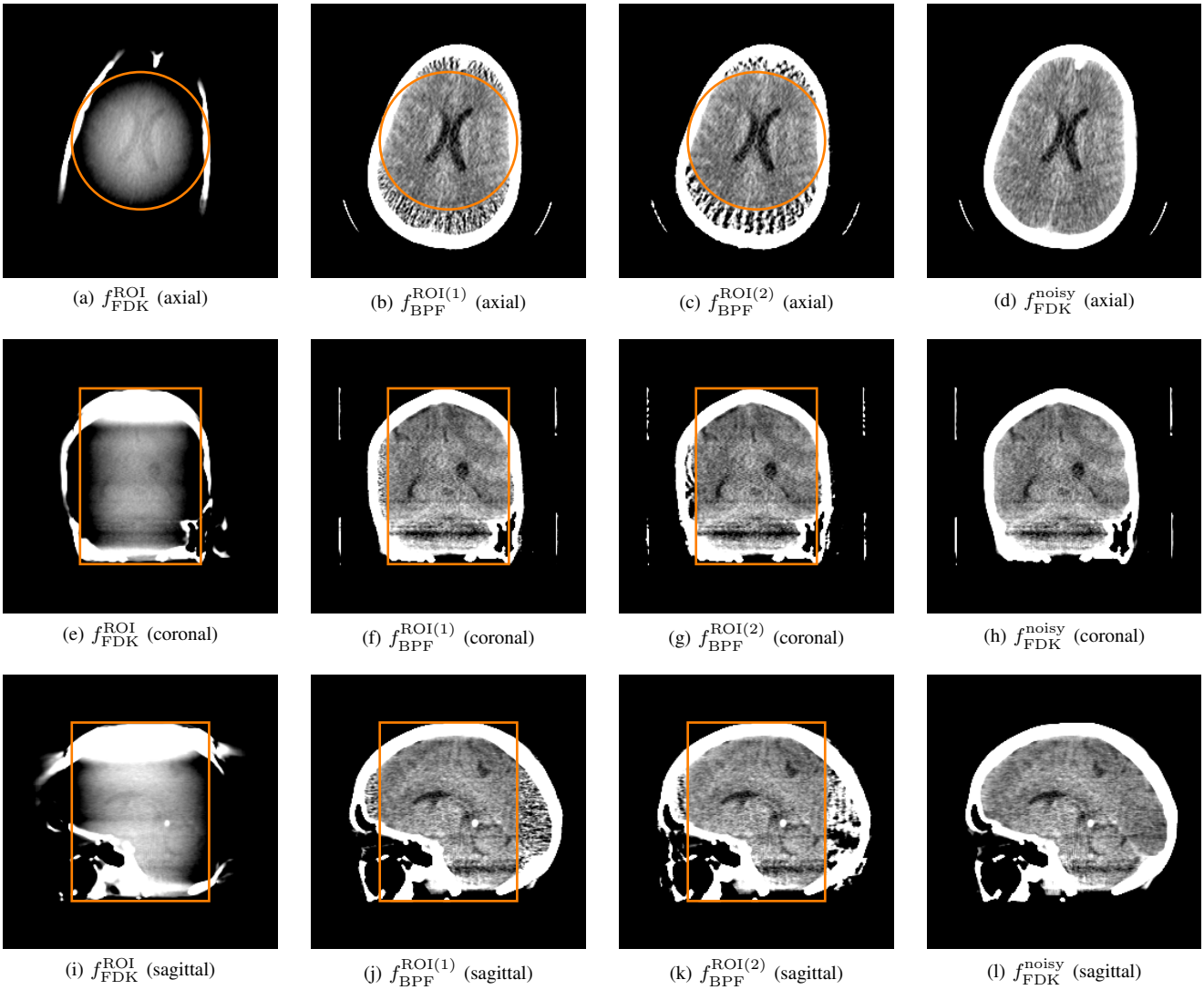
B. Region-of-interest reconstruction

Results of ROI reconstruction are shown in Fig. 3. The first column shows the FDK reconstruction from the truncated projections only using empirical projection extrapolation. As expected, such extrapolation cannot perform well when highly contrasted structures such as bones lie at the edge of the field of view. The image $f_{\mathrm{FDK}}^{\mathrm{ROI}}$ suffers from a shift in gray values and from low-frequency non-uniformities that prevent from using a narrow window display.

Results from our reconstruction method are shown in the second and third columns of Fig. 3. Both configurations yield images that are visually similar to the reference FDK reconstruction $f_{\mathrm{FDK}}^{\mathrm{noisy}}$ (fourth column) inside the ROI Ω'. Outside the ROI, image $f_{\mathrm{BPF}}^{\mathrm{ROI}(1)}$ shows a very noisy reconstruction of the head, while image $f_{\mathrm{BPF}}^{\mathrm{ROI}(2)}$ shows streaks characteristic of angular subsampling. However, neither the high noise contained in p_{F} in the first configuration, nor the subsampling streaks of the second configuration propagate inside Ω'. The values of the MRE inside region $\Omega = \Omega' \cap \Omega_0$ with respect to $f_{\mathrm{FDK}}^{\mathrm{noisy}}$ remain below 1%, at 0.44% for $f_{\mathrm{BPF}}^{\mathrm{ROI}(1)}$ and 0.50% for $f_{\mathrm{BPF}}^{\mathrm{ROI}(2)}$ (Table II).

VI. Discussion

A new BPF formula was described for CBCT reconstruction with flat-panel detectors, that is exact in the fan-beam geometry and provides a different approximate reconstruction from FDK in the cone-beam geometry. It coincides with the parallel-beam Hilbert-transformed differentiated



(a) $f_{\mathrm{FDK}}^{\mathrm{ROI}}$ (axial) (b) $f_{\mathrm{BPF}}^{\mathrm{ROI}(1)}$ (axial) (c) $f_{\mathrm{BPF}}^{\mathrm{ROI}(2)}$ (axial) (d) $f_{\mathrm{FDK}}^{\mathrm{noisy}}$ (axial)

(e) $f_{\mathrm{FDK}}^{\mathrm{ROI}}$ (coronal) (f) $f_{\mathrm{BPF}}^{\mathrm{ROI}(1)}$ (coronal) (g) $f_{\mathrm{BPF}}^{\mathrm{ROI}(2)}$ (coronal) (h) $f_{\mathrm{FDK}}^{\mathrm{noisy}}$ (coronal)

(i) $f_{\mathrm{FDK}}^{\mathrm{ROI}}$ (sagittal) (j) $f_{\mathrm{BPF}}^{\mathrm{ROI}(1)}$ (sagittal) (k) $f_{\mathrm{BPF}}^{\mathrm{ROI}(2)}$ (sagittal) (l) $f_{\mathrm{FDK}}^{\mathrm{noisy}}$ (sagittal)

Fig. 3: ROI reconstruction. The ROI Ω' is delineated in orange. Display windows: [-450 HU, -250 HU] (first column), [10 HU, 60 HU] (second to fourth columns).

backprojection method (DBP-HT) [1], when letting $(d, D) \to (+\infty, +\infty)$ and $K = 1$: in this case, image b_Θ becomes the parallel-beam DBP image, and filtration needs to be performed using finite Hilbert transform inversion. However, our approach differs from the fan-beam DBP-HT formula of [1]. In the fan-beam DBP-HT, the same DBP image is computed from fan-beam projections through a parallel-to-fan-beam change of variables, which requires a dense angular sampling. Instead, we propose to compute an alternative, intrinsically fan-beam DBP image, so that the whole backprojection step translates into a view-wise algorithm. Moreover, when $K > 1$, filtration is performed in the Fourier domain and does not require any finite Hilbert transform inversion. The method is thus expected to work as good as FDK with coarser angular sampling; it is also adapted to non-ideal circular geometries using calibrated projection matrices. As with FDK, the reconstructed images suffer from cone-beam artifacts, however, we anticipate faster iterative BPF reconstructions to reduce them [7]. Finally, excellent ROI reconstruction was obtained with only 6% of dose increase and flexible acquisition designs in terms of dose per view and angular sampling.

References

[1] F. Noo, R. Clackdoyle, and J. D. Pack, "A two-step Hilbert transform method for 2D image reconstruction," *Physics in Medicine and Biology*, vol. 49, no. 17, p. 3903, 2004.

[2] H. Kudo, M. Courdurier, F. Noo, and M. Defrise, "Tiny a priori knowledge solves the interior problem in computed tomography," *Physics in Medicine and Biology*, vol. 53, no. 9, p. 2207, 2008.

[3] A. Reshef, C. Riddell, Y. Trousset, S. Ladjal, and I. Bloch, "Dual-rotation C-arm cone-beam computed tomography to increase low-contrast detection," *Medical Physics*, vol. 44, no. 9, pp. e164–e173, 2017.

[4] H. K. Tuy, "An inversion formula for cone-beam reconstruction," *SIAM Journal on Applied Mathematics*, vol. 43, no. 3, pp. 546–552, 1983.

[5] L. Feldkamp, L. Davis, and J. Kress, "Practical cone-beam algorithm," *JOSA A*, vol. 1, no. 6, pp. 612–619, 1984.

[6] J. Hsieh, R. H. Armstrong, P. J. Arduini, and R. F. Senzig, "Methods and apparatus for truncation compensation," 2004, US Patent 6,810,102.

[7] H. Langet, C. Riddell, A. Reshef, Y. Trousset, A. Tenenhaus, E. Lahalle, G. Fleury, and N. Paragios, "Compressed-sensing-based content-driven hierarchical reconstruction: Theory and application to C-arm cone-beam tomography," *Medical Physics*, vol. 42, no. 9, pp. 5222–5237, 2015.

GCC and FBCC for linear tomosynthesis

Jérôme Lesaint, Simon Rit, Rolf Clackdoyle, Laurent Desbat

***Abstract*—Grangeat-based Consistency Conditions (GCC) and Fan-beam Consistency Conditions (FBCC) are two ways to describe consistency (or redundancy) between cone-beam projections. Here we consider cone-beam projections that are collected in the linear tomosynthesis geometry. We propose a theoretical comparison of these two sets of consistency conditions and illustrate the comparison with numerical simulations of a thorax Forbild phantom.**

***Index Terms*—Cone-beam computed tomography (CBCT), data consistency conditions (DCCs), tomosynthesis.**

I. Introduction

In Computed Tomography (CT), the 3D density map of a patient (or an object) is reconstructed from a set of 2D radiographs. Should the acquisition be realized in perfect conditions and neglecting physical side-effects (like scattering or beam-hardening), these radiographs (after their log-transform) follow the forward line-integral model. Unfortunately, these conditions are never fullfilled and systematic effects always degrade the projection data. One way to detect such effects is to quantify how inconsistent the data are, using the concept of data consistency conditions (DCC). DCC are equations that characterize the image of the forward operator. A large amount of research has been published on DCC. In parallel geometries, the well-known Helgason-Ludwig DCC [1], [2] provide necessary and sufficient conditions on the Radon transform. In divergent geometries, [3] and [4] also give necessary and sufficient conditions for a source moving along a linear trajectory (in 2D) or planar trajectory (in 3D) respectively. When a set of DCC is known to be complete (necessary *and* sufficient conditions), no additional information can be expected from another set of DCC. On the other hand, if two sets of DCC are known to be necessary but no information on the sufficiency is available, one may wonder which to use. Some works have introduced necessary-only conditions, like Grangeat-based DCC (GCC) [5], [6] and fan-beam DCC (FBCC) [3] extended to 3D projection data as suggested by [7]. As explained below, FBCC refers to what would be called zeroth-order conditions in [3].

Various applications of GCC and FBCC have been published. Geometric calibration in circular cone-beam CB [5] and in x-ray tomosynthesis [8] were described using GCC, and also jitter-correction in x-ray CT [9]. For FBCC, a circular cone-beam micro CT application appeared recently [10].

In this work, we focus on GCC and FBCC and propose a theoretical comparison of these two sets of DCC. We carry this work in the specific context of tomosynthesis, with an X-ray source moving along a line. It is proved that the FBCC are stronger than the GCC, in the sense that if the FBCC are satisfied, then so are the GCC, but not conversely. We also show that if all the projections are complete (non-truncated) then the FBCC and GCC are equivalent. Furthermore the hypothesis of complete projections is essential; we show that under particular circumstances (with truncated projections), the FBCC are more restrictive, i.e., the FBCC can fail even when the GCC are satisfied. We finally prove that neither of these two sets of conditions are sufficient. The theory is detailed in Section II. Numerical aspects are presented in Section III and finally, Section IV contains discussion and conclusion.

II. Theory

A. The forward X-ray model

We consider X-ray projection data g of an object function f, acquired along a 1D trajectory of the source parametrized by a scalar $\lambda \in \Lambda \subset \mathbb{R}$. The source position is denoted $\vec{s}_\lambda$. A projection g_λ associates to each unit vector $\vec{\alpha} \in \mathcal{S}^2$ the corresponding line integral:

$$g_\lambda(\vec{\alpha}) = \int_0^{+\infty} f(\vec{s}_\lambda + t\vec{\alpha})\,\mathrm{d}t. \tag{1}$$

B. The tomosynthesis geometry

In the following, we consider a tomosynthesis geometry, where an X-ray source moves along a line parallel to the plane of the detector. Let (O, x, y, z) be a 3D coordinate system. Without loss of generality, the detector, denoted $\mathcal{D}$, is assumed to lie in the $z = 0$ plane. It is equipped with 2D coordinates (O, u, v) where the detector origin O and the u- and v- axes coincide with the 3D origin and the x- and y- axes respectively. Note that the coordinates (u, v) are independent of the source position, as if the detector was large enough to capture every projection of the scan. (In the numerical simulations though, the detector was displaced horizontally.) The X-ray source moves along a line $\mathcal{L}$ such that $\vec{s}_\lambda$ has coordinates $(\lambda, 0, d)$ for $\lambda \in \Lambda$, where d is the source-to-detector distance and Λ is an interval. The object of interest is assumed to be entirely contained between $\mathcal{D}$ and the plane parallel to $\mathcal{D}$ and containing $\mathcal{L}$. See Figure 1.

C. The family of planes containing the source trajectory

We now focus on the family of planes which contain the source line. They will play an important role both in the

J. Lesaint, R. Clackdoyle and L. Desbat are with the TIMC-IMAG laboratory, CNRS UMR 5525 and Université Grenoble Alpes (e-mail : jerome.lesaint@univ-grenoble-alpes.fr).

S. Rit is with Univ. Lyon, INSA-Lyon, UCB Lyon 1, UJM-Saint Etienne, CNRS, Inserm, CREATIS UMR5220, U1206, Centre Léon Bérard, F-69373, LYON, France.

This work is partially supported by the Agence Nationale de la Recherche (France), Labex CAMI, number ANR-11-LABX-0004-01, Labex PRIMES, number ANR-11-LABX-0063 and project "DROITE", number ANR-12-BS01-0018.

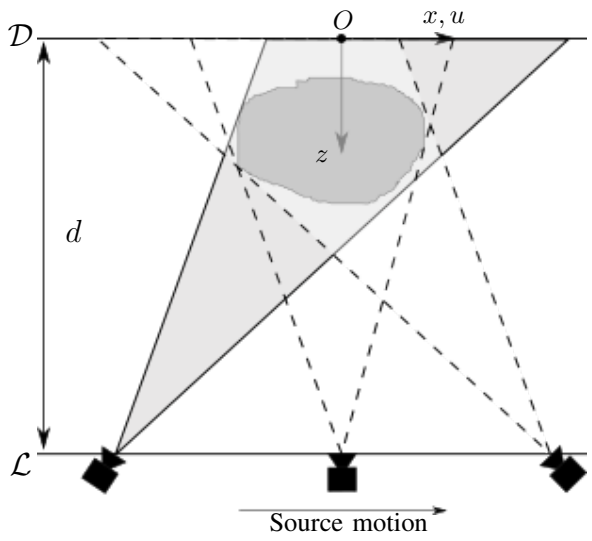


Fig. 1. Top-view of the tomosynthesis acquisition geometry. The X-ray source moves at constant distance d from the detector plane, along a horizontal line.

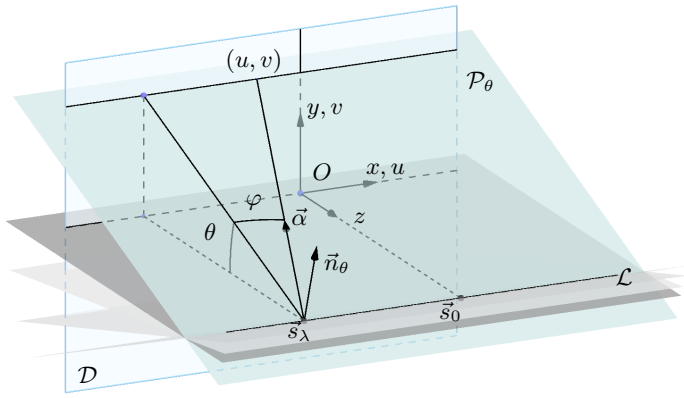


Fig. 2. The family of planes $\mathcal{P}_\theta$. Each plane is defined by a normal vector $\vec{n}_\theta$ which makes an angle θ with the central plane. To each detector pixel (u, v) is associated a unit vector $\vec{\alpha}$ together with its spherical coordinates (θ, φ).

Grangeat case and in the fan-beam case. For $\theta \in \left]-\frac{\pi}{2}; \frac{\pi}{2}\right[$, we let $\mathcal{P}_\theta$ denote the plane which contains the line $\mathcal{L}$ and makes an angle θ with the central plane ($y = 0$). We let $\vec{n}_\theta = (0, \cos\theta, \sin\theta)$ denote the unit vector orthogonal to this plane. There is a one-to-one correspondence between the detector's horizontal rows (with offset v) and the planes $\mathcal{P}_\theta$ via the relation $d\tan\theta = v$. See Figure 2.

We assume that the detector is fixed so from detector coordinates (u, v), the unit vector $\vec{\alpha}$ from Equation 1 is given by $\vec{\alpha} = ((u, v, 0) - \vec{s}_\lambda)/\|(u, v, 0) - \vec{s}_\lambda\| = (u - \lambda, v, -d)/\sqrt{(u-\lambda)^2 + v^2 + d^2}$. Furthermore $\vec{\alpha}$ can be expressed in spherical coordinates $\vec{\alpha}(\theta, \varphi) = (\sin\varphi, \cos\varphi\sin\theta, -\cos\varphi\cos\theta)$ where θ corresponds to the definition above, and φ is the latitude with respect to the polar axis x. (Note the unusual orientation of the spherical coordinates.) With a slight abuse of notation, we write $g_\lambda(u, v) = g_\lambda(\vec{\alpha}(\theta, \varphi))$. We also use $\tilde{g}_\lambda$ to denote the projection g_λ weighted by the cosine of the incidence angle of the ray. Given the coordinates of the source position $(\lambda, 0, d)$, we have:

$$\tilde{g}_\lambda(u, v) = g_\lambda(u, v)\frac{d}{\sqrt{(u-\lambda)^2 + v^2 + d^2}} \tag{2}$$

D. Grangeat-based consistency condition

Consistency conditions are equations derived from the forward model that ideal projection data must satisfy. Consistency conditions can be used to detect - and possibly correct for - inconsistencies introduced in the data by systematic effect such as mis-calibration, motion, scattering or beam-hardening.

Let $\mathcal{R}f(\mathcal{P}_\theta)$ denote the 3D Radon transform of f over the plane $\mathcal{P}_\theta$. We recall the well-known result of Grangeat [11] (expressed with our notation):

$$\frac{1}{\cos^2\theta}\frac{\partial}{\partial v}\int_{-\infty}^{+\infty}\tilde{g}_\lambda(u, v)\,\mathrm{d}u = \frac{\partial}{\partial s}\mathcal{R}f(\mathcal{P}_\theta)\Big|_{s=\vec{s}_\lambda\cdot\vec{n}_\theta} \tag{3}$$

where the derivative of $\mathcal{R}f$ is taken in the direction $\vec{n}_\theta$. Again, v and θ are related through $v = d\tan\theta$.

For fixed $\mathcal{P}_\theta$ (equivalently, for fixed $v = d\tan\theta$), we note that $\vec{s}_\lambda \cdot \vec{n}_\theta = d\sin\theta$ is constant throughout the trajectory, so the right-hand-side of Equation 3 does not depend on the source position (it only depends on v, not on λ). We let $G(\lambda, v) = \frac{\partial}{\partial v}\int\tilde{g}_\lambda(u, v)\,\mathrm{d}u$, so the left-hand-side of Equation 3, $(1/\cos^2\theta)G(\lambda, v)$, must be independent of λ. We thus obtain the following necessary DCC for this particular acquisition geometry:

The projection data $(g_\lambda)_{\lambda\in\Lambda}$ are consistent only if, for all v, $G(\lambda, v)$ is a constant function of λ.

This is what will be called in the sequel, the Grangeat-based Consistency Conditions, GCC for short.

Grangeat's result, Equation 3, is also the basis of Epipolar consistency conditions described in [6].

E. Fan-beam consistency conditions

Another way to quantify the redundancy/consistency of a set of projections is to fix the v coordinate and consider the corresponding plane $\mathcal{P}_\theta$ and the fan-beam projections therein. We then consider fan-beam consistency conditions (see [3] for a complete proof) extended to 3D cone-beam projections acquired in our particular tomosynthesis geometry: the set of projections $(g_\lambda)_{\lambda\in\Lambda}$ is consistent if and only if, for each $n = 0, 1, 2...$ and each $\theta \in \left]-\frac{\pi}{2}; \frac{\pi}{2}\right[$, the function:

$$J_n(\lambda, \theta) = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\frac{g_\lambda(\alpha(\theta, \varphi))}{\cos\varphi}\tan^n\varphi\,\mathrm{d}\varphi, \tag{4}$$

is a polynomial of order n in λ.

In particular, for $n = 0$, the quantity J_0 does not depend on the source position λ. It only depends on θ. And since there is a one-to-one correspondence between θ and v, we will write (with another small abuse of notation) $J_0(\lambda, v)$. These zeroth-order fan-beam consistency conditions will be referred to as FBCC. (When non-zero orders are being considered, we will state them explicitly.) These necessary conditions can be restated as:

The projection data $(g_\lambda)_{\lambda\in\Lambda}$ are consistent only if, for all v, $J_0(\lambda, v)$ is a constant function of λ.

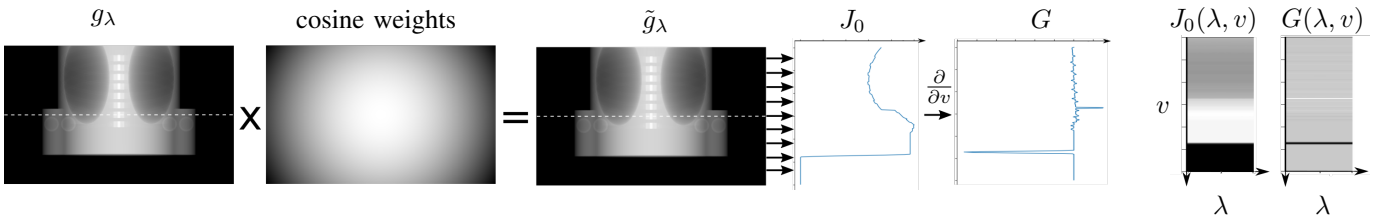


Fig. 3. a) Computation of both DCC for a single projection λ: projections g_λ are weighted with the cosine of the incidence angle. Integrals of $\tilde{g}_\lambda$ over detector rows are computed: up to a fixed factor of d, this is $J_0(\lambda, v)$. For each λ, $J_0(\lambda, v)$ is differentiated with respect to v to get $G(\lambda, v)$. The dotted line indicates the detector row used in Figures 5 and 6. b) The functions $J_0(\lambda, v)$ and $G(\lambda, v)$ plotted as gray-value images. Note that the gray scales are different in the two.

F. Relationships between FBCC and GCC

We can now prove the theoretical contribution of this abstract:

Proposition 1 (FBCC $\Rightarrow$ GCC)**.** *If the projection data satisfy FBCC, then they necessarily satisfy GCC.*

Before proving this statement, a few comments are in order. First, the proposition means that (order-0) FBCC carry at least as much information as GCC. Second, under some specific circumstances, we will prove that this assertion is strict, in the sense that we can design a projection example which satisfies GCC but breaks FBCC. Finally, neither GCC nor FBCC are sufficent consistency conditions because the zeroth-order fan-beam conditions alone are insufficient; it was shown in [3] that all orders $n = 0, 1, 2...$ are necessary and sufficient.

Proof of Proposition 1: We start with FBCC (Equation 4) and change the spherical variables (θ, φ) to the (u, v) detector coordinates with $v = d \tan\theta$ and $u = \sqrt{v^2 + d^2} \tan\varphi$. After elementary computations, we get:

$$J_0(\lambda, v) = \frac{1}{d} \int_{-\infty}^{+\infty} \tilde{g}(u, v)\, \mathrm{d}u. \tag{5}$$

The FBCC state that for each v, J_0 is a constant function of λ. Therefore, $J_0(\lambda, v)$ is independent of λ. Now, it is clear from the definition of G that:

$$G(\lambda, v) = d \frac{\partial}{\partial v} J_0(\lambda, v), \tag{6}$$

which proves that G is also independent of λ. Therefore, for each v, $G(\lambda, v)$ is a constant function of λ, so the GCC are satisfied and the proof is complete. ■

If the converse of this statement holds then the two sets of DCC would be equivalent.

Proposition 2 (FBCC $\Leftrightarrow$ GCC)**.** *The FBCC and GCC are equivalent in the ideal case of untruncated projections.*

Proof of Proposition 2: In light of Proposition 1, we only need to prove that if the GCC are satisfied, then the FBCC are satisfied. From Equation 6, we obtain:

$$J_0(\lambda, v) = \frac{1}{d} \int G(\lambda, v)\, \mathrm{d}v + \kappa_\lambda, \tag{7}$$

where κ_λ depends on λ but not on v. The GCC are assumed to be satisfied, so $G(\lambda, v)$ is independent of λ, and therefore the integral term in Equation 7 is also independent of λ. So for each λ, $J_0(\lambda, v)$ is the same function of v except for the additive constant κ_λ (which varies with λ). Now, since the projections are assumed to not be truncated, there exist some v^* (which corresponds to an actual row of the physical detector) such that:

$$\forall \lambda \in \Lambda, \forall v \geq v^* \quad g_\lambda(u, v) = 0. \tag{8}$$

In particular, $g_\lambda(u, v^*) = 0$ for all λ, and therefore $J_0(\lambda, v^*) = 0$ for all λ. For each λ, the functions (of v) $J_0(\lambda, v)$ all agree for $v = v^*$, and therefore the additive offsets κ_λ must be the same, so $\kappa_\lambda = \kappa$ is a constant (independent of λ). Therefore $J_0(\lambda, v)$ is independent of λ and the FBCC are satisfied. ■

We now show that the hypothesis of untruncated projections cannot be relaxed in Proposition 2.

Proposition 3. *FBCC and GCC are not equivalent; a counter-example.*

Proof of Proposition 3: Let us modify one single projection g_{λ_0} as follows:

$$\gamma_{\lambda_0}(u, v) = g_{\lambda_0}(u, v) + \frac{\sqrt{(u - \lambda_0)^2 + v^2 + d^2}}{Ld}. \tag{9}$$

Note that the modified projection is non-zero everywhere, so it would be considered a truncated projection and would not satisfy the hypotheses of Proposition 2. The added term is the inverse of the cosine of the incidence angle, weighted by the inverse of the width of the detector (L). It is easily seen that this modification adds a constant term to $J_0(\lambda_0, v)$ for all v, hence breaks the FBCC without affecting the GCC. ■

Note that this modification of the projection is equivalent (up to a constant) to filtering the X-ray beam with a flat filter of constant thickness, placed perpendicularly to the z-axis.

III. Numerical simulations

We simulated 40 projections of a thorax Forbild phantom (see Figure 3a for a sample projection). The acquisition geometry was described in Section II-B. The source-to-detector distance was fixed at 600 mm. The source positions were $(\lambda, 0, 600)$ where λ varies evenly from -200 mm to $+190$ mm. The detector size was 500×300 pixels, with pixel size 0.3 mm^2. In our reference projections, the phantom center of mass was placed at $(0, 12, 300)$ (centered in x, at mid-distance between the source and the detector plane and with a vertical offset of 12 mm). All simulations were carried out with the Reconstruction ToolKit

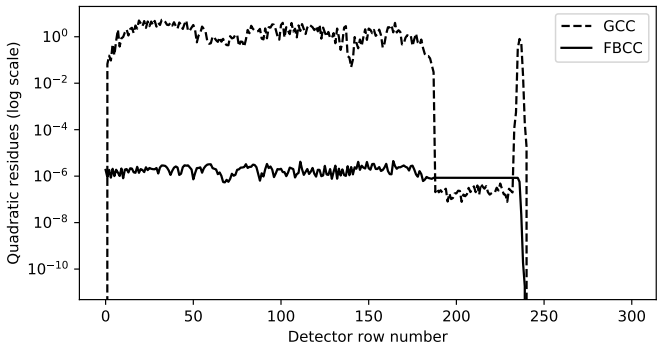


Fig. 4. Quadratic residues of the linear regression of the functions $G(\lambda, v)$ and $J_0(\lambda, v)$ as a function of v. On the last 60 detector rows, the signal is 0, so are G and J_0. The residues are not represented on this log-scale figure.

(RTK, [12]). The implementation of FBCC in this set up only requires a sum over the rows of the projections. The derivative in the v-direction involved in GCC is implemented with a central finite difference. Note that the two DCC involve computing line integrals in the u-direction of the projection images. Hence, projections must not be truncated in that direction.

A. Are G and J_0 constant ?

Saying that the projection data satisfy GCC or FBCC amounts to saying that for each v (i.e. each detector row), the 1D functions $G(\cdot, v)$ and $J_0(\cdot, v)$ are constant functions of λ. For each projection, we computed $G(\lambda, v)$ and $J_0(\lambda, v)$ for all v according to Equations 3 and 4 respectively (see Figure 3a) and concatenated those 1D signals to get 300 (number of detector rows) supposedly constant signals (see Figure 3b). To quantify how constant those signals were, we computed the slope of their linear regression. In both cases (FBCC and GCC), the mean over all planes and the standard deviation were numerically zero, as stated by the theory. To further investigate the respective behavior of each consistency measure, for each v we computed the sum of squared difference between the signals $G(\cdot, v)$ and $J_0(\cdot, v)$ and their regression line. These residues are presented in Figure 4.

Both the GCC and FBCC results are theoretical in the sense that the functions G and J_0 cannot be exactly constant in λ practically since they required numerical approximation of integrals. The errors in this approximation are amplified by the differentiation step in the computation of G. This amplification of the discretization errors is the probable explanation for why the residues for G are 10^6 times greater than J_0.

B. Illustration of inconsistency

We examined the effect on the GCC and FBCC of a small rigid motion of the object. After 20 projections, the phantom was displaced vertically and 15 projections were collected in the displaced position. The phantom was then returned to the original position for the last 5 projections. We repeated the study with 3 different magnitudes of displacement: 2.5 mm, 1.0 mm, and 0.2 mm. We restricted the study to one plane

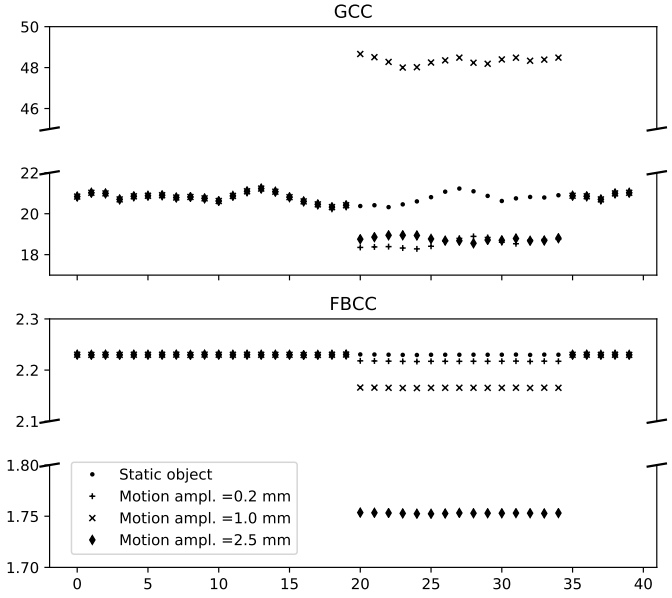


Fig. 5. Sensitivity of each DCC to a vertical motion of the object. The object is displaced between projections 20 and 35.

(row index 150) when calculating the consistency conditions. Figure 5 shows that $G(\lambda, v)$ and $J_0(\lambda, v)$ were both constant (to within small numerical errors) for the consistent collections of projections, but that the constant changed when the object was displaced to a new position. As the displacement became larger, the change in FBCC increased, as was expected. On the other hand, the GCC over-reacted to the 1.0 mm displacement, due to particular features of the phantom.

C. FBCC and GCC are not sufficient conditions

In order to illustrate the fact that both GCC and FBCC are not sufficient conditions, we simulated 40 projections of the same phantom which was displaced horizontally (in the x-direction) at projection 20 and left in the displaced position for the remaining projections. The first and last 20 projections alone were consistent, but not the full 40 projections. Since both GCC and FBCC involve computations of integrals along horizontal detector rows, it is intuitive that they will fail in detecting inconsistencies resulting from motion in the horizontal direction. In Figure 6, we show that FBCC and GCC erroneously indicate consistent data (both are constant as expected), while FBCC-1 detects inconsistency (severe discontinuity in $J_1(\lambda)$, the theoretical order-1 polynomial) when the object was displaced. This study was conducted in the plane corresponding to the row index 192. (This plane is such that J_0 and G have the same order of magnitude, which facilitates convenient plotting on a single figure.)

IV. Discussion and conclusion

In this abstract we proved that FBCC carry at least as much information as GCC. In the non-truncated case, FBCC and GCC are equivalent and in the truncated case, we designed projection

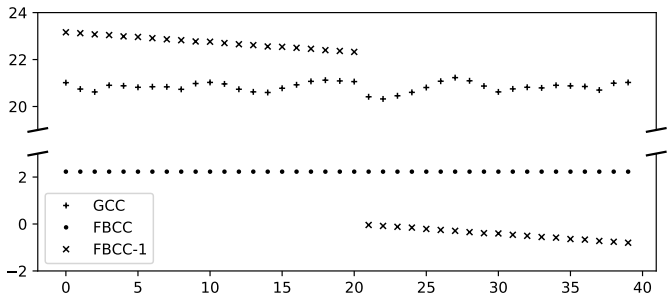


Fig. 6. From projection index 20, the object was moved horizontally by 5 mm. While GCC and FBCC fail to detect the inconsistency induced by the motion, FBCC-1 does.

data which satisfy GCC but not FBCC. We illustrated that both the GCC and FBCC are not sufficient conditions for this tomosynthesis geometry. In practice, neither of the DCC are perfectly satisfied, due to numerical errors in the quadrature methods involved in FBCC and GCC. Based on our simulation studies, the FBCC seem to be more robust to GCC in the presence of numerical errors.

REFERENCES

[1] S. Helgason, *The Radon transform*, ser. Progress in mathematics. Boston, Basel, Berlin: Birkhäuser, 1999. [Online]. Available: http://opac.inria.fr/record=b1095821

[2] D. Ludwig, "The radon transform on euclidean space," *Communications on Pure and Applied Mathematics*, vol. 19, no. 1, pp. 49–81, 1966. [Online]. Available: http://dx.doi.org/10.1002/cpa.3160190105

[3] R. Clackdoyle, "Necessary and Sufficient Consistency Conditions for Fanbeam Projections Along a Line," *IEEE Transactions on Nuclear Science*, vol. 60, no. 3, pp. 1560–1569, 2013.

[4] R. Clackdoyle and L. Desbat, "Full data consistency conditions for cone-beam projections with sources on a plane," *Physics in Medicine and Biology*, vol. 58, no. 23, p. 8437, 2013. [Online]. Available: http://stacks.iop.org/0031-9155/58/i=23/a=8437

[5] C. Debbeler, N. Maass, M. Elter, F. Dennerlein, and T. M. Buzug, "A New CT Rawdata Redundancy Measure applied to Automated Misalignment Correction," in *The 12th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, Lake Tahoe, USA, 2013, pp. 264–267. [Online]. Available: http://www.fully3d.org/2013/Fully3D2013Proceedings.pdf

[6] A. Aichert, M. Berger, J. Wang, N. Maass, A. Doerfler, J. Hornegger, and A. K. Maier, "Epipolar Consistency in Transmission Imaging," *IEEE Transactions on Medical Imaging*, vol. 34, no. 11, pp. 2205–2219, nov 2015.

[7] M. S. Levine, E. Y. Sidky, and X. Pan, "Consistency Conditions for Cone-Beam CT Data Acquired with a Straight-Line Source Trajectory," *Tsinghua science and technology*, vol. 15, no. 1, pp. 56–61, 2010. [Online]. Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2886312/

[8] T. Grulich, W. Holub, U. Hassler, A. Aichert, and A. Maier, "Geometric Adjustment of X-ray Tomosynthesis," in *Proceedings of the 13th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine.*, 2015, pp. 468–471.

[9] N. Maass, F. Dennerlein, A. Aichert, and A. Maier, "Geometrical Jitter correction in computed tomography," in *Proceedings of the third international conference on image formation in x-ray computed tomography*, Salt Lake City, USA, 2014, pp. 338–342.

[10] J. Lesaint, S. Rit, R. Clackdoyle, and L. Desbat, "Calibration for Circular Cone-Beam CT Based on Consistency Conditions," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 1, no. 6, pp. 517–526, 2017. [Online]. Available: http://ieeexplore.ieee.org/document/7999272/

[11] P. Grangeat, "Mathematical framework of cone beam 3D reconstruction via the first derivative of the radon transform," in *Mathematical Methods in Tomography*, ser. Lecture Notes in Mathematics, G. Herman, A. Louis, and F. Natterer, Eds. Springer Berlin Heidelberg, 1991, vol. 1497, pp. 66–97. [Online]. Available: http://dx.doi.org/10.1007/BFb0084509

[12] S. Rit, M. V. Oliva, S. Brousmiche, R. Labarbe, D. Sarrut, and G. C. Sharp, "The Reconstruction Toolkit (RTK), an open-source cone-beam CT reconstruction toolkit based on the Insight Toolkit (ITK)," *Journal of Physics: Conference Series*, vol. 489, no. 1, p. 12079, 2014. [Online]. Available: http://stacks.iop.org/1742-6596/489/i=1/a=012079

Differential Tomography: Influence of Sensitivity Direction and Noise-suppressing Windows

Sebastian Kaeppler, Andreas Maier, Christian Riess

***Abstract*—Phase-sensitive systems can deliver improved CT imaging performance for materials with low attenuation contrast, and, due to the differential nature of their images, when high spatial resolution is required. Typically, these systems only yield a directional derivative of a projection image along one sensitivity direction. The sensitivity direction is usually tangential to the CT trajectory. However, it is also conceivable to build systems with other sensitivity directions.**

The purpose of this paper is two-fold. First, we investigate the noise behavior of differential CT for sensitivity directions in z-direction and for bi-directional projections. Second, we use Wiener filtering to derive noise-suppressing window functions for all three sensitivity directions.

Our experiments indicate that the benefit that can be obtained from optimized window functions depends on the overlap between object spectrum and the Noise Power Spectrum, and thus on the sensitivity direction. We also find that sensitivity in z-direction yields a noise texture that is unsuitable for CT.

***Index Terms*—X-ray, phase-contrast, sensitivity direction, denoising**

I. Introduction

Recently, several measurement principles for X-ray phase contrast haven been proposed. Most notable of these are the Talbot-Lau interferometer [1], diffraction-enhanced systems [2] and coded apertures [3]. These systems share a common trait: they can only obtain differential phase-contrast contrast projection images along one sensitivity direction.

When performing CT acquisitions, these systems are typically operated such that the sensitivity direction is tangential to the CT trajectory, and thus radial to the rotation axis when rotating. For the Talbot-Lau interferometer (coded apertures) this means that the grating bars (the apertures) are parallel to the rotation axis. The noise behavior of this configuration has been studied extensively in [4], [5], [6].

In principle, the sensitivity axis can be chosen arbitrarily by rotating the phase-sensitive components of the system along the optical axis. For example, it has been shown that direction-dependent dark-field CT [7], [8] reconstruction can benefit from having the sensitivity direction in z-direction [9], i.e. parallel to the rotation axis. For a fan-beam-like geometry, z-sensitivity has the further advantage that the gratings (apertures) do not have to be bent to match the fan angle.

It is also conceivable to acquire two perpendicular directional derivatives. For geometries that can be approximated well using parallel beams, this can be realized relatively easily by using diagonal gratings (apertures) and has shown to yield reduced reconstruction noise when compared to the standard radial sensitivity approach [10].

In the this paper, we investigate the Noise Power Spectrum of a differential CT system depending on its sensitivity direction. We ignore the system-specific phase sensitivity and noise behavior to obtain a system-neutral analysis for CT reconstruction of differential projections. Specifically, we consider three configurations: the standard approach of radial sensitivity, sensitivity in z-direction and sensitivity in both directions (at half dose for each direction).

Additionally, we use the analytical Noise Power Spectrum and Wiener filtering to define noise-suppressing window functions for these configuration, as each shows a vastly different noise texture. To our knowledge, this is the first study of the noise behavior of non-standard sensitivity configuration and also the first paper to propose practicable window functions for differential CT.

All authors are with the the Pattern Recognition Lab, Department of Computer Science, Friedrich-Alexander-University Erlangen-Nuremberg. Contact: sebastian.kaeppler@fau.de
www5.cs.fau.de/%7Ekaeppler

II. Methods

In the following sections, we will derive the Noise Power Spectrum of differential CT and use this information to derive noise-suppressing window functions.

A. Noise Power Spectrum

We follow the simplified model of a parallel beam reconstruction with uniform projection Noise Power Spectrum and multiple slices with voxel and pixel size a as in [4]. In this case the total radial Noise Power Spectrum $\mathrm{NPS}_{\mathrm{tot}}$ is given by

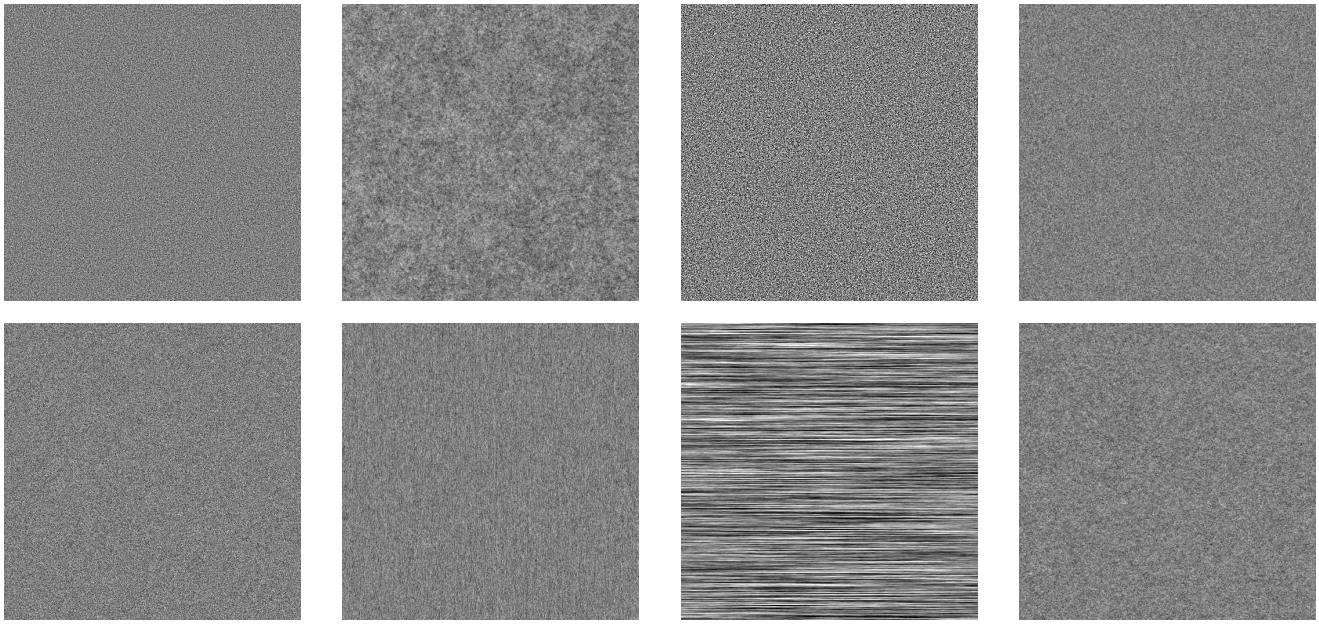
$$\mathrm{NPS}_{\mathrm{tot}}(\rho_{\mathrm{r}},\rho_{\mathrm{z}}) = f^2_{\mathrm{RamLak}} \cdot f_{\mathrm{invradon}} \cdot f^2_{\mathrm{interpol}} \cdot f^2_{\mathrm{integration}} \ . \tag{1}$$

Here, ρ_{r} denotes the radial frequency and ρ_{z} denotes the frequency in z-direction, i.e. across slices. For the RamLak filter we have $f_{\mathrm{RamLak}} = |\rho_{\mathrm{r}}| \cdot \mathrm{rect}(\rho_{\mathrm{r}}/2a)$, for Radon inversion we have $f_{\mathrm{invradon}} = 1/|\rho_{\mathrm{r}}|$ and for linear interpolation onto a grid of voxels with size a we have $f_{\mathrm{interpol}} = \mathrm{sinc}(\pi a \rho_{\mathrm{r}})^2 = [\sin(\pi a \rho_{\mathrm{r}})/(\pi a \rho_{\mathrm{r}})]^2$. This yields

$$\mathrm{NPS}_{\mathrm{tot}}(\rho_{\mathrm{r}},\rho_{\mathrm{z}}) = |\rho_{\mathrm{r}}| \cdot \mathrm{rect}(\rho_{\mathrm{r}}/(2a)) \cdot \mathrm{sinc}(\pi a \rho_{\mathrm{r}})^4 \cdot f^2_{\mathrm{integration}} \ . \tag{2}$$

The frequency response $f^2_{\mathrm{integration}}$ for integrating the differential projections depends on the sensitivity direction. For the standard approach of radial sensitivity we have for a $[-1;1]$ backward discretization of the derivative operator [4]

$$f^{\mathrm{rad}}_{\mathrm{integration}} = a/[2\sin(\pi a \rho_{\mathrm{r}})] \ . \tag{3}$$



a) b) c) d)

Fig. 1. Sample CT noise realizations displayed in arbitrary units, created as described in Section III. Top row: X-Y plane. Bottom row: X-Z plane. a) Non-differential projections. b) Differential (radial). c) Differential (z-direction) d) Differential (both directions). The linear scale is identical all images, except for the noise obtained for differential (z-direction), which was downscaled by a factor of five.

For sensitivity in z-direction we simply have to change the integration direction

$$f^{\mathrm{z}}_{\mathrm{integration}} = a/[2\sin(\pi a \rho_{\mathrm{z}})] \ . \tag{4}$$

Note that both integration steps assume that the object is non-truncated along the integration direction. Thus, having the sensitivity direction in z-direction requires not only a non-truncated object in radial direction (due to the inverse Radon transform) but also in z-direction (due to the integration across slices).

The frequency response of 2-D integration can be calculated by considering its closed form solution in Fourier domain, yielding

$$f^{\mathrm{2D}}_{\mathrm{integration}} = \sqrt{2}a/[2\sin(\pi a \rho_{\mathrm{r}}) + 2\sin(\pi a \rho_{\mathrm{z}})] \ . \tag{5}$$

Here, the prefactor $\sqrt{2}$ accounts for having only half of the dose for determining the image information for each direction. The non-truncation requirements of 2-D integration are relaxed compared to 1-D integration, here only one known pixel has to be outside of the object.

Each of the integration factors lead to the amplification of low frequency noise. For radial sensitivity $\mathrm{NPS}^{\mathrm{rad}}_{\mathrm{tot}} \to \infty$ for $\rho_{\mathrm{r}} \to 0$, although the frequency response of the ramp filter combined with the inverse Radon transform partially compensate the low-frequency noise induced by integration.

This is not the case for integration in z-direction, where $\mathrm{NPS}^{\mathrm{z}}_{\mathrm{tot}} \to \infty$ for $\rho_{\mathrm{z}} \to 0$. Here, ramp filtering acts as a high-pass in radial direction, while integration acts as a low-pass in z-direction, suggesting a noise pattern with a high autocorrelation in z-direction and negative autocorrelation in radial direction.

For 2-D integration $\mathrm{NPS}^{\mathrm{2D}}_{\mathrm{tot}} \to \infty$ only for $\rho_{\mathrm{r}}, \rho_{\mathrm{r}} \to 0$, indicating autocorrelated noise in both directions.

The quadratic scaling of the Noise Power Spectrum by the pixel size a is identical for all sensitivity directions and can thus be ignored when comparing the relative performance of sensitivity directions.

Sample CT noise realizations obtained by reconstructing the same amount of projection noise are shown in Fig. 1. The noise texture is in accordance with the noise autocorrelation behavior given by the Noise Power Spectrum. The approach with sensitivity in z-direction suffers from strong noise amplification due to integration being not aligned with ramp filtering.

B. Noise-Suppressing Windows

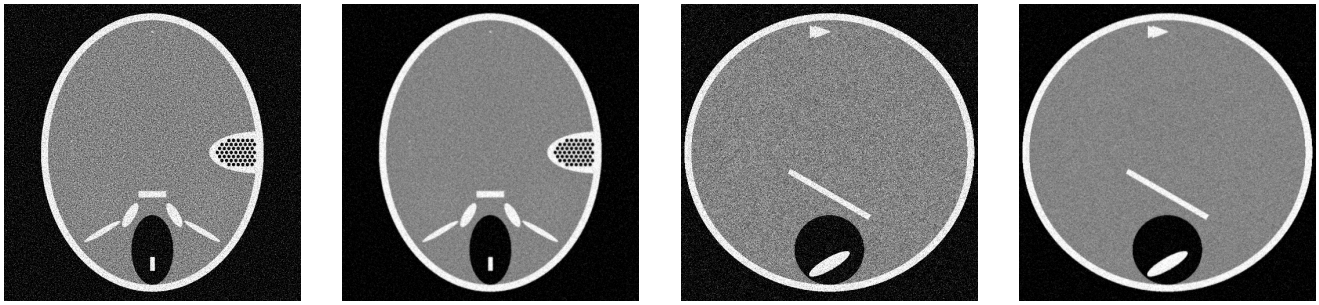
In the following, we will use Wiener Filtering to derive noise-suppressing window functions for differential CT. Note that in contrast to a conventional window function, our window not only operates in radial direction, but also in z-direction. Wiener filtering can be realized by applying window function $\mathbf{W}$ in Fourier domain. The frequency-dependent window function is given by the (also frequency-dependent) signal-to-noise ratio:

$$\mathbf{W} = \frac{1}{1 + \frac{1}{\mathrm{SNR}}} \ . \tag{6}$$

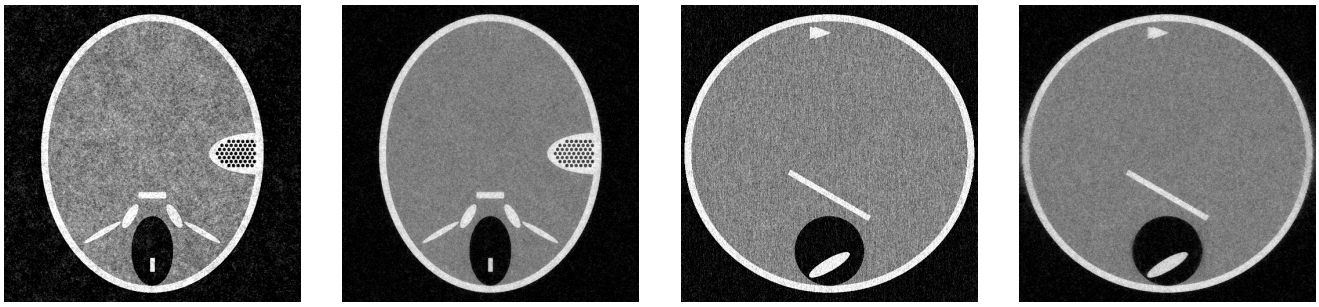
For CT, the signal-to-noise ratio depends on the object to be reconstructed and the noise. Both are unknown quantities, however the noise can be estimated using the analytical model of the Noise Power Spectrum (Eqn. 2). We thus require only an estimate of the Power Spectrum of the object, denoted as $\hat{\mathbf{O}}$, to estimate the SNR as $\hat{\mathbf{S}}$:

$$\hat{\mathbf{S}}(\rho_{\mathrm{r}}, \rho_{\mathrm{z}}) = \frac{\hat{\mathbf{O}}(\rho_{\mathrm{r}}, \rho_{\mathrm{z}})}{\mathrm{NPS}_{\mathrm{tot}}(\rho_{\mathrm{r}}, \rho_{\mathrm{z}})} \ . \tag{7}$$

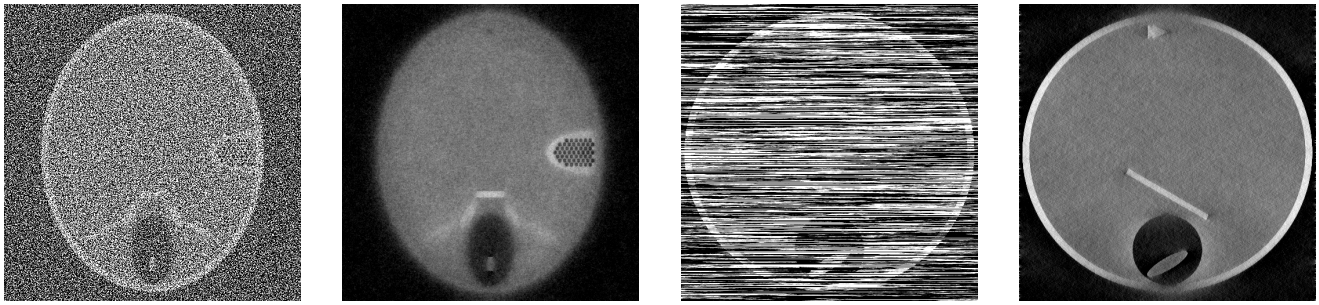
This step allows to define the noise-suppressing window function independent on the noise behavior of the system.



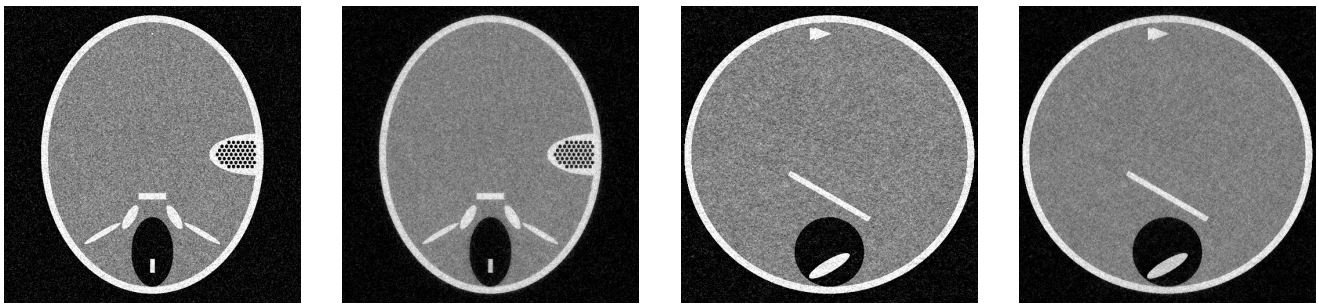
non-differential reconstruction, PSNR noisy: 11.7dB, PSNR denoised: 20.2dB (+8.5dB)



radial-differential reconstruction, PSNR noisy: 14.3dB, PSNR denoised: 18.0dB (+3.7dB)



z-differential reconstruction, PSNR noisy: -7.4dB, PSNR denoised: 12.4dB (+19.0dB)



bi-differential reconstruction, PSNR noisy: 14.5dB, PSNR denoised: 17.2dB (+2.7dB)

a) b) c) d)

Fig. 2. Reconstruction results of applying the proposed Wiener window to modified version of the head phantom. a) central X-Y plane, noisy. b) central X-Y plane, denoised. c) central X-Z plane, noisy. d) central X-Z plane, denoised. The same intensity window has been applied to all images.

Note that by choosing an appropriate $\hat{\mathbf{O}}$, the tradeoff between reconstruction fidelity and noise can be attuned to objects of different contrast, shape, or size, given that the object is radially symmetric.

Due to the interchangeability of filtering and backprojection, the window function can also be applied to projections. This results in negligible computational demand for applying the window, as it involves replacing multiple 1-D Fourier transforms (for ramp filtering) with a single 2-D Fourier transform (for ramp filtering combined with the window function).

Note that for differential CT our approach of using a 2-D window function is more powerful than the conventional approach of using a radial window function combined with increasing the slice thickness, because the implicit frequency response of this approach is linearly separable in ρ_{r} and ρ_{z}, which is not optimal for the Noise Power Spectrum encountered in differential CT.

III. Experiments & Results

Before evaluating the proposed window function, we simulated Noise Power Spectra for all configurations and compared them with their analytic representations. They were in good agreement but are omitted in this paper due to space constraints.

For evaluating the proposed window function, we performed reconstruction of a modified version of the FORBILD head phantom onto a 512^3 grid with voxel size $0.5\mathrm{cm}^3$. Reconstruction parameters were selected to match our assumption of a parallel beam geometry, rectangular apodization, and linear interpolation. The angular increment was 0.5° to avoid angular undersampling. Projection images were of size 730×512 pixels with a pixel size identical to the voxel size. Projections were corrupted with additive white Gaussian noise. Reconstruction was performed once without and with the proposed window function.

Defining the window function requires knowledge of the Noise Power Spectrum and the Object Power Spectrum. For the Noise Power Spectrum, we used the analytic definition. The Object Power Spectrum needs to estimated. To this end, we assumed a spherically symmetric Power Spectrum which decays exponentially towards high frequencies. We model this power spectrum as:

$$\hat{\mathbf{O}}(\rho_\mathrm{r}, \rho_\mathrm{z}) = a \cdot x^b, \;\; x = \sqrt{\rho_\mathrm{r}^2 + \rho_\mathrm{z}^2} \;\; . \tag{8}$$

The model parameters a and b are estimated from the head phantom.[1]

Results of the evaluation with the peak signal-to-noise ratio (PSNR) are shown in Fig. 2. Applying the proposed window reduces reconstruction error for all approaches, and yields a more even residual noise texture.

The reconstruction obtained with z-sensitivity benefits most from denoising, due to the high amount of noise. Still, it suffers from a poor signal-to-noise ratio which renders the resulting images unusable.

The second most benefit can be observed in the non-differential reconstruction. We attribute this to the low overlap between Noise Power Spectrum and Object Power Spectrum. Here, our window shows the well-known tradeoff between resolution and noise in conventional CT.

The least benefit is observed for the reconstruction using two sensitivity axes. This can again be attributed to the overlap between Noise Power Spectrum and Object Power Spectrum. Here the overlap is high, since both spectra are strongest in the low frequencies.

Reconstruction using the standard approach with radial sensitivity yields a higher PSNR than the bi-directional approach after applying the window function, while its PSNR is lower when no window is used. We attribute this to the noise structure of this approach, which can relatively easily be attenuated by frequency adaptive smoothing in z-direction that is strong at low radial frequencies, and decays towards higher radial frequencies.

[1] Strictly speaking, this step violates the guidelines for proper parameter selection. However, we argue that if the shape of the Noise Power Spectrum is known, the Object Power Spectrum could also be estimated from the noisy reconstruction, yielding possibly even better performance.

IV. Conclusions

We have investigated the noise behavior of CT reconstruction from directional differential projections. Our experiments indicate that reconstructions from differential projections in z-direction suffer from high noise and are unusable in practice. This leads to a design conflict, as dark-field reconstruction can benefit from this configuration.

Reconstruction from two directional derivatives yields a more even noise texture than the standard radial approach. However, our experiments indicate that this can be mitigated by choosing an appropriate reconstruction window for the radial direction. We will investigate this finding in further studies.

Our investigation yields many directions for other future work. We plan to further investigate the shape of our window function for the different configurations, and to quantify the gain in reconstruction quality depending on the overlap between Object and Noise Power Spectrum. Additionally, we are interested in the implication of our findings for designing regularizers for iterative phase-contrast CT reconstruction.

References

[1] F. Pfeiffer, T. Weitkamp, O. Bunk, and C. David, "Phase Retrieval and Differential Phase-contrast Imaging with Low-brilliance X-ray Sources," *Nature Physics*, vol. 2, no. 4, pp. 258–261, Apr. 2006.

[2] C. Parham, Z. Zhong, D. Connor, L. Chapman, and E. Pisano, "Design and Implementation of a Compact Low-Dose Diffraction Enhanced Medical Imaging System," *Academic Radiology*, vol. 16, no. 8, pp. 911–917, Aug. 2009.

[3] A. Olivo and R. Speller, "A coded-aperture technique allowing x-ray phase contrast imaging with conventional sources," *Applied Physics Letters*, vol. 91, no. 7, pp. 074106, 2007.

[4] R. Raupach and T. Flohr, "Analytical Evaluation of the Signal and Noise Propagation in X-ray Differential Phase-Contrast Computed Tomography," *Physics in Medicine and Biology*, vol. 56, no. 7, pp. 2219–2244, Apr. 2011.

[5] C. Hagen, P. Diemoz, and A. Olivo, "On the relative performance of edge illumination x-ray phase-contrast ct and conventional, attenuation-based ct," *Medical Physics*, vol. 44, no. 5, pp. 1876–1885, 2017.

[6] K. Li, N. Bevins, J. Zambelli, and G.-H. Chen, "Fundamental relationship between the noise properties of grating-based differential phase contrast ct and absorption ct: Theoretical framework using a cascaded system model and experimental validation," *Medical Physics*, vol. 40, no. 2, 2013.

[7] A. Malecki, G. Potdevin, T. Biernath, E. Eggl, K. Willer, T. Lasser, J. Maisenbacher, J. Gibmeier, A. Wanner, and F. Pfeiffer, "X-ray tensor Tomography," *Europhysics Letters*, vol. 105, no. 3, pp. 38002, 2014.

[8] F. Bayer, S. Hu, A. Maier, T. Weber, G. Anton, T. Michel, and C. Riess, "Reconstruction of scalar and vectorial components in X-ray dark-field tomography," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, no. 35, pp. 12699–12704, Sept. 2014.

[9] F. Schaff, F. Prade, Y. Sharma, M. Bech, and F. Pfeiffer, "Non-iterative directional dark-field tomography," *Scientific Reports*, vol. 7, 2017.

[10] S. Rutishauser, T. Donath, C. David, F. Pfeiffer, F. Marone, P. Modregger, and M. Stampanoni, "A tilted grating interferometer for full vector field differential x-ray phase contrast tomography," *Optics Express*, vol. 19, no. 25, pp. 24890–24896, 2011.

Model-based iterative reconstruction for propagation-based phase-contrast computed tomography, suitable for laboratory sources

Lorenz Hehn, Kaye Morgan, Martin Dierolf, Regine Gradl, Peter B. Noël, and Franz Pfeiffer

***Abstract*—We present a model-based iterative reconstruction algorithm for propagation-based phase-contrast computed tomography, suitable for laboratory sources. Thereby, we recover the three-dimensional distribution of the phase-shifting properties of the sample, by incorporating the relevant aspects of the image formation process directly into the tomographic reconstruction. Therefore, we model the extent of the source, which is the limiting factor for the spatial coherence of the x-rays. Furthermore, the influence of the homogeneous sample on the intensity as well as on the phase of the x-ray wave-field is described. In addition, we model the free-space propagation of the x-ray wave-field from the sample to the detector, which leads to intensity modulations crucial for the reconstruction of the phase-shifting properties of the sample. Finally, the detector response of a scintillator-based x-ray detector is included. We validate our reconstruction approach on simulated data and compare it to conventional reconstruction techniques. We find that our integrated approach yields distinct improvements in terms of resolution as well as contrast. As any blurring of the system diminishes the pivotal intensity modulations, we believe that accurately modeling the whole image formation process might even be more beneficial in phase-contrast imaging than for conventional attenuation imaging.**

I. Introduction

As opposed to conventional x-ray absorption imaging, which relies solely on the absorption of the x-rays in matter, phase-contrast imaging (PCI) is sensitive to x-ray phase shifts. This technique has become one of the most commonly used x-ray imaging methods in laboratory and preclinical studies, yielding distinct advantages for the visualization of weakly absorbing details that are common in biological and medical samples. By extending PCI to computed tomography (CT), it has become a valuable tool for three-dimensional visualization of thick and complex samples due to its high sensitivity and high contrast produced between materials with similar absorption properties [1].

The most intuitive way to detect phase contrast is via propagation-based methods [2], which rely on the fact that when a coherent wavefront traverses a sample and propagates sufficiently far to a detector, the phase shifts induced by the sample lead to distinct variations in the measured intensity. From these variations, the phase-shifting properties of the sample can be extracted when imposing certain conditions onto the sample.

All authors are with the Chair of Biomedical Physics, Department of Physics and Munich School of Bioengineering, Technical University of Munich, 85748 Garching, Germany. L. Hehn, P. B. Noël and F. Pfeiffer are with the Department of Diagnostic and Interventional Radiology, Klinikum Rechts der Isar, Technical University of Munich, 81675 Munich, Germany. K. Morgan, R. Gradl and F. Pfeiffer are with the Institute for Advanced Study, Technical University of Munich, 85748 Garching, Germany. K. Morgan is with the School of Physics and Astronomy, Monash University, Clayton VIC 3800, Australia.

II. Image formation process

An overview of the image formation in propagation-based phase-contrast imaging is depicted in Fig. 1. In the following, we derive the evolution of the intensity from the source to the detector. Thereby, the x-ray wave-field plane at each position along the optical axis can be written as a monochromatic forward propagation scalar wave according to

$$\psi_i(\mathbf{r}_\perp) = \sqrt{I_i(\mathbf{r}_\perp)}e^{\mathrm{i}\phi_i(\mathbf{r}_\perp)}, \tag{1}$$

where $\mathbf{r}_\perp$ denotes the coordinates orthogonal to the propagation direction, $I(\mathbf{r}_\perp)$ denotes the distribution of the intensity in the corresponding plane and $\phi(\mathbf{r}_\perp)$ the relative change in phase, compared to an undisturbed wave-field. The position of the x-ray wave-field according to Fig. 1 is indicated by the subscript i. Consequently, the measured intensity, which is experimentally accessible, is given by the squared modulus of the wave-field, namely $I_i(\mathbf{r}_\perp) = |\psi_i(\mathbf{r}_\perp)|^2$.

At the source depicted in blue in Fig. 1, partially coherent x-rays are generated. The degree of coherence is connected to the extent of the source, making the geometric parameters of the source spot crucial for this imaging technique, as discussed later. At the first stage, between the source and the sample, we assume a homogeneous illumination. Therefore, we can describe the intensity of the x-ray wave-field by a scalar $I_1(\mathbf{r}_\perp) = I_0$ and, as the wave-field is not disturbed by the free space propagation, the relative phase shift to an undisturbed wave-field vanishes corresponding to $\phi_1(\mathbf{r}_\perp) = 0$.

To describe the interactions of the x-ray wave-field with the sample, depicted in orange in Fig. 1, independently from propagation, we assume that the sample is sufficiently thin or scatters sufficiently weakly to neglect any propagation effects within the sample. This is known in literature as the projection approximation [3]. In order to describe the interaction of the x-ray wave-field with the sample, we impose the assumption that the sample consists of only one material, an assumption which has proven to be extremely

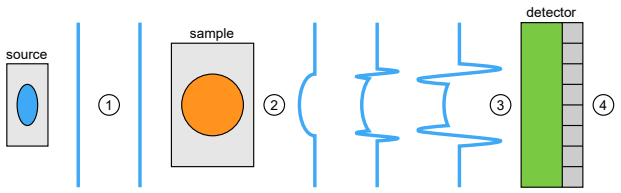


Fig. 1. Image formation in propagation-based phase-contrast imaging. Depicted from left to right is an extended x-ray source, a cylindrical homogeneous sample and an scintillator-based x-ray detector. The blue lines indicate the evolution of the intensity of the x-ray wave-field from the source to the detector at different stages.

successful in the propagation-based phase-contrast community [4]. Thereby the three-dimensional density distribution of the sample is denoted by $x(\mathbf{r})$. The projection of this density along the x-ray paths is modeled by a projection operator $\mathcal{P}$, yielding a two-dimensional map of projected thicknesses of the sample, which is referred to as the trace $t(\mathbf{r}_\perp)$ of the sample

$$t(\mathbf{r}_\perp) = \mathcal{P}\left[x(\mathbf{r})\right]. \tag{2}$$

Knowing the material's specific attenuation coefficient μ and the deviation of the real-part of the complex refractive index from unity δ, as well as the energy of the wave, denoted by the wave number k, we can determine the intensity and phase shift of the x-ray wave-field at the second stage behind the sample as indicated in Fig. 1. The attenuation is described by the Lambert-Beer law according to

$$I_2(\mathbf{r}_\perp) = I_0 e^{-\mu t(\mathbf{r}_\perp)} \tag{3}$$

and the total phase shift compared to the undisturbed reference wave-field is given by

$$\phi_2(\mathbf{r}_\perp) = -k\delta t(\mathbf{r}_\perp). \tag{4}$$

However, as we are only sensitive to the intensity of the wave-field, the phase information and therefore the phase-shifting properties of the sample are experimentally not accessible directly behind the sample.

Now, while the wave-field propagates from the sample to the detector, the phase shifts induced by the sample lead to distinct variations in intensity due to interference effects. These interference effects can then be used to determine the phase-shifting properties of the sample. By inserting the representation of the x-ray wave-field according to (1) into the paraxial Helmholtz equation and isolating the imaginary part [3], the transport-of-intensity equation [5] is obtained. If we further assume that the sample and detector are sufficiently close to each other such that the intensity evolves over this distance in a way that is linear with the propagation distance, we find

$$I_3(\mathbf{r}_\perp) = I_2(\mathbf{r}_\perp) - \frac{z}{k}\nabla_\perp \cdot \left(I_2(\mathbf{r}_\perp)\nabla_\perp\phi_2(\mathbf{r}_\perp)\right), \tag{5}$$

where z now denotes the propagation distance from the sample to the detector. Due to the limited coherence of laboratory sources, the linear dependency of the intensity on the propagation distance is well justified [3]. Finally, we further discard the typically small cross-term $[\nabla_\perp I_2]\cdot\nabla_\perp\phi_2$, assuming that there are no strong intensity gradients overlaying strong phase gradients and end up with the final expression that relates the phase and intensity behind the sample at the second stage to the intensity in front of the detector [6], as illustrated in Fig. 1 at the third stage

$$I_3(\mathbf{r}_\perp) = I_2(\mathbf{r}_\perp)\left(1 - \frac{z}{k}\nabla_\perp^2\phi_2(\mathbf{r}_\perp)\right). \tag{6}$$

So far, a point-like x-ray source has been assumed. In laboratory setups, the extent of the source eventually diminishes the interference effects necessary for recovering the phase-information of the sample due to blurring, which is also connected to the spatial coherence of the x-ray wave-field. The influence of the source is modeled by an operator $\mathcal{B}_s$. In addition, in order to be sensitive to the interference effects, scintillator-based x-ray detectors with comparably small pixel sizes are commonly used. However, the conversion step from x-ray photons to visible light inside the scintillator introduces another kind of blurring. This contribution is modeled by another operator $\mathcal{B}_d$. Finally, the measured intensity at stage four can be written as

$$I_4(\mathbf{r}_\perp) = \mathcal{B}_d\left[\mathcal{B}_s\left[I_3(\mathbf{r}_\perp)\right]\right]. \tag{7}$$

The reason for separating the influence of the source and the detector is that their response to noise could be modelled differently [7].

III. Reconstruction approaches for tomography

By acquiring an intensity map at different angles around the object, this imaging technique can be extended to CT. In the following, different reconstruction approaches are discussed.

The conventional approach for the reconstruction of tomographic data, without explicitly modeling the propagation of the x-rays, is identical to conventional CT. The most widely used filtered back-projection algorithm that recovers the three dimensional distribution of the sample can be written as

$$\mathbf{x} = \frac{1}{\mu}\mathrm{FBP}\left[-\ln\left(\frac{\mathbf{y}}{I_0}\right)\right], \tag{8}$$

where we denote the discrete representation of our measurement $I_4(\mathbf{r}_\perp)$ as $\mathbf{y}$ and the voxel representation of our volume $x(\mathbf{r})$ as $\mathbf{x}$. Note that the product $\mu\mathbf{x}$ recovers the three-dimensional distribution of the linear attenuation coefficient.

Recently, more advanced model-based iterative reconstruction techniques emerged, which include models for the source and the detector with different noise models [7]. In particular, an intensity-based approach [8] uses the following model

$$\bar{\mathbf{y}} = \mathbf{B}_d\mathbf{B}_s I_0 e^{-\mu\mathbf{A}\mathbf{x}}, \tag{9}$$

where $\bar{\mathbf{y}}$ denotes the expected intensity, the matrix $\mathbf{A}$ models the projection operation and the matrices $\mathbf{B}_s$, $\mathbf{B}_d$ account for

the blurring of source and detector. In addition, a diagonal covariance matrix is established, according to

$$\mathbf{K_y} = \mathbf{D}\left[\mathbf{y}\right]. \tag{10}$$

Thereby, $\mathbf{D}$ denotes the diagonal matrix, with the vector inside the parenthesis on its diagonal. The optimal distribution of the sample $\hat{\mathbf{x}}$ is found by minimizing a cost-function

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \frac{1}{2}\left(\mathbf{y}-\bar{\mathbf{y}}\right)^T \mathbf{K_y}^{-1}\left(\mathbf{y}-\bar{\mathbf{y}}\right) + \lambda \mathcal{R}_{\text{Quadratic}}, \tag{11}$$

where the first term originates from a Gaussian noise model and the second term denotes the regularization. Thereby, we use a quadratic penalty in order not to attribute the reconstruction quality on prior knowledge about the sample.

Explicitly using the interference effects for reconstructing the three-dimensional distribution of the sample has proven to increase contrast significantly. The conventional reconstruction approach consists of two steps. First, the phase information is recovered from the measured intensities. The most commonly-used phase-retrieval algorithm is the single-material algorithm [4]. This algorithm also imposes the homogeneity assumption and can be derived by inserting (3) and (4) into (5) and solving for the trace $t(\mathbf{r}_\perp)$,

$$\mathbf{t} = -\frac{1}{\mu}\ln\left(\mathcal{F}^{-1}\left[\frac{\mathcal{F}\left[\mathbf{y}/I_0\right]}{z\frac{\delta}{\mu}\mathbf{k}_\perp^T\mathbf{k}_\perp + 1}\right]\right), \tag{12}$$

where $\mathcal{F}$ denotes the two-dimensional Fourier-transform along the spatial dimensions and $\mathbf{k}_\perp$ are the respective spatial frequencies. We now denote the pixel representation of the traces by $\mathbf{t}$. The volume can be reconstructed from the traces acquired at different angles by means of the filtered back-projection, written as

$$\mathbf{x} = \mathrm{FBP}\left[\mathbf{t}\right]. \tag{13}$$

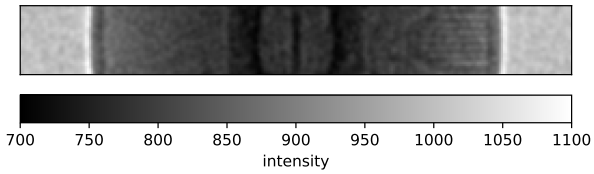
Finally, we can utilize the whole image formation formulation as derived previously. Its discrete form can be written as

$$\bar{\mathbf{y}} = \mathbf{B}_d\mathbf{B}_s I_0 \mathbf{D}\left[e^{-\mu\mathbf{A}\mathbf{x}}\right]\left(\mathbf{1} + z\delta\mathbf{L}\mathbf{A}\mathbf{x}\right), \tag{14}$$

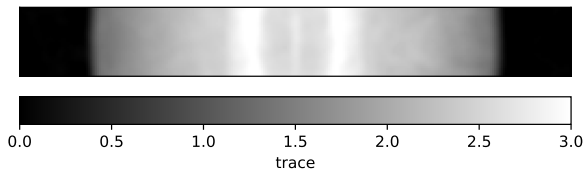
where $\mathbf{L}$ denotes the matrix for the Laplacian operation. This equation directly generates the detected intensity from the three-dimensional distribution of the sample. We also use the covariance matrix (10) and solve for $\hat{\mathbf{x}}$ according to (11).

IV. Results

For the simulation, the source flux was set to $I_0 = 1000$ photons per pixel. A symmetric Gaussian blur with a standard deviation of one pixel each is assumed for the source and detector blur respectively. The parameters defining the propagation of the x-rays are $z = 1$, $\mu = 0.1$ and $\delta = 0.0001$. The pixel size of the detector is 0.01. The dimensions of the phantom are 256×256 with 32 slices. Assuming parallel-beam geometry, the 256 acquired angles are equidistantly distributed between 0 and 180 degrees around the sample. For the generation of the measurements, the phantom is oversampled by a factor of four. The resulting intensity profiles are afterwards binned by the same factor.



(a) Simulated intensity



(b) Retrieved trace

Fig. 2. Projection-based images for the first angle. In (a) the intensity map as simulated according to (15) is depicted, which shows the edge-enhancement effects at the borders of the sample as well as the blurring from the source and detector. In (b), the trace of (a) recovered by (12) is shown.

The simulation is thereby performed according to

$$\mathbf{y} = \mathbf{B}_d\mathrm{Poisson}\left\{\mathbf{B}_s I_0 \exp\left[-\mu\mathbf{A}\mathbf{x}\right]\left(\mathbf{1} + z\delta\mathbf{L}\mathbf{A}\mathbf{x}\right)\right\}, \tag{15}$$

where Poisson draws one realization from a Poisson distribution with a mean value specified in the parenthesis. The first projection is depicted in Fig. 2a.

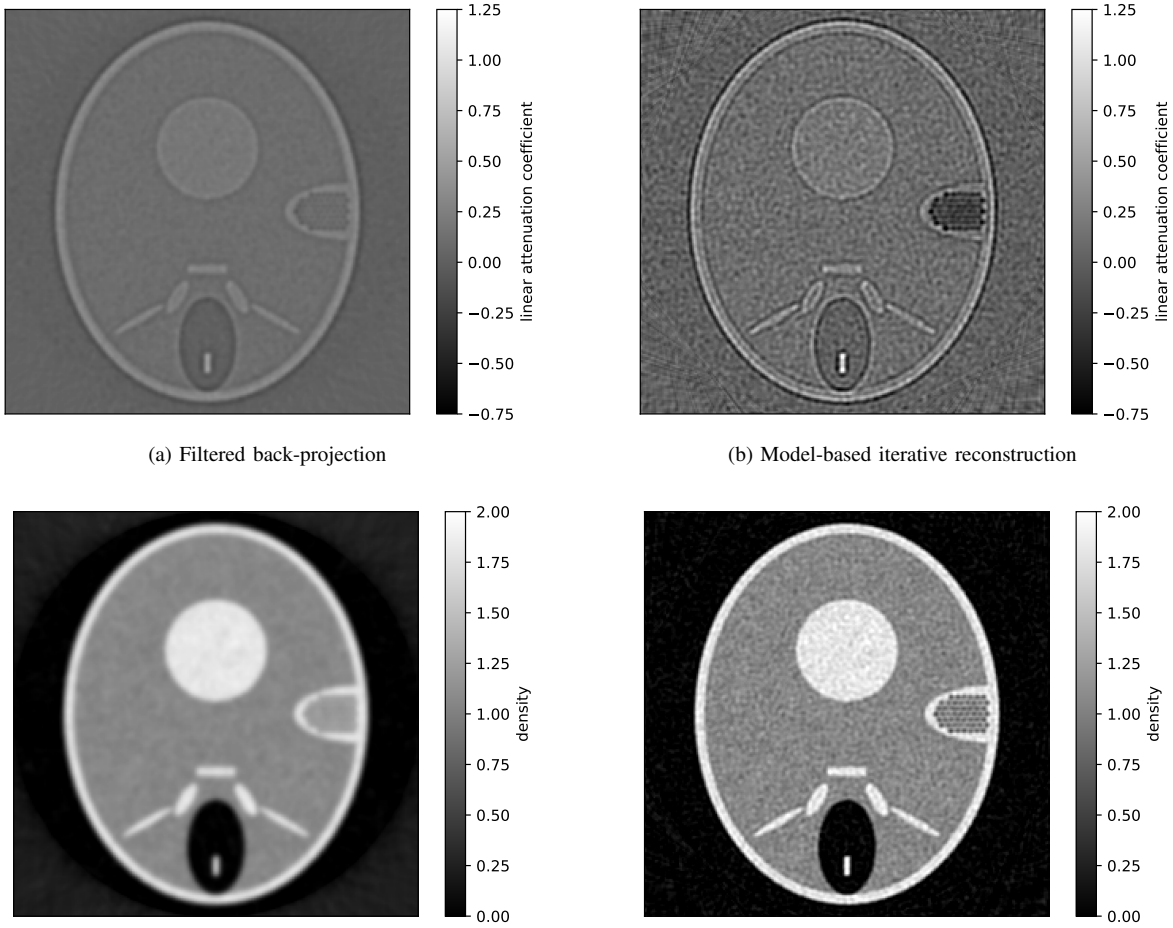
Phase-retrieval is performed according to (12) and the corresponding recovered trace is depicted in Fig. 2b.

The analytical reconstruction of the intensity values according to (8) is depicted in Fig. 3a. As expected, the edge-enhancement effects are visible in the reconstructed volume. The contrast between different densities is low. In addition, the small features on the right can hardly be identified.

The model-based iterative reconstruction approach according to (11) for attenuation can be seen in Fig. 3b. The edge-enhancement effects at the borders become more apparent, as the diminishing influence of the source and the scintillator spread are taken into account during the reconstruction. For instance, the feature at the bottom of the figure becomes clearly visible.

The conventional approach that uses the traces depicted in Fig. 2b for the analytical tomographic reconstruction is seen in Fig. 3c. The contrast is distinctively improved, however at the cost of spatial resolution. Not only the extent of the source and the scintillator blur, but also the phase-retrieval step leads to the blurred character of the reconstructed volume.

Designed to overcome these drawbacks, our model-based iterative approach with the model shown in (14), optimized according to (11), yields the result depicted in Fig. 3d. Thereby, the phase-retrieval is directly integrated into the tomographic reconstruction and the effects of the source and the detector are accounted for. This gives the best reconstruction quality in terms of contrast as well as resolution.



(a) Filtered back-projection

(b) Model-based iterative reconstruction

(c) Single material phase-retrieval and filtered back-projection

(d) Proposed method for phase-contrast tomography

Fig. 3. Comparison of different reconstruction approaches. The conventional filtered back-projection algorithm applied directly onto the measured intensities is depicted in (a). A model-based reconstruction approach, which includes models for source and detector into a statistical reconstruction algorithm, can be seen in (b). The conventional approach for propagation-based phase-contrast computed tomography is shown in (c). Thereby the traces under each angle are recovered with the single-material phase-retrieval algorithm. The retrieved traces are then tomographically reconstructed by means of the filtered back-projection algorithm. Finally the model-based algorithm, which additionally models the propagation of the x-rays, is shown in (d).

V. Conclusion

As the crucial information for propagation-based phase-contrast tomography is encoded in the edge-enhancement effects that are rapidly diminished by any blurring of the system, we believe that correctly modeling the whole image formation process might even be more beneficial than for conventional attenuation CT. We believe that for the ongoing transition of this imaging technique from synchrotron facilities to laboratory sources, our iterative reconstruction approach can recover phase information that would otherwise be lost due to the smearing of laboratory sources and high-resolution scintillator-based x-ray detectors. In addition, including a noise model will prove particularly beneficial for the limited flux available at laboratory sources capable of providing sufficient spatial coherence.

References

[1] A. Bravin, P. Coan, and P. Suortti, "X-ray phase-contrast imaging: from pre-clinical applications towards clinics," *Physics in Medicine and Biology*, vol. 58, no. 1, p. R1, 2013.

[2] P. Cloetens, R. Barrett, J. Baruchel, J.-P. Guigay, and M. Schlenker, "Phase objects in synchrotron radiation hard x-ray imaging," *Journal of Physics D: Applied Physics*, vol. 29, no. 1, p. 133, 1996.

[3] D. M. Paganin, *Coherent X-Ray Optics*. Oxford University Press, 2006.

[4] D. M. Paganin, S. C. Mayo, T. E. Gureyev, P. R. Miller, and S. W. Wilkins, "Simultaneous phase and amplitude extraction from a single defocused image of a homogeneous object," *Journal of Microscopy*, vol. 206, no. 1, pp. 33–40, 2002.

[5] M. R. Teague, "Deterministic phase retrieval: a Green's function solution," *Journal of the Optical Society of America*, vol. 73, no. 11, pp. 1434–1441, 1983.

[6] L. Hehn, K. Morgan, P. Bidola, W. Noichl, R. Gradl, M. Dierolf, P. B. Noël, and F. Pfeiffer, "Nonlinear statistical iterative reconstruction for propagation-based phase-contrast tomography," *APL Bioengineering*, vol. 2, no. 1, p. 016105, 2018.

[7] S. Tilley, J. H. Siewerdsen, and J. W. Stayman, "Model-based iterative reconstruction for flat-panel cone-beam CT with focal spot blur, detector blur, and correlated noise," *Physics in Medicine and Biology*, vol. 61, no. 1, pp. 296–319, 2016.

[8] S. Tilley, M. Jacobson, Q. Cao, M. Brehler, A. Sisniega, W. Zbijewski, and J. W. Stayman, "Penalized-likelihood reconstruction with high-fidelity measurement models for high-resolution cone-beam imaging," *IEEE Transactions on Medical Imaging*, vol. 37, no. 4, pp. 988–999, 2018.

Simulation of a propagation-based phase-contrast imaging system with a compact x-ray light source

Yongjin Sung, Rajiv Gupta, Brandon J. Nelson, Shuai Leng, William S. Graves, and Cynthia H. McCollough

***Abstract*— In this work we use simulations to compare a new inverse Compton scattering compact x-ray light source with a conventional microfocus x-ray source and demonstrate the magnitude of improvement in x-ray phase contrast imaging performance. The simulation framework is validated with an image acquired using a propagation-based x-ray phase contrast imaging system with a microfocus source. The design of the proposed x-ray phase contrast imaging system with the new source is also described. As a compact intermediary between synchrotron and x-ray tube performance, this compact light source could provide high-quality phase contrast imaging in a medical environment with considerably shorter exposure times than microfocus-sources.**

***Index Terms*—Phase-contrast, propagation, source, simulation**

I. Introduction

Current clinical and preclinical imaging using x-ray computed tomography (CT) measures the imaginary component (β) of the complex refractive index $n = 1 - \delta - i\beta$ to determine the x-ray attenuation properties of materials in three-dimensions (3-D). However, linear attenuation coefficients are unable to discriminate between materials of similar composition, as is often the case in imaging soft tissues. The real component (δ), characterizes the phase shifting properties of the material, which when converted into measurable signal in an x-ray phase contrast imaging (XPCI) system, can be used as an additional independent contrast mechanism that has been shown to improve contrast in soft tissues [1, 2]. This presents new opportunities for expanding both preclinical and clinical x-ray imaging applications with improved soft tissue contrast.

In XPCI, refractive index differences among tissue types are the result of small angular deviations (~10^{-4} to 10^{-5} deg) in the x-ray beam path. One requirement for capturing these small angular deviations as contrast in an x-ray image is a highly spatially coherent x-ray source, which is currently unavailable in a clinical setting. While, several successful laboratory XPCI setups have been reported [1, 3-7], a key factor for the lack of clinical adoption has been the unavailability of an x-ray source with beam properties suitable for XPCI that is compact enough for a hospital environment.

Y. Sung is with the College of Engineering and Applied Science, University of Wisconsin-Milwaukee, Milwaukee, WI 53211. R. Gupta is with the Department of Radiology, Massachusetts General Hospital, Boston, MA 02114. W. S. Graves is with the Department of Physics, Arizona State University, Tempe, AZ 85287. B. J. Nelson, S. Leng, and C. H. McCollough are with the Department of Radiology, Mayo Clinic, Rochester, MN 55902 USA. Please direct all correspondence to C. H. McCollough (mccollough.cynthia@mayo.edu).

Focal spot size, photon energy, energy bandwidth, beam divergence and flux are all important beam properties for XPCI and cannot all be adequately attained with the traditional rotating-anode x-ray tubes widely used in the clinic. Grating based XPCI can overcome some of these limitations with the use of an absorption grating immediately after a low-brilliance rotating anode source to split the beam into an array of sources with smaller focal spot sizes on the order of the grating period (20 ~ 50um) [6]. However, this technique requires high precision grating fabrication, which becomes more costly with increasing x-ray energy and size. Additionally, these gratings can limit the size of the imaging field of view. On the other hand, microfocus x-ray tubes with focal spot sizes on the order of 10 μm can provide relatively high spatial coherence, important for XPCI, without using a source grating, but have low photon flux, making them impractical for clinical use. Synchrotron radiation sources, which provide high output, coherent and monoenergetic x-rays, have successfully been used for XPCI [1-4, 7] and are more than a billion times brighter than rotating x-ray tubes. However, their large size, infrastructure cost, and complexity preclude synchrotron use for routine clinical imaging.

This work describes a new x-ray source based on inverse Compton scattering (ICS), referred to as a compact x-ray light source (CXLS), which has performance characteristics intermediate between a large synchrotron and a rotating anode x-ray tube. Similar to a large synchrotron, an ICS source uses a small linear accelerator (linac) to produce relativistic electrons. However, unlike a synchrotron, which uses magnetic undulators, an ICS source uses a laser to produce bursts of x-rays at much lower electron energies due to the short laser period, allowing for a much more compact device only spanning a few meters in length. The resulting x-ray beam shares many similarities with a synchrotron source x-ray beam including high average flux, a narrow cone of divergence, nearly monochromatic bandwidth, tunable photon energy, and a focal spot size of less than 10 μm.

This work describes the key features and performance metrics of the CXLS currently being built by teams at MIT and Arizona State University, along with a detailed simulation framework that demonstrates the effectiveness of a CXLS in XPCI.

II. Methods

A. X-Ray Production Via Inverse Compton Scattering

X-rays are emitted by ICS when a relativistic electron beam collides with a laser pulse at a shallow angle to

increase the rate of photon-electron interactions (Fig. 1). ICS sources are fractions of the size of synchrotron sources due to the lower electron energy required for x-ray production. This is done by using a laser's electromagnetic field as an undulator that has an undulating period more than ten thousand times smaller than that of synchrotrons. The electron energies required for ICS in our setup are produced by a small linear accelerator [8]. While small synchrotron approaches have also been used, their requirement to keep electrons accelerating in a circuit requires additional magnets and energy, and adds cost and complexity to the source compared to our linac-based method.

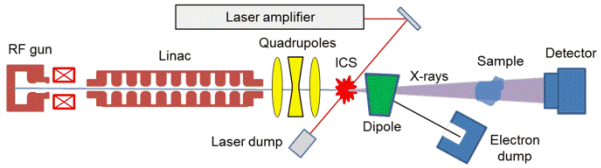


Fig. 1: Schematic of CXLS components including laser, sample, and detector. The CXLS accelerator is comprised of an RF waveguide and solenoid magnet enclosing the photoinjector to produce the electron beam followed by a short linear accelerator (linac). Quadrupole magnets focus the electron beam to a small spot where they collide with the laser to produce x-rays before being rerouted via a dipole magnet to an electron dump. The incident angle of the laser onto the electron beam is exaggerated in this schematic.

The x-ray wavelength is determined by the electron beam energy and laser wavelength and can be tuned over a wide range of energies quickly by changing the electron energy in the linac. The x-rays produced are nearly monochromatic, with a bandwidth of 5%, in contrast to the polychromatic bremsstrahlung x-rays produced by a conventional x-ray tube. The source size of the x-ray beam is approximately the same as the electron beam size (<10 μm), This small source size, combined with monochromatic output provide the coherence necessary for optimum phase-contrast imaging. Complete details regarding the CXLS can be found in Ref. [8]; a summary of the relevant parameters used in the CXLS simulation is presented in Table 1.

TABLE 1. X-PARAMETER FOR THE DESCRIBED CXLS

Parameter	Value	Units
Monochromatic x-ray energy	<45	keV
Time averaged flux	1×10^{15}	Photons/s/sr
Source diameter	10	μm
Source divergence	8	mrad
Photons per pulse	5×10^{7}	Photons per shot
Pulse length	1	Picosecond
Repetition rate	1000	Hz

B. Simulation of X-Ray Phase-Contrast Imaging

In XPCI, the phase alteration of x-rays induced by a sample serves as imaging contrast. Our XPCI simulations consisted of two steps: (i) modeling the interaction of x-rays with the 3-D structure of the sample, where the phase of the incident x-ray is altered, and (ii) modeling the conversion of the phase, which cannot be directly recorded, into modulations of the x-ray amplitude, which can be recorded using a conventional x-ray detector.

In XPCI, the interaction of x-rays with the 3-D structure of the sample can be calculated by solving the wave equation, however directly doing so for large complicated objects is computationally prohibitive. In this work, the first-order Rytov approximation is used to simplify the wave equation and serves as an accurate model for XPCI. This model is applied to both the x-ray-object interaction as well as to free-space propagation [9]. This full-wave approach can be used to simulate XPCI with a finite focal spot, either using a polychromatic or monochromatic beam.

For an imaging specimen, our full–wave model XPCI simulation incorporates the four-dimensional extended cardiac-torso (XCAT) phantom, that is represented with nonuniform rational B-splines (NURBS) [10]. This continuous, nonvoxelized NURBS-based full-wave model allows us to simulate XPCI with high accuracy, without the artifacts that arise in XPCI from having a discretized phantom. For the complex refractive index value, we used the data in "Photon, Electron, Proton, and Neutron Interaction Data for Body Tissues" from the International Commission on Radiation Units and Measurements.

Due to the small focal spot size (10 μm) of the CXLS, a plane wave is assumed incident on the imaged specimen. Additionally, due to the less than 5% energy bandwidth expected for the CXLS, we assume monoenergetic x-rays from the source in our simulation to reduce computation time. Large bandwidth polychromatic x-rays smooth the phase contrast signal, reducing contrast. However, with a less than 5% bandwidth in the CXLS this effect is small compared to that encountered with a polychromatic source. Thus, for the microfocus simulation, the full spectrum was used. This simulated microfocus source used a tungsten anode with 40 kV peak tube potential with a 200 μm beryllium window, typical in experimental microfocus tubes. The CXLS, due to its tunable narrow energy band is left unfiltered. Using a single graphics card (NVIDIA, Tesla K40C), the computation took about 3.3 days to generate a 2048x2048 XPCI image from a NURBS phantom with 8.8 million control points. Further details of the simulation method can be found here [11].

C. Simulation Validation

The simulation framework used in this study was validated with the Mie solution which is the exact solution with minimal approximations [9]. For further

validation, we acquired an image of a bead using a propagation-based XPCI (PB-XPCI) system and compared it with the simulation. Figure 2(a) shows the imaging geometry. The source-to-sample distance was 0.2 m and the sample-to-detector distance was 1.85 m. The sample was a polyethylene beam (Cospheric LLC) with a diameter of 600 μm. A microfocus source (Hamamatsu Corp., L8121-03) was used at a peak tube potential of 40 kV, tube current 100 μA, and an exposure time of 400 s. A custom-built x-ray detector from Radiation Monitoring Devices Inc. with a pixel size of 16 μm was used in the simulation and experimental system. To extract a representative one-dimensional (1-D) profile across the center of the bead in the raw image [Fig. 2(c)], we averaged the pixel intensities radially about the circumference of the bead. For the simulation, we assumed an x-ray source with a photon flux of 3.2×10^{11} photons/s/sr, which approximated the noise in the background region [square in Fig. 2(c) and 2(d)]. The mass attenuation coefficient of the polyethylene bead at different energies was derived from the NIST website of mass attenuation coefficients. However, a variation in the value is expected depending on the exact composition of the material. Thus, we performed a normalization step by adjusting the mass attenuation coefficient and electron densities of the simulated sample such that the magnitude of the simulated 1-D profile is in agreement with the profile derived from experimental data.

D. Comparison of XPCI Using the CXLS and a Microfocus X-Ray Tube

For this simulation, we extracted from the XCAT phantom the NURBS models that represented the right coronary artery, and one of its acute marginal branches. These coronary arteries were virtually embedded in a 3-cm-thick myocardial tissue simulant. We assigned the complex refractive index for water to the arteries and the value for muscle to the surrounding tissue and applied a median filter of size 3×3. This is a non-contrast-enhanced imaging task aimed at comparing soft-tissue contrast the imaging performance of the sources used. The region inside the artery was assumed to be composed of muscle and outside the artery was assumed to be water. This simulated anatomy was used to compare the XPCI performance of both the CXLS and microfocus sources.

III. RESULTS

A. Validation of Simulation Framework: Comparison with Experimental Data

Figure 2 shows the comparison of simulated and experimental images. The accurate match between the 1-D profiles [Fig. 2(b)], together with the similarity between Figs. 2(c) and 2(d), confirms the validity of our simulation framework.

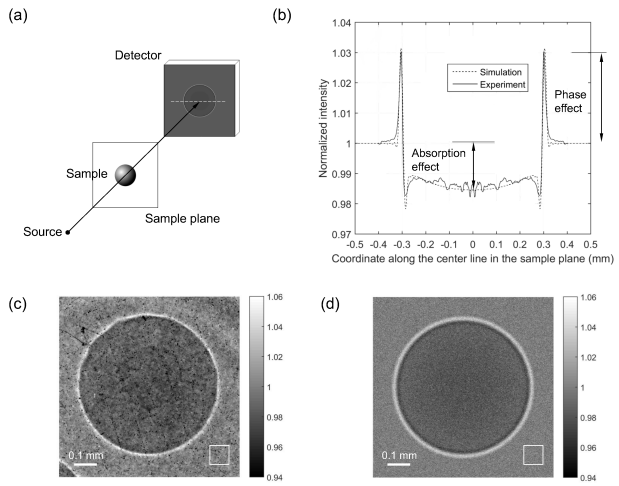


Fig. 2: Validation of the simulation framework used in this study: (a) cone-beam propagation-based x-ray phase-contrast imaging geometry; (b) comparison of the simulated cross section profile with the experimental data; (c) experimentally acquired image of a polyethylene bead; and (d) simulated bead image. The noise texture differs between (c) and (d) due to detector effects that were not modeled in the simulation, such as imperfections in the physical detector.

B. Design of a PB-XPCI System Using CXLS

Figures 3(a)-3(c) compare simulated XPCI images using the CXLS for different exposure times: 1, 3, and 10 s. Figure 3(d) shows the simulated XPCI image using the conventional microfocus x-ray source for an exposure time of 10 s. In Figs. 3(a) and 3(b), the arteries can be identified, although the inset figures show that the noise is as strong as the phase contrast signal. The region inside the artery is slightly brighter than the background because the linear attenuation coefficient of muscle is greater than that of water. Due to the differential phase contrast between artery and water, the boundary of the artery looks darker, assisting in its identification. The inset graph in each figure is a cross sectional profile along the dotted line. The diameters of the arteries are 1.7 and 1.1 mm at the intersections with the dotted line. When the exposure time increases to 10 s, the phase-contrast signals at the edges can be clearly distinguished from the fluctuating background noise. Of note, the microfocus source XPCI image with an exposure time of 10 s [Fig. 3(d)] does not show the arteries because of the high noise level of 39%. A microfocus tube XPCI image of similar noise characteristics to Fig. 3(c) would require > 30,000 s to acquire.

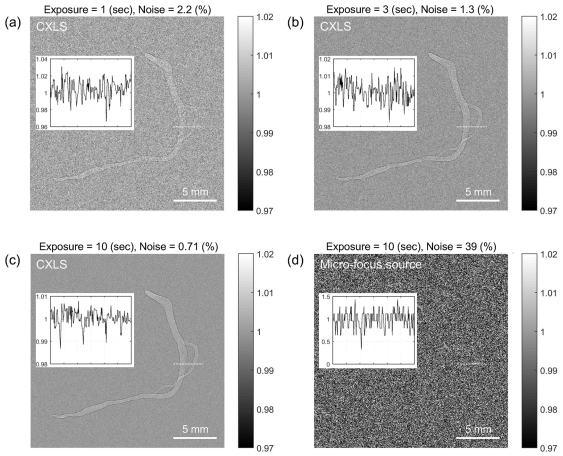


Fig. 3: Simulation of PB-XPCI using the design parameters for the system under construction. The sample is the right coronary artery included in the XCAT digital phantom. (a), (b), and (c) were simulated with CXLS, and (d) with a microfocus x-ray tube source.

IV. Discussion

There is a demonstrated need for improved soft tissue contrast in x-ray applications as shown by the poor specificity and sensitivity of current mammography techniques. While research teams around the world have searched over decades to address this clinical need with XPCI, their success has been limited due to a lack of a coherent source of reasonable size and cost for preclinical and clinical implementation of XPCI. Other compact laser-driven x-ray light sources have been developed, such as the Munich Compact Light Source (MuCLS) by Lyncean Technologies [12]. However, it uses both a linac and small synchrotron to generate and maintain electron energies, which adds size, complexity, and cost to the source. In this work, we present the specifications and operating characteristics of a linac-based CXLS currently being built by teams from MIT and Arizona State University with a PB-XPCI setup. This source generates a bright, monochromatic, coherent x-ray beam using ICS with several features that are attractive for XPCI. The small focal spot size of the CXLS results in minimal geometric blurring and the resultant high spatial coherence preserves the phase information of the object. The high photon flux of CXLS allows image formation to require less than 1/3000 the time required for a microfocus source [Figs. 3(c) and 3(d)]. To test the imaging characteristics of our CXLS, we constructed a simulation framework using the first-order Rytov approximation and validated our simulation technique by showing agreement between simulated and experimentally acquired microfocus images. This simulation method was then used to compare the CXLS and microfocus source. These results showed an improvement in contrast using the CXLS over a microfocus source and the benefit of a greatly reducing acquisition time. For more information on these simulations and their role in our system design see [13].

V. Conclusion

This work presents the key features of a new generation of compact light sources and demonstrated the high-quality phase-contrast images produced from this source via simulations. Further development of such simulations will assist us to better understand important image quality metrics of PB-XPCI, such as noise and contrast. This will allow us to better optimize system characteristics such as propagation distance, photon energy, and flux while maintaining image quality. Additionally, such models can assist us in developing improved forward models for iterative noise reduction techniques and others important in making XPCI more applicable to a clinical reality.
References

References

[1] U. Bonse and M. Hart, "An X-Ray Interferometer," *Applied Physics Letters,* vol. 6, pp. 155-156, 1965.

[2] A. Momose, S. Kawamoto, I. Koyama, Y. Hamaishi, K. Takai, and Y. Suzuki, "Demonstration of X-Ray Talbot interferometry," *Japanese Journal of Applied Physics Part 2-Letters & Express Letters,* vol. 42, pp. L866-L868, Jul 15 2003.

[3] M. Ando, A. Maksimenko, H. Sugiyama, W. Pattanasiriwisawa, K. Hyodo, and C. Uyama, "Simple X-ray dark- and bright-field imaging using achromatic lane optics," *Japanese Journal of Applied Physics Part 2-Letters,* vol. 41, pp. L1016-L1018, Sep 15 2002.

[4] D. Chapman, W. Thomlinson, R. E. Johnston, D. Washburn, E. Pisano, N. Gmur*, et al.*, "Diffraction enhanced x-ray imaging," *Phys Med Biol,* vol. 42, pp. 2015-25, Nov 1997.

[5] C. David, B. Nohammer, H. H. Solak, and E. Ziegler, "Differential x-ray phase contrast imaging using a shearing interferometer," *Applied Physics Letters,* vol. 81, pp. 3287-3289, Oct 21 2002.

[6] F. Pfeiffer, T. Weitkamp, O. Bunk, and C. David, "Phase retrieval and differential phase-contrast imaging with low-brilliance X-ray sources," *Nature Physics,* vol. 2, pp. 258-261, Apr 2006.

[7] S. W. Wilkins, T. E. Gureyev, D. Gao, A. Pogany, and A. W. Stevenson, "Phase-contrast imaging using polychromatic hard X-rays," *Nature,* vol. 384, pp. 335-338, Nov 28 1996.

[8] W. S. Graves, J. Bessuille, P. Brown, S. Carbajo, V. Dolgashev, K. H. Hong*, et al.*, "Compact x-ray source based on burst-mode inverse Compton scattering at 100 kHz," *Physical Review Special Topics-Accelerators and Beams,* vol. 17, Dec 1 2014.

[9] Y. J. Sung, C. J. R. Sheppard, G. Barbastathis, M. Ando, and R. Gupta, "Full-wave approach for x-ray phase imaging," *Optics Express,* vol. 21, pp. 17547-17557, Jul 29 2013.

[10] W. P. Segars, G. Sturgeon, S. Mendonca, J. Grimes, and B. M. W. Tsui, "4D XCAT phantom for multimodality imaging research," *Medical Physics,* vol. 37, pp. 4902-4915, Sep 2010.

[11] Y. J. Sung, W. P. Segars, A. Pan, M. Ando, C. J. R. Sheppard, and R. Gupta, "Realistic wave-optics simulation of X-ray phase-contrast imaging at a human scale," *Scientific Reports,* vol. 5, Jul 14 2015.

[12] E. Eggl, S. Schleede, M. Bech, K. Achterhold, R. Loewen, R. D. Ruth*, et al.*, "X-ray phase-contrast tomography with a compact laser-driven synchrotron source," *Proc Natl Acad Sci U S A,* vol. 112, pp. 5567-72, May 05 2015.

[13] Y. Sung, R. Gupta, B. Nelson, S. Leng, C. H. McCollough, and W. S. Graves, "Phase-contrast imaging with a compact x-ray light source: system design," *J Med Imaging (Bellingham),* vol. 4, p. 043503, Oct 2017.

New X-ray Small-angle Scattering Radiography System Design

Guang Li, Wenxiang Cong, James S. Michaelson, Hong Liu, Ge Wang†

***Abstract*—Abstract X-ray radiography has been used as an effective diagnosis tool in clinical application for decades owing to its high resolution and high sensitivity to specific features such as calcification and hemorrhage. However, the innate disadvantage of x-ray radiography and tomography is the poor soft tissue contrast. Small-angle scattering has been found to provide some special information about the abnormality of some soft tissues. At present, there is no effective small-angle scattering radiography system. In this paper, we propose a new system design dedicated to capturing small-angle scattering radiography images, which carry critical pathological information for differentiation between normal and malignant tissues, in theory. Our design consists of two interlaced gratings so that both the primary flux and Compton scattering photons are effectively blocked to leave the apertures mainly open to small-angle scattering photons. By theoretical analysis and Monte Carlo simulation, it is demonstrated that a small-angle scattering radiography image can be captured directly and effectively with our proposed radiography system.**

Index Terms—Small-angle scattering, grating based x-ray imaging, radiography, pathological studies.

I. Introduction

A variety of commercial x-ray imaging devices have been developed for clinical and preclinical applications [1], [2], [3]. With the advancement of the x-ray source, detector, and reconstruction techniques, the spatial and temporal resolution of these scanners have been significantly improved. Because of these improvements, x-ray imaging devices become more powerful and more popular.

At present, most of these scanners utilize only one contrast mechanism, which is known as x-ray linear attenuation. Consequently, the resulting images have poor soft tissue contrast. Indeed, the x-ray attenuation coefficient is roughly proportional to the fourth power of the atomic number, and biological soft tissues are rather similar in their compositions of chemical elements. To address this disadvantage, researchers have been investigating alternative contrast mechanisms to reveal soft tissue features. One popular option to improve the contrast resolution is to introduce contrast agents containing high-Z elements. Unfortunately, there are many cases in which a contrast agent is not easy to be transferred to the target area and take effect, such as in-vitro cancer specimen imaging.

Guang Li, Wenxiang Cong, and Ge Wang† are with Biomedical Imaging Center, BME/CBIS, Rensselaer Polytechnic Institute, New York, USA (e-mail: {lig10, congw, wangg6}@rpi.edu).

James S. Michaelson is with Department of Pathology, Harvard Medical School, Massachusetts, USA (email: JamesMichaelsonPhD@gmail.com)

Hong Liu is with Center for advanced medical imaging, University of Oklahoma, USA (email: liu@ou.edu)

A large body of literature demonstrated that small-angle scattering imaging provides positive correlation with the presence of breast cancer [4], [5], [6], and different tissues have quite distinct coherent scattering profiles [7] and separate scattering peaks [4]. For example, scattering profiles from normal tissues have a sharper peak at position $x = 0.11\mathring{A}^{-1}$which indicates the presence of adipose component [8], the scattering profiles from malignant tissues present a broad scattering peaks at $x = 0.17\mathring{A}^{-1}$ [9], and the scattering peaks from fibrous collagens are positively related to fibril stacking period and collagen fiber alignment [10]. Because of these findings, small-angle scattering imaging techniques have been used in preclinical studies with a synchrotron source for many years. The emergence of both Talbot interferometers [11] with synchrotron and micro-focus x-ray sources, and Talbot-Lau interferometers [12] with hospital-grade x-ray sources and recent single shot methods [13], [14], proved that small-angle scattering imaging can provide some supplemental information compared with traditional absorption-based CT. We call these indirect detection methods since the small-angle scattering signals collected are obtained indirectly by decomposing the projection data, and we do not know where these signals come from and how these signals are distributed.

Direct coherent scattering systems [14], [15], [16], [17] have a great potential to differentiate subtle differences due to coherent scattering between different tissues. Current scattering-based imaging techniques perform slot scanning aided by a collimated detector. These systems can collect scattering signals at characteristic angles and obtain a 2D material map point by point, or line by line, and hence the scanning efficiency is very low. In the associated imaging geometry, the collimated detector should be deviated from the direction along which the primary x-ray travels. However, in practice it cannot capture all the small-angle scattering signals concurrently with this geometric design because the stronger primary beam should be completely blocked or it will cause serious impact on the small-angle scattering detection.

Here we propose a new dedicated small-angle scattering radiography system composed of two interlaced gratings, which can detect the small-angle scattering signals effectively and efficiently. The key idea behind our design is that both the primary x-rays and Compton scattering photons are basically blocked so that only small-angle scattering x-rays are admitted to go through the paired gratings and reach the detector behind the gratings. In the remainder of this paper, we first describe the theory and method in the second section. Then, we report our numerical results in the parallel beam geometry. Finally, we discuss relevant issues and conclude the paper.

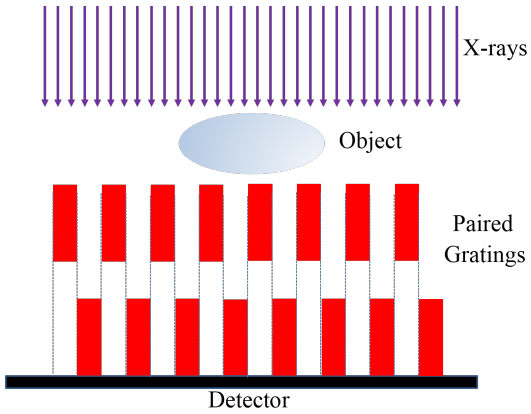


Fig. 1: Paired gratings for small-angle scattering imaging.

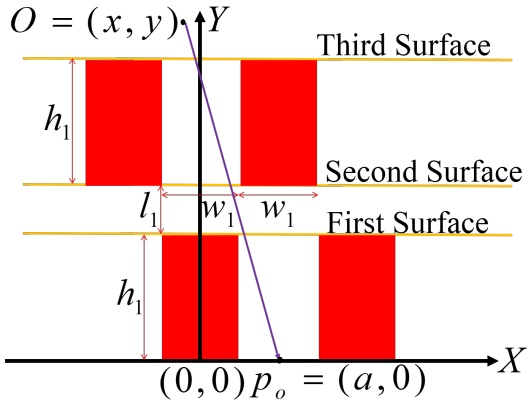


Fig. 2: Geometrical parameters of the paired gratings.

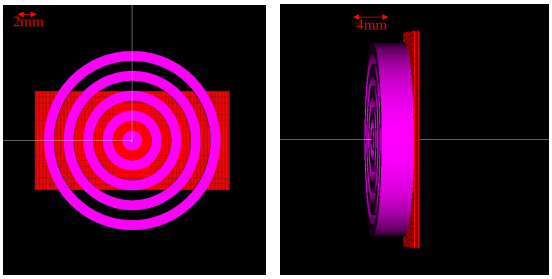
II. METHODS

A. Overall Setup

As our proposed small-angle scattering imaging scheme is new, we start with the parallel beam geometry to verify the feasibility. In the parallel-beam geometry, the schematic design is illustrated in Fig. 1. The most important component in this system is the combination of two absorption gratings (red columns in Fig. 1) made of tungsten or lead or gold. The duty cycle of each grating is 50%. The two gratings are complementary, and all the primary parallel x-rays are completely blocked, and only those rays that pass through the object and scattered at small/specific angles can be detected by the detector behind the gratings. Between the two layers, the gap can be used to modulate the incident angle.

B. System Design and Analysis

Scattered photons are often considered as noise in practice as they randomly diffuse over the field of view, and cannot be effectively analyzed or removed. Here our proposed detection scheme can specifically extract small-angle scattering signals with appropriate geometrical parameters. All the key parameters are shown in Fig. 2. Each unit of the grating contains two cells: solid cell (red component) and hollow cell. The height and width of each unit are h_1 and $2w_1$ respectively, and the widths of the solid cell and the hollow cell are both equal to w_1. The length of the gap between two layers is l_1.



(a) (b)

Fig. 3: Numerical phantom used in our simulation studies. (a) The sagittal view and (b) the coronal view of the phantom.

In order to determine these parameters optimally, we first set up a rectangular coordinate system as shown in Fig. 2. For a scattered photon from a point $O = (x, y)$ to reach a point $p_o = (a, 0)$ on the detector, it must go through the three surfaces shown in Fig. 2 without being attenuated by the grating metal. Hence, the following four equations should be simultaneously satisfied:

$$-\frac{w_1}{2} < a + \frac{(2h_1 + l_1)(x-a)}{y} < \frac{w_1}{2}, \tag{1}$$

$$-\frac{w_1}{2} < a + \frac{(h_1 + l_1)(x-a)}{y} < \frac{w_1}{2}, \tag{2}$$

$$\frac{w_1}{2} < a + \frac{h_1(x-a)}{y} < \frac{3w_1}{2}, \tag{3}$$

$$\frac{w_1}{2} < a < \frac{3w_1}{2}. \tag{4}$$

Equation (1) to (3) respectively means that a scattered photon from point $O = (x, y)$ can pass through the third to first surface, and (4) means that the photon can successfully reach the detector. By simplifying (1) to (4), we can draw the conclusion that if $a \geq h_1 w_1/(h_1 + l_1) + w_1/2$, the following relationship must hold

$$\frac{-\frac{w_1}{2} - a}{2h_1 + l_1} y + a < x < \frac{\frac{w_1}{2} - a}{h_1 + l_1} y + a. \tag{5}$$

Otherwise, we have

$$\frac{\frac{w_1}{2} - a}{h_1} y + a < x < \frac{\frac{w_1}{2} - a}{h_1 + l_1} y + a. \tag{6}$$

Based on (5) and (6), we can derive the energy distribution function as follows:

$$E(\theta) = E(\frac{\tau}{h_1 + l_1}) = \begin{cases} \frac{l_1}{h_1 + l_1}|\tau| & 0 \leq |\tau| < w_1 \\ \frac{l_1}{h_1 + l_1}|w_1| & |\tau| = w_1 \\ 2w_1 - \frac{2h_1 + l_1}{h_1 + l_1}|\tau| & w_1 < |\tau| \leq \frac{2(h_1 + l_1)}{2h_1 + l_1} w_1 \end{cases} \tag{7}$$

From this function, it is seen that the acquired small-angle scattered photons from different angles approximately form a triangular distribution of left-right symmetry. The principal energy cluster is around the angle

$$\theta_p = \frac{w_1}{h_1 + l_1}. \tag{8}$$

Equation (7) can be treated as a response function of the proposed system. Assuming that the scattering signal from a position $M = (x, z)$ along a direction θ is $f(x, z, \theta)$, the captured scattering signal can be represented as

$$F(x') = \sum_{z} \sum_{x} f(x, z, \theta) A(x, z, \theta) E(\theta)|_{x'=x \pm z\theta}. \quad (9)$$

where $A(x, z, \theta)$ is attenuation factor. By modifying the gap l_1, the incident angle can be controlled as we demand. The response function can also be easily to extend to 2-dimensional case as a triangular distribution of circular symmetry. According to (9), we can conclude that if we limit the incident angle to be a small value, we can get a high-resolution small angle scattering radiography image directly.

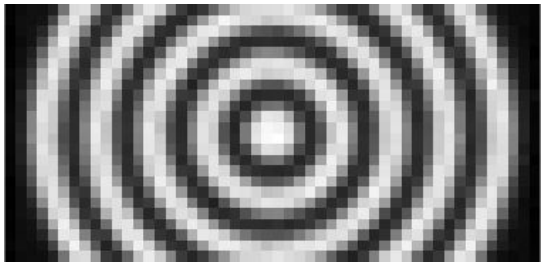
III. RESULTS

EGSnrc is a well-known simulation software whose function is to model the propagation of electrons and photons through matters. Because it relies on the Monte Carlo simulation, it is highly accurate and can be used to validate the feasibility of our small angle scattering radiography system. Our numerical simulations were performed in parallel beam geometry with x-rays of mono-energy at 20 keV. In the simulation, the total number of photons was set to 1.0×10^{10}. The detector had a 0.4 mm pixel size with 50×50 pixels per frame, and each pixel covering 16×16 cells.

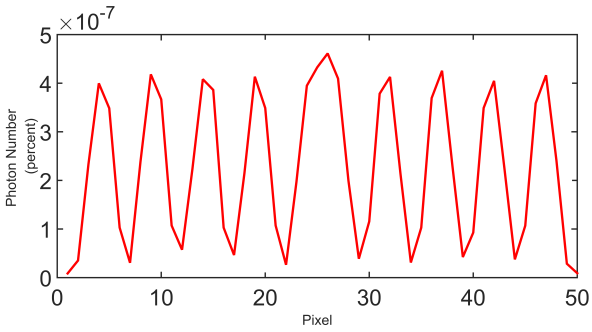
In order to demonstrate that the proposed scheme can be effectively used to detect small angle scattering radiography image, the experimental phantom composed of many concentric cylinders as shown in Fig. 3. The solid parts (pink parts) of the phantom were made of fibro-glandular tissue whose intensity was 1.04 $\mathrm{g/cm^3}$ with the mass fraction of carbon, hydrogen and oxygen being 0.185:0.094:0.68 [18]. The hollow parts were filled with nothing. The solid parts and hollow parts were alternately placed. The radii from the innermost to outermost were 1 mm to 9 mm with common difference of 1 mm. The red component behind the phantom was the double grating layers made of tungsten. The height and width of each cell were 25 μm and 120 μm respectively. The gap between two layers was adjustable and can be used to modify the incident angle of scattering x-rays.

In the following experiments, the length of the gap was 280 μm, meaning that the principal energy angle was about 3.6 degrees. As there were tens of thousands of separate regions involved and the simulation based on EGSnrc was very time-consuming, the grating was not set to the same size as the phantom and it covered the whole phantom in the horizontal direction but only half in the vertical direction. The parallel beam was confined to the same size as the grating. According to the theoretical analysis described above, if the phantom is thin enough and put close enough to the detector (the phantom in this experiment was 4 mm thick, and put 0.5 mm away from the grating), we can get a small resolution.

The acquired scattering image is shown in Fig. 4a and the profile along its central horizontal line is shown in Fig. 4b. It is evident that the image in Fig. 4a is approximately isotropic and very similar to the pattern inside the phantom shown in Fig. 3a, and they are positively related, which means that the

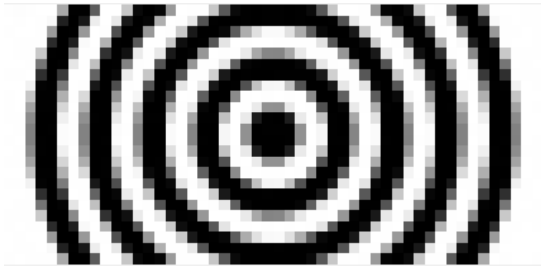


(a)

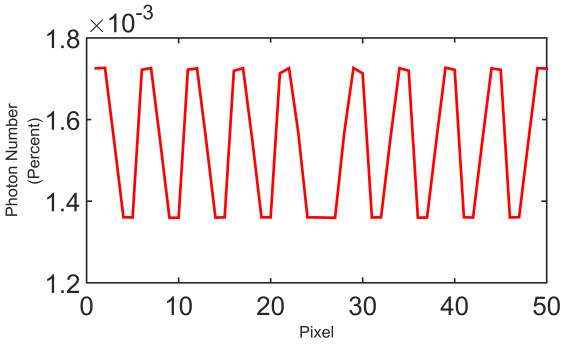


(b)

Fig. 4: Representative scattering results. (a) The acquired scattering signals from the phantom shown in Fig. 4 and (b) the profile along the central horizontal line.



(a)



(b)

Fig. 5: Representative absorption results. (a) The absorption image from the phantom and (b) the profile along the central horizontal line.

solid parts of the phantom correspond to the bright parts of the acquired scattering image.

Projective raw data from the conventional CT should be negatively correlated with the line integrals through the phantom, which means solid parts of the phantom correspond to the dark parts of the projective raw data. In our simulation, projective

views were simulated with the same number of x-ray photons. A typical projection view is shown in Fig. 5a, and it is intuitive that this image contains a pattern very similar to the scattering image shown in Fig. 4a except that their corresponding pixels were on reversed gray scales. This difference demonstrates that the acquired image from our proposed scheme comes from scattering signals. By comparing the profiles in Fig. 4b with Fig. 5b, it is seen that the profile from the absorption image is more like a piecewise constant function which is reasonable in parallel beam geometry, but the profile from the acquired scattering image has sharp peaks and valleys which is a little different from the absorption image. It is because each voxel in the phantom is diffused over a small neighboring area by multiplying the response function.

IV. DISCUSSION AND CONCLUSIONS

By theoretical analysis and Monte Carlo simulation, we have demonstrated that the proposed small angle radiography system can get small angle scattering radiography image directly and effectively in parallel-beam geometry. In the experiments, the height of each unit is set to 120 μm to block all the primary x-ray of 20 keV. We can adjust the gap to set the incident principal angle. For example, in the experiment we want to the principal angle to be about 3.6 degrees, the gap was set to 280 μm. The resolution of the directly detected scattering data depend on the angle range and the distance from the sample to the grating. If we put the sample close enough to the grating, we can get the similar high-resolution scattering image in comparison with attenuation image which means the proposed scheme can only be used to scan thin samples for now. Fortunately, we have also derived the response function of the system which means we can build a new reconstruction model to get the tomography images of the small angle scattering in the future. There are several dark-field forward models in the literature [19], [20] which have provided feasible ways to address this problem. These existing methods will be helpful to our further study for small scattering tomography based on our proposed detection scheme. Eventually, small-angle scattering tensor tomography should be of our interest [21].

Though all the experiments is based on parallel-beam, it is not very complicated to extend the scheme to fan-beam and cone-beam geometry as we can modify the geometrical parameters to control the incidence range of incoming scattering x-ray. Theoretical analysis demonstrates that the response function should be isotropic, but the experiment result showed that isotropy is only approximate. That is because the two-dimensional grating is not isotropic either. The better design for the single unit should be hexagonal or circular. Pathologically, small-angle scattering imaging can provide sensitive and specific information about surgical specimens, and guide surgical procedures on site. Our proposed technology should be robust and cost-effective. Hence, we are interested in developing the technology further and prototyping it as a novel tool for niche surgical and pathological applications.

In summary, we have proposed a new small-angle scattering radiography system featured by a pair of complementary x-ray gratings. This approach is focused on small-angle scattering signal detection, involves no mechanical stepping, and should be practical if it is fully developed.

REFERENCES

[1] M. J. Paulus, S. S. Gleason, S. J. Kennel, P. R. Hunsicker, and D. K. Johnson, "High resolution x-ray computed tomography: an emerging tool for small animal cancer research," *Neoplasia*, vol. 2, no. 1, pp. 62–70, 2000.

[2] M. Ding, A. Odgaard, and I. Hvid, "Accuracy of cancellous bone volume fraction measured by micro-ct scanning," *Journal of biomechanics*, vol. 32, no. 3, pp. 323–326, 1999.

[3] C. T. Badea, S. M. Johnston, Y. Qi, and G. A. Johnson, "4d micro-ct for cardiac and perfusion applications with view under sampling," *Physics in Medicine & Biology*, vol. 56, no. 11, p. 3351, 2011.

[4] D. Cunha, O. Oliveira, C. Pérez, and M. Poletti, "X-ray scattering profiles of some normal and malignant human breast tissues," *X-Ray Spectrometry*, vol. 35, no. 6, pp. 370–374, 2006.

[5] O. R. Oliveira, A. L. Conceicao, D. M. Cunha, M. E. Poletti, and C. A. Pelá, "Identification of neoplasias of breast tissues using a powder diffractometer," *Journal of radiation research*, vol. 49, no. 5, pp. 527–532, 2008.

[6] M. Fernández, J. Keyriläinen, R. Serimaa, M. Torkkeli, M.-L. Karjalainen-Lindsberg, M. Leidenius, K. von Smitten, M. Tenhunen, S. Fiedler, A. Bravin, *et al.*, "Human breast cancer in vitro: matching histo-pathology with small-angle x-ray scattering and diffraction enhanced x-ray imaging," *Physics in Medicine & Biology*, vol. 50, no. 13, p. 2991, 2005.

[7] G. Kidane, R. Speller, G. Royle, and A. Hanby, "X-ray scatter signatures for normal and neoplastic breast tissues," *Physics in Medicine & Biology*, vol. 44, no. 7, p. 1791, 1999.

[8] A. Tartari, E. Casnati, C. Bonifazzi, and C. Baraldi, "Molecular differential cross sections for x-ray coherent scattering in fat and polymethyl methacrylate," *Physics in Medicine & Biology*, vol. 42, no. 12, p. 2551, 1997.

[9] M. Poletti, O. Gonçalves, and I. Mazzaro, "Coherent and incoherent scattering of 17.44 and 6.93 kev x-ray photons scattered from biological and biological-equivalent samples: characterization of tissues," *X-Ray Spectrometry*, vol. 31, no. 1, pp. 57–61, 2002.

[10] S. Sidhu, G. Falzon, S. Hart, J. Fox, R. Lewis, and K. Siu, "Classification of breast tissue using a laboratory system for small-angle x-ray scattering (saxs)," *Physics in Medicine & Biology*, vol. 56, no. 21, p. 6779, 2011.

[11] A. Momose, T. Takeda, Y. Itai, and K. Hirano, "Phase–contrast x–ray computed tomography for observing biological soft tissues," *Nature medicine*, vol. 2, no. 4, pp. 473–475, 1996.

[12] F. Pfeiffer, M. Bech, O. Bunk, P. Kraft, E. F. Eikenberry, C. Brönnimann, C. Grünzweig, and C. David, "Hard-x-ray dark-field imaging using a grating interferometer," *Nature materials*, vol. 7, no. 2, p. 134, 2008.

[13] E. E. Bennett, R. Kopace, A. F. Stein, and H. Wen, "A grating-based single-shot x-ray phase contrast and diffraction method for in vivo imaging," *Medical physics*, vol. 37, no. 11, pp. 6047–6054, 2010.

[14] H. Wen, E. E. Bennett, M. M. Hegedus, and S. C. Carroll, "Spatial harmonic imaging of x-ray scatteringinitial results," *IEEE transactions on medical imaging*, vol. 27, no. 8, pp. 997–1002, 2008.

[15] K. Kern, L. Peerzada, L. Hassan, and C. MacDonald, "Design for a coherent-scatter imaging system compatible with screening mammography," *Journal of Medical Imaging*, vol. 3, no. 3, p. 030501, 2016.

[16] G. Harding, "X-ray diffraction imaginga multi-generational perspective," *Applied Radiation and Isotopes*, vol. 67, no. 2, pp. 287–295, 2009.

[17] A. Castoldi, C. Ozkan, C. Guazzoni, A. Bjeoumikhov, and R. Hartmann, "Experimental qualification of a novel x-ray diffraction imaging setup based on polycapillary x-ray optics," *IEEE Transactions on Nuclear Science*, vol. 57, no. 5, pp. 2564–2570, 2010.

[18] M. Antoniassi, A. Conceicao, and M. E. Poletti, "Compton scattering spectrum as a source of information of normal and neoplastic breast tissues' composition," *Applied Radiation and Isotopes*, vol. 70, no. 7, pp. 1451–1455, 2012.

[19] W. Cong, F. Pfeiffer, M. Bech, and G. Wang, "X-ray dark-field imaging modeling," *JOSA A*, vol. 29, no. 6, pp. 908–912, 2012.

[20] S. Pang, Z. Zhu, G. Wang, and W. Cong, "Small-angle scatter tomography with a photon-counting detector array," *Physics in Medicine & Biology*, vol. 61, no. 10, p. 3734, 2016.

[21] A. Malecki, G. Potdevin, T. Biernath, E. Eggl, K. Willer, T. Lasser, J. Maisenbacher, J. Gibmeier, A. Wanner, and F. Pfeiffer, "X-ray tensor tomography," *EPL (Europhysics Letters)*, vol. 105, no. 3, p. 38002, 2014.

High sensitivity and fast full-field X-ray fluorescence CT imaging method and its experimental results

Siyuan Zhang, Liang Li, Zhiqiang Chen

***Abstract*–X-ray fluorescence computed tomography (XFCT), which is a quantitative imaging technique that collects x-ray fluorescent photons emitted from the target element, has higher sensitivity when detecting high Z elements compared with conventional transmission CT. XFCT experiments were first carried out on the first generation CT system with limited detection efficiency which resulted in a long scanning time. To overcome this problem, recent studies focus more on the XFCT implement with fan-beam (or cone-beam) source and the linear-array (or flat panel) detector, which significantly accelerate the data acquisition procedure. Another method to reduce the scanning time is to collect the projection data by sparse sampling strategy and to reconstruct the image based on the few-view projection data and appropriate priori hypotheses. In this study, we presented a fast few-view XFCT imaging method with polychromatic source and pinhole collimation. A linear-array photon counting detector operating on the multi-energy-bin mode was used for data acquisition and corrections. A phantom containing Gd insertions was irradiated by a polychromatic cone beam produced by a conventional x-ray tube (125kVp, 25mA). To cover a 360° angular range, the phantom was rotated in 8° intervals to obtain 45 projections with an integral time of 20s per angle. An expectation maximization-total variation (EM-TV) iteration algorithm was applied for image reconstruction. The purpose of this study is to demonstrate a rapid XFCT imaging method with sparse sampling strategy and few-view reconstruction algorithm. The experimental results indicate that it is a promising method to reduce the scanning time and total dose of XFCT while maintaining its sensitivity and image quality.**

I. Introduction

X-ray fluorescence computed tomography (XFCT) is a quantitative imaging technique that collects x-ray fluorescent photons emitted from the target element. Compared with conventional transmission CT techniques, XFCT has higher sensitivity when determining the high Z element in vivo and can produce the image of the target element with higher contrast.

It is generally agreed that the first XFCT was performed in 1986 using the x-ray sources produced by NSLS [1]. However, although the synchrotron can produce monochromatic and polarized x-ray source which is ideal to XFCT [2]-[4], it is still impractical for clinical research due to its high cost and limited access. In recent years, XFCT experiments have been carried out more on the benchtop systems with polychromatic x-rays produced by conventional x-ray tubes [5]-[8].

One of the biggest challenges current XFCT meets is the low signal intensity caused by the limited detection efficiency. At first, XFCT experiments were carried out on first generation CT devices which scan a whole phantom by translating the pencil-beam source. This imaging technique enables high signal quality but results in long scanning time (several hours). Therefore, recent studies focus more on the XFCT implement with fan-beam (or cone-beam) source and the linear-array (or flat panel) detector that significantly accelerate the data acquisition procedure [9]-[11].

In our recent studies, we presented a set of simulations and experiments based on a full-field fan-beam XFCT device with pinhole collimation and photon counting detectors [13]-[14]. In that works, we obtained a 256*360 projection data in a 30 min scan procedure with the current of the x-ray tube relatively low (500μA), and then successfully reconstructed 1% gadolinium (Gd) in an acrylic phantom which was 3 cm in diameter. However, it is noticed that the signal intensity and the signal-to-noise-ratio (SNR) of the raw data of each projection angle is not such sufficient for image reconstruction in case of low dose and high sampling density. Therefore, a few-view data acquisition strategy with sparse angle sampling and higher dose in scanning each projection angle might be more suitable for current XFCT experiments.

In this study, we presented a few-view XFCT imaging experiment with polychromatic source and pinhole collimation. A linear-array photon counting detector operating on the energy bin mode was used for data acquisition and scatter correction. The phantom containing Gd insertions was irradiated by a polychromatic cone beam produced by a conventional x-ray tube (125kVp, 25mA). To cover a 360° angular range, the phantom was rotated in 8° intervals to obtain 45 projections with an integral time of 20s per angle. An expectation maximization-total variation (EM-TV) iteration algorithm was applied for image reconstruction. Results show that 0.25% Gd solutions can be successfully reconstructed by using only 45 projections. The purpose of this study is to demonstrate a rapid XFCT imaging method with sparse sampling strategy and few-view reconstruction algorithm based on the priori hypotheses. The experimental results indicate that it is a promising method to reduce the scanning time and total dose of XFCT while maintaining its sensitivity and image quality, which may further improve its practicability in clinical research.

Manuscript received January 5, 2018. This work was partially supported by the grants from NSFC 11775124, 61571256, 11525521, and The National Key Research and Development Program of China, 2017YFC0109100.

Siyuan Zhang, Liang Li and Zhiqiang Chen are all with Department of Engineering Physics, Tsinghua University, Beijing, 100084, China & Key Laboratory of Particle & Radiation Imaging (Tsinghua University), Ministry of Education, 100084, China.

Corresponding author: Liang Li, lliang@tsinghua.edu.cn

II. Materials and Methods

A. XFCT System with Pinhole Design

The experimental setup in this work (shown in Fig. 1) is based on a Full-field fan-beam XFCT imaging system. The data acquisition procedure can be divided into the following steps:

i) One incident beam $I_0(E)$ with the energy higher than the k-edge of the fluorescent material reaches arbitrary point *B* in the phantom after entering the phantom at point *A*.

ii) Assume that the concentration of fluorescence material at point *B* is $\rho(B)$ and the fluorescence yield is ω, then several fluorescent photons may be stimulated and emitted from point B and the emitting angle is isotropic;

iii) A part of fluorescent photons emitted from point *B* will finally reach the *p*-th pixel of the detector after leaving the phantom at point *C* and then crossing through the pinhole.

Therefore, according to our previous work, the number of the fluorescent photons collected by the *p*-th pixel that are emitted from point B can be expressed as:

$$\begin{aligned} N_{B\to p} &= \int_{Aug} N_{B\to D}\, d\Omega \\ &= \int_{Aug}\int_0^{\infty}\left[I_0(E) e^{-\int_A^B \mu(E,l)dl}\, \omega\rho(B)\mu_{mpe}(E,B) \right. \\ &\left. \times e^{-\int_B^C \mu(E_{xrf},l)dl} \right] dE d\Omega \end{aligned} \tag{1}$$

where $\mu(E,l)$ is the linear attenuation coefficient, $\mu_{mpe}(E,B)$ is the photoelectric mass absorption coefficient at point *B* and $\Omega_{B\to D}$ is the solid angle from *B* to *D*.

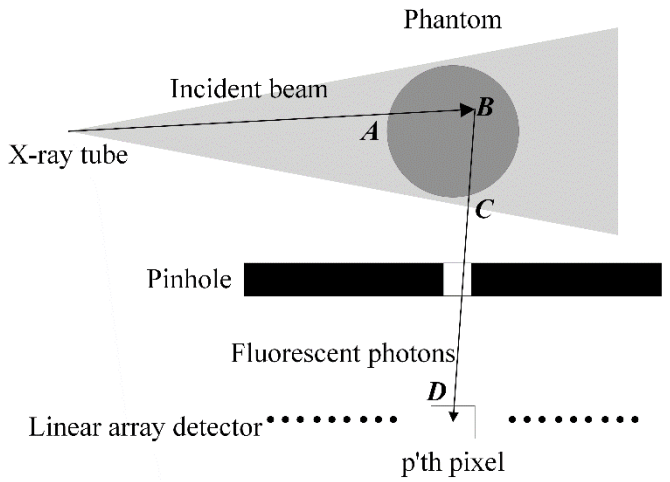


Fig. 1. The geometry of the XFCT imaging system.

B. Data Acquisition

As the phantom is irradiated by the polychromatic source, the fluorescent signal will be contaminated by the continuous Compton scatter background. In order to extract the fluorescent signal from the scatter background, the photon counting detector is working at the energy bin mode and a triple-energy window method is applied for the scatter correction [12].

In this work, Gd (k-edge=50.23 keV) was used as the fluorescent material. As the energy of its Kα fluorescence is about 42.5 keV, the energy bin was set at 39-45 keV for fluorescence detection. Another two energy bins were respectively set at 33-39 keV and 45-51 keV. The middle energy bin records both fluorescent photons and scattered photons with energies between 39 keV and 45 keV, while the other two bins only record the scattered photons. Therefore, the number of scattered photons counted in the middle energy bin can be estimated as

$$Csca_{middle} = \frac{\frac{Nsca_{middle}}{Nsca_{above}} Csca_{above} + \frac{Nsca_{middle}}{Nsca_{below}} Csca_{below}}{2} \tag{2}$$

where *Nsca* denotes the relative number of scattered photons in each energy bin which can be calculated according to KN formula and the incident spectrum. More details about this scatter correction method has been presented in our recent work [13].

C. Image Reconstruction

As XFCT is a quantitative imaging technique that obtain the distribution of a certain element, a pixel value in the reconstructed image can be viewed as the concentration of the target element at corresponding position. According to the system geometry and the transport process of the fluorescent photons mentioned above, the relationship between the projection data and the pixel value of the reconstructed image can be written as

$$p_i = \sum_j M_{ij} f_j \tag{3}$$

where *p* is the projection data after scatter correction, *f* is the pixel value, *M* is the system matrix which can be calculated according to eq. (1). It should be explained that the attenuation map is necessary for accurate calculation of the system matrix and it can be obtained by transmission CT images. In this work, the attenuation map of the phantom was assumed to be known. Further research about the attenuation correction method will be discussed in our future work.

Consider that the fluorescence signal is affected by the statistical noise which obeys Poisson distribution, the well-known maximum likelihood expectation maximization (MLEM) algorithm is applied for image reconstruction [15]. Each iteration step of MLEM algorithm can be written as

$$f_j^{EM(k+1)} = \frac{f_j^{EM(k)}}{\sum_{i=1}^n M_{ij}} \sum_{i=1}^n \frac{g_i}{\sum_{j'=1}^n M_{ij'} f_{j'}^{EM(k)}} M_{ij} \tag{4}$$

As the projection data is sparse, the solution of eq. (3) is not unique. Therefore, the total-variation (TV) minimization [16] was added as an extra constraint and a TV gradient descent iteration step was added after each MLEM iteration step:

$$f_j^{TV(k+1)} = f_j^{EM(k+1)} - \alpha \left\| \vec{f}^{TV(k)} - \vec{f}^{EM(k)} \right\|_2 \times \nabla \left\| f_j^{EM(k+1)} \right\|_{TV} \tag{5}$$

where α is the parameter that controls the speed of the gradient descent.

III. Experiments

A. Phantom

The phantom used in the experiment is a PMMA cylinder with 50mm in diameter containing Gd insertions (shown in Fig. 2). The insertions consist of Gd solutions with concentrations from 0.25% to 2%. Two of the insertions were respectively made of water or air as a contrast.

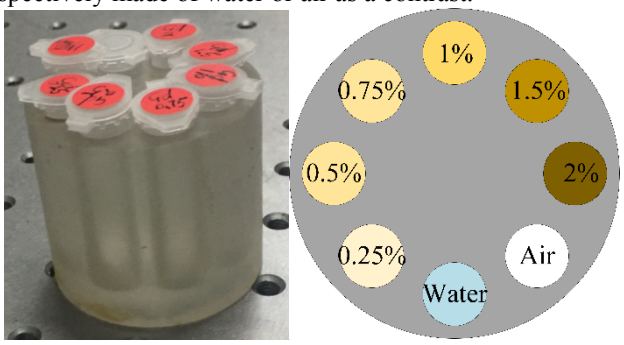


Fig. 2. Geometry of the phantom used in the experiments.

B. Experimental Device

The schematic of the top view of the experimental setup is shown in Fig. 3. The x-ray beam was generated by a conventional x-ray tube (125kVp, 25mA) and then collimated by lead bricks to produce a cone beam. The distance between the x-ray source and the rotation center is about 400mm. A linear-array photon counting detector covered by lead shields operating on the energy bin mode with single pinhole collimation was placed on one side of the rotation stage for data acquisition and scatter correction. The distance between the pinhole and the rotation center is about 80mm and the distance between the pinhole and the center of the detector is 64mm. In order to cover a 360° range, the phantom was rotated in 8° intervals to obtain 45 projections with scanning time set at 20s per angle.

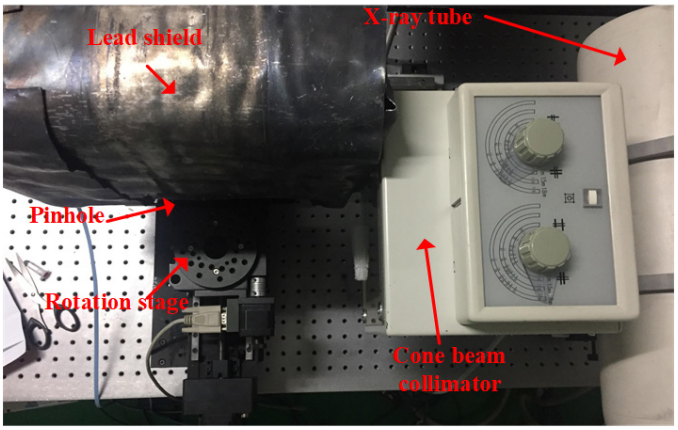


Fig. 3. Schematic of the experimental setup.

C. Results

The projection data of 45 angles collected by different energy bins and the data after scatter correction are shown in Fig. 4. It can be seen that the scattered photons in the middle energy bin are well subtracted after the scatter correction. However, this scatter correction method does not decrease the statistical noise caused by the scattered background. Therefore, further methods may focus on reducing the number of scattered photons emitted from the phantom by spectrum optimization or application of polarized x-ray source.

The images reconstructed by 15 projections or 45 projections are shown in Fig. 5. The reconstruction results indicate that 0.25% Gd solutions can be well reconstructed using only 45 projections. And if only 15 of the 45 projections are used for image reconstruction, although the image is contaminated by higher statistical noise and more artifacts, 0.5% Gd solutions can still be seen clearly in the images, which validates the feasibility of the few-view scan strategy.

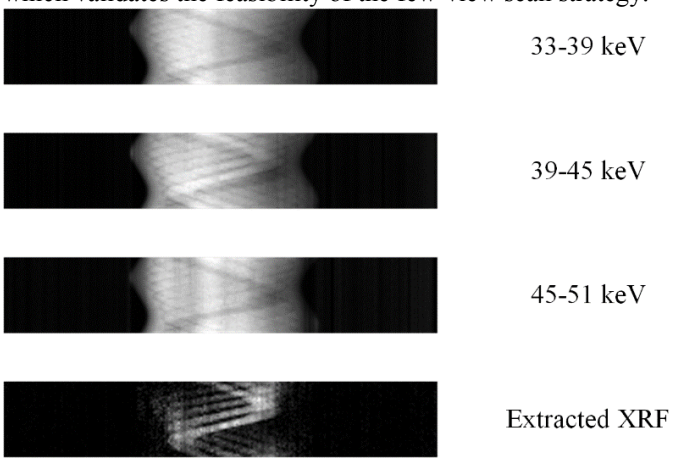


Fig. 4. Projection data detected by different energy bins and the fluorescent signal in the middle energy bin after scatter correction.

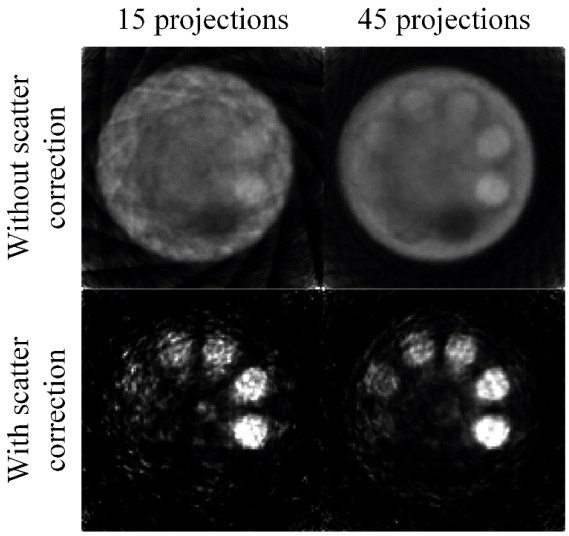


Fig. 5. XFCT images reconstructed from 15 projections and 45 projections.

To further investigate the image quality, the relationship between the Gd concentration and the corresponding pixel values in the reconstructed images (shown in Fig. 6). The pixel values are converted to concentration and 2% Gd solutions are used for calibration. It can be seen that the signal intensity in the reconstructed image after scatter correction is linear with the concentration and the linear relationship fits better with more projections.

The contrast-to-noise ratio (CNR) of Gd insertions with different concentrations was also calculated and is shown in Fig. 7. The CNR is calculated as follows:

$$CNR_{ROI} = \frac{Mean_{ROI} - Mean_{BG}}{\sigma_{BG}} \quad (6)$$

where *ROI* and *BG* respectively denote the sampling points of signal and background.

It can be seen from Figs. 5 and 7 that the scatter correction method effectively improves the contrast of the target element, which helps a lot when detecting the contrast agent with very low concentrations.

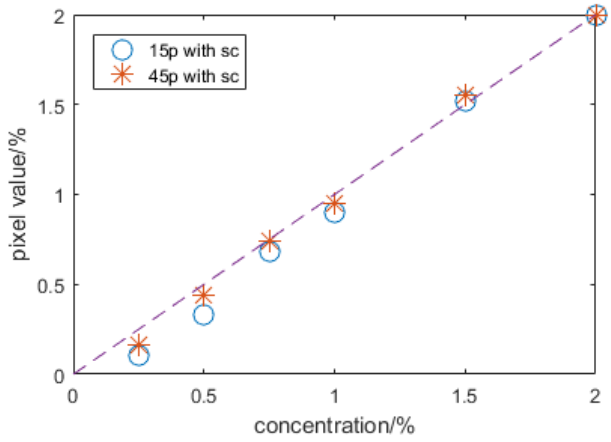


Fig. 6. Relationship between the Gd concentration and the corresponding pixel values.

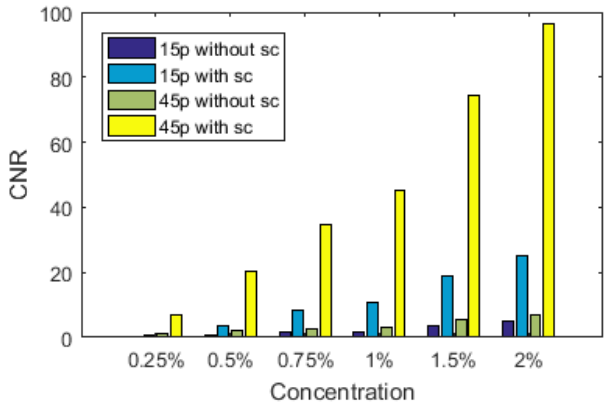


Fig. 7. CNR of Gd insertions of reconstructed images shown in Fig. 5.

IV. Conclusion

In this study, we presented and discussed our recent experimental results about the few-view XFCT imaging with pinhole collimation. The phantom was irradiated by the cone beam polychromatic source and was rotated in 8° intervals to obtain 45 projections with scanning time set at 20s per angle, which means the total scan time is only 15 min. The triple-energy-bin method was applied for scatter correction and the EM-TV iteration algorithm was applied for image reconstruction. Results show that 0.25% Gd solutions can be successfully reconstructed by using only 45 projections. This study indicates that the few-view cone-beam XFCT imaging strategy can significantly reduce the scanning time and total dose while maintaining its sensitivity and image quality. Further research about the system optimization and the scatter reduction algorithm will be discussed in our future works.

References

[1] Boisseau P, Grodzins L. Fluorescence tomography using synchrotron radiation at the NSLS[J]. Hyperfine Interactions, 1987, 33(1-4):283-292.

[2] Takeda T, Maeda T, Yuasa T, et al. Fluorescent scanning x‐ray tomography with synchrotron radiation[J]. Review of Scientific Instruments, 1995, 66(2):1471-1473.

[3] Sasaya T, Sunaguchi N, Thetlwin T, et al. Dual-energy fluorescent x-ray computed tomography system with a pinhole design: Use of K-edge discontinuity for scatter correction[J]. Scientific Reports, 2017, 7:44143.

[4] Rust G F, Weigelt J. X-ray fluorescent computer tomography with synchrotron radiation[J]. IEEE Transactions on Nuclear Science, 1998, 45(1):75-88.

[5] Cheong S K, Jones B L, Siddiqi A K, et al. X-ray fluorescence computed tomography (XFCT) imaging of gold nanoparticle-loaded objects using 110 kVp x-rays[J]. Physics in medicine and biology, 2010, 55(3): 647.

[6] Kuang Y, Pratx G, Bazalova M, et al. Development of XFCT imaging strategy for monitoring the spatial distribution of platinum-based chemodrugs: instrumentation and phantom validation.[J]. Medical Physics, 2013, 40(3):S132-S133.

[7] Manohar N, Reynoso F J, Diagaradjane P, et al. Quantitative imaging of gold nanoparticle distribution in a tumor-bearing mouse using benchtop x-ray fluorescence computed tomography.[J]. Scientific Reports, 2016, 6:22079.

[8] Kuang Y, Pratx G, Bazalova M, et al. First demonstration of multiplexed x-ray fluorescence computed tomography (XFCT) imaging[J]. IEEE transactions on medical imaging, 2013, 32(2): 262-267.

[9] Manohar N, Jones B L, Cho S H. Improving x‐ray fluorescence signal for benchtop polychromatic cone‐beam x‐ray fluorescence computed tomography by incident x‐ray spectrum optimization: A Monte Carlo study[J]. Medical physics, 2014, 41(10).

[10] Jones B L, Cho S H. The feasibility of polychromatic cone-beam x-ray fluorescence computed tomography (XFCT) imaging of gold nanoparticle-loaded objects: a Monte Carlo study[J]. Physics in Medicine & Biology, 2011, 56(12):3719-3730.

[11] Cong W, Shen H, Cao G, et al. X-ray fluorescence tomographic system design and image reconstruction.[J]. Journal of X-ray science and technology, 2013, 21(1):1.

[12] Ichihara T, Ogawa K, Motomura N, et al. Compton scatter compensation using the triple-energy window method for single- and dual-isotope SPECT.[J]. Journal of Nuclear Medicine Official Publication Society of Nuclear Medicine, 1993, 34(12):2216.

[13] Li L, Zhang S, Li R, et al. Full-field fan-beam x-ray fluorescence computed tomography with a conventional x-ray tube and photon-counting detectors for fast nanoparticle bioimaging[J]. Optical Engineering, 2017, 56(4): 043106-043106.

[14] Zhang S, Li L, Li R, et al. Full-field fan-beam x-ray fluorescence computed tomography system design with linear-array detectors and pinhole collimation: a rapid Monte Carlo study[J]. Optical Engineering, 2017, 56(11):1.

[15] Lange K, Carson R. EM reconstruction algorithms for emission and transmission tomography[J]. Journal of Computer Assisted Tomography, 1984, 8(2):306.

[16] Sidky E Y, Pan X. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization[J]. Physics in Medicine & Biology, 2008, 53(17):4777.

A multi pass approach to reduce cone beam artifacts in a circular orbit cone beam CT system

Chulhee Han, and Jongduk Baek*

***Abstract*—We propose a multi pass approach to reduce cone beam artifacts in a circular orbit cone beam computed tomography (CBCT) system. Employing a large 2D detector array reduces the scan time, however, produces cone beam artifacts in FDK reconstruction. While the two-pass algorithm proposed by Hsieh is effective, when the bone density is moderate and cone angle is large, the correction performance is degraded. In this work, we treated the cone beam artifacts generated from bone and soft tissue as those from less dense bone objects, and corrected them iteratively. The proposed method was validated using an XCAT phantom data, and compared with two-pass algorithm. The results show that the proposed method is superior to the two-pass algorithm and reduces the overcorrection of the two-pass algorithm near bone regions effectively. Qualitative evaluation with mean square error (MSE) is also performed, demonstrating the effectiveness of the proposed method.**

***Index Terms*—Cone beam computed tomography (CBCT), cone beam artifacts, large cone angle, FDK, Two-pass algorithm**

I. Introduction

THE cone beam computed tomography (CBCT) system has been widely used in diagnostic imaging. While the CBCT system provides more morphological information of patients with 3D volumetric data, cone beam artifacts generated in FDK reconstruction may degrade the diagnostic performance [1,2].

To reduce the cone beam artifacts, several approaches have been proposed [3]-[7], which are classified into two categories. Since the cone beam artifacts are originated by insufficient sampling of the object in CBCT geometry, approaches in the first category change the data acquisition scheme (e.g., circle-and-line trajectory [3], saddle orbit [4], and utilizing multiple X-ray sources [5]) to acquire full sample data of the object. One drawback is a long scan time, which may introduce motion artifacts in the reconstructed image.

This research was supported by Ministry of Science, ICT and Future Planning(IITP-2017-2017-0-01015) and National Research Foundation of Korea(2015R1C1A1A01052268,2017M2A2A4A01070302,2017M2A2A6A01019663). *Asterisk indicates corresponding author*.

C. Han is with the School of Integrated Technology and Yonsei Institute of Convergence Technology, Yonsei University, Incheon 406-840, South Korea (e-mail: hanch0232@gmail.com).

*J. Baek is with the School of Integrated Technology and Yonsei Institute of Convergence Technology, Yonsei University, Incheon 406-840, South Korea (e-mail: jongdukbaek@yonsei.ac.kr).

The methods in the second category reduce the cone beam artifacts by algorithmic way. Grass et al. suggested modified FDK algorithms (i.e., Parallel-FDK (P-FDK) and Tent-FDK (T-FDK) [6]), which rearranged the cone beam rays into parallel (P-FDK) and tent (T-FDK) geometry. A two-pass algorithm proposed by Hsieh tries to reduce the cone beam artifacts iteratively by assuming the high density materials are dominant factors of the cone beam artifacts [7]. While these approaches are effective with a moderate cone angle (i.e., less than 5 degrees), the correction performance becomes poor as the cone angle increases.

In this work, we propose a new method to reduce the cone beam artifacts with a large cone angle scheme. Since the cone beam artifacts generated by soft tissue become severe in a large cone angle scheme, the proposed method incorporates the soft tissue effect in the cone beam artifacts reduction. The proposed method is validated using an XCAT phantom, and performance comparison with two-pass algorithm is conducted.

II. Methods

A. Brief review of the two-pass algorithm and its limitation

The basic assumption of two-pass algorithm is that the cone beam artifacts are generated dominantly by the high density

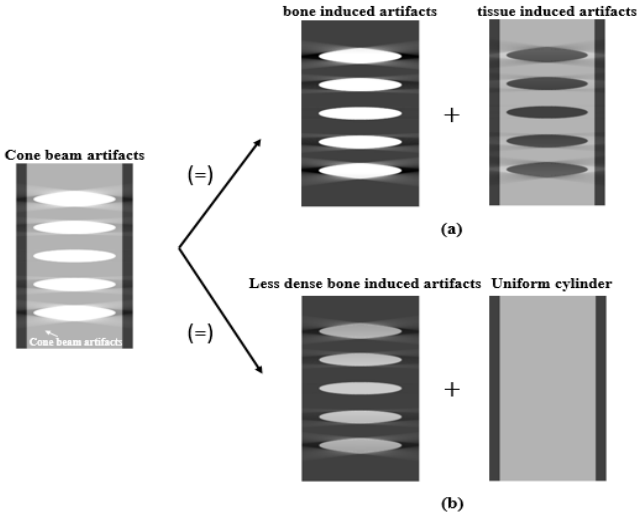


Figure 1. Decomposition of the cone beam artifacts into (a) bone and soft tissue induced artifacts, and (b) the less dense bone induced artifacts.

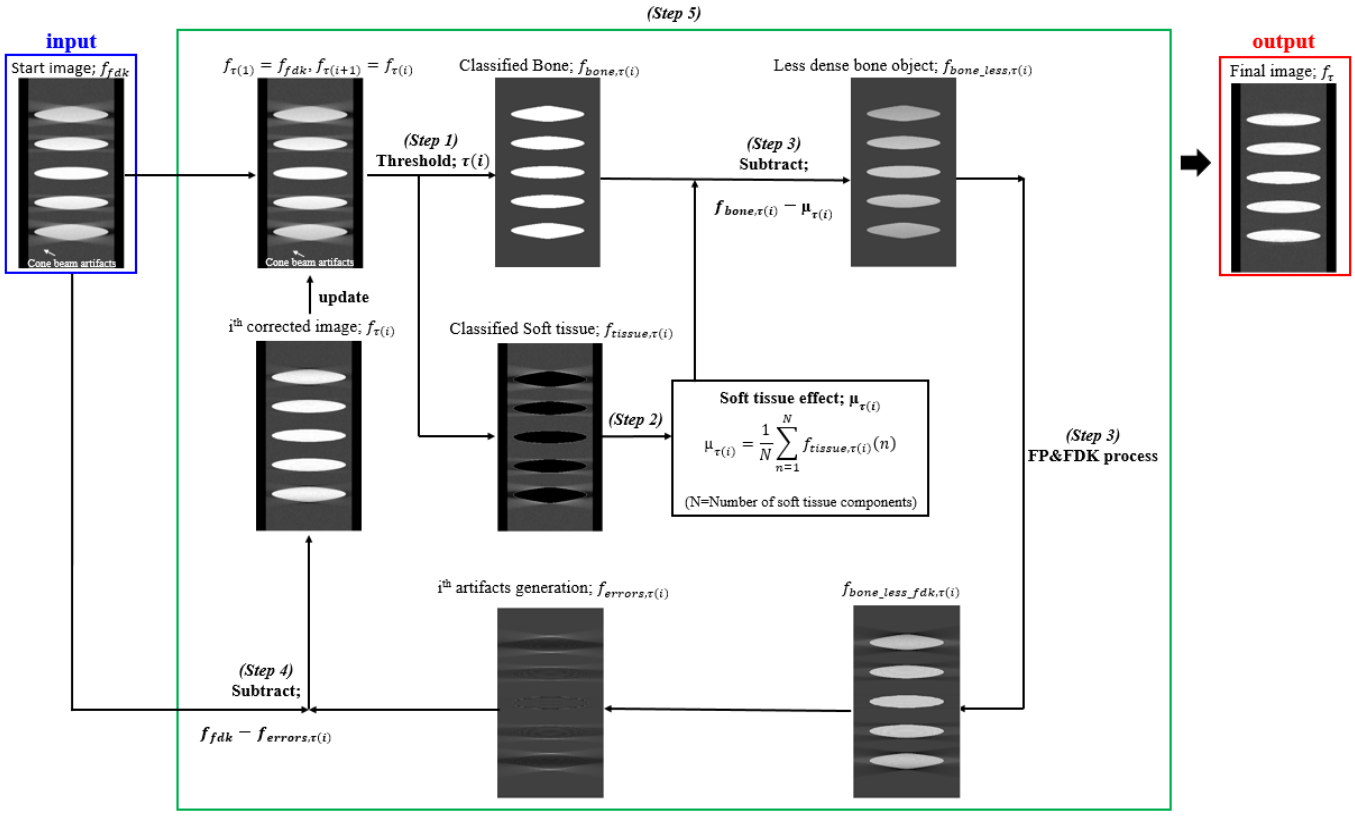


Figure 2. Schematic diagram of proposed method.

materials (e.g., high density bone). Thus, two-pass algorithm segments the high density materials from FDK image by simple thresholding, and regenerates the cone beam artifacts using the segmented high density materials. Then, corrected image is acquired by subtracting the generated cone beam artifacts from the original FDK image [8]. However, when the bone density is low or the X-ray energy used for data acquisition is high, the density difference between bone and soft tissue is reduced. Thus, the cone beam artifacts generated by the bone and soft tissue should be considered together in the artifact reduction, which would be more significant when the cone angle is large.

B. *Proposed Method*

The basic idea of the proposed method is to regenerate the cone beam artifacts from less dense bone objects. Consider a defrise phantom embedded within a cylinder phantom. Since the reconstruction process is linear, the cone beam artifacts can be decomposed as a summation of the bone and soft tissue induced cone beam artifacts (shown in figure 1 (a)) or purely generated from the less dense bone object since the uniform cylinder phantom does not produce cone beam artifacts (shown in figure 1 (b)). By assuming the density of soft tissues around bone objects is not varying rapidly, this approach would help to prevent the overcorrection of conventional two-pass algorithm, especially when the bone density is moderate.

The schematic diagram of the proposed method is depicted in figure 2 and detail of each step is as follows.

(Step 1) Set FDK reconstructed image f_{fdk} as a start image. Then, classify bone, $f_{bone,\tau(1)}$, and tissue, $f_{tissue,\tau(1)}$, images using a simple threshold value, $\tau(1)$.

(Step 2) Calculate a mean value of soft tissue component, $\mu_{\tau(1)}$, in $f_{tissue,\tau(1)}$.

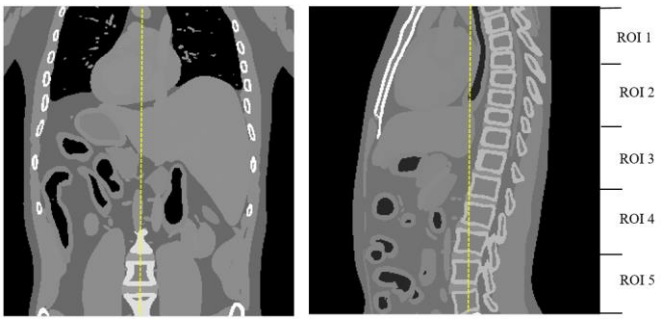
(Step 3) Subtract $\mu_{\tau(1)}$ from $f_{bone,\tau(1)}$ to obtain a less dense bone object, $f_{bone\_less,\tau(1)}$. Conduct a forward projection (FP) on $f_{bone\_less,\tau(1)}$, and perform FDK reconstruction to generate $f_{bone\_less\_fdk,\tau(1)}$.

(Step 4) Errors, $f_{errors,\tau(1)}$, are calculated by subtracting $f_{bone\_less\_fdk,\tau(1)}$ from $f_{bone\_less,\tau(1)}$, and artifacts corrected image, $f_{\tau(1)}$, is acquired by subtracting $f_{errors,\tau(1)}$ from the original f_{fdk} image.

(Step 5) Repeat *(Step 1)* ~ *(Step 4)* until mean square error (MSE) of the $f_{errors,\tau(i)}$ is converged.

C. *Simulations*

To validate the proposed method, we used an XCAT phantom developed by Duke OLV. We generated the XCAT thoracic phantom at monochromatic 60kV energy as shown in Figure 3, where the attenuation coefficients of bone and soft tissue range from 0.36 to 0.5 cm^{-1}, and 0.22 to 0.26 cm^{-1}, respectively.



(a) (b)

Figure 3. (a) Coronal, (b) sagittal slice of XCAT phantom at 60kV. Unit of the values is cm^{-1}, and display window is [0.1 0.4].

The XCAT phantom was constructed with a 300×300×320 matrix with a voxel size of 1.2 mm^3 and 4 subvoxels per voxel were allocated to avoid discretization errors. Projection data were acquired using a 3D forward projector developed by H.Gao [9]. We added Poisson random noise to the projection data, assuming ideal bow-tie filter [10]. The number of detected X-ray photons per detector cell was 2500. After that, a Hanning weighted ramp filter was applied on the projection data. FDK algorithm based on voxel-driven back projection using a linear interpolation was used for the reconstruction. The simulation parameters are summarized in table 1.

Table 1. Simulation parameters

Source to iso-center distance	800 mm
Detector to iso-center distance	400 mm
Detector cell size	1.552 mm × 1.552 mm [576 (row) × 512 (column)]
Number of views	720
Cone angle	-17° through 17° (34°)
Reconstructed volume size	360 × 360 × 384 mm^3
Reconstructed matrix size	300 × 300 × 320
Number of detected X-ray photons	2500
Reconstruction algorithm	FDK

D. Image Quality Evaluation

To compare image quality of FDK, two-pass, and proposed method, we set 5 region of interests (ROIs) as shown in figure 3. For quantitative analysis, the MSE between FDK, two-pass, proposed algorithm, and the reference XCAT phantom were calculated. Averaged MSE was calculated from 20 independent noise realizations. The central vertical profiles in figures 3 (a) and (b) are also presented.

III. Results

Figure 4 (a) shows the coronal and sagittal slices of the reference XCAT phantom. For visual inspection, we zoomed in 116×116 voxel blocks in bony structures of ROIs 4 and 5, corresponding 10 to 17 degrees cone angles. The cone beam artifacts are clearly observed in FDK, showing distorted anatomical structure and intensity in both slices.

Figure 4 (c) shows corrected image from two-pass algorithm, where 0.3 cm^{-1} was selected as a threshold to segment the bone materials. It is shown that two-pass algorithm improves the sharpness of the bones, but generates additional streak artifacts near the bone regions. Note that the intensity drop of the bone object with 0.36 cm^{-1} would be significant within ROI5, and thus simple thresholding would not be effective to segment the bone objects.

Figures 4 (d) and (e) show the corrected image of the proposed method with 1st and 3rd iteration, respectively. In each iteration, we increased the threshold value linearly from 0.3 cm^{-1} to 0.32 cm^{-1}. It is shown that the correction performance is improved with more iterations. We found that MSE converged after 3rd iteration. Table 2 summarizes MSE of ROIs between reference and other algorithms, demonstrating the effectiveness of the proposed method quantitatively. Figures 5 (a) and (b) shows the vertical profiles of the coronal and sagittal slices in Figure 4. It is shown the overcorrection of the two-pass algorithm is reduced significantly with the proposed method.

IV. Discussion and Conclusion

In this work, we proposed a new method to reduce cone beam artifacts with a large cone angle. With the presence of the moderate bone density objects, two-pass algorithm introduces additional streak artifacts near the bone regions due to the inappropriate bone segmentation. The proposed method solved this problem using different threshold values for corrected image at each iteration. The limitation of the proposed method is the computation time due to the iterative procedures, which would be solved using GPU based parallel computing.

V. References

[1] B.D. Smith, "Cone-beam tomography: Recent advances and a tutorial review," Opt. Eng. (Bellingham), vol.29, no.5, pp. 524-534, 1990.
[2] L.A. Feldkamp, "Practical cone-beam algorithm," J.Opt.Soc.Am.A.,vol.1, no.6, 1984.
[3] G.L. Zeng, "A cone-beam tomography algorithm for orthogonal circle-and-line orbit," Phys. Med. Biol., vol. 37, no. 3, pp. 563-577, Oct. 1991.
[4] H. Yang, "Exact cone beam reconstruction for a saddle trajectory," Phys. Med. Biol., vol. 51, pp. 1157-1172, Sep. 2005.
[5] J. Baek, "A multi-source inverse-geometry CT system: initial results with an 8 spot x-ray source array," Phys. Med. Biol., vol. 59, pp. 1189-1202, Jan. 2014.
[6] M Grass, "3D cone-beam CT reconstruction for circular trajectories," Phys. Med. Biol., vol. 45, pp. 329-347, Oct. 1999.
[7] Hsieh, J. (2000, June). A two-pass algorithm for cone beam reconstruction. In Proc. SPIE (Vol. 3979, pp. 533-540).
[8] Hsieh, J. (2000). A practical cone beam artifact correction algorithm. In Nuclear Science Symposium Conference Record, 2000 IEEE (Vol. 2, pp. 15-71). IEEE.
[9] H. Gao, "Fast parallel algorithms for the x-ray transform and its adjoint," Med. Phys., vol. 39, no. 11, Nov. 2012.
[10] SS.Hsieh, "The feasibility of a piecewise-linear dynamic bowtie filter," Med. Phys., vol. 40, no. 3, Mar. 2013.

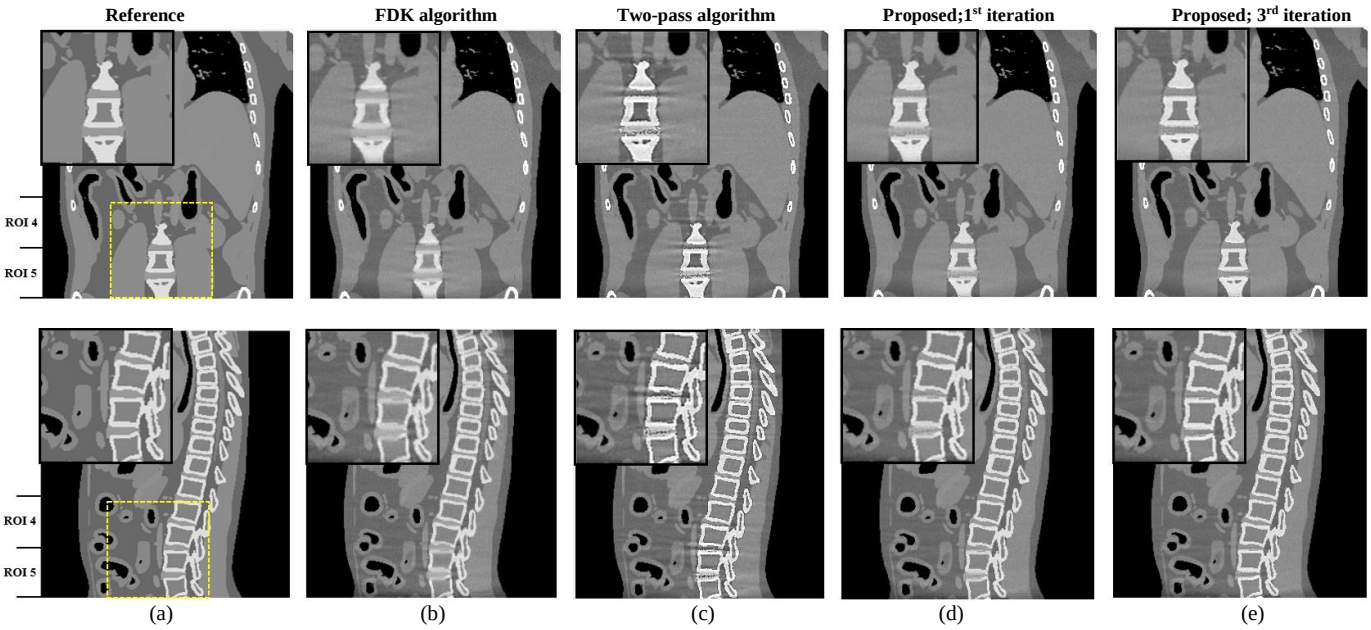


Figure 4. Coronal (upper) and sagittal (lower) XCAT phantom. (a) Reference and (b) FDK images, and corrected images by (c) two pass algorithm and proposed algorithm with (d) 1st iteration and (e) 3rd iteration. Unit of the values is cm^{-1}, and display window is [0.1 0.4].

Table 2. MSE of coronal and sagittal slice

ROI	FDK algorithm		Two-pass algorithm		Proposed; 1st iteration		Proposed; 3rd iteration	
	Coronal	Sagittal	Coronal	Sagittal	Coronal	Sagittal	Coronal	Sagittal
1	2.22×10^{-4}	2.93×10^{-4}	2.12×10^{-4}	3.40×10^{-4}	1.76×10^{-4}	1.99×10^{-4}	1.68×10^{-4}	1.86×10^{-4}
2	0.88×10^{-4}	0.80×10^{-4}	1.37×10^{-4}	1.94×10^{-4}	0.67×10^{-4}	0.56×10^{-4}	0.66×10^{-4}	0.57×10^{-4}
3	0.91×10^{-4}	7.21×10^{-4}	0.96×10^{-4}	1.39×10^{-4}	0.70×10^{-4}	0.47×10^{-4}	0.67×10^{-4}	0.47×10^{-4}
4	1.09×10^{-4}	1.18×10^{-4}	1.27×10^{-4}	2.07×10^{-4}	0.71×10^{-4}	0.61×10^{-4}	0.63×10^{-4}	0.55×10^{-4}
5	2.07×10^{-4}	3.05×10^{-4}	3.27×10^{-4}	4.32×10^{-4}	1.38×10^{-4}	2.28×10^{-4}	1.29×10^{-4}	2.21×10^{-4}

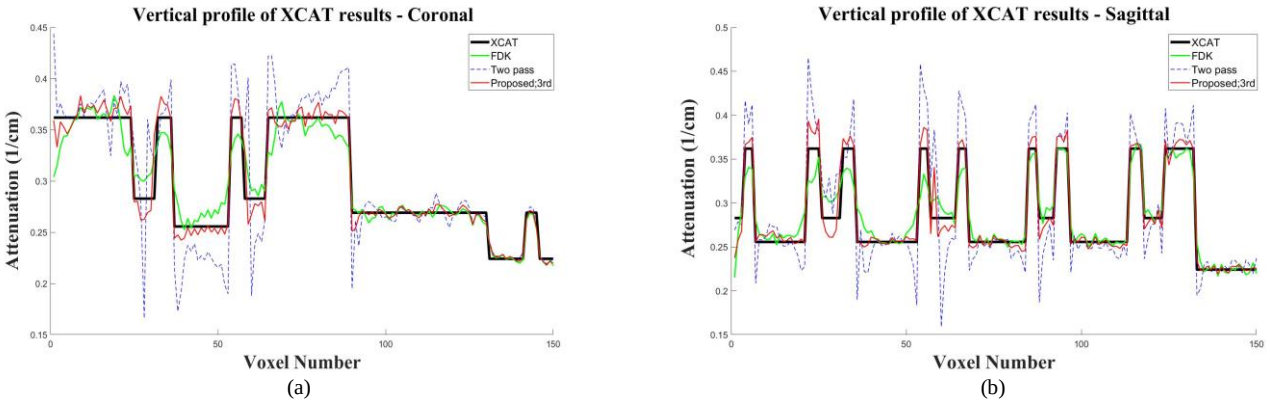


Figure 5. Vertical profiles of XCAT phantom image, original FDK image, and corrected images by two pass algorithm and proposed method in (a) coronal and (b) sagittal slice.

Setting up a low-cost C-arm for its use as a tomograph: preliminary results

M. Abella, C. de Molina, N. Ballesteros, A. García, I. García, A. Martínez, M. Desco

***Abstract*—In scenarios where the use of a CT is cumbersome, such as intraoperative or ICU, using a C-arm system as a tomograph would provide interesting additional clinical information. Recent days are seeing the development of the so-called cone-beam CT based on advanced motorized isocentric C-arm systems. To use more basic non-motorized C-arm systems, apart from the geometric non-idealities common to any CBCT, we need to solve other difficulties: (1) the trajectory of source-detector pair may differ from a circular path, (2) the system may have mechanical strains changing the relative positions of the source and detector for each projection angle, and (3) the exact position of the source and detector elements may not be repeatable for consecutive rotations due to low mechanical precision.**

In this work, we present a protocol to adapt a standard C-arm, originally designed for planar imaging, to be used as a tomograph. The key parts of the new acquisition protocol are a geometrical calibration method that deals with high mechanical tolerance that prevents accurate repetition of source-detector position between acquisitions, and an advanced image reconstruction method able to deal with limited angle data and non-circular trajectories. Both methods make use of the surface information of the patient.

The feasibility of the proposed method was ensured with an in-house prototype system based on a flat panel detector.

***Index Terms*— C-arm, tomography, surface, geometrical calibration, image reconstruction.**

I. Introduction

A C-arm is a fluoroscopic system comprising two units, an X-ray generator and a detector (image intensifier or flat panel), mounted in an arc-shaped gantry, together with a workstation unit used to visualize, store, and manipulate the images. Its main purpose is to acquire real-time planar images, demonstrating to be a useful qualitative assessment tool to guide surgical procedures thanks to its open design, which allows to set the C-shape around the patient lying in the bed [1]. In situations, where a CT system is not available, the use of a C-arm as a tomograph would raise the possibility of tomographic information, with the potential of improving surgical performance and precision [2]. The so-called cone-beam CT, based on advanced isocentric motorized C-arm systems (generally attached to a gantry), have been applied in surgical and radiological scenarios.

When cost is an important issue, it could be helpful to obtain tomographic information using the most basic C-arm systems. The use of a standard non-isocentric non-motorized C-arm for computed tomography, presents several additional difficulties apart from the geometric non-idealities common to any CBCT. On the one hand, the trajectory of source-detector pair may differ from a circular path and the system may have mechanical strains that change the relative positions of the source and detector for each projection angle, hindering the use of calibration methods, widely used in CT systems, that obtain global parameters for all projection angles, such as the one proposed in [3]. On the other hand, the exact position of the source and detector elements may not be repeatable for consecutive rotations due to low mechanical precision, thus the geometrical parameters obtained with periodical calibration cannot be used for consecutive acquisitions.

In this work, we develop a method to adapt a standard C-arm, originally designed for planar imaging, to be used as a tomograph. The key parts of the new method are: 1) a geometrical calibration procedure that deals with high mechanical tolerances, and 2) an advanced image reconstruction algorithm able to deal with limited angle data and non-circular trajectories. Both methods make use of the surface information of the patient that could be obtained with a 3D surface scanner.

II. Methods

The workflow of the proposed method is shown in Fig. 1.

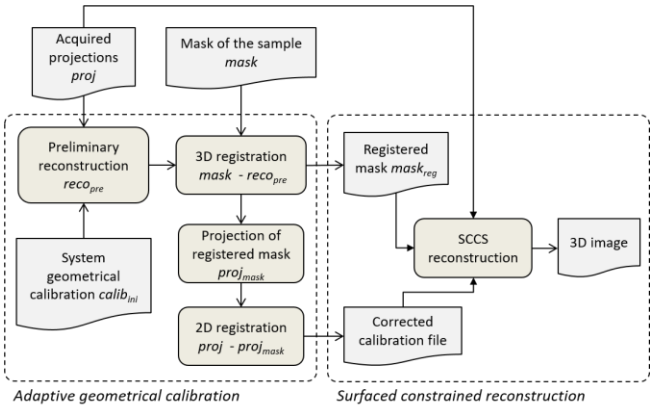


Fig. 1. Workflow of the acquisition/reconstruction protocol.

Manuscript submitted April 2nd 2018.

All authors are with the Dpto. Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid and with the Instituto de Investigación Sanitaria Gregorio Marañón (IiSGM), Madrid, Spain (e-mail: mabella@ing.uc3m.es).

M. Abella and M. Desco are with the Centro Nacional de Investigaciones Cardiovasculares Carlos III (CNIC), Madrid, Spain.

M. Desco is also with the Centro de Investigación en Red de Salud Mental (CIBERSAM) (e-mail: desco@hggm.es).

The mask of the sample is generated from data acquired with a 3D surface scanner. For the preliminary reconstruction we used the FDK-based method proposed in [4]. Projections were obtained with the simulation tool described in [5]. Finally, both 2D and 3D registrations are done based on mutual information.

A. Adaptive geometrical calibration

We generate the system geometrical calibration, $calib_{ini}$, based on the method proposed by Cho et al. for a cone-beam systems [6]. This method was specifically designed to obtain individual calibration parameters for each projection angle using a phantom with two circular patterns (Fig. 2, left). The calibration method estimates the ellipses formed by the balls on each projection and finds out the system parameters based on geometrical relationships between the ellipses: detector rotation (skew), inclination angles (pitch and roll), horizontal and vertical detector shifts, and source to detector distance. This calibration needs to be repeated periodically, as in standard CT systems.

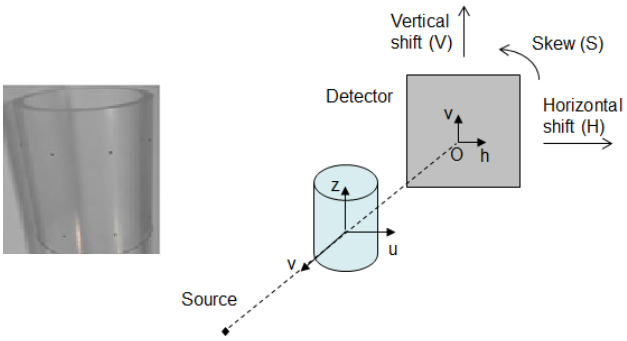


Fig. 2. Left: Calibration phantom. Right: Geometry of the C-arm system showing the parameters that are corrected during the fine tuning: H, V, and S.

We refine the value of geometrical parameters horizontal shift (H), vertical shift (V), and skew (S) (Fig 2 right) following the process outlined in Fig 1. First, we generate a preliminary reconstruction, $reco_{pre}$, which is used to orientate and adjust the mask to the field of view of the C-arm. The registered mask is then projected using $calib_{ini}$. The misalignments between the projections of the mask, $proj_{mask}$, and the acquired data, $proj$, reflect the errors in the values H, V, and S of $calib_{ini}$. Therefore, the parameters of a 2D registration between $proj_{mask}$ and $proj$, ΔH_{reg}, ΔV_{reg} and ΔS_{reg}, are used to generate refined values U_{corr}, V_{corr}, and S_{corr} as:

$$H_{corr} = H + \Delta H_{reg} + O_h - \left(O_h \cdot cos(\Delta S_{reg}) - O_v \cdot sin(\Delta S_{reg})\right) \quad (1)$$

$$V_{corr} = V + \Delta V_{reg} + O_v - \left(O_h \cdot sin(\Delta S_{reg}) + O_v \cdot cos(\Delta S_{reg})\right) \quad (2)$$

$$S_{corr} = S - \Delta S_{reg} \quad (3)$$

where O_x and O_y are the coordinates of the detector center.

B. Surface constrained image reconstruction

Image reconstruction of the limited data is done with SCCS, a limited data surface-constrained reconstruction method [7]. The reconstruction problem follows the idea of the Total Variation (TV) minimization subject to a support constraint, which contains the *a priori* surface information and the data penalty function, formulated as:

$$\min_u \mathrm{TV}(u) \text{ s.t. } \|\mathrm{Au} - \mathrm{f}\|_2^2 < \sigma^2,\ u \geq 0,\ u \in \Omega \quad (4)$$

where u is the reconstructed image, Ω the subspace that corresponds with the surface support of the sample, A is the system matrix, f is the acquired data and σ^2 is the image noise. The L_1-constrained optimization problem (shown in Eq. 4) is efficiently solved using the Split Bregman formulation [8] and formulated as the following unconstrained problems, which are solved at each iteration k:

$$(u^{k+1}, d^{k+1}) = \min_{u,d} \left\|\left(\nabla_x u, \nabla_y u\right)\right\|_1 + \frac{\mu}{2}\left\|Au - f^k\right\|_2 + \\ + \frac{\lambda}{2}\left\|d_x^k - \nabla_x u - b_x^k\right\|_2^2 + \frac{\lambda}{2}\left\|d_y^k - \nabla_y u - b_y^k\right\|_2^2 + \\ + \frac{\gamma}{2}\left\|v - u - b_v^k\right\| \quad (5)$$

$$f^{k+1} = f^k + f - Au^{k+1} \quad (6)$$

$$b_v^{k+1} = b_v^k + u^{k+1} - v^{k+1} \quad (7)$$

Equation (5) leads to two sub-problems: the first one contains only L_2 norm terms and is solved iteratively using a Krylov space solver, and the second one (with the L_1 terms) is solved using analytical formulas. Equations (6) and (7) are the Bregman iterations that impose the data constraint and surface constraint, respectively.

C. Hardware

We evaluated the proposed algorithm using an in-house built C-arm prototype and the hand of the PBU-60 anthropomorphic phantom manufactured by Kyoto Kagaku (Fig 3).

Fig. 3. Anthropomorphic hand phantom acquisition with an in-house C-arm device.

The detector was a wireless, light-weight flat panel, the XRpad 4336, with a 35 cm × 43 cm imaging area and a pixel size of 0.1 mm. The X-ray generator was also a light-weight integrated system, Transportix, with 125 kVp, 100

mA. The useful FOV for reconstruction is 17 cm, without truncation. As in other conventional C-arms, not originally designed for tomography, the movement is completely manual. Experiments were done using the rotation along C-arm plane acquiring 42 projections with an angle span of 120 degrees.

The phantom was also acquired in a Toshiba Aquilion/LB helical scanner and reconstructed as a CT volume of 512×512×1645 voxels, with 0.931×0.931×0.5 mm pixel size. The mask of the sample was simulated from this CT by thresholding.

III. Results

Fig 4 shows values of two of the main parameters for two calibrations obtained at different moments. After parameter refining, both calibrations coincide.

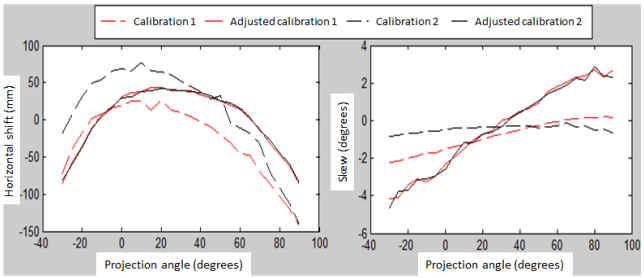


Fig 4. Values of horizontal shift and skew for two system calibrations performed at different moments before and after fine tuning.

Fig. 5 shows a zoom of the image reconstructed using the FDK-based method before and after parameter refinement. The direct use of the initial system calibration shows severe artifacts (gray arrows) in the reconstruction, which are compensated when using refined geometrical parameters U_{corr}, V_{corr}, and S_{corr} (white arrows).

Fig. 5. Coronal and sagittal views of the reconstructed image before (left) and after correcting the calibration parameters (right).

Finally, Fig 6 shows the result of the reconstructed image using the refined geometrical calibration with FDK and SCCS.

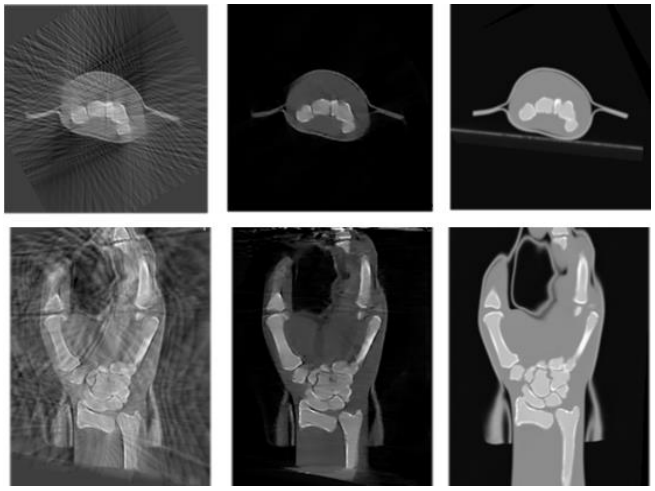


Fig. 6. Axial and coronal views of the reconstruction of the limited data obtained on the C-arm with FDK (left) and SCCS (center). Axial and coronal views of the CT volume acquired on the helical scanner (right).

IV. Discussion and Conclusions

We propose a new method to incorporate tomographic capabilities in a standard non-motorized C-arm, originally designed for planar imaging, by taking advantage from the knowledge of the sample surface.

One of the main problems of these systems is their high mechanical tolerance that prevents accurate repetition of source-detector position between acquisitions. We have proven that standard system calibration is not enough to prevent misalignment artifacts in the reconstructed image. The first algorithm in our proposed method, *adaptive geometrical calibration*, enables the refinement of three geometrical parameters: horizontal shift, vertical shift and skew. Although possible errors in the rest of the geometrical parameters are neglected, results showed that the correction is accurate enough to obtain images free of misalignment artifacts.

The other two main problems to tackle when using a standard non-motorized C-arm as a tomograph are the limitations in the angular span and the difficulty of obtaining a high number of projections. The second algorithm in the proposed method, *surface constrained reconstruction*, removed the artifacts derived from the low number of projections and limited span angle.

The results using data from a real C-arm system based on a flat panel detector showed the feasibility of the proposal. The surface of the sample could be obtained with a structure light 3D scanner. Nevertheless, since the issues related to the acquisition of surface information are out of the scope of this paper, for the evaluation of the method presented here, we have used a simulated surface.

Acknowledgment

This work was partially supported by the Ministry of Economy, Industry and Competitiveness (DPI2016-79075-R) and ISCIII-FIS grants DTS17/00122, co-financed by

ERDF (FEDER) Funds from the European Commission, “A way of making Europe”. The CNIC is supported by the Ministry of Economy, Industry and Competitiveness (MEIC) and the Pro CNIC Foundation, and is a Severo Ochoa Center of Excellence (SEV-2015-0505).

REFERENCES

[1] S. Amiri, D. R. Wilson, B. A. Masri, and C. Anglin, "A low-cost tracked C-arm (TC-arm) upgrade system for versatile quantitative intraoperative imaging," *International Journal of Computer Assisted Radiology and Surgery*, vol. 9, pp. 695-711, Jul 2014.

[2] J. H. Siewerdsen, M. J. Daly, G. Bachar, D. J. Moseley, G. Bootsma, K. K. Brock, *et al.*, "Multi-Mode C-Arm Fluoroscopy, Tomosynthesis, and Cone-Beam CT for Image-Guided Interventions: From Proof of Principle to Patient Protocols," *Proc. SPIE*, vol. 6510, pp. 434-444, 2007.

[3] F. Noo, R. Clackdoyle, C. Mennessier, T. A. White, and T. J. Roney, "An analytic method based on identification of ellipse parameters for scanner calibration in conebeam tomography," *Phys. Med. Biol.*, vol. 45, pp. 3489-3508, 2000.

[4] J. Garcia Blas, M. Abella, F. Isaila, J. Carretero, and M. Desco, "Surfing the optimization space of a multiple-GPU parallel implementation of a X-ray tomography reconstruction algorithm," *The Journal of Systems and Software*, vol. 95, pp. 166–175, 2014.

[5] M. Abella, E. Serrano, J. Garcia-Blas, I. Garcia, C. De Molina, J. Carretero, *et al.*, "FUX-Sim: Implementation of a fast universal simulation/reconstruction framework for X-ray systems," *PLoS ONE*, vol. 12, p. e0180363, 2017.

[6] Y. Cho, D. J. Moseley, J. H. Siewerdsen, and D. A. Jaffray, "Accurate technique for complete geometric calibration of cone-beam computed tomography systems," *Med. Phys.*, vol. 32, pp. 968-83, 2005.

[7] C. de Molina, J. F. P. J. Abascal, M. Desco, and M. Abella, "Study of the possibilities of Surface-Constrained Compressed Sensing (SCCS) Method for Limited-View Tomography in CBCT systems, (in press)," presented at the CT meeting, Bamberg, Germany, 2016.

[8] T. Goldstein and S. Osher, "The Split Bregman Method for L1 Regularized Problems," *SIAM Journal on Imaging Sciences*, vol. 2, pp. 323-343, 2009.

On the Influence of Acquisition Angle and Slice Thickness on the in-plane Spatial Resolution of Calcifications in Digital Breast Tomosynthesis

Christoph Luckner, Frank Schebesch, Christopher Syben, Thomas Mertelmeier, Andreas Maier, and Ludwig Ritschl

***Abstract*—Digital breast tomosynthesis (DBT) is a three-dimensional (3-D) X-ray imaging modality that allows the breast to be viewed in a 3-D format, minimizing the effect of overlapping breast tissue. DBT is commonly known for its high in-plane spatial resolution allowing to detect very small structures inside the breast which makes it a powerful tool in the clinical environment. However, since DBT is a limited angle tomography, artifacts are inevitable. In this paper, we investigate the influence of the angular scanning range as well as the slice thickness, i. e. the distance between two adjacent slices, on the in-plane spatial resolution of calcifications and present an analytic model to describe the imaging process. For the validation of the analytic model, 54 datasets with varying calcification diameter, slice thickness, and angular scanning range, were used and compared to a ray-casting simulation. It could be shown that the overall relative mean error between the analytic model and the generated ground truth is $\bar{\varepsilon} = 0.0137$. The results indicate that both investigated parameters affect the in-plane spatial resolution in a non-linear fashion which yields that they have to be considered in a cascaded system analysis.**

I. Introduction

Digital breast tomosynthesis (DBT) is a three-dimensional (3-D) X-ray imaging modality that allows the breast to be viewed in a 3-D format, minimizing the effect of overlapping breast tissue. Therefore, multiple low-dose X-ray projection images are acquired in an arc trajectory. However, due to the system geometry, only a limited angular range can be covered, leading to limited-angle artifacts in the reconstructed volume. DBT has a variety of tunable parameters, like dose, dose distribution, the number of projections and the total angular range of those projections [1] as well as various reconstruction techniques and filters [2], [3], which affect the image quality. We distinguish between the depth resolution, which is mainly dependent on angular range [1], [4], the in-plane resolution, i. e. the ability to resolve adjacent objects and the slice thickness, which is a user-defined parameter within the reconstruction and depicts the distance between two reconstructed planes [5]. In general, DBT images have an exceptionally high spatial in-plane resolution which allows the detection of small calcifications and lesions in the breast [6], [7]. The in-plane spatial resolution was already subject to a lot of research which has shown, that the resolution is mainly determined by the reconstruction filter as well as imaging geometry [8], [9], [10]. However, these investigations have in common, that they either rely on simulated images, e. g. of a point or a wire, or on experimental phantom measurements.

In this work, we will - based on the observations in literature - investigate how the acquisition angle and the reconstructed slice thickness affect the in-plane spatial resolution of a calcification in the tomographic volume and will develop and validate an analytic framework.

Christoph Luckner, Frank Schebesch, Christopher Syben, and Andreas Maier are with the Pattern Recognition Lab, Friedrich-Alexander University Erlangen-Nürnberg, Germany; Email: christoph.luckner@fau.de
Christoph Luckner, Thomas Mertelmeier, and Ludwig Ritschl are with Siemens Healthcare GmbH, Forchheim, Germany

II. Materials and Methods

Figure 1 shows the acquisition setup. The detector is located in the x-y-plane and the slices are reconstructed parallel to the detector plane in the z-direction with slice thickness d. The scan is performed in the x-direction and the corresponding acquisition angle is depicted by θ. The angular spacing, i.e. the rotation between two subsequent projections is denoted as $\triangle\theta$. For simplification, we omit the y-coordinate orthogonal to the scanning (x) and depth (z) direction and assume a parallel projection geometry. Thus, the calculus, as well as the simulations, will be done in 2-D. The images are acquired using an arc trajectory with simultaneous moving source and detector. The object at location (x, z) will be denoted as $f(x, z)$ and its reconstruction as $\widehat{f}(x, z)$. The 1-D slice of thickness d which is used to assess the in-plane resolution w. r. t. the slice thickness and the acquisition angle is denoted as $s_{d,\theta}(x)$.

A. Theoretical Derivation

In the following we assume that a calcification can be modeled as a circular disk [11] with radius R centered at the origin

$$f(x,z) = \begin{cases} 1, & \text{if } \sqrt{x^2+z^2} \leq R \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

Under the assumption of a parallel beam geometry, each projection $p(t, \theta)$ of an object $f(x, z)$ can be written as Radon transformation $\mathcal{R}$ of the object along a

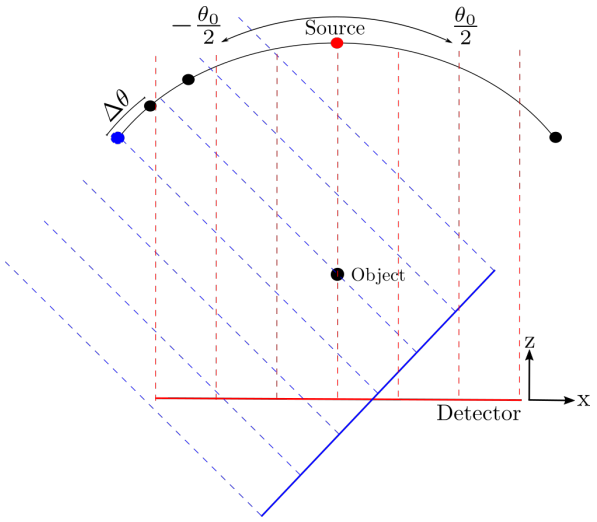


Fig. 1. For image acqusition an arc trajectory with simultaneous rotating source and detector is used. The angular range is denoted as $\left[-\frac{\theta_0}{2}, \frac{\theta_0}{2}\right]$, the angular spacing, i. e. the rotation between to subsequent projections, as $\triangle\theta$.

line $l = x\cos\theta + z\sin\theta$ with acquisition angle θ

$$p(t,\theta) = \mathcal{R}f(x,z) = \int_{-\infty}^{+\infty}\!\!\!\int f(x,z)\cdot\delta(x\cos\theta + z\sin\theta - t)\,\mathrm{d}x\,\mathrm{d}z, \quad (2)$$

where δ denotes the Dirac delta function and t the orthogonal distance of the line l to the origin. Equation 2 evaluates to

$$p(t,\theta) = \begin{cases} 2\sqrt{R^2 - t^2}, & \text{if } |t| \le R \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

Assuming a complete scan trajectory covering the angular range $\theta \in [0,\pi]$, the reconstruction $\widehat{f}(x,z)$ of a scanned object $f(x,z)$ using unfiltered backprojection [11] can be formulated as

$$\widehat{f}(x,z) = \int_0^{\pi} p(x\cos\theta + z\sin\theta,\theta)\,\mathrm{d}\theta = \int_0^{\pi} 2\sqrt{R^2 - (x\cos\theta + z\sin\theta)^2}\,\mathrm{d}\theta. \quad (4)$$

We deliberately chose an unfiltered backprojection approach to exclude any influence of the filtering process on the results.

However, since the angular scanning range θ is usually limited in tomosynthesis, the trajectory does not cover the full 180° but will be denoted as $\theta = \left[-\frac{\theta_0}{2}, \frac{\theta_0}{2}\right]$. Changing the limits of the integral accordingly yields

$$\widehat{f}(x,z) = \int_{-\frac{\theta_0}{2}}^{\frac{\theta_0}{2}} 2\sqrt{R^2 - (x\cos\theta + z\sin\theta)^2}\,\mathrm{d}\theta. \quad (5)$$

Now, a reconstructed slice $s_{d,\theta}(x)$ of slice thickness d and scanning angle θ can be defined as simple integral of $\widehat{f}(x,z)$ over $z = \left[-\frac{d}{2}, \frac{d}{2}\right]$

$$s_{d,\theta}(x) = \int_{-\frac{d}{2}}^{\frac{d}{2}}\int_{-\frac{\theta_0}{2}}^{\frac{\theta_0}{2}} 2\sqrt{R^2 - (x\cos\theta + z\sin\theta)^2}\,\mathrm{d}\theta\,\mathrm{d}z. \quad (6)$$

Since the integration limits are independent of the integration variables we may change the integration order. After the integration w. r. t. to z we eventually obtain

$$s_{d,\theta}(x) = \left[\int_{-\frac{\theta_0}{2}}^{\frac{\theta_0}{2}} \frac{R^2}{\sin(\theta)}\arctan\left(\frac{x\cos(\theta) + z\sin(\theta)}{\sqrt{R^2 - (x\cos\theta + z\sin\theta)^2}}\right) + \left(z + \frac{x}{\tan\theta}\right)\sqrt{R^2 - (x\cos\theta + z\sin\theta)^2}\,\mathrm{d}\theta\right]_{-\frac{d}{2}}^{+\frac{d}{2}}. \quad (7)$$

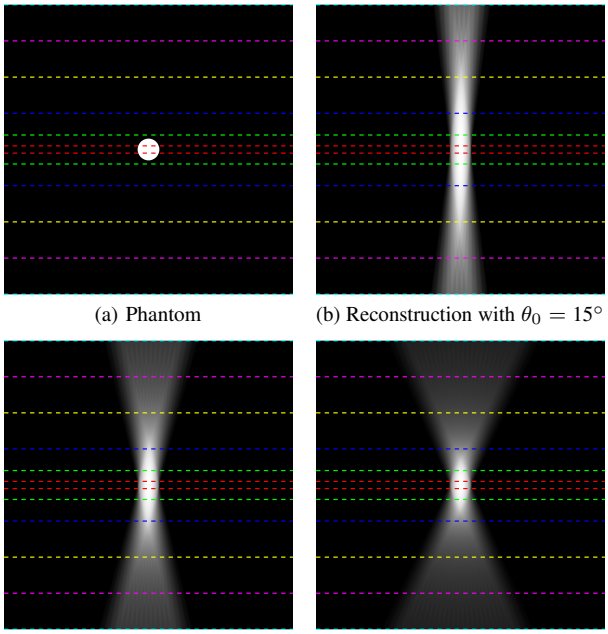
Due to the complexity of the integrand, an analytic evaluation of Equation 7 is not feasible, hence the remaining integral was solved numerically.

B. Experimental Validation

In order to evaluate and validate the analytic model (Equation 7) we compared $s_{d,\theta}(x)$ to a simulation study which was carried out in Matlab. Therefore, various calcification radii, acquisition angles as well as slice thicknesses were modeled and an unfiltered backprojection algorithm was used to generate ground truth (GT) data. The corresponding range of each parameter is shown in Table I.

TABLE I
OVERVIEW OF THE INVESTIGATED PARAMETERS

Variable		Range of values	Unit
Calcification radius	R	{50, 75, 100}	μm
Acquisition angle	θ_0	{15, 30, 50}	°
Slice thickness	d	{50,200,500,1000,1500,2000}	μm



(a) Phantom (b) Reconstruction with $\theta_0 = 15°$

(c) Reconstruction with $\theta_0 = 30°$ (d) Reconstruction with $\theta_0 = 50°$

Fig. 2. Phantom data of a calcification with radius $R = 75\,\mu$m and its reconstructions for the studied angular ranges. The colored lines (red: $50\,\mu$m, green: $200\,\mu$m, blue: $500\,\mu$m, yellow: $1000\,\mu$m, magenta: $1500\,\mu$m, turqoise: $2000\,\mu$m) indicate the used slice thicknesses.

In total 54 simulations were created as ground truth and used for evaluation. To avoid aliasing artifacts, a sampling grid of 1 px $\equiv 1\,\mu$m was used, which allowed us to ignore the detector transfer function, since the grid is very dense compared to the commonly used detector pixel spacing (1 px $\equiv 85\,\mu$m) in mammography. Furthermore, to simulate a common mammography system the angular spacing was set to $\triangle\theta = 2°$.

For evaluation, we compared the results of the derived analytic model (Equation 7) to the GT simulations by visual inspection and evaluated the relative mean error ε over all pixels in x-direction for each dataset individually

$$\varepsilon = \frac{1}{N}\sum_{x=0}^{N-1}\left|\frac{s_{d,\theta}^{\text{Model}}(x) - s_{d,\theta}^{\text{GT}}(x)}{s_{d,\theta}^{\text{GT}}(x)}\right|. \tag{8}$$

III. Results

Figure 2 shows the used phantom for a calcification with radius $R = 75\,\mu$m, as well as the reconstructions for all studied acquisition angles. Additionally, the underlying slice thicknesses used for evaluation are indicated as colored lines. In Figure 3 the normalized profile plots for all slice thicknesses d are presented. It can be seen that with increasing slice thickness the blurring of the calcification is increased for all angular ranges. However, for small slice thicknesses up to $200\,\mu$m only a minor blurring effect in the profile plots can be observed. The individual relative error ε between the GT simulations and the derived model for all datasets ranges from $\varepsilon_{\text{min}} = 0.003$ to $\varepsilon_{\text{max}} = 0.032$ with an overall mean error of $\bar{\varepsilon} = 0.0137$ and thus may be considered as a numerical error.

IV. Conclusion

We were able to present an analytic model that describes the backprojection process based on the angular scanning range as well as the reconstructed slice thickness. In the evaluation it was shown that the results of the derived analytic model are consistent with the simulated ground truth results with a very small mean error of $\bar{\varepsilon} = 0.0137$ We found that up to a slice thickness of $200\,\mu$m the angular range has basically no influence on the blurring of calcifications. This phenomenon is also considered in dedicated tomosynthesis reconstruction algorithms [12]. Moreover, it was found that the angular range and the slice thickness affect the in-plane spatial resolution in a non-linear fashion, which yields that they have to be considered in a cascaded system model [13]. For future work and in order to improve and to make the derived analytic model more realistic noise has to be considered.

Disclaimer

The presented method is not commercially available. Due to regulatory reasons, its future availability cannot be guaranteed.

References

[1] T. Deller, K. N. Jabri, J. M. Sabol, X. Ni, G. Avinash, R. Saunders, and R. Uppaluri, "Effect of acquisition parameters on image quality in digital tomosynthesis," in *Proc. SPIE*, vol. 6510, p. 65101L, 2007.

[2] T. Wu, R. H. Moore, E. A. Rafferty, and D. B. Kopans, "A comparison of reconstruction algorithms for breast tomosynthesis," *Medical physics*, vol. 31, no. 9, pp. 2636–2647, 2004.

[3] Y. Zhang, H.-P. Chan, B. Sahiner, J. Wei, M. M. Goodsitt, L. M. Hadjiiski, J. Ge, and C. Zhou, "A comparative study of limited-angle cone-beam reconstruction methods for breast tomosynthesis," *Medical physics*, vol. 33, no. 10, pp. 3781–3795, 2006.

[4] B. Li, G. B. Avinash, R. Uppaluri, J. W. Eberhard, and B. E. Claus, "The impact of acquisition angular range on the z-resolution of radiographic tomosynthesis," in *International Congress Series*, vol. 1268, pp. 13–18, Elsevier, 2004.

[5] I. Sechopoulos, "A review of breast tomosynthesis. part i. the image acquisition process," *Medical physics*, vol. 40, no. 1, 2013.

[6] G. Lauritsch and W. H. Härer, "A theoretical framework for filtered backprojection in tomosynthesis," in *Proc. SPIE*, vol. 3338, pp. 1127–1137, 1998.

[7] R. J. Acciavatti and A. D. Maidment, "Observation of super-resolution in digital breast tomosynthesis," *Medical physics*, vol. 39, no. 12, pp. 7518–7539, 2012.

[8] B. Li, R. Saunders, and R. Uppaluri, "Measurement of slice thickness and in-plane resolution on radiographic tomosynthesis system using modulation transfer function (mtf)," in *Proc. SPIE*, vol. 6142, p. 61425D, 2006.

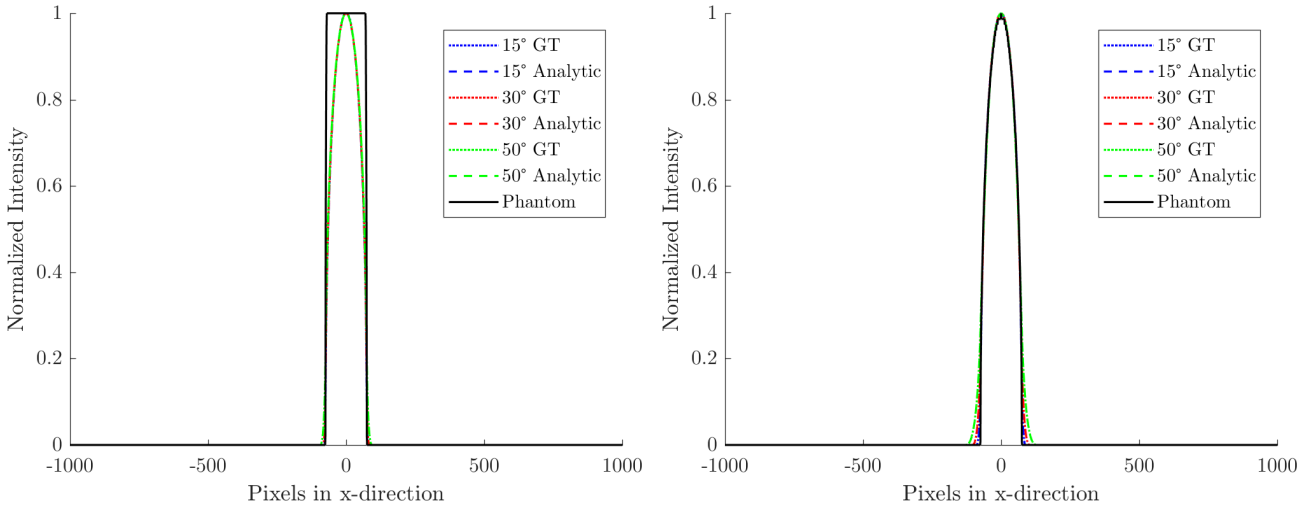
[9] Y. Chen, J. Y. Lo, and J. T. Dobbins III, "Impulse response analysis for several digital tomosynthesis mammography reconstruction algorithms," in *Proc. SPIE*, vol. 5745, pp. 541–549, San Diego, CA, USA, 2005.

[10] J. Zhou, B. Zhao, and W. Zhao, "A computer simulation platform for the optimization of a breast tomosynthesis system," *Medical physics*, vol. 34, no. 3, pp. 1098–1109, 2007.

[11] T. G. Feeman, "The mathematics of medical imaging," *Springer*, 2010.

[12] S. Abdurahman, F. Dennerlein, A. Jerebko, A. Fieselmann, and T. Mertelmeier, "Optimizing high resolution reconstruction in digital breast tomosynthesis using filtered back projection," in *International Workshop on Digital Mammography*, pp. 520–527, Springer, 2014.

[13] Y.-H. Hu, B. Zhao, and W. Zhao, "Image artifacts in digital breast tomosynthesis: Investigation of the effects of system geometry and reconstruction parameters using a linear system approach," *Medical physics*, vol. 35, no. 12, pp. 5242–5252, 2008.



(a) Slice $s_{d,\theta}$ with d = 50 μm

(b) Slice $s_{d,\theta}$ for d = 200 μm

(c) Slice $s_{d,\theta}$ for d = 500 μm

(d) Slice $s_{d,\theta}$ for d = 1000 μm

(e) Slice $s_{d,\theta}$ for d = 1500 μm

(f) Slice $s_{d,\theta}$ for d = 2000 μm

Fig. 3. Plots for various slice thicknesses d for all acquisition angles for a calcification with radius $R = 75\,\mu$m. The dotted lines indicate the ground truth, the dashed lines the evaluation of the analytic model presented in Equation 7. For comparison, a forward projection of the used phantom is shown as a black solid line. The pixel size in the x-direction is 1 μm.

3D-2D *Known-Component* Registration for Metal Artifact Reduction in Cone-Beam CT

A Uneri,[1] T Yi,[1] X Zhang,[1] JW Stayman,[1] P Helm,[2] GM Osgood,[3] N Theodore,[4] JH Siewerdsen[1]

Abstract—Intraoperative cone-beam CT (CBCT) is increasingly used for surgical navigation and validation of surgical device placement. In spine surgery, for example, CBCT provides visualization of spinal pedicle screws relative to target anatomy and adjacent structures. In the surgical settings, however, high attenuation objects in the field of view are often the norm, producing severe streak artifacts that can confound visualization in precisely the area of interest. In this work, a new method for metal artifact reduction (MAR) is introduced that uses prior information of the shape of surgical instruments to reduce or eliminate metal artifacts.

The approach leverages concepts from conventional MAR (often limited by segmentation error [1]) and more advanced *known-component* (KC) reconstruction (KC-Recon) [2], maintaining the speed and simplicity of simple MAR with the power of prior information as in KC-Recon. The proposed "KC-MAR" approach uses 3D-2D registration of the component model to precisely identify the component in the projection domain, thus overcoming conventional pitfalls associated with (3D or 2D) segmentation in conventional MAR. The result (projected region of the registered component) is then inpainted as in conventional MAR using 2D interpolation or more advanced polyenergetic inpainting. Image reconstruction is performed either by 3D filtered back-projection (FBP, the nominal approach in this work) or iterative model-based iterative reconstruction (MBIR) with corresponding benefits to low-dose performance.

The KC-MAR method was investigated in phantom and cadaver studies presenting a range of challenging metal artifact. Algorithm parameters were investigated in phantom experiments using simple (sphere) components ranging in size and composition. The results were translated to a cadaver study involving spinal pedicle screw placement imaged using an interventional O-arm (Medtronic, Littleton MA). KC-MAR images were assessed in terms of visualization of the screw within cortical margins adjacent to the spinal cord, showing strong reduction in metal artifacts without sensitivity to conventional segmentation errors, maintaining the speed and simplicity of FBP (but also compatible with MBIR), and providing confident visualization adjacent to instrumentation.

Index Terms—Metal artifact reduction, 3D-2D registration, cone-beam CT imaging

I. Introduction

Metal instrumentation within the field of view (FoV) of CT / CBCT systems can cause severe streak artifacts that degrade image quality and confound confirmation of device placement. Such artifacts are attributable to several effects, including beam-hardening, scatter, and photon starvation [1]. Metallic objects such as prosthetic implants or surgical instruments are routinely introduced in the FoV in image-guided interventions, where they challenge visualization of the device relative to target anatomy and adjacent critical anatomy. In spine surgery, for example, such artifacts arising from pedicle screws can confound visualization of screw placement adjacent to the spinal cord – precisely the region of interest to identify possible breach of the pedicle cortex. Clear visualization right up to the boundary of the screw is required to confidently assure safe delivery of the implant.

Metal artifacts originate from data inconsistency and/or photon starvation caused by strong energy-dependent attenuation. A fairly broad range of "metal artifact reduction" (MAR) methods have been proposed to correct or compensate for such errors. Algorithms can be considered in two broad categories: i) those that model the physics of beam hardening, noise, etc., implemented within a form of iterative model-based reconstruction; and ii) methods that simply modify the measured projection data within affected regions of strong attenuation (e.g., interpolation / inpainting) [3], [4]. A fundamental problem underlying such approaches involves accurate delineation of the metallic object and strong sensitivity to segmentation error. Segmentation from the initial (uncorrected) 3D image is significantly challenged by the very artifacts the method aims to reduce. Alternatively, segmentation directly in the 2D projection / sinogram domain is challenged by overlapping structures and the need for geometric consistency across the scan orbit.

In image-guided surgery, prior knowledge (e.g. the 3D shape) of components introduced in the FoV is often available, even though the level of information may vary from a simple description (e.g., a deformable, cylindrical tube) to exact, vendor-specific specifications (e.g., CAD drawings for a particular surgical screw). Prior work by Stayman et al. [2] integrated the registration of such known components (KC) within a model-based iterative reconstruction algorithm ("KC-Recon") in a powerful joint registration and reconstruction approach that demonstrated major improvement in image quality compared to FBP (with or without conventional MAR). The KC-Recon approach, however, carries a fairly large computational burden and can challenge time constraints of intraoperative workflow.

In this work, a known-component metal artifact reduction algorithm ("KC-MAR") method is presented that overcomes the sensitivity to segmentation error and preserves the simplicity of conventional MAR (i.e., is compatible with 3D FBP). Recent analogous work by Ruth et al. [5] used the CAD model of knee implants to improve upon an initial,

1. Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD 21205, USA
2. Medtronic Inc., Littleton, MA 01460, USA
3. Department of Orthopaedic Surgery, Johns Hopkins Medicine, Baltimore, MD 21287, USA
4. Department of Neurosurgery, Johns Hopkins Medicine, Baltimore, MD 21287, USA

error-prone 3D segmentation, thereby leading to improved artifact reduction. Prior information of component shapes (e.g., CAD models specific to a particular device) are used to carry out 3D-2D registration and produce extremely precise and accurate localization of the metallic objects directly in the measurement domain, thus avoiding the potential pitfalls of methods that rely on segmentation.

II. Known-Component Metal Artifact Reduction

The KC-MAR algorithm is illustrated in Figure 1, highlighting the consecutive stages for 3D-2D registration, inpainting of the region in projection of the registered component (by simple interpolation or more sophisticated replacement methods), and 3D image reconstruction (by 3D FBP or other MBIR techniques that usually enjoy improved tradeoffs in noise, resolution, and dose).

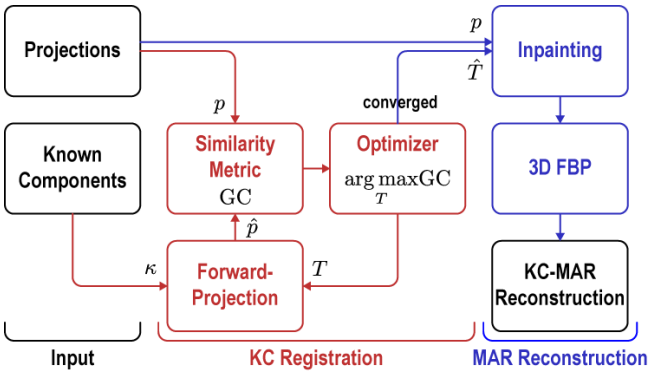


Figure 1. Flowchart for the KC-MAR algorithm depicting the consecutive stages for KC registration and MAR reconstruction with 3D FBP. The registration provides exact localization of the metal components, which are then used to correct (inpaint) projections prior to reconstruction.

A. KC Registration

The details of the 3D-2D registration process for known components ("KC-Reg") have been previously reported by Uneri et al. [6]. A few simplifications to the original approach are possible for KC-MAR, namely: (i) a prior 3D patient image is not necessary, since the components are directly registered to the projections acquired for reconstruction; and (ii) while the KC-Reg method typically operates on just 2–3 projection views, the 3D-2D registration can be rendered extremely precise by using more views – e.g., *all* of the projection views acquired for the 3D reconstruction. To keep runtimes within <1 min in the current work, 6 views with 30° of separation were used, with future work to include a more fully parallelized implementation that suffers little or no additional time penalty in registering more (or all) views.

3D-2D registration of the device components is iterative, beginning with a digitally reconstructed radiograph (DRR) from the input component mesh model:

$$\hat{p}(\kappa, T) = \int_{\vec{r}} \kappa(T)\, \mathrm{d}\vec{r} \tag{1}$$

where each pixel is computed according to the line integral along ray $\vec{r}$ incident on the transformed component κ. The framework was shown to support different component models (e.g., exact technical drawings or simplified parametric models when exact models are not available [6]), and transformation models (e.g. rigid or deformable [7]).

The DRRs are then compared against the actual projections (p) using the gradient correlation (GC) similarity metric [8], defined by the sum of normalized cross-correlation (NCC) of orthogonal image gradients:

$$\mathrm{GC}(p, \hat{p}) = \frac{1}{2}\left\{\mathrm{NCC}(\nabla_x p, \nabla_x \hat{p}) + \mathrm{NCC}\left(\nabla_y p, \nabla_y \hat{p}\right)\right\}$$
$$\text{s.t. } \mathrm{NCC}(a, b) = \frac{\sum_i (a_i - \bar{a})(b_i - \bar{b})}{\sqrt{\sum_i (a_i - \bar{a})^2}\sqrt{\sum_i (b_i - \bar{b})^2}} \tag{2}$$

The GC metric favors the high-intensity gradients from the device component, improving robustness against gradients associated with anatomical structures (e.g., bones).

The objective function for the registration can then be formulated as:

$$\hat{T} = \arg\max_T \sum_\theta \mathrm{GC}(p_\theta, \hat{p}_\theta(\kappa, T)) \tag{3}$$

which can be iteratively solved to obtain the component transform ($\hat{T}$) yielding the greatest similarity to the measurements. A stochastic, derivative-free optimization method referred to as the covariance matrix adaptation evolution strategy (CMA-ES) was used, chosen due to its robust convergence and amenability to parallelization [9].

B. MAR Reconstruction

The second stage of KC-MAR uses the registered component to demarcate the region of metal and facilitate sinogram inpainting. Specifically, the DRR of the registered component is computed such that $\hat{p}(\kappa, \hat{T}) > 0$ demarks the metal regions within the original projections, p. The masked region can be optionally dilated to improve robustness against minor errors in registration that can be introduced by manufacturing variation, geometric calibration of the imaging system, or floating precision errors in computing ray-component intersections.

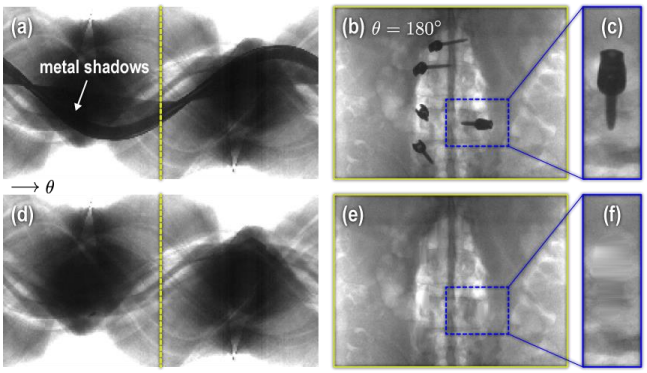


Figure 2. Inpainting of pedicle screws in a cadaver study. (a) The original sinogram with screws, (b) a sample projection, and (c) close-up of an example screw. KC-MAR inpainted counterparts are in d–f, respectively.

The 3D-2D registration yields the transformation by which the shape of the component is forward-projected to precisely define the region to be inpainted. Among the variety of potential inpainting methods, a simple linear interpolation was chosen in order to emphasize the benefits of the registration-based approach in isolation of other potential

enhancements. This is achieved by first producing a Delaunay triangulation over the convex hull of the masked regions, followed by barycentric interpolation on each triangle. Repeating this for all θ projections, the region of metal shadow is inpainted as shown in Figure 2. The images are then reconstructed on a 512×512×385 voxel grid with 0.415×0.415×0.415 mm^3 spacing using 3D FBP based on the Feldkamp-Davis-Kress (FDK) algorithm [10].

III. Experimental Evaluation

A. O-Arm CBCT Imaging

CBCT images were acquired using a research prototype (not-for-clinical-use) implementation of the Medtronic O-arm (Medtronic, Littleton MA). Projection data (1536×1536 pixels at 194 μm pitch) were acquired in dual-gain mode with 2×4 pixel binning. Imaging was performed over a 360° orbit using the manufacturer's high-definition (HD) protocol, giving ~720 projections with ~0.5° radial separation. The geometry was calibrated using a BB phantom [11] to give the projective transforms used in the forward / back-projection operations of the algorithm.

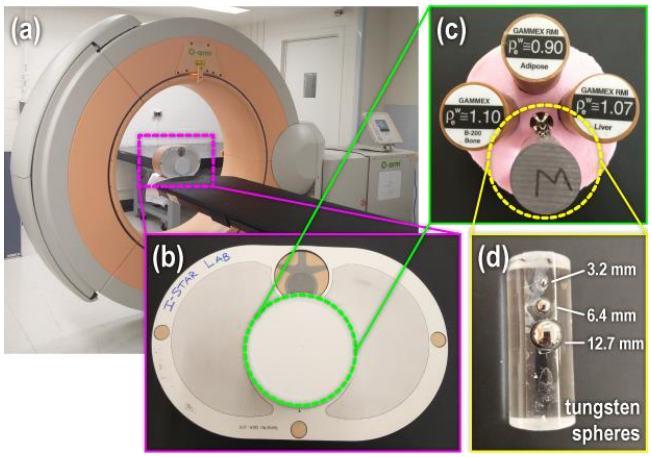


Figure 3. Phantom study. (a) O-arm setup with (b) abdomen phantom, (c) centerpiece containing tissue-equivalent inserts, and (d) 3 metal spheres of varying diameters attached to an acrylic insert.

B. Phantom Experiments with Simple Components

Phantom studies were performed using chest and phantom models (QRM, Möhrendorf, Germany) shown in Figure 3c including tissue-equivalent inserts (adipose, liver, and bone) and an assortment of metal spheres. The insert assembly was encased in water and placed in the center of the chest or abdomen QRM models. Three spheres of varying diameter (12.7, 6.4, and 3.2 mm) were attached to the acrylic insert (Figure 3d) for three types of metal (tungsten, steel, and titanium). Scans were acquired for each phantom configuration at four dose levels: x-ray tube current of 10, 12, 16, and 25 mA (all at 110 kV).

A simple parametric model of spherical components was employed for KC registration, where each metal sphere was modeled according to 3+1 degrees-of-freedom representing the 3D centroid and diameter. The 3 spheres were registered simultaneously with collision avoidance enforced by bounding boxes as in [12].

The magnitude of the metal artifacts was quantified in terms of a simple "Artifact Magnitude" metric given by the standard deviation (σ) in a homogeneous background region about each metal sphere. Artifact reduction was analyzed as a function of material type (×3), diameter (×3), and dose (×4). The sensitivity to segmentation error as typically evident in conventional MAR was evaluated in a manner that was agnostic / irrespective of a particular segmentation method (and recognizing ongoing work in improving the accuracy of such segmentation): specifically, the boundary of the forward-projected component registration was eroded or dilated over a range ($\Delta\emptyset$) ±3 mm to simulate an arbitrarily small or large segmentation error.

C. Cadaver Study

A second set of experiments was conducted to translate the robustness of algorithm parameters identified in the phantom studies and test the performance of KC-MAR under realistic conditions using a cadaveric human torso (77 year old male, medium body habitus). Pedicle screw placement was performed on the cadaver using a total of 5 pedicle screws in

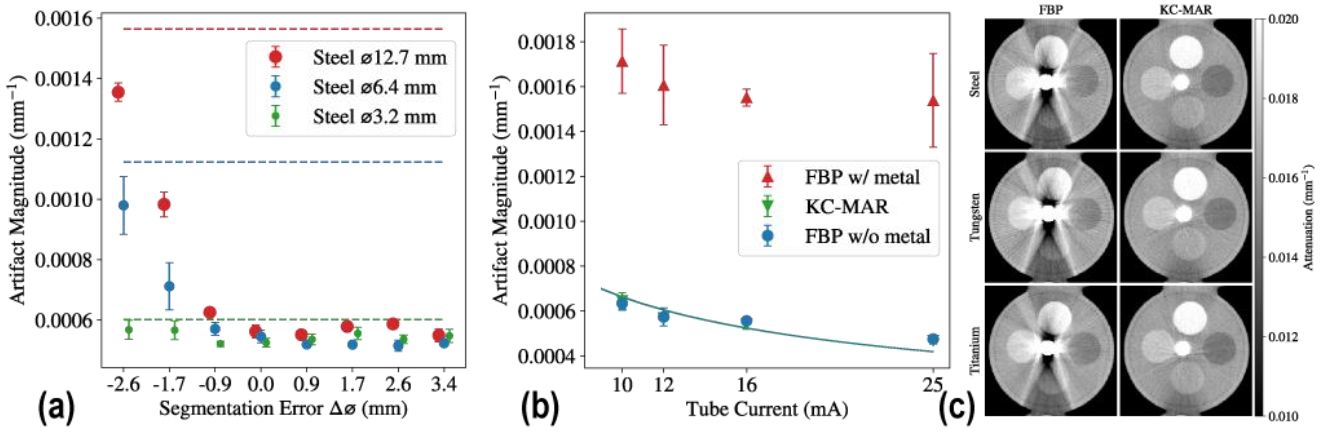


Figure 4. Phantom studies. (a) Evaluation of sensitivity to model mismatch (horizontal lines signify uncorrected values), (b) Artifact magnitude measured as a function of dose with and without metal vs. KC-MAR. (c) Axial image reconstructions with and without KC-MAR for metal spheres of varying material type. The largest of three diameters is shown for each material (and smaller spheres demonstrated similar or better levels of artifact reduction).

the lumbar spine. Manufacturer-specific CAD models (Solera pedicle screws, Medtronic, Littleton MA) were used as the known components. The component models included a rigid polyaxial screw and its articulating tulip head.

IV. Results

A. Phantom Experiments

The performance of KC-MAR in the phantom studies are shown in Figure 4, showing the magnitude of metal artifact as a function of material type, component size (sphere diameter), and dose. Figure 4a shows the sensitivity to segmentation error as often evident in simple MAR – showing the expected result in which underestimation of the true region (simulated here by erosion of the 3D2D-registered region) suffers steep increase in artifact severity, whereas overestimation (dilation) offers a greater degree of stability and robustness to error. KC registration errors were observed to be less than 1 pixel (< 0.5 mm), but a single pixel dilation of the projected component was found to better handle small errors arising from possible variability in device manufacturing or geometric calibration.

Figure 4b shows the performance of KC-MAR as a function of dose, showing the method to be robust even under low dose 3D imaging protocols, whereas the methods relying on segmentation in the 3D image would presumably suffer increased segmentation error.

Finally, Figure 4c shows axial images of the phantom insert (adipose, liver, bone, and water in proximity to Ti, W, and Fe spheres). In each case, the largest diameter (12.7 mm) sphere is shown, and similar (or better) results were observed for smaller components. In each case, the KC-MAR approach yielded strong reduction in artifacts and improved visualization of adjacent structures.

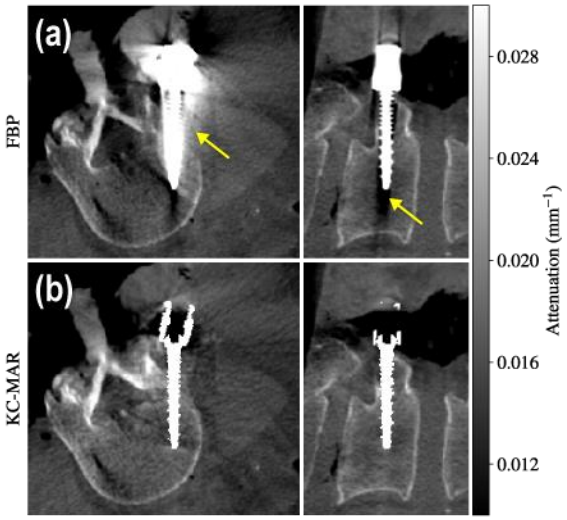


Figure 5. Cadaver study: pedicle screw placement in the L4 vertebra, showing FBP reconstruction (a) without metal artifact reduction and (b) with KC-MAR. The corrected image shows significant reduction in artifacts and allows more reliable visualization of the instrument and adjacent anatomy such as the pedicle cortex and spinal cord.

B. Cadaver Study

The cadaver study investigated KC-MAR performance in real anatomy using a complex component model. Figure 5a shows images of a pedicle screw delivered to the L4 vertebra, where artifacts are seen to confound visualization of adjacent cortical bone. Specifically, the uncorrected image is one in which a surgeon would be challenged to identify whether the screw was delivered with a lateral breach of the pedicle. The KC-MAR image exhibits clear improvement in image quality by reduction of metal artifacts, enabling visualization right up to the edge of the screw – in this case, demonstrating that the lateral pedicle cortex is intact.

V. Discussion

An algorithm for metal artifact reduction was presented based on 3D-2D registration of "known components," offering a variety of potential advantages over existing techniques. The KC-MAR approach maintains the simplicity of conventional MAR (e.g., consistent with 3D FBP and other artifact corrections that may be incorporated in the reconstruction process) but yields a near-perfect localization of the metal object – i.e., does not suffer segmentation errors that are often a limiting factor in conventional MAR. It leverages the power of prior information (i.e., the component model) in a manner similar to KC-Recon [2] but is not bound to joint model-based registration / reconstruction approaches that often carry large computation time that can challenge surgical workflow. In its current implementation (a single GPU GTX TITAN Black [Nvidia, SantaClara CA]), KC-MAR registration required < 1 min with FBP reconstruction in < 2 min. The method benefits from precise models of the metal device but can be extended to simplified parametric models as in [6], with dilation of the resulting region shown to provide robustness to error.

References

[1] M. L. Kataoka, M. G. Hochman, E. K. Rodriguez, P.-J. P. Lin, S. Kubo, and V. D. Raptopolous, "A review of factors that affect artifact from metallic hardware on multi-row detector computed tomography.," *Curr. Probl. Diagn. Radiol.*, vol. 39, no. 4, pp. 125–36, 2010.

[2] J. W. Stayman, Y. Otake, J. L. Prince, A. J. Khanna, and J. H. Siewerdsen, "Model-based tomographic reconstruction of objects containing known components.," *IEEE Trans. Med. Imaging*, vol. 31, no. 10, pp. 1837–48, Oct. 2012.

[3] M. Bal and L. Spies, "Metal artifact reduction in CT using tissue-class modeling and adaptive prefiltering.," *Med. Phys.*, vol. 33, no. 8, pp. 2852–9, Aug. 2006.

[4] J. Wei, L. Chen, G. A. Sandison, Y. Liang, and L. X. Xu, "X-ray CT high-density artefact suppression in the presence of bones.," *Phys. Med. Biol.*, vol. 49, no. 24, pp. 5407–18, Dec. 2004.

[5] V. Ruth, D. Kolditz, C. Steiding, and W. A. Kalender, "Metal artifact reduction in X-ray computed tomography using computer-aided design data of implants as prior information," *Invest. Radiol.*, vol. 52, no. 6, pp. 349–359, 2017.

[6] A. Uneri *et al.*, "Known-component 3D–2D registration for quality assurance of spine surgery pedicle screw placement," *Phys. Med. Biol.*, vol. 60, no. 20, pp. 8007–24, Oct. 2015.

[7] A. Uneri *et al.*, "Deformable 3D-2D registration of known components for image guidance in spine surgery," *Med. Image Comput. Comput. Assist. Interv.*, vol. 19, no. 3, pp. 124–32, 2016.

[8] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. Hill, and D. J. Hawkes, “A comparison of similarity measures for use in 2-D-3-D medical image registration.,” *IEEE Trans. Med. Imaging*, vol. 17, no. 4, pp. 586–95, Aug. 1998.
[9] N. Hansen and A. Ostermeier, “Completely derandomized self-adaptation in evolution strategies.,” *Evol. Comput.*, vol. 9, no. 2, pp. 159–95, Jan. 2001.
[10] L. A. Feldkamp, L. C. Davis, and J. W. Kress, “Practical cone-beam algorithm,” *J. Opt. Soc. Am. A*, vol. 1, no. 6, pp. 612–9, Jun. 1984.
[11] Y. Cho, D. J. Moseley, J. H. Siewerdsen, and D. a. Jaffray, “Accurate technique for complete geometric calibration of cone-beam computed tomography systems,” *Med. Phys.*, vol. 32, no. 4, p. 968, Apr. 2005.
[12] A. Uneri *et al.*, “Intraoperative evaluation of device placement in spine surgery using known-component 3D-2D image registration,” *Phys. Med. Biol.*, Feb. 2017.

Autocalibration of cone beam CT projection matrices based on arbitrary traceable features within a regular tomographic scan

Jonas Dittmann, Simon Zabler

***Abstract*—As an essential prerequisite to tomographic reconstruction, the relative positions of X-ray source, rotary stage and detector of a cone beam computed tomography setup need to be known. Due to the relatively large amount of mechanical degrees of freedom, it is for practical reasons desirable to obtain this information directly from the imaging system, e.g. by scanning a reference phantom. While such approaches relying on specifically designed phantoms do exist, sufficiently well defined phantoms become much harder to manufacture when considering the small field of view of high resolution micro CT systems. In the extreme case of sub-µm resolutions and sub-mm fields of view, it is desirable to get along with completely undefined assemblies of opaque markers or even just traceable features within a sample itself. The most that can be assumed about these tracers is that they follow circular trajectories about the rotary stage's axis. In the following, a self consistent determination of the acquisition geometry based solely on the assumption of circular tracer trajectories is presented, facilitating geometrically consistent tomographic reconstructions. The technique is demonstrated using a simple assembly of four ball pen balls glued to a plastic rod, serving here both as phantom and sample. Qualitative confirmation of successful calibration is given by visual inspection of details within the sample.**

I. Introduction

Calibration of cone beam computed tomography setups, i.e. the determination of the relative positions of X-ray source, rotary stage and detector, is a common problem to be solved prior to tomographic reconstruction. The problem may be partitioned into two separate problems: The problem of relative position and orientation of the detector with respect to rotational and optical axes as well as the absolute scale of the sample, which is related to its actual position along the optical axis (i.e. the orthogonal connecting line between X-ray source and rotational axis). While the latter information is relevant for a metrological quantification of the results, only the former information is actually essential for the tomographic reconstruction process. Due to the relatively large amount of mechanical degrees of freedom of a cone beam tomography device, it is for practical reasons desirable to do the calibration by means of the imaging system itself, i.e. by scanning a reference phantom.

J. Dittmann and S. Zabler are with the Lehrstuhl für Röntgenmikroskopie at the University of Würzburg, Germany and with the Fraunhofer NCTS group of the Fraunhofer IIS/EZRT/MRB in Erlangen/Fürth/Würzburg, Germany. (email: jonas.dittmann@physik.uni-wuerzburg.de)

Given the ubiquity of the problem, the topic of pose determination based on reference phantoms in the context of cone beam tomography has been covered by several authors in the past of which only a very limited number of references is given here. Many approaches found in literature are based on specifically designed, i.e. fully known phantoms, see e.g. [1], [2], [3]. Although this is a preferable situation and allows to calibrate to absolute scales and for some approaches also viewing angle dependent effects (particularly relevant when the whole system, rather than only a lightweight sample, is moved), it obviously requires the availability of a respective phantom both at the scale of the field of view and at the precision of the resolution of the system of interest. While this on the one hand becomes a dominating complication for high resolution micro CT systems, viewing angle dependent pose determination is usually less important for these systems on the other hand, i.e. the amount of to-be-calibrated degrees of freedom can be reduced to those of a single projection view. And although the knowledge of the absolute scale of a scan is of course also of interest for micro CT applications, it is not strictly necessary for the volume reconstruction and can also be deferred to complementary techniques independent of the CT scan.

Calibration techniques based on simple phantoms with little or no requirements apart from a error free rotary stage have been previously proposed e.g. in [4], [5], [6], [7], using various levels of complexity to selectively extract individual geometric parameters such as detector tilt, projection of the rotational axis on the detector, etc. while partially assuming some degrees of freedom fixed (e.g. [4], [5]).

The following contribution presents a simple geometry calibration approach solely based on the assumption of circular tracer trajectories perpendicular to and centered at the rotational axis of the CT setup, i.e. based on a stacked-rings-of-markers phantom virtually generated by a regular CT scan of an uncalibrated assembly of arbitrary traceable features. The presented technique avoids separate treatment of individual degrees of freedom. Instead, the final result of the procedure will be a homogeneous 3×4 projection matrix that can be directly used for tomographic reconstruction (cf. [8]). An example will given using a simple assembly of four ball pen balls glued to a plastic rod, serving here both as phantom and sample. Visual confirmation of successful auto-calibration is provided by means of reconstructed volume

sections showing fine details within the sample.

II. Method

Based on a general forward model comprising the unknown projection matrix projecting circular orbits of as well unknown radii and vertical positions, an iterative scheme will be used to self consistently reconstruct both the projection matrix *and* the orbit parameters simultaneously up to some unknown affine transform.

The forward model, restricting without loss of generality the rotational axis to the z-axis and the start (and end) of the cyclic marker trajectories to the x-axis, is:

$$\begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \end{bmatrix} \begin{bmatrix} r'_i \cos\phi \\ r'_i \sin\phi \\ z'_i \\ w_i \end{bmatrix}$$
$$= \begin{array}{l} r'_i(P_{11}\cos\phi + P_{12}\sin\phi) + P_{13}z'_i + P_{14}w_i \\ r'_i(P_{21}\cos\phi + P_{22}\sin\phi) + P_{23}z'_i + P_{24}w_i \\ r'_i(P_{31}\cos\phi + P_{32}\sin\phi) + P_{33}z'_i + P_{34}w_i \end{array}$$
$$= \begin{array}{l} r'_i P_{1\mathrm{a}}\sin(\phi - \phi_{\mathrm{h}0}) + P_{13}z'_i + P_{14}w_i \\ r'_i P_{2\mathrm{a}}\sin(\phi - \phi_{\mathrm{v}0}) + P_{23}z'_i + P_{24}w_i \\ r'_i P_{3\mathrm{a}}\sin(\phi - \phi_{\mathrm{w}0}) + P_{33}z'_i + P_{34}w_i \end{array}$$

with the unknown projection matrix $\boldsymbol{P}$ (represented by its components P_{mn}) and multiple marker trajectories (indexed by i) characterized by their respective distance from the rotational axis $r_i = r'_i/w_i$, their axis-parallel coordinate $z_i = z'_i/w_i$ and their homogeneous coordinate's scaling component w_i. The matrix-vector product can be separated into ϕ-dependent sinus curves with amplitudes $r'_i P_{X\mathrm{a}} = r'_i\sqrt{P_{X1}^2 + P_{X2}^2}$ and ϕ-independent offsets $P_{X3}z'_i + P_{X4}w_i$. The observable projections' horizontal (detector rows) and vertical (detector columns) components on the detection plane are then given by

$$h_{ij} = \frac{r'_i P_{1\mathrm{a}}\sin(\phi_j - \phi_{\mathrm{h}0}) + P_{13}z'_i + P_{14}w_i}{r'_i P_{3\mathrm{a}}\sin(\phi_j - \phi_{\mathrm{w}0}) + P_{33}z'_i + P_{34}w_i}$$
$$v_{ij} = \frac{r'_i P_{2\mathrm{a}}\sin(\phi_j - \phi_{\mathrm{v}0}) + P_{23}z'_i + P_{24}w_i}{r'_i P_{3\mathrm{a}}\sin(\phi_j - \phi_{\mathrm{w}0}) + P_{33}z'_i + P_{34}w_i}$$

respectively, with j indexing positions ϕ_j on the circular trajectories.

Given that nothing more than a number of projection samples (h_{ij}, v_{ij}) is known, the equivalent reduced model

$$h_{ij} = \frac{a_{\mathrm{h},i}\sin(\phi_j - \phi_{\mathrm{h}0}) + o_{\mathrm{h},i}}{a_{\mathrm{w},i}\sin(\phi_j - \phi_{\mathrm{w}0}) + 1}$$
$$v_{ij} = \frac{a_{\mathrm{v},i}\sin(\phi_j - \phi_{\mathrm{v}0}) + o_{v,i}}{a_{\mathrm{w},i}\sin(\phi_j - \phi_{\mathrm{w}0}) + 1}$$

will be used to get started with the solution of the inverse problem. Using a least squares method to fit this model to the available projection samples, the derived observables $a_{\mathrm{h},i}$, $a_{\mathrm{v},i}$, $a_{\mathrm{w},i}$, $o_{\mathrm{h},i}$, $o_{\mathrm{v},i}$, $\phi_{\mathrm{h}0}$, $\phi_{\mathrm{v}0}$, $\phi_{\mathrm{w}0}$ are obtained.

The calibration problem can now be reduced to the solution of the following system of equations arising from the comparison of the full forward model and the observable parameters of the fit model for all i:

$$\begin{aligned} r'_i P_{1\mathrm{a}} &= a_{\mathrm{h},i} \\ r'_i P_{2\mathrm{a}} &= a_{\mathrm{v},i} \\ r'_i P_{3\mathrm{a}} &= a_{\mathrm{w},i} \\ P_{13}z'_i + P_{14}w_i &= o_{\mathrm{h},i} \\ P_{23}z'_i + P_{24}w_i &= o_{\mathrm{v},i} \\ P_{33}z'_i + P_{34}w_i &= 1 \quad . \end{aligned}$$

After initializing the unknown trajectory parameters with $w_i = 1$, $r'_i = a_{\mathrm{h},i}$ and $z'_i = o_{\mathrm{v},i}$, common iterative techniques for the solution of linear systems of equations are applied to alternatingly solve for P_{nm} and (r'_i, z'_i, w_i) respectively until consistency is reached. Intermediate regularization steps incorporating additional knowledge on either the projection matrix or on e.g. distances between markers (establishing an absolute scale) may as well be incorporated within this alternating scheme.

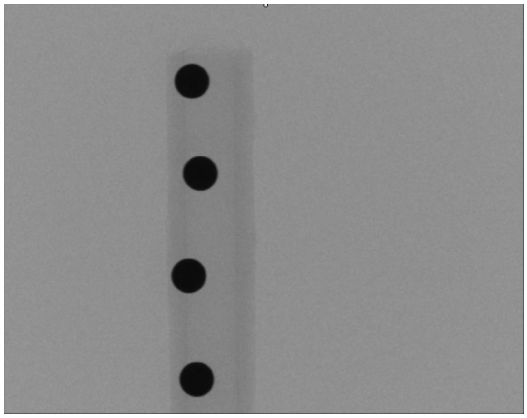


Figure 1. X-ray projection of the four ball pen balls on a hollow plastic rod sample. The balls have a diameter of 1mm.

III. Experiment and Results

In order to apply the proposed calibration method, a simple metal spheres phantom was assembled from ball pens' balls attached to a Q-tip using superglue. The spheres are roughly equidistantly distributed over a range of about 1cm and the whole assembly is mounted eccentrically on the rotary stage of a micro CT setup such that their trajectories when rotating will similarly cover a horizontal range of roughly 1cm. An exemplary projection view is shown in Fig. 1. A total of 360 projections over the full range of 360° have been acquired in fly-by mode. A superposition of a subset of these is shown in Fig. 2 providing a first visual impression of the projected circular trajectories.

Using standard threshold segmentation and subsequent determination of the centers of mass of the spheres' projections, point traces are extracted as depicted in Fig. 3. Based on these projection samples (h_{ij}, v_{ij}), the corresponding projection matrix is determined as described in the previous section.

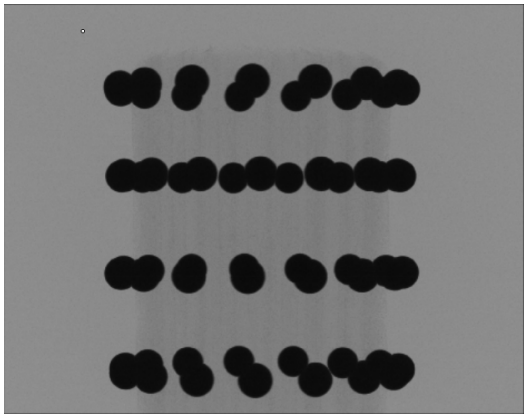


Figure 2. Superposition of several projections showing the emerging virtual stacked-rings-of-markers phantom.

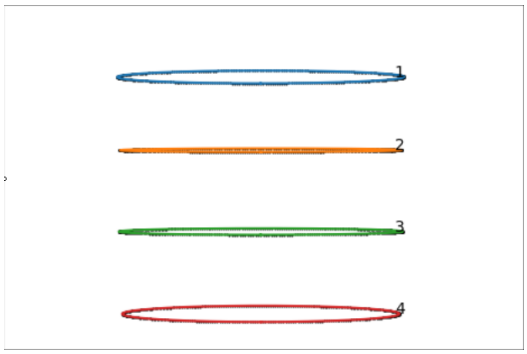


Figure 3. Extracted marker trajectories (h_{ij}, v_{ij}) based on classic segmentation of the X-ray projections and subsequent center-of-mass analysis.

The successful calibration based on these marker traces shall be demonstrated with a tomographic reconstruction of the calibration phantom itself. The reader is kindly asked to excuse the low signal to noise ratio and the limited amount of projections which were originally motivated by the needs of marker segmentation rather than volume reconstruction. Nevertheless, Figures 4 and 5 show reasonable Feldkamp type reconstructions that were based solely on the previously determined projection matrix without further intervention. In particular, no double outlines or streak artifacts as typical signs of imprecisely estimated geometry can be found. Further, sharply defined edges and voxel-scale details within the glue surface can be observed. Notable is a slight difference in edge contrast between tangential and radial directions which can be attributed to the employed fly-by acquisition mode causing blurring in the tangential direction already within the acquired projections.

IV. Discussion and conclusion

Motivated by the difficulty to manufacture precise calibration phantoms for high resolution micro CT systems, a self consistent calibration method has been proposed managing to work with arbitrary traceable features. Similar to other techniques using phantoms of reduced complexity, the assumption of a sufficiently precise rotary stage is central

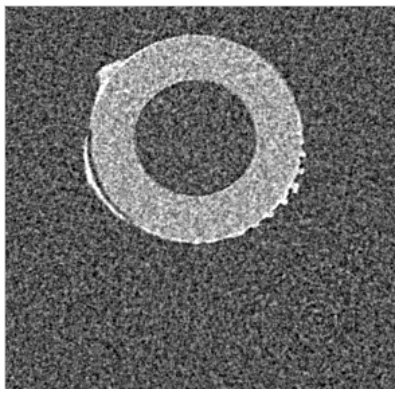


Figure 4. Section of an axial reconstruction slice showing the hollow plastic rod with a layer of glue on the outer surface exhibiting some detailed shapes. On the left, part of the glue detached from the rod. The slice's center can be identified by the subtle ring artifacts in the lower right corner of the displayed section. The slight blurring tangential to the rotational axis is caused by the employed fly-by acquisition mode.

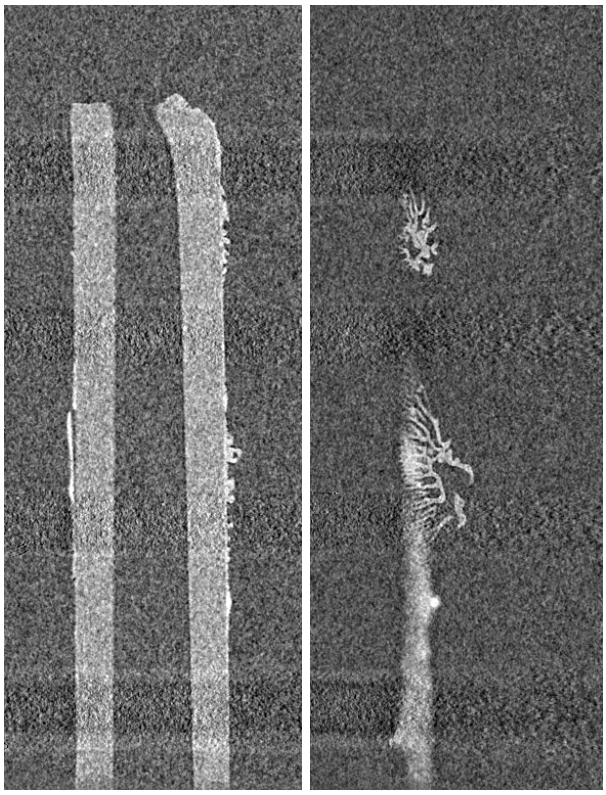


Figure 5. Coronal and sagittal sections of the reconstructed volume. Both sections show details within the glue on the outer surface of the hollow rod. On the left, the same detachment already shown in Fig. 4 is visible. The notable horizontal stripes are caused by the highly absorbing metal spheres, and their increased visibility at the top and bottom of the sample is a consequence of classic Feldkamp cone beam artifacts.

to the present approach so that the considered features actually follow circular trajectories throughout a CT scan. Based on a reduced set of derived observables completely describing the projected trajectories, the calibration problem is formalized as the self consistent solution of a small system of linear equations containing both unknowns with respect to the calibration tracers as well as the projection geometry. In contrast to e.g. the recent publication by [7], all alignment parameters are treated simultaneously in form of the projection matrix similar to [6]. In contrast to the latter, the present scheme starts with a fraction-of-sinusoids model of the projection data rather than ellipsoid fits. The possibility to work with unknown tracers allows in particular also for retrospective calibration of tomographic scans provided that some eccentric traceable features are contained, or similarly for the auto-calibration of time intensive or hard to reproduce scans. A first proof of concept was demonstrated using a metal sphere assembly within a roughly 1cm field of view and achieving a visually artifact-free (with respect to geometry) reconstruction. Further quantitative analyses are currently in progress at the time of writing and will be presented as part of a more extensive article later on.

REFERENCES

[1] Y. Cho, D. J. Moseley, J. H. Siewerdsen, and D. A. Jaffray, "Accurate technique for complete geometric calibration of cone-beam computed tomography systems," *Med. Phys.*, vol. 32, no. 4, pp. 968–983, 2005.

[2] S. Hoppe, F. Noo, F. Dennerlein, G. Lauritsch, and J. Hornegger, "Geometric calibration of the circle-plus-arc trajectory," *Phys. Med. Biol.*, vol. 52, pp. 6943–6960, 2007.

[3] C. Mennessier, R. Clackdoyle, and F. Noo, "Direct determination of geometric alignment parameters for cone-beam scanners," *Phys. Med. Biol.*, vol. 45, pp. 1633–1660, 2009.

[4] F. Noo, R. Clackdoyle, C. Mennessier, T. A. White, and T. J. Roney, "Analytic method based on identification of ellipse parameters for scanner calibration in cone-beam tomography," *Phys. Med. Biol.*, vol. 45, pp. 3489–3508, 2000.

[5] K. Yang, A. L. C. Kwan, D. F. Miller, and J. M. Boone, "A geometric calibration method for cone beam ct systems," *Med. Phys.*, vol. 33, no. 6, pp. 1695–1706, 2006.

[6] D. Gross, U. Heil, R. Schulze, E. Schoemer, and U. Schwanecke, "Auto calibration of a cone-beam-ct," *Med. Phys.*, vol. 39, no. 10, pp. 5959–5970, 2012.

[7] J. Xu and B. M. W. Tsui, *IEEE T. Med. Imaging*, vol. 32, no. 9, pp. 1731–1744, 2013.

[8] N. Navab, A. Bani-Hashemi, M. S. Nadar, K. Wiesent, P. Durlak, T. Brunner, K. Barth, and R. Graumann, "3d reconstruction from projection matrices in a c-arm based 3d-angiography system," *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 119–129, 1998.

ACKNOWLEDGEMENTS

The authors thank Kilian Dremel for the acquistion of the projection data.

Data-Fidelity Impact on Cone-beam Artifact Reduction in Diagnostic CT

Dan Xia, Yan Liu , Zhou Yu, Buxin Chen, Zheng Zhang, Richard Thompson, Emil Y. Sidky, Xiaochuan Pan

***Abstract*—Adequate correction for cone-beam artifacts remains an issue of practical interest in diagnostic CT imaging with a set of detectors that form a large cone angle. It seems that algorithms that were developed from the standard FDK algorithm may have limited success in correction for the cone-beam artifacts especially when soft tissue imaging is of concern. Evidence exists suggesting that optimization-based reconstruction may be more effective in terms of artifact reduction than the FDK algorithm or its variants. In the work, we investigate the effectiveness of the different designs of optimization-based reconstruction in reduction of cone-beam artifacts in diagnostic CT. Specifically, we investigate the performance of optimization-based reconstructions employing three different designs of data fidelity in minimizing the cone-beam artifacts. The three data fidelities are considered because they have been used widely for image reconstruction. Numerical studies from physical phantom data show that, for the imaging system and data conditions considered, the optimization-based reconstructions with the three different data fidelities can correct, at a comparable level of effectiveness, for the cone-beam artifacts observed in the FDK reconstruction.**

I. INTRODUCTION

In advanced diagnostic CT with a large number of detector rows, cone-beam artifacts can be observed in images reconstructed from circular-trajectory data by use of the FDK algorithm or its variants especially for an imaged subject with longitudinally strong contrast variation [1]. Such artifacts may obscure low-contrast soft-tissue anatomy of clinical interest. Empirical schemes have been developed for minimizing the cone-beam artifacts by extrapolating data longitudinally and/or by applying different weighting functions to data. While these schemes may be useful for some cases, their success especially in reconstruction of low-contrast soft-tissue anatomy embedded in a longitudinally varying high-contrast region remains limited.

There exists a high level of interest in optimization-based reconstruction because evidence accumulated in the past several years shows that optimization-based reconstruction may yield images with improved quality over the FDK algorithm and can be more effective in compensating for reconstruction artifacts than existing algorithms, such as the FDK algorithm. In the design of optimization-based reconstruction, the data fidelity plays a key role in impacting image quality and/or artifact correction. In this work, we investigate optimization-based reconstruction employing different data-fidelity terms in the context of correcting for cone-beam artifacts in image reconstruction from circular cone-beam data collected by use of an advanced diagnostic CT with a set of detectors forming a large cone angle. The investigation of the effect of different data-fidelity terms on cone-beam artifact correction is enabled by a CP algorithm, which is tailored from the generic CP algorithm. We use the CP algorithm because it can solve convex optimization problems formed with different data-fidelity terms [2], [3], [4], [5].

D. Xia, B. Chen, Z. Zhang, E. Y. Sidky, and X. Pan are with the Department of Radiology, The University of Chicago, Chicago, IL 60637, USA. Y. Liu, Z. Yu, and R. Thomas are with Canon Medical Research USA, Inc., Vernon Hills, IL 60061, USA.

II. METHODS AND MATERIALS

A. Optimization-based reconstruction

We model the CT-imaging process by using a discrete-to-discrete, linear system,

$$\mathbf{g} = \mathcal{H}\mathbf{f},$$

where vectors $\mathbf{g}$ and $\mathbf{f}$ denote discrete data vector of size M and image vector of size N, and system matrix $\mathcal{H}$ of size $M \times N$ describes the cone-beam X-ray transform. In this work, element h_{ij} of matrix $\mathcal{H}$ is the intersection length of ray j with voxel i.

Using this imaging model, we formulate the image reconstruction as the solution to a constrained total variation (TV)-minimization program containing a data-fidelity term, i.e., [3]

$$\mathbf{f}^{\star} = \underset{\mathbf{f}}{\operatorname{argmin}} D(\mathbf{g}_m, \mathcal{H}\mathbf{f}) \quad \text{s.t.} \quad ||\mathbf{f}||_{TV} \leq t_1 \qquad \text{and} \qquad f_j \geq 0, \tag{1}$$

where $D(\mathbf{f})$ denotes the data-fidelity term that gives rise to measure of inconsistency between measured data $\mathbf{g}_m$ and model data $\mathcal{H}\mathbf{f}$; $||\mathbf{f}||_{TV}$ the image TV, and $t_1 > 0$ a pre-selected image-TV-constraint parameter. Clearly, the optimization program, and consequently the final image, depends strongly upon the specific form of data fidelity $D(\mathbf{g}_m, \mathcal{H}\mathbf{f})$. In the work, we investigate the impact of three different data-fidelity terms, which have been considered for image reconstruction:

a) data-ℓ_2 fidelity:

$$D_{\ell_2}(\mathbf{g}_m, \mathcal{H}\mathbf{f}) = ||\mathbf{g}_m - \mathcal{H}\mathbf{f}||_2^2, \tag{2}$$

where $||\cdot||_2^2$ denotes the square of ℓ_2-norm;

b) data-KL fidelity:

$$D_{KL}(\mathbf{g}_m, \mathcal{H}\mathbf{f}) = \sum_{i=1}^{M} [-\mathcal{H}\mathbf{f} + \mathbf{g}_m + \mathbf{g}_m \ln(\mathbf{g}_m) - \mathbf{g}_m \ln(\mathcal{H}\mathbf{f})]_i, \tag{3}$$

where $D_{KL}(\mathbf{g}_m, \mathcal{H}\mathbf{f})$ denotes the Kullback-Leibler divergence and i denotes the index of the ith entry;

c) data-ℓ_1 fidelity:

$$D_{\ell_1}(\mathbf{g}_m, \mathcal{H}\mathbf{f}) = ||\mathbf{g}_m - \mathcal{H}\mathbf{f}||_1, \tag{4}$$

where $||\cdot||_1$ denotes the ℓ_1-norm.

The optimization programs mentioned above are non-smooth and convex, and they can be solved by use of a primal-dual algorithm. In the work, we focus on using a specific primal-dual algorithm, i.e. the Chambolle-Pock (CP) algorithm [2], [6].

B. Data acquisition

The cone-beam projection data were collected in Fujita Health University Hospital from a physical head phantom by using an advanced Canon Medical Systems 320-slice diagnostic CT scanner with a circular trajectory. In this scanner, the source-to-isocenter distance and source-to-detector distance are 600 mm and 1073 mm, respectively. The curved detector consists of 896 $\times$ 320 pixels, thus forming a fan-angle of 49.2° and a cone angle of 15°. Data were acquired at 1200 projection views uniformly distributed over an angular range of 2π. Prior to reconstruction, scatter and beam-hardening have been corrected for with the Canon's well-established algorithm. The FDK reconstruction with a Hanning filter and a cutoff at 0.6 of the Nyquist frequency from the data set was performed and used as a reference for benchmarking the CP reconstructions.

III. RESULTS

A. Visualization study

We have performed image reconstruction from the head-phantom data by using the CP algorithm to solve the optimization program in Eq. (1) with one of three different data-fidelity terms in Eqs. (2)-(4). The reconstructed images within 2D transverse, sagittal, and coronal slices were displayed in Figs. 1 and 2, respectively. In order to reveal the details of the reduction of the cone-beam artifacts, a narrow display window [0, 100] HU is applied to visualizing the images. Cone-beam artifacts can be observed in FDK reconstructions, highlighted by circles and arrows, which obscure low-contrast structures within the region. Conversely, all of optimization-based reconstructions employing different data-fidelity terms show generally comparable performance in terms of cone-beam artifact reduction. However, remaining streak artifacts near the skull top can also be observed in the reconstruction with data-KL fidelity, as indicated by the arrow in the third column in Fig. 2, suggesting a reduced effectiveness of the data-KL fidelity suppressing some of the cone-beam artifacts.

B. Quantitative analysis

In addition to visual inspection described above, we have also performed numerical studies to evaluate quantitatively the properties of the optimization-based reconstructions, including spatial resolution and contrast resolution, for the regions in which the cone-beam artifacts can be observed in FDK reconstruction.

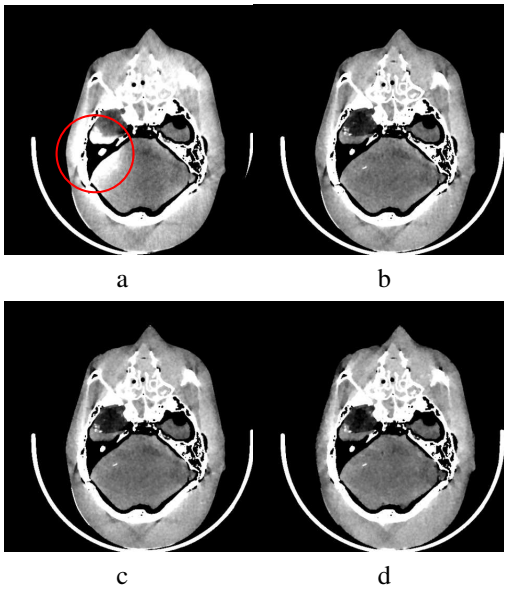


Figure 1. Images within a transverse slice reconstructed from 1200-view data by use of the FDK-type algorithm (a)and the CP algorithms for optimization programs with data-ℓ_2 fidelity (b), data-KL fidelity (c) and data-ℓ_1 fidelity (d). A narrow display window, [0, 100] HU, is used to reveal the details of the cone-beam artifacts enclosed by the red circle.

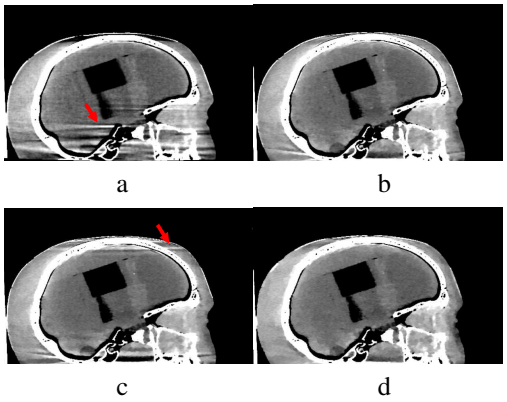


Figure 2. Images within a sagittal slice reconstructed from 1200-view phantom data by use of the FDK-type algorithm (a) and the CP algorithms for optimization programs with data-ℓ_2 fidelity (b), data-KL fidelity (c), and data-ℓ_1 fidelity (d). A narrow display window, [0, 100] HU, is used to reveal the details of the cone-beam artifacts, which are indicated by the red arrow in panel (a).

1) Numerical phantom design: In this study, a numerical phantom including bar patterns and a low-contrast sphere was generated on a fine image grid of $1024 \times 1024 \times 640$ with a pixel size of $0.23 \times 0.23 \times 0.25$ mm^3, and was used to evaluate both spatial and contrast resolution of the optimization-based reconstructions. The eight bar patterns have varying line-pair densities (i.e., 10.7 cm^{-1}, 7.1 cm^{-1}, 5.3 cm^{-1}, 4.3 cm^{-1}, 3.6 cm^{-1}, 3.0 cm^{-1}, 2.7 cm^{-1}, 2.4 cm^{-1}), whereas the low-contrast sphere possesses a contrast level of 20 HU. From this numerical phantom, we used the same scanning configuration as that in Sec. II-B to generate the cone-beam projection data. In an attempt to incorporate the partial volume effect which may affect the spatial resolution, we subdivided uniformly the detector element into 4×4 subelements. Therefore, for each detector element, there exist 16 straight lines that are used

for computing projection data. For each detector element, we computed the integration of the object function over these lines connecting the source and detector subelements, then took the average of these integrations as the projection data. By adding the generated cone-beam projection data to the collected head-phantom data, we obtained cone-beam projection data for the physical head phantom embedded with the bar patterns and low-contrast sphere designed. The optimization-based reconstructions with different data-fidelity terms were subsequently from the data.

2) Spatial-resolution study: The spatial resolution was first determined qualitatively by visual inspection of the bar patterns. We displayed the optimization-based reconstructions within 2D transverse and coronal slices containing the bar patterns in Fig. 3, along with the corresponding FDK reconstructions. It can be observed that the reconstructed bar patterns in the FDK reconstruction are blurred that bar patterns of only up to 3.6 cm^{-1} can be resolved. However, in the optimization-based reconstructions, the visually resolvable bar patterns of 7.1 cm^{-1} can be obtained.

In an attempt to quantitatively characterize the reconstruction performance in terms of spatial resolution, the modulation transfer function (MTF) of the bar patterns was calculated. The MTF values of a CT system can be obtained by use of the standard deviation of pixel values in a region of interest (ROI) containing a bar pattern [7]. Using the reconstructed images within the 2D transverse slice, we calculated the MTF values, which are shown in Fig. 4. It can be observed that, in general, the MTF values obtained from the optimization-based reconstructions are higher than those obtained from the FDK reconstruction, suggesting optimization-based reconstructions yield spatial resolution higher than does the FDK reconstruction. The MTF values obtained from the optimization-based reconstructions were reduced to 20% at the spatial frequency around 7.1 cm^{-1}, while the MTF value obtained from the FDK reconstruction was reduced to 20% at the spatial frequency 3.6 cm^{-1}; this observation is consistent with our visual inspection.

3) Contrast-resolution study: Again, we first performed visually an inspection of the contrast resolution of the low-contrast sphere reconstructed, as shown in Fig. 5. It can be observed that the low-contrast sphere was obsecured by the cone-beam artifacts in the FDK reconstruction. Conversely, the cone-beam artifacts appear to be considerably suppressed in optimization-based reconstructions with data-ℓ_2 fidelity, data-KL fidelity, and data-ℓ_1 fidelity, resulting in a low-contrast sphere that can readily be discerned.

In an attempt to quantitatively characterize the contrast resolution for the optimization-based reconstructions, we define a contrast-resolution metric below,

$$\text{CNR} = \frac{|\mathbf{f}_s - \mathbf{f}_b|}{\sqrt{\sigma_s^2 + \sigma_b^2}},$$

where $\mathbf{f}_s$ and σ_s denote the mean and standard deviation within a region of interest (ROI), and $\mathbf{f}_b$ and σ_b denote the mean and standard deviation within a selected background region. It can be expected that the standard deviation, σ_s, within a reconstructed ROI with the cone-beam artifacts can be large due to the gray-level non-uniformity. CNRs were

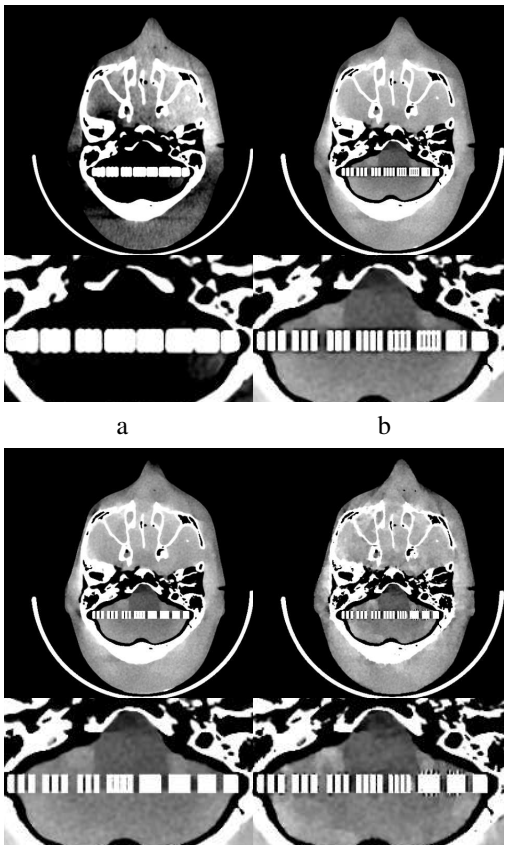


Figure 3. Reconstructions and the corresponding zoomed-in views of ROI images within the transverse slice reconstructed from 1200-view data by use of the FDK algorithm (a) and the CP algorithms with data-ℓ_2 fidelity (b), data-KL fidelity (c), and data-ℓ_1 fidelity (d). A narrow display window [0, 100] is used to reveal the details of the cone-beam artifacts.

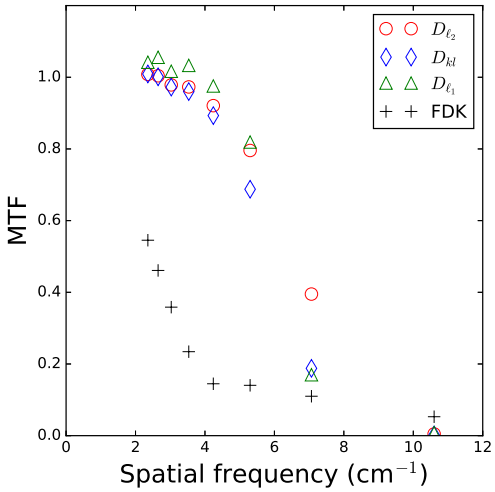


Figure 4. MTF values obtained from the images reconstructed by use of the FDK algorithm (+) and the CP algorithms for optimization programs with data-ℓ_2 fidelity (o), data-KL fidelity ($\Diamond$), and data-ℓ_1 fidelity ($\triangle$).

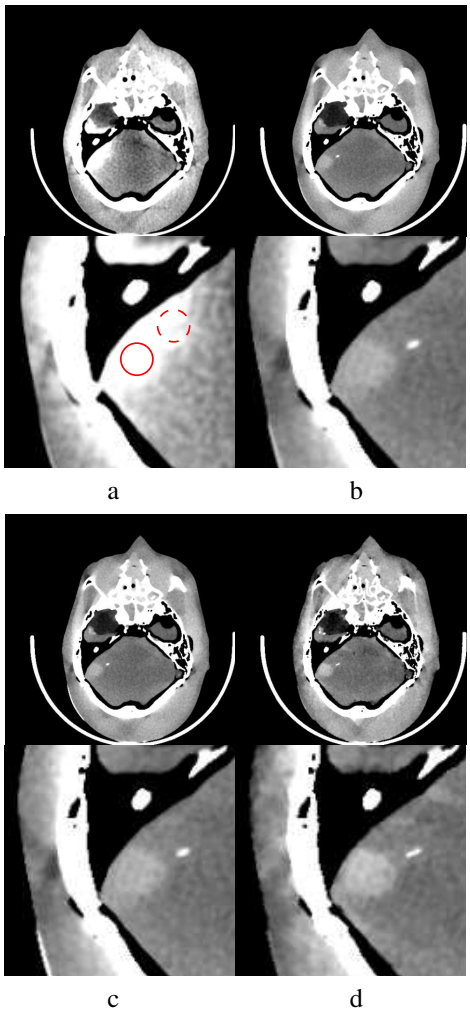


Figure 5. Reconstructions and the corresponding zoomed-in views of ROI images within the transverse slice reconstructed from 1200-view data by use of the FDK algorithm (a) and the CP algorithms with data-ℓ_2 fidelity (b), data-KL fidelity (c), and data-ℓ_1 fidelity (d). The narrow display window [0, 100] HU is used to reveal the details of the cone-beam artifacts.

computed in ROIs within the sphere in the reconstructed 2D transverse, sagittal, and coronal slices, indicated by the solid circle in the first columns of Fig. 5, and the corresponding background regions are indicated by the dashed circle. We show in Fig. 6 the CNRs calculated from images reconstructed by use of these three different data fidelity terms: data-ℓ_2 fidelity, data-KL fidelity, and data-ℓ_1 fidelity, along with the CNRs computed from the corresponding FDK images. It can be observed that cone-beam artifacts lower the CNRs in the FDK image whereas the optimization-based reconstruction yields an enhanced CNR. The CNR results corroborated visual inspection results of Fig. 5.

IV. CONCLUSION

In this work, we investigated the effectiveness of optimization-based image reconstructions in correction for cone-beam artifacts in diagnostic CT with a set of detectors forming a large cone angle. The study was carried out with real data of a physical head phantom collected with a Canon Medical Systems 320-slice diagnostic CT. In an attempt to

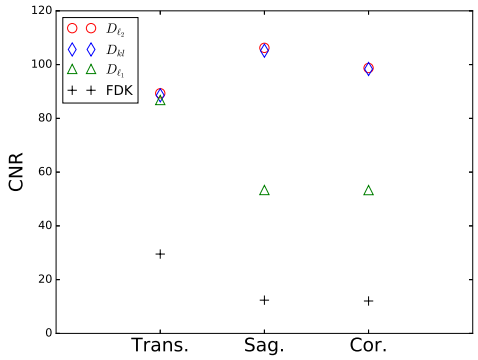


Figure 6. CNRs obtained from the 2D slice images reconstructed by use of the FDK algorithm (+) and the CP algorithms for optimization programs with data-ℓ_2 fidelity (o), data-KL fidelity (◇), and data-ℓ_1 fidelity (△).

investigate the impact of a data-fidelity term in an optimization program on the cone-beam artifact reduction, we have studied three different data-fidelity terms that are used widely for image reconstruction. The results of our study show that optimization-based reconstructions have the potential to reduce cone-beam artifacts observed in the FDK reconstructions and that the three data-fidelities appear to possess a comparable level of effectiveness for reducing the cone-beam artifacts.

In order to quantitatively evaluate the performance of the optimization-based reconstruction, two quantitative metrics, MTF and CNR were used for characterizing the spatial and contrast resolution. Results of quantitative study also indicate that the optimization-based reconstruction improve image quality over the FDK reconstruction in terms of spatial and contrast resolution.

ACKNOWLEDGMENT

This work was supported in part by NIH R01 Grants Nos. CA182264 and EB018102. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official NIH views.

REFERENCES

[1] L. A. Feldkamp, L. C. Davis, and J. W. Kress, "Practical cone-beam algorithm," *J. Opt. Soc. Am.*, vol. A1, pp. 612–619, 1984.

[2] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Math. Imag. Vis.*, vol. 40, pp. 120–145, 2011.

[3] E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Phys. Med. Biol.*, vol. 53, pp. 4777–4807, 2008.

[4] Z. Zhang, X. Han, E. Pearson, C. Pelizzari, E. Y. Sidky, and X. Pan, "Artifact reduction in short-scan CBCT by use of optimization-based reconstruction," *Phys. Med. Biol.*, vol. 61, pp. 3387–3406, 2016.

[5] D. Xia, D. A. Langan, S. B. Solomon, Z. Zhang, B. Chen, H. Lai, E. Y. Sidky, and X. Pan, "Optimization-based image reconstruction with artifact reduction in C-arm CBCT," *Phys. Med. Biol.*, vol. 61, pp. 7300–7333, 2016.

[6] E. Y. Sidky, J. H. Jørgensen, and X. Pan, "Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm," *Phys. Med. Biol.*, vol. 57, pp. 3065–3091, 2012, PMCID:PMC3370658.

[7] R. T. Droege and R. L. Morin, "A practical method to measure the MTF of CT scanners," *Med. Phys.*, vol. 9, pp. 758–760, 1982.

C-arm CT imaging using the extended line-ellipse-line trajectory: first implementation and initial results

Zijia Guo, Günter Lauritsch, Andreas Maier, Patrick Kugler, Mohammad Islam, Florian Vogt, Frédéric Noo

***Abstract*—In previous work, we proposed a novel data acquisition geometry, called the Extended LEL trajectory, for C-arm CT imaging in interventional radiology. This novel geometry aims at enabling larger axial field-of-view coverage without cone-beam artifacts for imaging with a full X-ray beam as well as with a collimated X-ray beam used for scatter reduction purposes. In this work, we report on a first implementation of the Extended LEL trajectory on a state-of-the-art C-arm system. Highly satisfactory results are shown in terms of trajectory fidelity and repeatability. Suitability of the data for head imaging is also demonstrated using a Rando head phantom without and with 50% beam collimation.**

I. Introduction

C-arm Computed Tomography (CT) is a popular imaging tool in interventional radiology. Currently, the circular short-scan is the preferred data acquisition geometry for C-arm CT. However, the circular short-scan presents two major shortcomings: data incompleteness in terms of Tuy's condition [1], and limited axial coverage. Another important issue that affects image quality is scatter, particularly because the anti-scatter grid used in interventional C-arm systems is suboptimal due to the large variations in source-to-detector distances required by the clinical demands for the system. An alternative option to reduce scatter is axial collimation of the beam, but this further reduces the already limited axial coverage.

To overcome the issues mentioned above, novel data acquisition geometries should be investigated. Recently, we proposed using the Line-Ellipse-Line (LEL) trajectory and its extended version, called the Extended LEL trajectory. This trajectory is designed to provide complete data, as well as extended axial coverage both without and with beam collimation. The Extended LEL trajectory is a continuous curve consisting of tilted ellipses joined together by segments of line. The number of ellipses and the distance separating them define the axial coverage, while the line segments ensure data completeness. The angular tilt applied to the ellipses is critical for image reconstruction at each location from a minimum amount of contiguous projections while allowing uninterrupted data acquisition without double coverage of view positions, as observed with the extended arc-line-arc trajectory [2].

The theoretical properties of the Extended LEL trajectory were thoroughly studied in [3], [4]. Also, experiments from computer simulated data demonstrated accurate image reconstruction with no cone-beam (CB) artifacts and strong robustness to data sampling [3]. In this work, we report on a first implementation of the Extended LEL trajectory on a state-of-the-art C-arm system. The implemented trajectory consists of two ellipses and three lines. The quality of this first implementation was evaluated in terms of geometry fidelity and reproducibility, and also in terms of consistency with theory for CB image reconstruction of an anthropomorphic head phantom without and with beam collimation.

Z. Guo and F. Noo are with the Department of Radiology and Imaging Sciences, University of Utah, USA. G. Lauritsch, P. Kugler and F. Vogt are with Siemens Healthcare GmbH, Germany. A. Maier and Z. Guo are with the University of Erlangen-Nürnberg, Germany. M Islam is with SYNG4 GmbH, Germany.

II. Trajectory Implementation

We implemented the Extended LEL trajectory on a Siemens ARTIS pheno system (Siemens Healthcare, GmbH, Forchheim, Germany), which is a multi-axis robotic floor-mounted C-arm with flat panel detector. The trajectory was pre-defined through the use of 47 control points, using upsampling between the control points to enable data acquisition at 320 positions. The pre-defined trajectory was loaded to the C-arm system via a newly-developed programmable patch. X-ray exposure was triggered by angulations rather than time stamps, to ensure lower impact of motor accelerations and decelerations on geometrical repeatability of measurements.

Technical details regarding the trajectory configuration are listed in Table I. The limited number of projections was due to current software limitations. The relative number of projections between lines and ellipses was based on the requirement of keeping the physical distance between source positions nearly constant. The axial coverage parameters were chosen so as to cover a full head phantom without collimation as well as with a beam collimation of up to 50%.

Source-to-isocenter	$785\ mm$
Source-to-detector	$1300\ mm$
Field-of-view radius	$114\ mm$
Length of each line	$71.6\ mm$
Angular length of each ellipse	210^o
Ellipse-to-ellipse distance	$60\ mm$
Axial tilt of each ellipse	$12\ mm$
Total axial coverage	$192\ mm$
# of control points	47
# of views	320
# of views per ellipse	140-140
# of views per line	13-14-13
Detector pixel size	$0.308 \times 0.308\ mm^2$
Detector size	1248×928

TABLE I: Trajectory configuration

III. Evaluation tests

To assess the quality of the implementation, we performed tests evaluating the fidelity of the geometry and its repeatability, and also tests evaluating the suitability of the data for CB reconstruction of a head phantom.

A. Geometry fidelity and repeatability

The physical position of the source and detector during data acquisition was determined using a geometrical phantom, which we call the PDS-4 phantom. This phantom is specifically designed for calibration of trajectories with large axial coverage. See [5] for details on this phantom. By analyzing each CB projection of the PDS-4 phantom, we obtain full information on the position of the source and the detector relative to the phantom for each location where the X-ray source is triggered.

Geometry fidelity of the data acquisition was performed by comparing the calibrated source positions with the pre-defined trajectory. Geometry repeatability was performed through comparison of source and detector placements from one repetition of the data acquisition to another; the comparison was in terms of accuracy in backprojection for voxels within the field-of-view. The repeatability was assessed over 5 immediate repetitions and also over 3 short-term repetitions. The immediate repetitions amount to immediately repeating the protocol. The short-term repetition includes the utilization of a different protocol prior to repetition of the Extended LEL scan to force the system to recall it from another position. The phantom placement remained unchanged between all scan repetitions to avoid dealing with changes in the world coordinate system, which is attached to the phantom.

B. Image reconstruction

A Rando head phantom was used for this test. The phantom was positioned on a foam holder in such a way that transversal truncation is avoided. Axial positioning of the Extended LEL trajectory was such that the first ellipse was centered on the base of the skull and the second one was centered on the middle of the brain.

Prior to image reconstruction, the projection data was corrected for scatter and beam-hardening, and transformed into line integrals using the conventional log transformation with air scan providing information on the incoming fluence.

Reconstruction was formulated as a penalized least-square optimization problem with the penalty term applied to difference between neighbor voxels. This optimization problem was solved in an iterative manner using the GISTA method [6]. Both quadratic regularization and total variation were considered. The quadratic regularizer was used for its ability to produce results similar to a filtered-backprojection reconstruction, which was previously applied to ideal data [3] and which we plan to apply to real data in the future. Total variation was used to mitigate artifacts due to few view sampling, which could be anticipated to be significant given the limited number of measurements over each ellipse.

To assess the suitability of the implemented trajectory with beam collimation, we performed reconstruction from full projections as well as from axially cropped projections mimicking a 50% collimation of the beam.

All reconstructions were performed on a volume of $270 \times 270 \times 300$ cubic voxels covering the field-of-view with a voxel size of $1\,mm$. The number of iterations was 1000. This number was selected on the basis that the difference between using 2000 versus 1000 iterations was negligible on one test case.

IV. Evaluation results

A. Geometry fidelity

Figure 1 shows the Extended LEL trajectory in a 3-D view, as well as its projections on the (x, y)-plane, on the (x, z)-plane, and on the (y, z)-plane. The pre-defined trajectory is indicated by the solid curve, whereas the real trajectory, acquired from the C-arm, is shown with dots. The pre-defined trajectory was generated by interpolating the control points, and performing afterwards a registration to the real trajectory using the iterative closest point method.

Figure 1 shows strong fidelity in the trajectory shape. The ellipses appear as ellipses with the desired tilt, and the line segments appear fairly linear. Also, the transitions between ellipses and lines essentially occur at the desired places. Small deviations do exist. These are best appreciated using zoom plots.

Figure 2 shows a zoom on the connection between the middle line and the second (top) ellipse, and also shows a zoom on the connection between the top ellipse and the last line segment. The regions for these two zooms are indicated as boxes 1 and 2 in Figure 1. More deviations can be observed for the source positions on the lines than for the source positions on the ellipses, particularly near the transition from line to ellipse. However, the deviations are fairly small, on the order of 2-3 mm, so that we can still say that, overall, the geometry fidelity is highly satisfactory.

B. Geometry repeatability

As mentioned earlier, geometry repeatability was evaluated in terms of backprojection accuracy. For each view and each voxel, we computed the detector position to be used for backprojection and assessed how this position changes from one scan repetition to the next. This change in position was scaled by the CB geometry magnifaction factor to reflect an error near the center of the field-of-view. Our results are shown in Figure 3 in the form of box plots that display, over the views, the mean variation that was observed over the voxels.

For the immediate repetitions, the average backprojection error was less than $0.05\,mm$ for at least 75% of the views. The short-term repeatability had a larger average backprojection error, but this error was still less than $0.15\,mm$ for at least 75% of the views, with errors below $0.4\,mm$ for the outlier views. Compared with the standard circular short-scan, the backprojection error for the Extended LEL trajectory is largely comparable (not shown here) and thus promising for this first implementation.

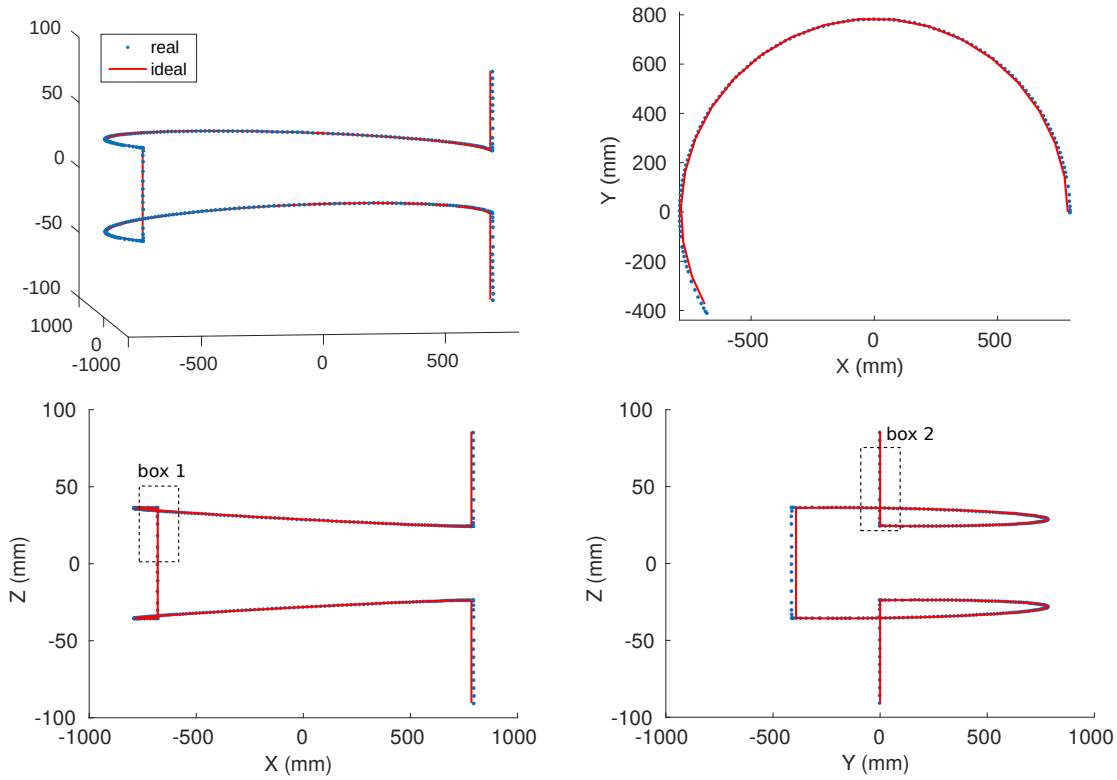


Fig. 1: Plots of the Extended LEL trajectory: 3-D view (top left) and projections on the (x, y)-plane (top right), on the (x, z)-plane (bottom left), and on the (y, z)-plane (bottom right) are shown. The pre-defined trajectory is indicated by the solid curve, whereas the real one, obtained from the C-arm, is shown with dots. All axes are scaled in mm.

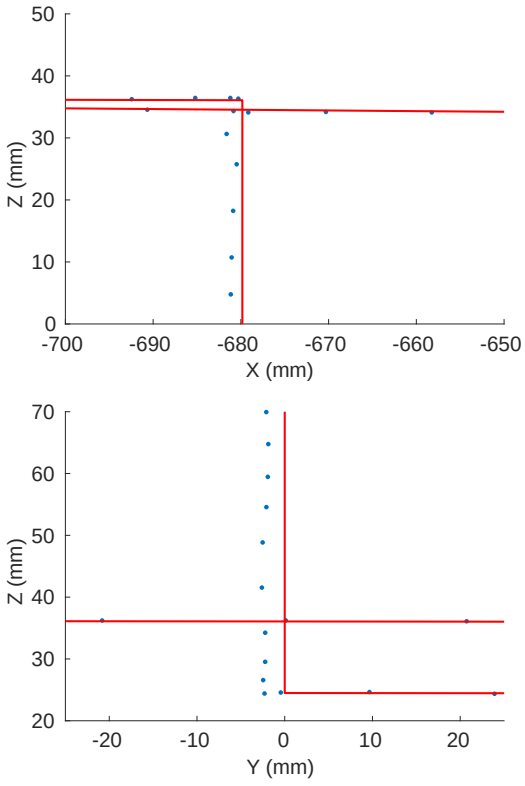


Fig. 2: Zooms on portions of the Extended LEL trajectory indicated as boxes 1 and 2 in Figure 1. All axes are scaled in mm.

C. Image reconstruction

Figure 4 shows a few representative projections along the Extended LEL trajectory. Our first reconstruction of the head phantom showed unexpected artifacts. By inspecting the projection data, we identified that the edges of the collimator cause data inconsistencies that are irrelevant for a circular short-scan reconstruction but matter for the Extended LEL trajectory. After cropping out the inconsistent part from the measurement (about 30 detector rows at both top and bottom of the detector), a much better image quality was obtained. Figure 5 and 6 show the reconstruction with data cropping, using a narrow grayscale window width of 400 HU. Fine details of the bony structures appear clear and no severe CB artifacts are observed. Due to the limited number of projection views and due possibly also to a few select views that showed less repeatability, streak artifacts are observed. As anticipated, more streaks are observed with quadratic regularization than with total variation, which perfoms well at reducing the strength of these streaks.

The data acquisition process with a simulated 50% collimation of the beam is illustrated in Figure 7, and Figure 8 shows reconstructions from this collimated data using the total variation penalty. In this collimated set-up, most voxels are only seen by one ellipse. Although fewer measurements are used for reconstruction, there is no significant difference in noise, primarily because the total variation strength was increased to maintain good mitigation of streak artifacts. This experiment demonstrated that the trajectory still delivered

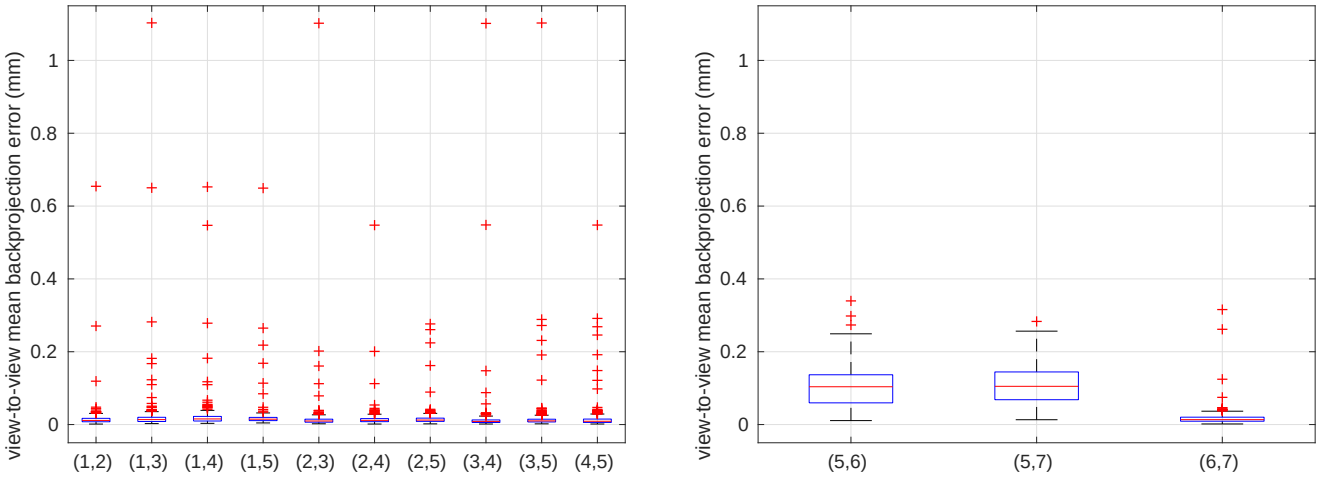


Fig. 3: Box plots showing the view-to-view variations in mean backprojection error over voxels covering the field-of-view, when considering immediate scan repetitions (left) and short-term scan repetitions (right). The horizontal axis indicates the pairs of repeated scans under comparison. The vertical axis shows the distribution of the mean backprojection error over the 320 views. Each box accounts for 75% of the views. The crosses mark outliers.

sufficient data for reconstruction, as anticipated by the theory. (Naturally, the upper and lower portion of the phantom are now truncated.)

V. Conclusion and Discussion

We reported on a successful first implementation of the Extended LEL trajectory on a state-of-the-art C-arm system used in interventional radiology. Promising results were shown in terms of trajectory fidelity, data acquisition repeatability, and data suitability for head imaging. This work demonstrates the feasibility of data acquisition on a C-arm system with a trajectory other than the standard circular short scan for axially extended field-of-view imaging, without and with beam collimation for scatter reduction purposes. Future work will focus on refining the software component to enable the utilization of more projections. Once more projections are available, analytical reconstruction, which would be clinically more practical, will be investigated.

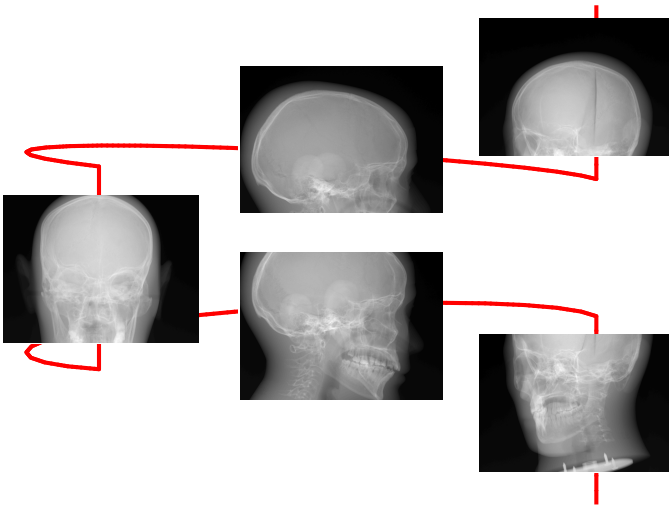


Fig. 4: Illustration of a few representative projections along the Extended LEL trajectory obtained when scanning the Rando head phantom.

Disclaimer

The concepts and information presented in this paper are based on research and are not commercially available.

Acknowledgment

This work was supported by a research contract with Siemens Medical Solutions USA, Inc. This work was also partly supported by the Erlangen Graduate School in Advanced Optical Technologies (SAOT); the authors gratefully acknowledge funding of SAOT by the German Research Foundation (DFG) in the framework of the German excellence initiative.

References

[1] Tuy, H. K., "An inversion formula for cone-beam reconstruction," *SIAM J. Appl. Math.* **43**, 546–552 (June 1983).

[2] Yu, Z., Wunderlich, A., Dennerlein, F., Lauritsch, G., and Noo, F., "Line plus arc source trajectories and their R-line coverage for long-object cone-beam imaging with a C-arm system," *Phys. Med. Biol.* **56**, 3447–3471 (May 2011).

[3] Yu, Z., Noo, F., Lauritsch, G., and Hornegger, J., "Extended-volume image reconstruction using the ellipse-line-ellipse trajectory for a C-arm system," in [*Proceedings of the 12th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*], 245–248 (June 2013).

[4] Guo, Z., Noo, F., Maier, A., and Lauritsch, G., "Properties of the ellipse-line-ellipse trajectory with asymmetrical variations," in [*Medical Imaging 2016: Physics of Medical Imaging*], **9783**, 97832Z, International Society for Optics and Photonics (2016).

[5] Yu, Z., Maier, A., Lauritsch, G., F., V., Schönhorn, M., C., K., Hornegger, J., and Noo, F., "Axially extended-volume C-arm CT using a reverse helical trajectory in the interventional room," *IEEE Trans. Med. Imaging* **34**, 203–215 (Jan 2015).

[6] Loris, I. and Verhoeven, C., "On a generalization of the iterative soft-thresholding algorithm for the case of non-separable penalty," *Inverse Problems* **27**(12), 125007 (2011).

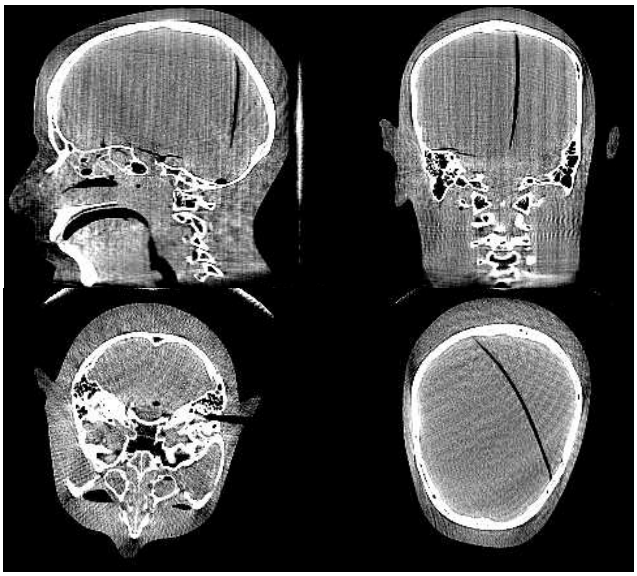


Fig. 5: Reconstruction of the Rando head phantom with quadratic regularization, using all acquired data. Grayscale: $[-200, 200]$ HU.

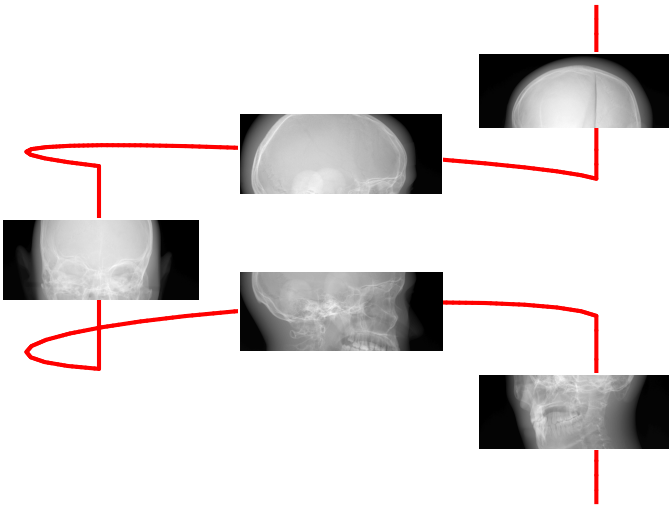


Fig. 7: Illustration of a few representative projections along the Extended LEL trajectory obtained when scanning the Rando head phantom with a simulated 50% collimation of the beam.

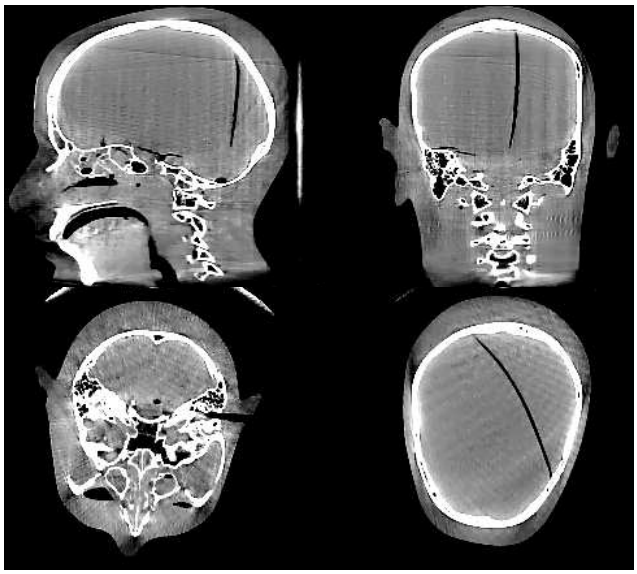


Fig. 6: Reconstruction of the Rando head phantom with total variation regularization, using all acquired data. Grayscale: $[-200, 200]$ HU.

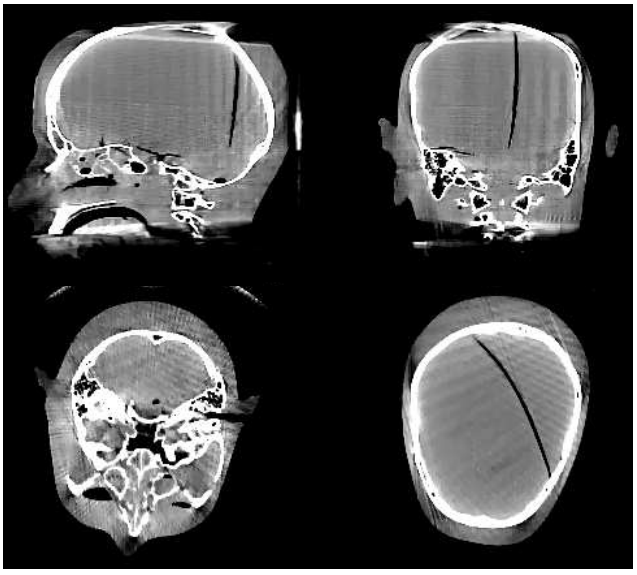


Fig. 8: Reconstruction of the Rando head phantom with total variation regularization, using only the central 50% of the data acquired at each source position, to simulate a 50% beam collimation. Grayscale: $[-200, 200]$ HU.

A Practical Dose Efficient Reconstruction Algorithm for Circular Cone Beam X-ray Tomography

Hongbin Guo, Abdelaziz Ikhlef, Chuang Miao, and Xuelin Cui

***Abstract*—It is well-known that circular cone beam X-ray scan trajectory does not provide complete data for exact image reconstruction. Most pixels in the end slices even do not receive enough projection angle for basic two dimensional reconstruction. That causes difficulties for the popular analytical approximation reconstruction method FDK. Current reconstruction methods usually generate a small volume with z-coverage shorter than actual beam illumination width at z-axis, which is the claimed nominal z-coverage of CT products. In this study we propose a new algorithm for sequential axial scans, which reconstruct the full z-coverage volume without discarding any projection data or introducing any overlap between consecutive rotations. It is dose efficient and can reconstruct larger volume than volume shorten or overlapping algorithms. In this algorithm the projection data of consecutive rotations are jointed as a whole to complement each other. The sampling gap, "missing" data, between two rotations are filled in with interpolation from the data across two rotations. The proposed algorithm has been validated with simulated data. The data set is quite challenge. With large cone angle data, the reconstructed images do not show strong artifacts at joint slices no matter the high contrast regions or the off center regions.**

***Index Terms*—Computer Tomography, Image reconstruction, Step and Shoot, Circular Cone Beam, Artifacts.**

I. Introduction

X-ray Tomography is widely used clinically as a tool for disease diagnosis. However, the fundamental circular cone beam (CCB) scans suffers from artifacts due to data incompletion, i.e. projection data does not satisfy the data sufficiency condition (DSC), which hinders the exact reconstructions. In clinical practice, the product users and developers are more in favor of the analytical approximation reconstruction algorithms, which is faster and less sensitive to noise as compared with the exact reconstruction algorithms. FDK (Feldkamp - Davis - Kress) type algorithms are usually adopted in current medical scanners. FDK type methods are extended from two-dimensional reconstruction methods which requires projections cover angle more than $\pi+\Gamma$ for fan beam scanning, here Γ is the full fan angle. Unfortunately, if our reconstruction volume z-coverage is set to the full illumination range at z-axis, some regions at the end slices of the volume does not satisfy this weaker requirement either even for a full circular, 360 degree, scan. see Figure 1. For FDK type reconstruction the largest reconstruction region contains the voxels which have projections with angle span more than $\pi + \Gamma$. Figure 1 shows that there are 4 slices at each end of the volume do not have enough projections to reconstruct the full slice with FDK if we set the FOV to 500 mm, where the geometry of Minfound 64 row scanner, see Section III, is used. The two end slices only have the ONE center voxel satisfy FDK reconstruction condition! If apply extrapolation to estimate data beyond the maximum cone angle strong artifacts may be introduced. Due to cone angle effects even the region with $\pi+\Gamma$ projections may have artifacts. Some artifacts reduction algorithms are introduced by adding a forward projection, [1], [2], [3], or adding a partial exact reconstruction to compensate the basic FDK reconstruction, [4], [5]. Various weighting methods are also proposed for FDK to reduce the artifacts of the reconstructed region, [6], [7], [8], [9].

In this study we are interested in sequential CCB scans, which is also called step-and-shoot (SAS) scanning. For this type scanning the reconstruction volume z-coverage can be set to less than the full illumination range at z-axis, [10], [11], so that the joint slices also have good quality. Though each individual axial scan still do not have complete data for exactly reconstruction of each independent scan, the neighbouring circular scans can provide compensation data each other within the region between two circular trajectory. Thus we can obtain better images for the central region with FDK-type methods. We propose algorithm to reconstruct larger volume as compared with the method presented in [10], [11], where shorten or overlapping is applied.

The rest of the paper is organization as follows. In Section II we details our algorithm step by step. The proposed algorithm is validated by simulated data in Section III. Conclusions are presented in Section IV.

II. METHODS

The fundamental FDK algorithm is extended from two-dimensional reconstruction algorithm. To mitigate cone beam artifacts various weighting strategies have been developed to be incorporated in the basic FDK. In practical industry field the rays in fan geometry are usually rebinned to data with cone parallel geometry. Algorithm 1 list the FDK reconstructions step with the use of rebinning and weighting.

Mathematically the last three step of Algorithm 1 can be expressed as follows

$$f(\mathbf{x}) = \frac{1}{2}\int_0^{2\pi} cos(v)\mathbf{w}(v)\int_{-\infty}^{\infty} p(u',v,\theta)h(u-u')\mathrm{d}u'\mathrm{d}\theta \tag{II.1}$$

here $p(u,v,\theta)$ is the rebinned cone parallel projection data, $h(u)$ is the ramp-filtering kernel, detector panel coordinates u and v are actually functions of $\mathbf{x}$ and θ: $u = u(\mathbf{x},\theta)$ and

All authors are affiliated with FMI Medical Systems, Inc. Solon, OH 44139, USA. Communication email address: Hongbin.Guo@fmimaging.com

Algorithm 1 Weighted rebinning FDK Algorithm.

1: Starting with the circular scanning data of a full scan, 360 degree, with multi-row cylindrical detector panel $p(\gamma, v, \beta)$. Here γ and β represent fan angle of each channel and source position angle of each projection respectively, variable v stands for the z-coordinate of the detectors at iso axis.
2: Rebinning the projection data $p(\gamma, v, \beta)$ from fan geometry to parallel geometry for each row v to produce $p(u, v, \theta)$. Here (u, v) is local coordinate for virtual flat detector plane passing iso of the CT system. Variable u stands for the distance to iso of each channel.
3: Filtering the rebinned projection by ram filtering kernel $h(u)$.
4: Applying cosine weight.
5: Performing weighted backprojection.

$v = v(\mathbf{x}, \theta)$. 0 If we reconstruct a volume for the full z-coverage at z-axis, the end images reconstructed by Algorithm 1 will have artifacts since truncation in z direction some voxels does not have 180 degree parallel projections which is the basic requirements for FDK reconstruction. Application dependent information compensation may relieve this issue at some extent though. This is out of the range of this study.

We are interested in the middle region of the volume for SAS scans. At joint region between two consecutive circular scans these two rotations can provide partial/complete complementary information each other. For example, for the rays passing z-axis we have complete complementary as shown in Figure 2. Let us look Figure 2a, the diamond ($\diamond$) point belongs to the subvolume of rotation 1, which does not illuminate this point at this view while the second rotation at the opposite view can provide a projection passing this point. It seems for the joint region between two consecutive rotations we have perfect 360 degree data everywhere. Unfortunately, this is not the case. For the voxels away from the z-axis there exists a gap between the projection data of the two rotations as shown by the star($\star$) point in Figure 2b. The far the voxel from the z-axis the large the gap. In Figure 3, we plot a plane cross section of the volume passing z-axis, the parallel projection view is perpendicular to this plane. The green solid curve enclosed regions are illuminated. It is clear that the voxels at z-axis are fully illuminated while there is a gap between two rotations for off-center positions. In Figure 3, the size of the gap is exaggerated instead of proportional to actual size to emphasize the phenomena.

Köhler *etal* proposed to use the data enclosed inside the dashed black rectangles only, see Figure 3, by dropping the rest data, [10]. This strategy implies to shorten the step size of patient table movement in practical clinical implementation. Hsieh *etal* proposed overlapping between consecutive scans, [11], thus the patient table movement step size is also shorten. To reduce the dose to patients we propose to perform interpolation to fill in the gap between two rotations. The complete algorithm is present in Algorithm 2

Remark 2.1: During rebinning step the parallel projection data of different circular trajectory are aligned to have the same projection angles. This is crucial so that the back projection

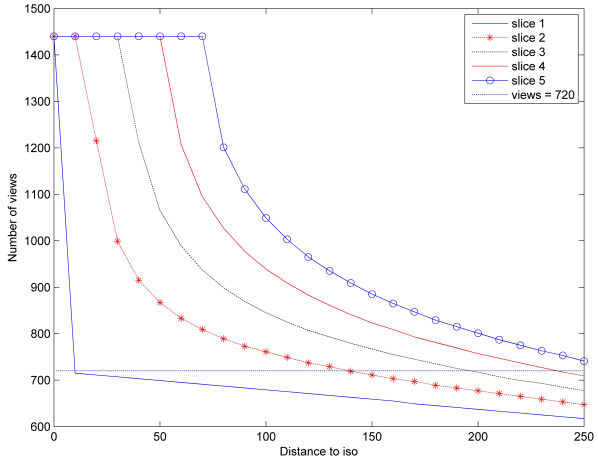
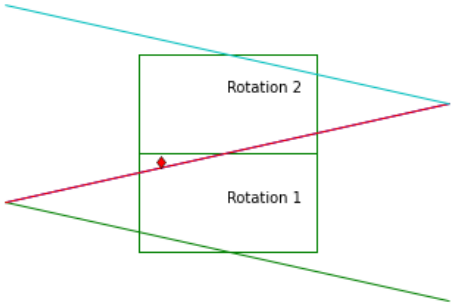


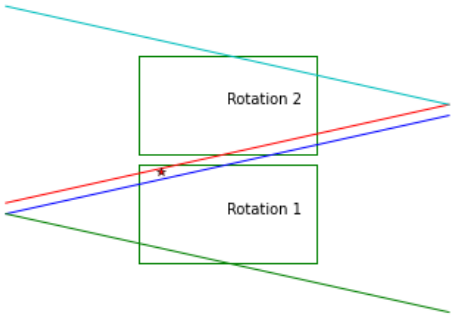
Fig. 1: Number of illuminated views vs voxel distances to z-axis for first 5 slices. In the computation, the geometry of Minfound 64 row scanner, see Section III, is used and the reconstruction volume z-coverage is set to the full illumination range at z-axis, i.e. 40mm. Slice 1 to 4 do not have enough projections to reconstruct the full slice with FDK if we set the FOV to 500 mm. The first slices only have the ONE center voxel satisfy FDK reconstruction condition.

Algorithm 2 SAS reconstruction by Weighted rebinning FDK Algorithm using complementary projection data between rotations.

1: Rebinning the CCB projection data $p(\gamma, v, \beta)$ from fan geometry to parallel geometry for each row v to produce $p(u, v, \theta)$.
2: Filtering the rebinned projection by kernel $h(u)$.
3: Applying cosine weight to obtain $q(u, v, \theta)$.
4: Performing weighted backprojection to reconstruct image volume from complementary projection data. Pseud-code is listed below:
   1) For every voxel $\mathbf{x}$,
   - For views $\theta_k \in [0, \pi), k = 1, 2, \cdots, K$
     - For half turns $m = 0, 1$
       If $\mathbf{x}$ is illuminated by some axial rotation at this view θ_k, do regular interpolation within this axial rotation data and compute the weight, otherwise use complementary data, $\theta_k + \pi$ from next or previous rotation, to do interpolation across neighbouring two scans to estimate the projection data and determine the associated weight.
     - End For half turns
     - Normalize weights over half turns.
     - $f(\mathbf{x}) = f(\mathbf{x}) + \sum_{m=0}^{1} \mathbf{w}(m, k, \mathbf{x}) * q(u, v, m, k)$
   - End For views
   2) End For voxel $\mathbf{x}$



(a) Rays passing z-axis at one view for the 1st rotation and its complementary at the 2nd rotation



(b) Rays away from z-axis with the same distance at one view for the 1st rotation and its complementary at the 2nd rotation

Fig. 2: Example of opposite rays of two SAS rotations. **(a) Top Figure:** For the rays passing z-axis the projection data from two rotations provide perfect complementary information each other. The diamond ($\diamond$) point belongs to the bottom half subvolume associated with the rotation 1, however rotation 1 does not illuminate this point at this view while the second rotation at the opposite view can provide a projection to cover this point. **(b) Bottom Figure:** The opposite rays, which are away from z-axis, from two rotations has gap between two opposite projection. The star ($\star$) point belongs to the bottom half subvolume associated with the rotation 1. Rotation 1 does not illuminate this point at this view neither the second rotation at the opposite view.

step can accurately locate the opposite projection from the neighbouring rotation when it is needed.

Remark 2.2: The number of the patient table movement step size is required to be accurate at certain tolerance so that the assumed geometry in the backprojection is consistent with the merged data of consecutive rotations.

Remark 2.3: The row weight before normalization should be supported beyond the physical row range. We gives an example weight as below (II.2). Let $\tilde{v} = v - v_{center}$, $q = \frac{|\tilde{v}|}{\tilde{v}_{max}}$,

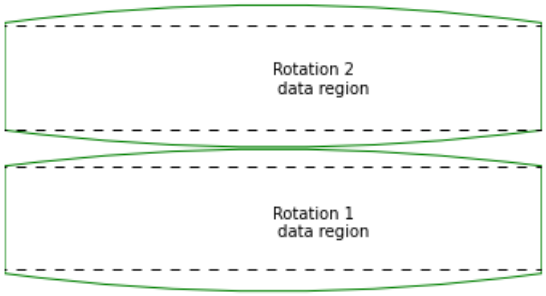


Fig. 3: An example of parallel projection map of one SAS view at a plane passing z-axis, the plot plane is perpendicular to the projection view. The green solid curve enclosed regions are illuminated regions. The voxels at z-axis are fully illuminated while there is a gap between two rotations for off-center regions.

and the constants Q, p and c satisfy $0 < Q < 1$, $p > 1$, $c > 10$.

$$\mathbf{w}(v) = \begin{cases} 1 & q \le Q, \\ \cos^2\left(\frac{\pi}{2} \cdot \frac{q-Q}{p-Q}\right) & Q < q \le 1 \\ \cos^2\left(\frac{\pi}{2} \cdot \frac{1-Q}{p-Q}\right) e^{c(1-q)} & q > 1 \end{cases} . \quad \text{(II.2)}$$

III. EXPERIMENTS AND RESULTS

Since the interpolation is used to estimate the gap between two rotations in the proposed method we have verified the accuracy by simulated data. Experimental data are obtained with the geometry of the Minfound Scinticare 64 row scanner, which has 64 row detectors each having height 0.625mm along z-axis, thus the z-coverage at z-axis is 40mm, and 840 channels at x-axis direction with arc angle 54.4^o. The scanner has maximum FOV 500mm, cone angle about $\pm 2^o$. 1440 projections per rotation is used.

Simulated FORBILD, [12], head phantom was projected and then reconstructed with FOV 350 mm. This phantom has complicated high contrast structure, it is a challenge object thus it is good for reconstruction performance evaluation.
The reconstruction backprojection weight parameters are set as: $Q = 0.5, p = 1.1$, and $c = 25$. The total number of slices of the reconstructed images is 128 slices with separation 0.625mm. The end slice of the first rotation, i.e. 64th slice, is presented in Figure 4 with window WL 50HU and WW 200HU. This most end slice has high resolution, and it does not have any distortion, black shading, or other strong artifacts caused by data missing or extrapolation related inaccuracy.

IV. CONCLUSION

We proposed a new algorithm for sequential circular axial CT reconstruction. The proposed algorithm does not discard any projection data or introduce any overlapping between consecutive rotations. Thus, it is dose efficient and it can

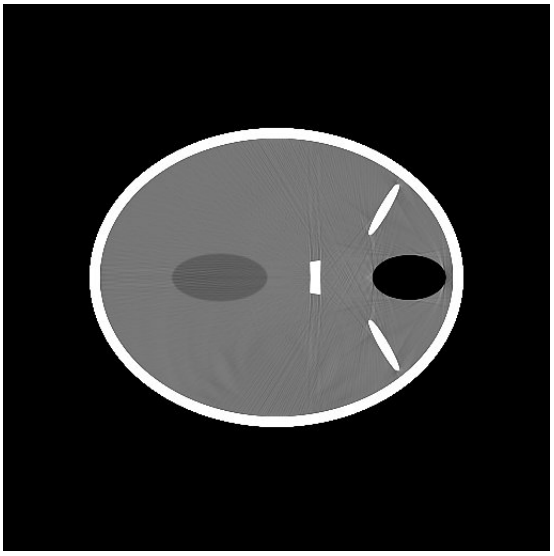


Fig. 4: The 64th slice of the reconstructed 128 slices simulated FORBILD Head phantom. It is the end slice of the first rotation. The slice separation is 0.625mm and displayed with window WL 50HU and WW 200HU.

reconstruct larger volume than existing algorithms. In this algorithm the projection data of consecutive rotations are jointed as a whole to complement each other. The sampling gap between to rotations are filled in with interpolation by using the data across two rotations. The algorithm has been validated with a simulated data. This data set is quite challenge. With the large cone angle geometry and the high contrast object the reconstructed images do not show strong artifacts at joint slices. This suggest that the proposed algorithm has potential to be used in clinical practice. Further verification and evaluation will be performed.

REFERENCES

[1] J. Hsieh, "A practical cone beam artifact correction algorithm," in *Nuclear Science Symposium Conference Record, 2000 IEEE*, vol. 2. IEEE, 2000, pp. 15–71.

[2] K. Zeng, Z. Chen, L. Zhang, and G. Wang, "An error-reduction-based algorithm for cone-beam computed tomography." *Medical physics*, vol. 31, pp. 3206–12, 2004.

[3] F. T. Sunnegardh J, Schoendube H, "Reduction of cone artifacts in ct with incomplete source trajectories," in *Proceedings of the third International Conference on image formation in X-ray computed tomography*, 2014, pp. 384–387.

[4] H. Hu, "An improved cone-beam reconstruction algorithm for the circular orbit," *Scanning*, vol. 18, no. 8, pp. 572–581, 1996.

[5] H. Yang, M. Li, K. Koizumi, and H. Kudo, "A fbp-type cone-beam reconstruction algorithm with radon space interpolation ability for axially truncated data from a circular orbit," in *Proceedings of the 8th International Conference on Fully 3D Reconstruction in Radiology and Nuclear Medicine*, 2005, pp. 401–404.

[6] X. Tang, J. Hsieh, R. A. Nilsen, S. Dutta, D. Samsonov, and A. Hagiwara, "A three-dimensional-weighted cone beam filtered backprojection (cb-fbp) algorithm for image reconstruction in volumetric ct神elical scanning," *Physics in Medicine & Biology*, vol. 51, no. 4, p. 855, 2006.

[7] R. Grimmer, M. Oelhafen, U. Elström, and M. Kachelriess, "Cone-beam ct image reconstruction with extended z range." *Medical physics*, vol. 36, no. 7, pp. 3363–70, Jul 2009.

[8] J. Pack, Z. Yin, K. Zeng, and B. Nett, "Mitigating cone-beam artifacts in short-scan ct imaging for large cone-angle scans," in *Proceedings of the International Meeting on Fully 3D Image Reconstruction in Radiology and Nuclear Medicine*, 2013, pp. 307–310.

[9] P. v. d. H. Dirk Schäfer and M. Grass, "Modified parker weights for super short scan cone beam ct," in *Proceedings of the International Meeting on Fully 3D Image Reconstruction in Radiology and Nuclear Medicine*, 2017, pp. 49–52.

[10] T. H. Köhler, R. Proksa, and M. Grass, "A fast and efficient method for sequential cone-beam tomography." *Medical physics*, vol. 28, no. 11, pp. 2318–27, Nov 2001.

[11] J. Hsieh, J. Londt, M. Vass, J. Li, X. Tang, and D. Okerlund, "Step-and-shoot data acquisition and reconstruction for cardiac x-ray computed tomography," *Medical Physics*, vol. 33, p. 4236, 2006.

[12] G. Lauritsch and H. Bruder. (2003) FORBILD head phantom. [Online]. Available: http://www.imp.uni-erlangen.de/phantoms/head/head.html

Respiratory gating method for low-dose small-animal CT studies: preliminary results

N. Ballesteros, M. Desco, M. Abella

***Abstract*– The quality of chest CT images in small-animal studies is often limited by respiratory movement, which causes a loss of resolution in the reconstructed images. Respiratory gating techniques allow to reduce this effect by selecting CT data at specific moments of the respiratory cycle. In retrospective gating, this sorting is done based on the processing of the acquired projections, with no need of external hardware. Nevertheless, the algorithms reported in the literature provide solutions that require high exposure doses for the animal. We present a new retrospective respiratory gating algorithm for low-dose small-animal CT. Preliminary evaluation in rodents under free breathing showed a good sorting even with a radiation dose halving the minimum reported in the literature, leading to gated CT reconstructions with enhanced resolution.**

Index Terms —X ray, small-animal CT, respiratory gating.

I. Introduction

Computed tomography (CT) systems for small animals have become an important imaging tool in preclinical and basic research due to the increasing use of animal models for the study of different human diseases. In particular, chest CT scans are used to monitor a variety of lung abnormalities and diseases, including emphysema [1], asthma [2] or acute respiratory distress syndrome [3]. However, the movement of the chest of the animal during *in-vivo* studies can reach 5 mm, which results in a reduced spatial resolution in the reconstructed image.

Several methods have been proposed to constrain breathing motion, such as tapes around the chest to restrict thorax movements or intubation with mechanical ventilation to ensure a stable breathing pattern. However, these strategies may cause physical damage to the animal or compromise its normal breathing, possibly resulting in a confounding factor in preclinical research with respiratory disease models. Gating strategies aim to synchronize the CT image acquisition with the respiratory cycle to reconstruct only images taken within a certain breathing phase. These techniques can be performed under free breathing and enable the structural and functional characterization of the lungs in a non-invasive way.

Manuscript submitted April 2nd 2018.

All authors are with the Dpto. Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid and with the Instituto de Investigación Sanitaria Gregorio Marañón (IiSGM), Madrid, Spain (e-mail: mabella@ing.uc3m.es).

M. Abella and M. Desco are with the Centro Nacional de Investigaciones Cardiovasculares Carlos III (CNIC), Madrid, Spain.

M. Desco is also with the Centro de Investigación en Red de Salud Mental (CIBERSAM) (e-mail: desco@hggm.es).

Gating methods have been categorized into prospective and retrospective techniques. In prospective gating, only data corresponding to a phase of interest are acquired by synchronizing the CT acquisition with the breathing of the animal in real time, thus involving long scan times and laborious setups [4-6]. Retrospective gating simplifies the acquisition process by acquiring data continuously throughout the respiratory cycle and then assigning the projections to different phases of the respiration. This sorting can be done based on a respiratory signal monitored during the acquisition or obtained from the processing of projection images themselves. Retrospective approaches found in the literature require oversampling for the sorting process of the projections, thereby increasing the exposure time for the animal and the dose [7-12].

We present a new retrospective respiratory gating algorithm for small-animal cone-beam CT scanners suitable for low-dose radiation studies.

II. Methods

A. Algorithm proposed

Fig. 1 shows the workflow of the proposed algorithm implemented with MATLAB (The MathWorks Inc., Natick, MA, USA).

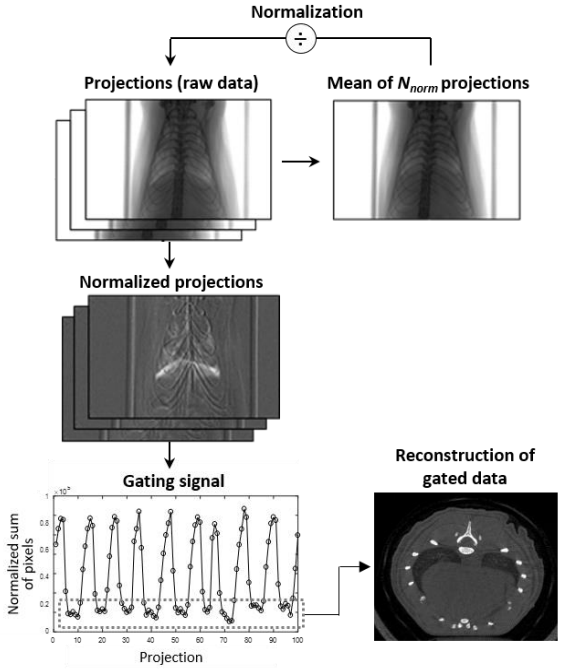


Fig. 1. Algorithm workflow.

The input of the algorithm is a complete raw dataset with a total number of projections given by:

$$N_{proj} = P \cdot F \quad (1)$$

where P is the number of projection angles and F is the number of projections acquired per projection angle in the case of step-and-shoot mode. In the case of a continuous rotation of the gantry during the acquisition, F would be 1.

The first step is the normalization of the acquired CT projections in order to minimize the angle-dependent attenuation effect, as previously described in the literature [13]. To this end, each projection is divided by a mean image generated from a subset of consecutive projections along different projection angles. The size of this subset, N_{norm}, is selected considering two facts: (1) it has to be large enough so as there is an equivalent contribution of projections for all respiratory phases and (2) it has to be small enough so that the projections belong to a reduced span of angles.

The attenuation differences remaining after the normalization step in a region of interest (ROI) covering the lungs are related with the volume of air contained in the lungs and the position of the diaphragm along the different respiratory phases. Projections taken in the inspiration phase show higher intensity within the ROI, while during expiration have lower pixel values, as shown in Fig. 2. The sum of the pixels in this ROI for each projection is used as a surrogate of the respiratory signal, referred to as *gating signal*. Thus, minimum values of the *gating signal* correspond to the expiration phase, which is the longer phase in the respiratory cycle of rodents.

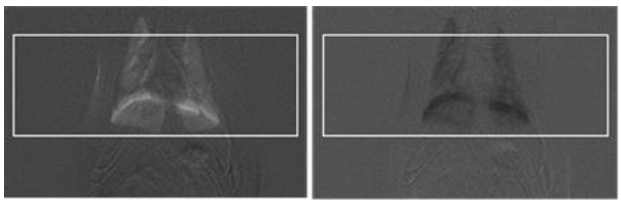


Fig. 2. Normalized projections taken during inspiration (left) and expiration phases (right).

We select the minimum points of each respiratory cycle in the *gating signal* and calculate the mean, μ_{min} and the standard deviation, σ_{min}. The gating interval for the expiratory breathing phase is defined as $\mu_{min} \pm \sigma_{min}$. This interval can be calculated using the complete data set or a subset of projections, to allow generating the gating signal simultaneously with the acquisition. In the latter case, a good sorting requires at least four complete respiratory cycles within the interval in order to have enough samples and avoid bias due to unsteady breathing patterns.

B. Evaluation

We evaluated our proposal in two scenarios, with high and low radiation doses respectively. Data were acquired with the scanner ARGUS CT (SEDECAL) in step-and-shoot scanning mode over an angular range of 360 degrees with an angular step of 1 degree. Voltage and current were 65 kVp and 200 µA, with an exposure time per projection of 39 ms. Projection size was 768×486 pixels with a pixel size of 0.2 mm. The number of images per projection angle (F in Eq. 1) was 32 and 4 for the high-dose and low-dose scenario respectively.

We reconstructed the complete dataset and the gated data using an analytical reconstruction method based on the FDK algorithm [14, 15], and the PICCS-based method proposed in [16] using the FDK reconstruction of the non-gated data as a prior. In all cases, the reconstruction matrix was 768×768×486 pixels with an isotropic voxel size of 0.115 mm.

The value of N_{norm} (size of the projection subset for the normalization step) was 32 for both cases.

III. Results

Fig. 3 shows reconstructions of the non-gated and gated data of the high-dose study. We can see enhanced details in the vasculature and a blurring reduction in the regions adjacent to the heart and diaphragm.

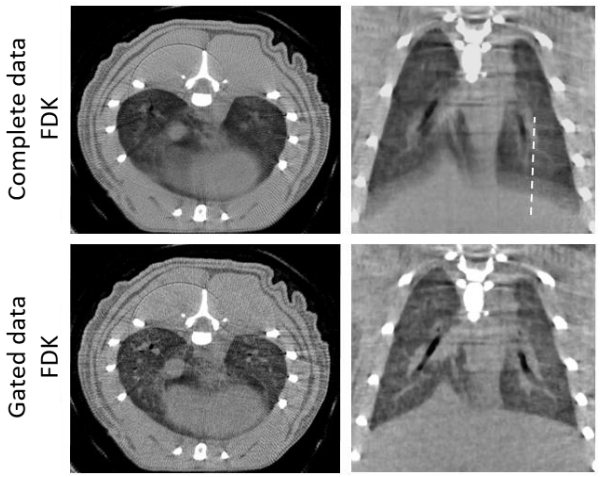


Fig. 3. Axial and coronal views of the reconstructed images for the high radiation dose study. White dashed line in the coronal view indicates where the profiles were taken.

The resolution improvement between gated and non-gated data translated into a higher slope between the lung parenchyma and adjacent tissues along the profiles outlined in Fig. 4.

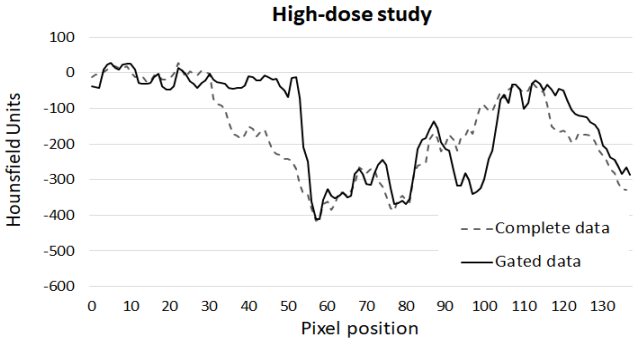


Fig. 4. Profile across the diaphragm for the high radiation study (white dashed line in coronal view, Fig. 3).

The reduction of the blurring in the diaphragm is still noticeable in the FDK reconstruction for the low-dose case (Fig. 5), but the low number of projections results in severe streak artifacts. The PICCS-based reconstruction method compensates this lack of data reducing the artifacts present in the FDK reconstruction.

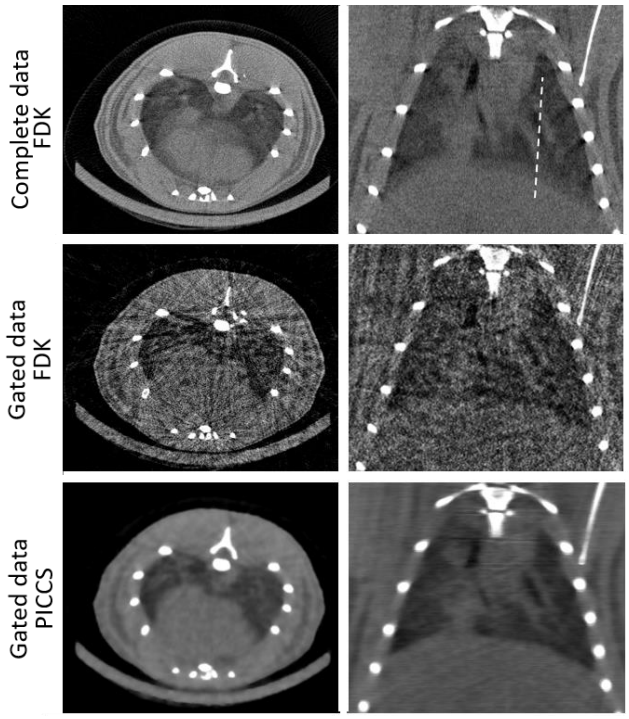


Fig. 5. Axial and coronal views of the reconstructed images for the study with low radiation dose. White dashed line in the coronal view indicates where the profiles were taken.

Fig. 6 shows the profile taken in the coronal view for the complete data and the gated data reconstructed with the PICCS-based method in the case of low radiation dose. Gated data show a better definition of the diaphragm and lung vessels.

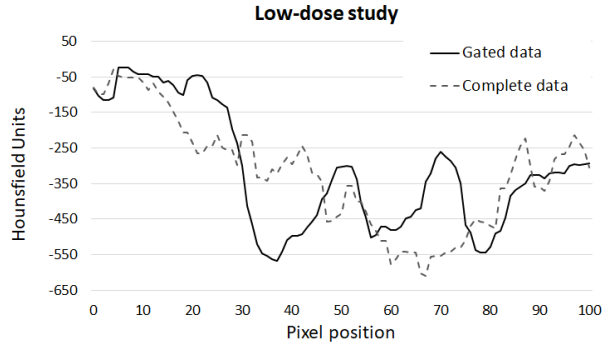


Fig. 6. Profile across the diaphragm for the low radiation study (white dashed line in coronal view, Fig. 5).

IV. Discussion

Gating techniques are required for preventing movement blurring in the reconstruction of chest micro-CT *in vivo* studies. Retrospective gating methods generally need high sampling involving high exposure doses for the animal.

In this work, we proposed a respiratory gating method for small animals CT systems suitable for low radiation doses. The gating signal is obtained directly from the acquired data, avoiding the need of any additional hardware.

Although the proposed method has only been tested with the volume corresponding to the expiration phase, the generation of volumes corresponding to other phases of the respiratory cycle is straightforward.

The evaluation of the method in rodent studies under free breathing proved its capacity to compensate the respiratory blurring, improving resolution even in situations with a radiation dose halving the minimum reported in the literature [8, 10]. For low-dose studies, the reduced number of noisy projections for each respiratory phase leads to severe streak artifacts in the FDK-reconstructed images. In these scenarios, the use of advanced iterative reconstruction methods is necessary to reduce the noise and artifacts in the reconstructed images.

Acknowledgment

This work was partially supported by the Ministry of Economy, Industry and Competitiveness (DPI2016-79075-R) and ISCIII-FIS grants DTS17/00122, co-financed by ERDF (FEDER) Funds from the European Commission, “A way of making Europe”. The CNIC is supported by the Ministry of Economy, Industry and Competitiveness (MEIC) and the Pro CNIC Foundation, and is a Severo Ochoa Center of Excellence (SEV-2015-0505).

References

[1] E. De Langhe, G. Vande Velde, H. Jeroen, U. Himmelreich, B. Nemery, F. P. Luyten, *et al.*, "Quantification of Lung Fibrosis and Emphysema in Mice Using Automated Micro-Computed Tomography," *Plos One*, vol. 7, p. e43123, 2012.

[2] S. H. Paik, W. K. Kim, J. S. Park, C. S. Park, and G. Y. Jin, "A Quantitative Study of Airway Changes on Micro-CT in a Mouse Asthma Model: Comparison With Histopathological Findings," *Allergy, Asthma & Immunology Research*, vol. 6, pp. 75-82, 2014.

[3] M. T. Voelker, F. Fichtner, M. Kasper, M. Kamprad, U. Sack, U. X. Kaisers, *et al.*, "Characterization of a double-hit murine model of acute respiratory distress syndrome," *Clin Exp Pharmacol Physiol*, vol. 41, pp. 844-853, 2014.

[4] N. L. Ford, H. N. Nikolov, C. J. Norley, M. M. Thornton, P. J. Foster, M. Drangova, *et al.*, "Prospective respiratory-gated micro-CT of free breathing rodents," *Med Phys*, vol. 32, pp. 2888-2898, 2005.

[5] T. Sera, H. Yokota, K. Fujisaki, K. Fukasaku, H. Tachibana, K. Uesugi, *et al.*, "Development of high-resolution 4D in vivo-CT for visualization of cardiac and respiratory deformations of small animals," *Phys Med Biol*, vol. 53, pp. 4285-4301, 2008.

[6] S. H. Bartling, W. Stiller, M. Grasruck, B. Schmidt, P. Peschke, W. Semmler, *et al.*, "Retrospective motion gating in small animal CT of mice and rats," *Invest Radiol*, vol. 42, pp. 704-714, 2007.

[7] J. Kuntz, J. Dinkel, S. Zwick, T. Bäuerle, M. Grasruck, F. Kiessling, *et al.*, "Fully automated intrinsic respiratory and cardiac gating for small animal CT," *Phys Med Biol*, vol. 55, pp. 2069–2085, 2010.

[8] N. L. Ford, A. Jeklin, K. Yip, D. Yohan, D. W. Holdsworth, and M. Drangova, "Optimization of Image Quality in Retrospective Respiratory-Gated Micro-CT for Quantitative

Measurements of Lung Function in Free-Breathing Rats," *J Bio Sci Eng*, vol. 7, pp. 157-172, 2014.

[9] A. Hahn, S. Sauppe, M. Lell, and M. Kachelrieß, "Automatic intrinsic cardiac and respiratory gating from cone-Beam CT scans of the thorax region," *Proc. SPIE 9783 Med. Imaging Phys. Med. Imaging*, vol. 9783, pp. 1–6, 2016.

[10] T. H. Farncombe, "Software-based respiratory gating for small animal conebeam CT," *Med Phys*, vol. 35, pp. 1785-1792, 2008.

[11] C. Chavarrías, J. J. Vaquero, A. Sisniega, A. Rodríguez-Ruano, M. L. Soto-Montenegro, P. García-Barreno, *et al.*, "Extraction of the respiratory signal from small-animal CT projections for a retrospective gating method," *Phys Med Biol*, vol. 53, pp. 4683-4695, 2008.

[12] M. Brehm, S. Sawall, J. Maier, S. Sauppe, and M. Kachelrie, "Cardiorespiratory motion-compensated micro-CT image reconstruction using an artifact model-based motion estimation," *American Association of Physicists in Medicine*, vol. 42, 2015.

[13] A. Kavanagh, P. M. Evans, V. N. Hansen, and S. Webb, "Obtaining breathing patterns from any sequential thoracic x-ray image set," *Phys Med Biol*, vol. 54, pp. 4879-4888, 2009.

[14] M. Abella, J. J. Vaquero, A. Sisniega, J. Pascau, A. Udías, V. García, *et al.*, "Software Architecture for Multi-Bed FDK-based Reconstruction in X-ray CT Scanners," *Comput Methods Programs Biomed*, vol. 107, pp. 218-32, 2012.

[15] J. Garcia Blas, M. Abella, F. Isaila, J. Carretero, and M. Desco, "Surfing the optimization space of a multiple-GPU parallel implementation of a X-ray tomography reconstruction algorithm," *The Journal of Systems and Software*, vol. 95, pp. 166–175, 2014.

[16] J. Abascal, A. Sisniega, C. Chavarrías, J. J. Vaquero, M. Desco, and M. Abella, "Investigation of different Compressed Sensing Approaches for Respiratory Gating in Small Animal CT," *IEEE Nucl Sci Symp & Med Ima Conf Rec*, pp. 3344-3346, 2012.

Fully Automatic Intrinsic Respiratory and Cardiac Gating of Cone-Beam CT Scans of the Thorax Region

Andreas Hahn, and Marc Kachelrieß

***Abstract*—In image-guided radiation therapy (IGRT) a cone-beam CT (CBCT) is often used as an on-board imaging device. Due to the limited acquisition time of about 60 s, breathholding techniques are not feasible and motion artifacts from respiratory and cardiac motion occur. All algorithms that aim for motion compensation require accurate information about the respiratory and cardiac phase. When there is no external respiratory/cardiac signal available or the signal ist corrupted, it has to be retrieved intrinsically. In this work we present a new method to retrieve a respiratory and a cardiac surrogate signal intrinsically in a fully automatic way. The method is rawdata–based and tested on seven CBCT scans of the thorax region. It was able to determine the correct number of respirations for all patients and the correct number of heart beats for six patients. For one patient it missed one heart beat.**

I. Introduction

In image-guided radiation therapy a cone-beam CT (CBCT) is often used as an on-board imaging (OBI) device. Due to the limited acquisition times of about 30 s breathholding is not feasible and respiratory and cardiac motion result in blurred images. Algorithms that compensate for this motion need accurate information of the respiratory/cardiac phase. When no external information of the respiratory/cardiac phase, e.g. recorded by a respiratory belt or an ECG, is available, the information has to be retrieved intrinsically. There are a lot of existing methods dedicated to retrieving a respiratory surrogate signal intrinsically in a fully automatic fashion [1]–[5] or with the help of some kind of user input [6]–[9]. For the cardiac signal there is a fully automatic method [10] that is, however, limited to spiral scans. There is a fully automatic method for respiratory and cardiac gating of small animals [11], that requires multiple rotations for the automization process. It is not applicable for our case since there is only one rotation.
A new method [12] showed promising results to extract a cardiac and respiratory surrogate signal for a simulation and a patient. The idea behind this method is that the best place to extract the motion surrogate in the raw data is at the diaphragm or ventricular wall for respiratory or cardiac gating, respectively. A point at the diaphragm and a point at the ventricular wall is selected automatically and tracked in the rawdata. A point at the ventricular wall might however not provide a good cardiac surrogate signal for all source angles. If its trace in the rawdata is at the ventricular wall for an angle α (Fig. 1a), it is not necessarily at the ventricular wall for an angle $\alpha+\pi$ (Fig. 1b). This is partly compensated by the size of the ROI around the grid point trace. The smaller the circumference of the heart in the axial plane of the selected grid point, the smaller this geometric effect gets.
In this work we present a new method that takes the adressed geometrical problem into account. We will call our new method IGC, since for cardiac gating a circle is used instead of the point, and [12] IGP, since a point is used.

Andreas Hahn and Prof. Dr. Marc Kachelrieß: Division of X–Ray Imaging and CT, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, Heidelberg, Germany. Andreas Hahn: Department of Physics and Astronomy, Ruprecht-Karls-University, Im Neuenheimer Feld 226, Heidelberg, Germany. Prof. Dr. Marc Kachelrieß: Medical Faculty, Ruprecht-Karls-University, Im Neuenheimer Feld 672, Heidelberg, Germany.
Corresponding author: andreas.hahn@dkfz.de

II. Materials and Methods

The goal of IGC is a fully automatic rawdata–based extraction of a respiratory and a cardiac surrogate signal. It can be divided into three parts which are explained in the following sections: Data preprocession, surrogate signal acquisition and identification of optimal grid point. Data preprocession and the identification of the optimal grid point are identical to IGP but are explained for convenience.

A. Data Preprocessing

Since the method aims to identify the regions in the rawdata that show respiratory/cardiac motion, everything that is not subjected to motion is subtracted from the original rawdata $\boldsymbol{q}$ in the following way. Since a standard FDK reconstruction gives a blurred volume, one might be tempted to think that its forward projection gives blurred rawdata that do not show any motion any more. This is however not the case, since filtered backprojection X^{-1} and forward projection X are inverse opertions. Even though the reconstructed volume is blurred, motion information is encoded in the volume in the form of noise. This information can be reduced by applying a Median filter M before forward projection. These new rawdata $\boldsymbol{q}_{\text{stat}}$ show a blurred average motion state and are nearly identical to the original rawdata in regions without motion. By subtracting them from the original rawdata, only regions that are subjected to motion remain, i.e. large dark or bright areas around the diaphragm depending on the respiratory state and small dark or bright areas at the ventricular wall depending on the heart phase. All further steps are carried out on these dynamic rawdata $\boldsymbol{q}_{\text{dyn}}$.

$$\boldsymbol{q}_{\text{dyn}} = \boldsymbol{q} - \boldsymbol{q}_{\text{stat}} = \boldsymbol{q} - \mathsf{X}M\mathsf{X}^{-1}\boldsymbol{q} \tag{1}$$

B. Signal Acquisition

Grid points are distributed regularly in the volume. The goal is to identify a grid point that is at a place that is best suited to obtain a surrogate signal. In [12], all grid points are forward projected and tracked in the rawdata (see Fig. 1). In a ROI around the trace of the grid point, the average gray value is calculated for every projection and serves as a scalar surrogate signal $s(n)$, where $n \in [0, N)$, with N being the number of projections:

$$s(n) = \sum_{(u,v)\in \mathrm{ROI}} p_{\mathrm{dyn}}(u,v) / \sum_{(u,v)\in \mathrm{ROI}} 1, \qquad (2)$$

where u and v are the detector pixel indices. A bandpass filter in the range of 10-30 rpm and 50-120 bpm is applied depending if one wants to get the respiratory or cardiac surrogate, respectively. Peaks are determined automatically and the signal is converted to a saw-tooth phase signal.
While this works well for the respiratory signal, it can result in problems with the cardiac signal. If a grid point at the ventricular wall is identified and it is also located at the ventricular wall in projection n_1 at an angle α_1 (Fig. 1a), it can be in the middle of the heart in projection n_2, with $\alpha_2 = \alpha_1 + \pi$ (Fig. 1b). In order to avoid this issue, not the grid point itself is tracked in the rawdata. Instead, each point represents the center of a circle with a fixed radius r and the tangential rays are tracked in the rawdata (Fig. 1c + 1d). To find the best radius, we interate over radii between 3 cm 5 cm in steps of 2 mm.

C. Identifying the Optimal Grid Point

Among all the grid points distributed over the volume, the best point to obtain a respiratory or cardiac surrogate has to be chosen. This is done by calculating the standard deviation of the distance between the peaks. If the grid point is at a reasonable spot, e.g. the center of the heart or the diaphragm, the peaks are very regular (heart beats) or fairly regular (respiration). If it is at a spot that is not subjected to motion, the position of the peaks will be very random. This is why the grid point with the most regular peaks/with the smallest standard deviation in the peak distance is chosen for the final motion surrogate.

D. Patient Data

In order to evaluate the proposed algorithm, it was tested on 7 patient scans of the thorax region. The patients were scanned using a Varian True Beam scanner at 11 fps with 660 projections/60 s acquisition time. The patients had between 11-26 rpm and 65-86 bpm. As a ground truth for the respiratory phase, the signal from the Varian real-time position management (RPM) system was available for 2 patients. For the other patients, the points of maximum inhale were determined manually from the projections. For respiratory gating $10\times10\times10$ grid points are distributed regularly over the volume. The ROI size is set to [300 px, 300 px]=[116.4 mm, 116.4 mm]. Since no ECG was available for all patients, the number of heart beats was determined by carefully inspecting the projections/sinograms. For the cardiac surrogate $50\times50\times50$ grid points are distributed in the volume. The ROI size is set to [200 px, 100 px]=[77.6 mm, 38.8 mm].
The number of grid points and ROI sizes are empiric values. The result, however, did not depend much on these parameters.

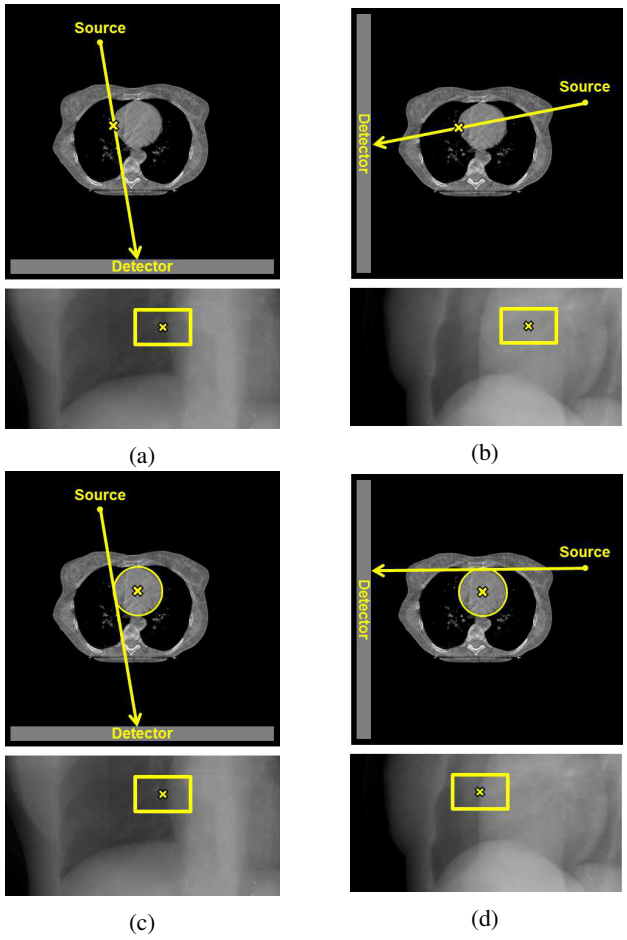


Fig. 1: The images on the top show the source at angle α ((a) and (c)) and at angle $\alpha + \pi$ ((b) and (d)). The yellow cross indicates a grid point. The images on the bottom show the corresponding projection with the position of the grid point in the rawdata. For IGP, this point is at the ventricular wall at α but in the middle of the heart at $\alpha+\pi$. In our new approach the grid point represents the center of a circle and tangential rays are considered instead of the ray intersecting the grid point. This way the ventricular wall is tracked at α (c) and $\alpha + \pi$ (d).

III. Results

A. Respiratory Gating

Since the proposed method is only suited to obtain the respiratory phase but not the amplitude, the positions of the detected peaks have to be compared. A constant offset between the intrinsically determined peaks and the ground truth is not important in this case, since this effect can be removed by just relabeling the phases. Instead, the standard deviation of the difference between the peak positions of intrinsic surrogate and the ground truth has to be considered. Let $p_{\mathrm{intr}}(n)$ and $p_{\mathrm{GT}}(n)$

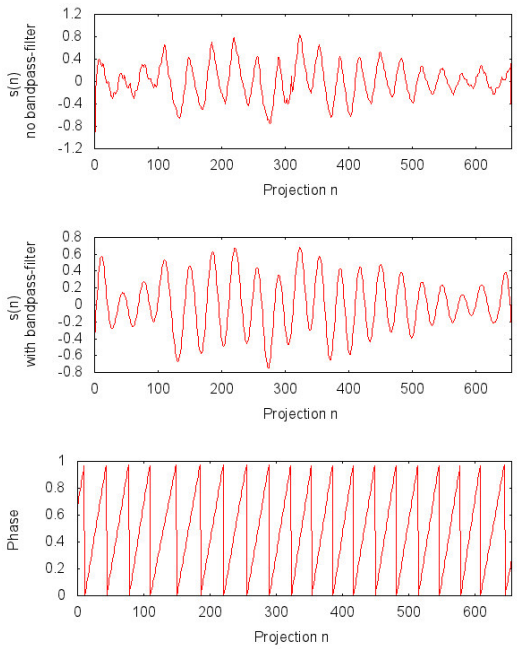


Fig. 2: A scalar signal s(n) is calculated for a grid point (first row). After a bandpass-filter (second row) the peaks are determined automatically and a saw-tooth phase signal is calculated (third row).

be the position of the n-th peak of the intrinsically determined signal and the ground truth, respectively. The measure σ_{intr} for the accuracy of the result is than determined by

$$\sigma_{\text{intr}} = \sqrt{\frac{1}{N_{\text{max}} - 1} \sum_{n=1}^{N_{\text{max}}} (\Delta p(n) - \overline{\Delta p})^2}, \tag{3}$$

with

$$\Delta p(n) = p_{\text{intr}}(n) - p_{\text{GT}}(n) \tag{4}$$

being the difference of the position of the maxima of the intrinsic surrogate and the ground truth, $\overline{\Delta p}$ being the mean of $\Delta p(n)$ and N_{max} being the number of maxima of the respiratory signal. This evaluation only works if the intrinsic gating determines the correct number of maxima.
The results are shown in table I. Since the improvement from IGC compared to IGP only concerns cardiac gating, respiratory gating was done just as a proof of concept. The focus of this work is on cardiac gating, which is why the respiratory gating was not compared to any other method.
IGC was able to determine the correct number of respirations per minute (rpm) for all patients. The highest value for σ_{intr} is 405 ms, which corresponds to the time between 4.455 projections.

B. Cardiac Gating

The results for the intrinsic cardiac gating are shown in table II. The beats per minute (bpm) detected by IGP/IGC is indicated by N_P/N_C respectively while the ground truth is indicated by N_{GT}. The differences to the ground truth are given by $\Delta\text{N}_\text{P} = \text{N}_\text{P} - \text{N}_{\text{GT}}$ and $\Delta\text{N}_\text{C} = \text{N}_\text{C} - \text{N}_{\text{GT}}$. The proposed

TABLE I: Results respiratory gating

Patient	N_{GT}	N_{IG}	σ_{intr} in ms
1	20	20	157
2	20	20	405
3	11	11	231
4	23	23	321
5	26	26	123
6	26	26	255
7	23	23	318

method worked well for all cases, while the IGP worked well for 3 patients ($\Delta\text{N}_\text{P} \leq 1$), showed minor deviations for 2 patients ($\Delta\text{N}_\text{P} = 3$) and was completely wrong for 2 patients ($\Delta\text{N}_\text{P} \geq 20$).

TABLE II: Results cardiac gating

Patient	N_{GT}	N_{IGP}	N_{IGC}	ΔN_P	ΔN_C
1	80	81	80	1	0
2	65	85	65	20	0
3	70	70	70	0	0
4	78	81	78	3	0
5	78	101	78	23	0
6	86	89	85	3	-1
7	74	74	74	0	0

IV. Conclusion and Discussion

We proposed a method, IGC, that is able to obtain a surrogate respiratory or cardiac motion signal in CBCT scans of the thorax region. It is able to do so in a fully automatic way and works rawdata–based. It is based on IGP but is able to deliver more reliable results for cardiac gating for patients with reduced visibility of the heart in the rawdata. IGC and IGP were compared on 7 patients. For cardiac gating, IGC outperformed IGP and was able to retrieve the correct bpm for every patient with a maximum error of 1 heart beat. IGP worked well for 3 patients, showed mediocre results for 2 patients and failed completely for 2 patients. Respiratory gating is identical between IGC and IGP and was done as a proof of concept. The correct number of rpm was determined for every patient and the error in the position of the maximum, i.e. maximum inhale, was below or equal to 405 ms, which corresponds to 4.455 projections.

Acknowledgements

This study was supported by DFG KA 1678/13. Parts of the reconstruction software were provided by RayConStruct GmbH, Nürnberg, Germany.

References

[1] L. Zijp, J.-J. Sonke, and M. van Herk, "Extraction of the respiratory signal from sequential thorax cone–beam x–ray images," *Proc. of the 14th ICCR, Seoul, Korea*, pp. 507–509, 2004.
[2] M. Van Herk, L. Zijp, P. Remeijer, J. Wolthaus, and J. Sonke, "On-line 4d cone beam CT for daily correction of lung tumour position during hypofractionated radiotherapy," *ICCR, Toronto, Canada*, 2007.
[3] A. Kavanagh, P. M. Evans, V. N. Hansen, and S. Webb, "Obtaining breathing patterns from any sequential thoracic x-ray image set," *Phys. Med. Biol.*, vol. 54, no. 16, p. 4879, 2009.

[4] I. Vergalasova, J. Cai, and F.-F. Yin, "A novel technique for markerless, self-sorted 4d-CBCT: Feasibility study," *Med. Phys.*, vol. 39, no. 3, pp. 1442–1451, 2012.

[5] S. Dhou, Y. Motai, and G. D. Hugo, "Local intensity feature tracking and motion modeling for respiratory signal extraction in cone-beam CT projections," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 2, pp. 332–342, 2013.

[6] J. Dinkel, S. H. Bartling, J. Kuntz, M. Grasruck, A. Kopp-Schneider, M. Iwasaki, S. Dimmeler, R. Gupta, W. Semmler, H.-U. Kauczor, and F. Kiessling, "Intrinsic gating for small–animal computed tomography: A robust ECG–less paradigm for deriving cardiac phase information and functional imaging," *Circulation: Cardiovascular Imaging*, vol. 1, no. 3, pp. 235–243, 2008.

[7] T. Farncombe, "Software-based respiratory gating for small animal cone-beam CT," *Med. Phys.*, vol. 35, no. 5, pp. 1785–1792, 2008.

[8] S. H. Bartling, J. Dinkel, W. Stiller, M. Grasruck, I. Madisch, H.-U. Kauczor, W. Semmler, R. Gupta, and F. Kiessling, "Intrinsic respiratory gating in small–animal CT," *European Radiology*, vol. 18, no. 7, pp. 1375–1384, Jul. 2008.

[9] J. Hu, S. T. Haworth, R. C. Molthen, and C. A. Dawson, "Dynamic small animal lung imaging via a postacquisition respiratory gating technique using micro-cone beam computed tomography," *Academic radiology*, vol. 11, no. 9, pp. 961–970, 2004.

[10] M. Kachelrieß, D.-A. Sennst, W. Maxlmoser, and W. A. Kalender, "Kymogram detection and kymogram–correlated image reconstruction from subsecond spiral computed tomography scans of the heart," *Med. Phys.*, vol. 29, no. 7, pp. 1489–1503, Jul. 2002.

[11] J. Kuntz, J. Dinkel, S. Zwick, T. Bäuerle, M. Grasruck, F. Kiessling, R. Gupta, W. Semmler, and S. Bartling, "Fully automated intrinsic respiratory and cardiac gating for small animal CT," *Phys. Med. Biol.*, vol. 55, no. 7, p. 2069, 2010.

[12] A. Hahn, S. Sauppe, and K. M., "Automatic intrinsic cardiac and respiratory gating from cone-beam CT scans of the thorax region," *SPIE Medical Imaging Proc.*, pp. 97 830S–97 830S, 2016.

Volumetric Blood Flow Estimation for 4D Digital Subtraction Angiography

Stephan Seitz[1, 2], Jürgen Endres[1, 2], Stefan Lang[2], Philip Hoelter[2], Harald Köstler[3], Andreas Maier[1], Arnd Dörfler[2]

***Abstract*—4D Digital Subtraction Angiography (4D DSA) provides a powerful new diagnosis tool in neuroangiography. Complex vascular structures can be visualized at different time instances from arbitrary angles. However, the resulting reconstructions might be compromised by artifacts and time-resolution is only approximate due to insufficient projection data. Computational fluid dynamics (CFD) is a promising option to resolve arising ambiguities by the enforcement of physical constraints. Yet, a proper patient-specific initialization of the simulation remains a complex open research topic. In this paper, we investigate the capabilities of volumetric optical flow to estimate flow patterns in 4D DSA. The proposed Farnebäck algorithmic is evaluated on a 4D reconstruction of a numerical vessel phantom in order to assess its usefulness for future incorporation of prior flow information for 4D reconstruction.**

***Index Terms*—angiography, 4D DSA, iterative reconstruction, flow estimation, volumetric optical flow**

I. Introduction

Cerebrovascular diseases like aneurysms and stenosises are among the major reasons for death in western industrialized countries. For their diagnosis and treatment, C-arm-based digital subtraction angiography (DSA) systems are used as a gold standard for the visualization of vascular structures, as they provide not only highly time-resolved fluoroscopy from flexible angles (2D DSA), but also static 3D reconstructions with high spatial resolution, generated from gantry rotations (3D DSA).

Recent efforts [1], [2] to combine both temporal and spatial information by a fusion of single 2D DSA projections with a static volume (4D DSA) have been proven to be of additional diagnostic use by visualizing the contrast agent flow at arbitrary time instances of the acquisition protocol and from arbitrary view angles [2]–[5]. This is especially beneficial for the assessment of complex vascular structures like arteriovenous malformations: an appropriate choice of the visualized time instance can avoid vascular occlusions and allow a clear view onto the respective regions interest [3].

However, 4D DSA was also found to be prone to artifacts in such applications. Foreshortened vessels, orthogonal to the detector plane, and vascular overlaps may cause unrealistic distributions of contrast agent in the reconstructed volumes. The reason for this are ambiguities inherent to the reconstruction of a volume from a single view. Fig. 1 illustrates an example.

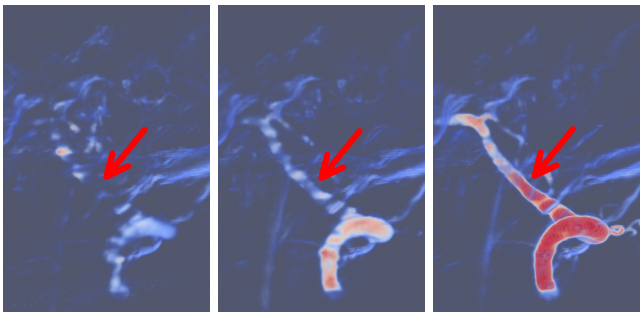


Fig. 1: Example of three frames of a 4D DSA reconstruction showing unrealistic contrast agent inflow. Arrows indicate artifacts caused by occluding bone structures.

Various approaches proposed to use information from neighboring views instead of solely one per time-instance. Clinical prototypes use angular searches like the ones presented in [2], [6] to find projections that are less affected by overlaps. [7] suggests a SART-like approach that iteratively refines neighboring volumes after correlating them by spatial and temporal smoothing, [8] a tomosynthesis-based approach that disregards rays that intersect multiple vessels.

The question of how the arising ambiguities should be a addressed remains an ongoing topic in research. Another solution to this open question could be the enforcement of physical constraints. The conservation of mass and momentum, for instance, could prevent aforementioned inconsistencies, by enforcing a steady and smooth advance of the injection bolus. Unavoidable ambiguities could be resolved by a reproduction of a natural flow pattern between the volumes of successive time instances.

Ideally, such an analysis could be provided by a computational fluid dynamics (CFD) simulation. Nevertheless, this requires the estimation of proper boundary and initial conditions for the simulation. Important parameters for CFD simulations like inflow velocity, heart rate and phase, directly can be estimated from a 4D DSA reconstruction [9]. This admits the assumption, that also the extraction of coarse flow pattern in some parts of 4D DSA could be possible by simple approaches. If we can embed CFD into an iterative algorithm,

[1]Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany
[2]Department of Neuroradiology, University Hospital, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany
[3]Chair for System Simulation, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

also an imperfect simulation could help to refine the existing 4D reconstruction.

One efficient method to extract flow information without expensive CFD simulation is the optical flow. In this preliminary work, we are interested in its capabilities to extract flow patterns. Therefore, we investigate its use on complete volumetric information as well as on intermediate reconstructions.

II. Methods

A. DSA Acquisition Protocol

In this paper, we assume a reconstruction from a 4D acquisition protocol with a duration of 12 s for the mask and contrast-enhanced fill scan and an angular range of 260°. 304 image pairs of projections are captured by the two runs, properly registered and log-subtracted. A static volume can be obtained by a standard filtered backprojection algorithm and successive vessel segmentation. We assume that both the projections and the static volume are already available at the beginning of our 4D reconstruction pipeline.

B. Iterative 4D Reconstruction

Our 4D reconstruction algorithm follows the proposal of [7]. The approach uses an iteration technique based on the well-known simultaneous algebraic reconstruction technique (SART) [10] to refine an already existing 4D volume series. For this purpose, the time-resolved volumes are forward projected onto the acquired project data at the corresponding time instances at the beginning of one iteration. This allows the generation of an error image. A modified backprojection step distributes the obtained correction values equally over the complete path length within vessels. Temporal and spatial Gaussian filtering is used as a regularization step after each iteration, in order to favor smooth solutions. The process can be repeated until convergence is observed.

Iterative reconstruction is well-suited for our proposition to integrate a CFD-based regularizer into the reconstruction process as it allows the flexible refinement of the solution by different regularizers. For the evaluations in this paper, we forgo the use of initial estimate and start with an all-zero volume.

C. Optical Flow Estimation

We suggest an optical flow estimation to determine initial velocity fields for CFD simulations based on the volumetric result of 4D reconstruction. In contrast to [11] who used a finite element approach to quantify flow rate in vessels, we do not have the goal to measure exact physical quantities. The focus is to determine coarse directions of the flows over the whole time series based on two successive volume frames without the need of any manual or automatic preprocessing like the generation of a voxelized or meshed vessel segmentation. Furthermore, the algorithm should generate a smooth dense optical flow field that resembles well the characteristics of the data as a fluid compound. Sharp edges or flow vectors normal to vessel boundaries could cause instabilities on the initialization or the regularization of a CFD simulation.

Ready-to-use approaches for volumetric flow or deformation estimation did not meet the aforementioned criteria. The Farnebäck optical flow algorithm [12], [13] showed promising flow patterns upon visual inspection of two-dimensional slices of the used data. Since an extension to the third dimension is possible and was already successful used in the field of 3D echocardiography for rigid registration [14], but not publicly available, we opted for an own implementation of this algorithm.

In contrast to many other optical flow estimators, that are based on the tracking of edges, the chosen algorithm is based on a polynomial expansion. This might be an advantage for the smooth characteristics of the captured contrast agent. Furthermore, the method is numerically very stable as the estimation is performed redundantly in a local neighborhood.

III. Evaluation

A ground truth for clinical 4D DSA data does inherently not exist. In the following, we wish to validate our optical flow algorithm on a numerical phantom that resembles a simple vascular geometry. Fig. 2 depicts three regions of interest of this phantom that will be distinguished in analysis.

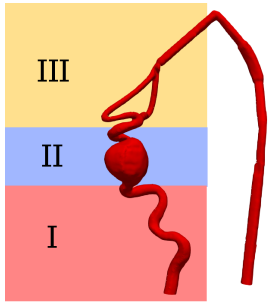


Fig. 2: Numerical phantom with three regions of interest used for evaluation

Similar to [7], a fluid simulation with a simulated heart rate of 60 beats per minute provides known velocity fields, needed for a subsequent particle-based advection and diffusion of virtual contrast agent. At the beginning of a simulated DSA acquisition, the virtual contrast agent is injected at constant rate into region I. Our simulation comprises only the inflow phase of a DSA acquisition until the bolus front starts to leave region III. For that time interval, we use projective geometry to determine ideal, noise-free projections, corresponding to the trajectory of a real 4D DSA protocol.

It should be noted that the availability of this ground truth data enables two possible evaluation scenarios for optical flow calculation. First, the displacements can be determined directly on successive volume frames of the simulated contrast agent transport. This is suited to show limits and general properties of the optical flow algorithm as calculation from a reconstruction can only imply an information loss. Second, the geometry of the phantom and the simulated projections can serve as an input for the presented iterative 4D reconstruction to generate the volume for flow estimation as the actual case of

(a) Ground truth velocity field (b) Optical flow calculated on simulation (c) Optical flow calculated on reconstruction

Fig. 3: Exemplary visualizations of the calculated optical flow for one time instance

application for our algorithm. Both variants will be evaluated by measuring the mean angular error in comparison with the velocity field provided by the CFD simulation. Since the optical flow can only be obtained in some locations. We define a threshold for the evaluations and consider only velocity estimates with higher magnitude.

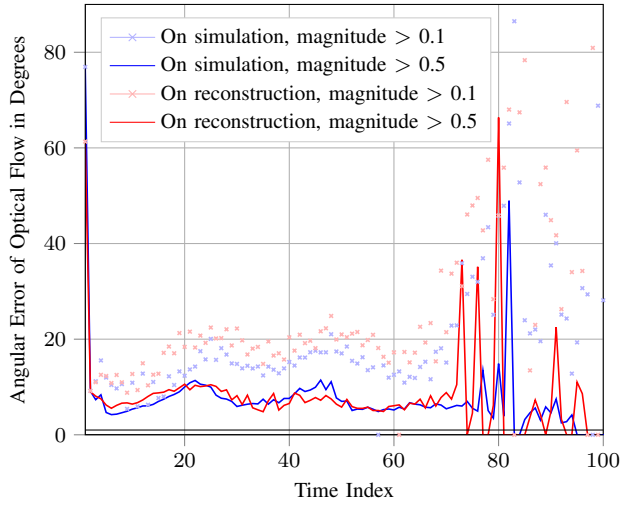
IV. Results

Figs. 4a and 4 show the mean angular error in dependence of the time index and a chosen threshold (0.1 or 0.5) for region I and II.

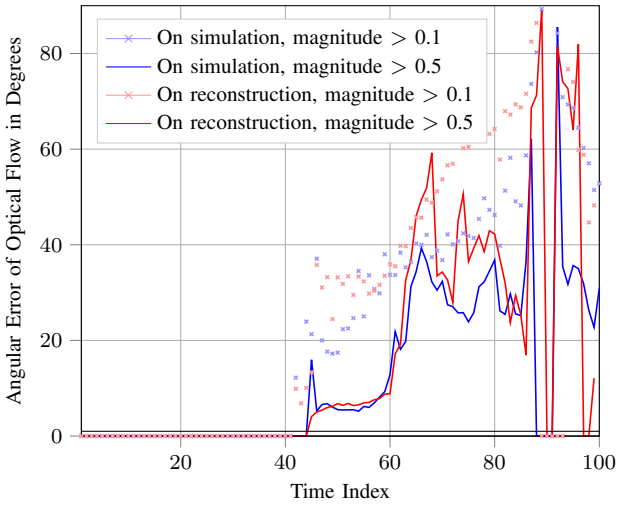
Two tendencies can be observed in both figures: First, the mean angular error is lower if only vectors with a magnitude greater 0.5 are considered. Second, the performance of the optical flow calculated on the 4D reconstruction is only slightly lower compared to the situation that the simulation is directly used as an input (e.g. 9.5° vs 8.0° for magnitude > 0.5 in region I or 32.6° vs 27.9° in region II).

The local performance is as follows: we obtain low error rates in the wide vessel of region I until time index 75. This coincides with time interval that the first bolus front passes through the wide vessel segment in region I. After that intensities merely stay the same and no optical flow can be estimated. For region II, confident flow directions can only be determined up on the entrance of the bolus in the modeled aneurysm. Then, turbulent flow prevents precise estimations. For region III with a bifurcation into two smaller vessels, no reliable figure of the mean angular error could be given as the numbers were strongly fluctuating. An explanation can be given by Fig. 5c: many flow vectors are estimated with opposite direction. Visual inspection of the affected area suggests that the flow in these smaller vessels was to fast to be captured by a frame-to-frame optical flow estimation.

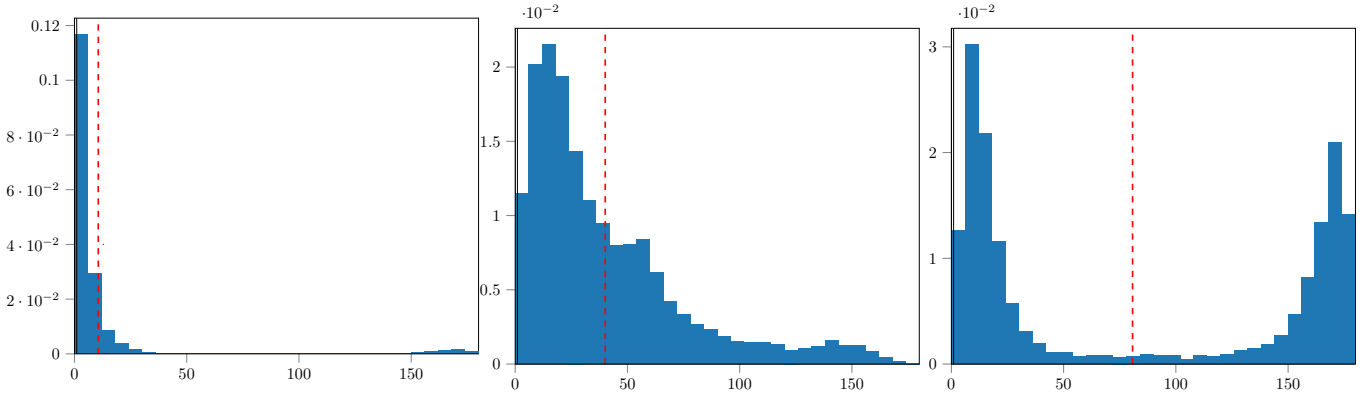
For all time-instances, the obtained flow field was very smooth without any coarse outliers (see Fig. **??** for an example. This would be a necessary condition for applying this



(4a) Average angular error in degrees over time in region I



(4b) Average angular error in degrees over time in region II



(a) Region I (b) Region II (c) Region III

Fig. 5: Normalized histograms of the distribution of the angular error of a typical representative time instance per region of interest (optical flow estimated on simulation, threshold 0.1)

method on real data affected by noise. Also in the rather problematic region II where only imprecise estimation could be performed, the averaging approach of the Farnebäck optical flow algorithm created a solution that was consistent without edges or abrupt changes over time and neighboring regions. The precision of the estimate could be further improved by a subsequent CFD simulation that could profit from regions where the optical flow performs well. It remains unclear in what degree reconstruction artifacts proceeding from vascular overlaps affect the optical flow estimation. In the presented experiment, artifacts from a short overlap phase could be completely avoided by the iterative reconstruction algorithm.

V. Conclusion

In this paper, we proposed the use of Farnebäck optical flow for the purpose of determining proper initial or regularization conditions for CFD simulation. The algorithm was evaluated on a numerical vessel phantom and reproduced well the orientation of the underlying ground truth vectors. Similar results could be achieved if the input data for flow estimation was generated by an iterative 4D DSA reconstruction. This feeds our prospects to determine coarse flow patterns also on clinical data that is affected by noise and possible reconstruction artifacts. It remains future work to investigate whether this approach is suited for integration of CFD into a 4D DSA framework.

The source code of the implemented optical flow algorithm has been made available as a ready-to-use Python package on Github (`https://github.com/theHamsta/farneback3d`, Nvidia CUDA required).

References

[1] A. D. Copeland, R. S. Mangoubi, M. N. Desai, S. K. Mitter, and A. M. Malek, "Spatio-temporal data fusion for 3d+t image reconstruction in cerebral angiography," *IEEE Transactions on Medical Imaging*, vol. 29, no. 6, pp. 1238–1251, June 2010.

[2] B. Davis, K. Royalty, M. Kowarschik, C. Rohkohl, E. Oberstar, B. Aagaard-Kienitz, D. Niemann, O. Ozkan, C. Strother, and C. Mistretta, "4D digital subtraction angiography: Implementation and demonstration of feasibility," *American Journal of Neuroradiology*, vol. 34, no. 10, pp. 1914–1921, 2013. [Online]. Available: http://www.ajnr.org/content/34/10/1914

[3] C. Sandoval-Garcia, K. Royalty, P.-F. Yang, D. Niemann, A. Ahmed, B. Aagaard-Kienitz, M. K Başkaya, S. Schafer, and C. Strother, "4D DSA a new technique for arteriovenous malformation evaluation: A feasibility study," vol. 8, 01 2015.

[4] S. Lang, P. Gölitz, T. Struffert, J. Rösch, K. Roessler, M. Kowarschik, C. Strother, and A. Doerfler, "4D DSA for dynamic visualization of cerebral vasculature: A single-center experience in 26 cases." *AJNR. American journal of neuroradiology*, vol. 38 6, pp. 1169–1176, 2017.

[5] C. Sandoval-Garcia, P. Yang, T. Schubert, S. Schafer, S. Hetzel, A. Ahmed, and C. Strother, "Comparison of the diagnostic utility of 4D-DSA with conventional 2D- and 3D-DSA in the diagnosis of cerebrovascular abnormalities," vol. 38, 03 2017.

[6] B. Davis, E. Oberstar, K. Royalty, S. Schafer, C. Strother, and C. Mistretta, "Volumetric limiting spatial resolution analysis of four dimensional digital subtraction angiography (4D-DSA)," 03 2015.

[7] J. Endres, C. Rohkohl, S. Schafer, K. Royalty, A. Maier, M. Kowarschik, and J. Hornegger, "4D DSA Iterative Reconstruction," in *Proceedings of the Fully3D*, 2015, pp. 276–279.

[8] M. Buehler, J. M. Slagowski, C. A. Mistretta, C. M. Strother, and M. A. Speidel, "4D DSA reconstruction using tomosynthesis projections," pp. 10 132 – 10 132 – 12, 2017. [Online]. Available: http://dx.doi.org/10.1117/12.2255197

[9] M. Boegel, S. Gehrisch, T. Redel, C. Rohkohl, P. Hoelter, A. Doerfler, A. Maier, and M. Kowarschik, "Patient-individualized boundary conditions for cfd simulations using time-resolved 3d angiography," *International Journal of Computer Assisted Radiology and Surgery*, vol. 11, no. 6, pp. 1061–1069, Jun 2016.

[10] A. H. Andersen and A. C. Kak, "Simultaneous algebraic reconstruction technique (sart): A superior implementation of the art algorithm," *Ultrasonic Imaging*, vol. 6, no. 1, pp. 81–94, 1984, pMID: 6548059. [Online]. Available: https://doi.org/10.1177/016173468400600107

[11] P. Maday, M. Kowarschik, S. Demirci, and N. Navab, "Towards blood flow quantification using dense flow volumes," in *MICCAI CVII-STENT proceedings*, 2014.

[12] G. Farnebäck, "Polynomial expansion for orientation and motion estimation," Ph.D. dissertation, Linköping University, Sweden, SE-581 83 Linköping, Sweden, 2002, dissertation No 790, ISBN 91-7373-475-6.

[13] G. Farneback, *Two-Frame Motion Estimation Based on Polynomial Expansion*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 363–370.

[14] A. Danudibroto, O. Gerard, M. Alessandrini, O. Mirea, J. D'hooge, and E. Samset, *3D Farnebäck Optic Flow for Extended Field of View of Echocardiography*. Cham: Springer International Publishing, 2015, pp. 129–136.

Performance Evaluation of Motion Artifact Correction for Myocardial Dual-Energy CT Perfusion Imaging

Zhye Yin, Xue Rui, Jed Pack, Kimberly Elmore, Peter M. Edic, James K. Min

***Abstract*— Myocardial Computed Tomography Perfusion (CTP) imaging provides hemodynamic information related to underlying Coronary Artery Disease and may facilitate its diagnosis when combined with anatomical information from Cardiac CT Angiography. Since the quantitative accuracy of myocardial CTP is often affected by motion artifacts introduced during data acquisition, we proposed a motion correction algorithm for the entire heart [1], which was augmented with an additional low-frequency correction algorithm [2] to remove residual artifacts. We collaborated with Weill Cornell Medical College to identify a subset of DECIDE-Gold exams [3], where cardiac motion limits the diagnostic capability. We applied the proposed algorithms to the identified subset of patient scans and demonstrated the image quality improvement qualitatively. During initial evaluation, we identified cases with suboptimal contrast-to-noise ratio limited our ability to correct for motion artifacts and proposed enhancements to the current motion correction algorithms. Furthermore, we proposed an entropy-based metric to quantify motion artifact reduction in clinical images provided using Dual-Energy CT Perfusion imaging.**

Keyword: Computed Tomography, Motion Artifacts, Dual-Energy CT, CT Perfusion

I. INTRODUCTION

RECENT multi-center studies have demonstrated the efficacy of combining anatomical and physiological information to diagnose Coronary Artery Disease using invasive techniques such as Invasive Coronary Angiography (ICA), which provides anatomical visualization and quantification of a suspected coronary stenosis, and Fractional Flow Reserve (FFR), which computes the ratio of the arterial pressure distal and proximal to a lesion to assess its physiological significance[4]. Since ICA and FFR are invasive techniques, there is increased risk to the patient, with FFR not used routinely in clinical practice. A similar study using non-invasive techniques such as coronary Computed Tomography Angiography (CCTA) and Computed Tomography Perfusion (CTP) has been recently proposed [3]. [1]

While early studies of conventional single-energy CTP have been promising, novel Dual-Energy CT Perfusion (DECTP) imaging provides additional information. DECTP enables estimation of the iodine concentration within the myocardium, providing a surrogate for blood volume. As with CCTA, cardiac motion during the acquisition of projection data for cardiac imaging impacts the performance of DECTP. Existing cardiac motion correction software such as SnapShot Freeze™ (GE Healthcare, Waukesha, WI - SSF) has focused on correcting for motion of coronary arteries with demonstrated improvement in image quality and interpretability [5,6]. We adapted the SSF algorithm to correct for motion in the whole heart, including myocardium and valves [1] and added low-frequency artifact correction to resolve subtle residual shading, which may confound the identification of a perfusion deficit [2].

With help of Weill Cornell Medical College (WCMC), we identified a subset of exams from the DECIDE-Gold (Dual-Energy CT for Ischemia Determination Compared to "Gold Standard" Non-Invasive and Invasive Techniques) clinical study [3] where cardiac motion during data acquisition may impact diagnostic capability and applied the proposed algorithm. We demonstrated the qualitative improvement in image quality in those identified cases. Among those cases, we found that insufficient contrast due to inadequate acquisition can affect the performance of the proposed algorithm. Additional improvements were proposed and implemented. Finally, we explored entropy-based quantitative metric to evaluate image quality improvement in motion corrected images.

II. METHOD

A. Data acquisition and process

As part of the DECIDE-Gold clinical study [7], patients suspected of coronary artery disease were

Zhye Yin, Jed Pack and Peter M Edic are with GE Global Research. Kimberly Elmore and James K. Min are with Weill Cornell Medical College.

scanned using a dual-energy CT scanner (Discovery™ CT750 HD, GE Healthcare, Waukesha, WI). CT projection data were acquired when the patient was nominally at rest and at pharmacologically-induced stress, i.e., after administration of a vasodilator such as adenosine, if so indicated. Dual-energy CT (DECT) data processing enabled quantification of the iodine density distribution in the patient's myocardium; we aim to use the iodine density map as a surrogate for blood volume in the myocardium utilizing a single acquisition – denoted as static myocardial perfusion imaging. Furthermore, material decomposition using DECT provided multi-material beam hardening correction, mitigating shading artifacts in the myocardium between highly-attenuating structures, such as between the contrast-enhanced descending aorta and left ventricle. The energy-independent basis material density distributions were used to compute monochromatic volumetric CT images, typically represented at 70 keV, although 60-keV images were used to improve contrast-to-noise ratio between the contrast-enhanced blood pool within the ventricle and adjoining myocardium, when necessary (details are provided in section B). Algorithms were developed to allow feathering of data at slab boundaries within the heart, since the CT scanner nominally provided 40 mm of axial coverage per gantry revolution. To mitigate residual shading artifacts due to motion of the heart during data acquisition, protocols similar to those used for coronary artery motion correction (SnapShot Freeze™, GE Healthcare, Waukesha, WI) were employed; whole-heart motion estimation techniques identified locations of significant motion (motion control points) during data acquisition, and determined corresponding motion vectors. We then corrected for motion in regions of interest (ROIs) surrounding the motion control points within the myocardium using the monochromatic images and the motion vectors [8]. To quantify the benefit of the motion estimation/correction techniques, we estimated the image entropy within ROIs about the motion control points, and compared values computed from original images to those in motion-corrected images (detailed provide in section C).

B. Improvement in motion correction

The motion correction algorithm was applied to the cases that we received. Motion artifacts were reduced significantly for a subset of the cases. However, we also noticed some challenging cases, where the motion artifacts remain uncorrected and improvements were subtle; we further investigated these cases. The contrast on these challenging cases appeared to be suboptimal, especially the contrast-to-noise ratio (CNR) between cardiac chambers and neighboring myocardium. Those cases were often acquired with inadequate acquisition protocols and/or imperfect contrast agent delivery. For dual-energy CT imaging, we can generate images at any given monochromatic energy using the basis material density estimates. Typically, lower-energy monochromatic images inherently show higher contrast; therefore, we decided to use 60-keV images for the estimation of the motion vectors, instead of the typical direct estimation of motion vectors from 70-keV images. As shown in Figure 1, a 60-keV images have more noise, but has significantly higher contrast between cardiac chambers and proximal myocardium compared to 70-keV images. The motion vectors estimated from the 60-keV images were then used to correct the motion artifacts for the basis material density images as well as the 70-keV images.

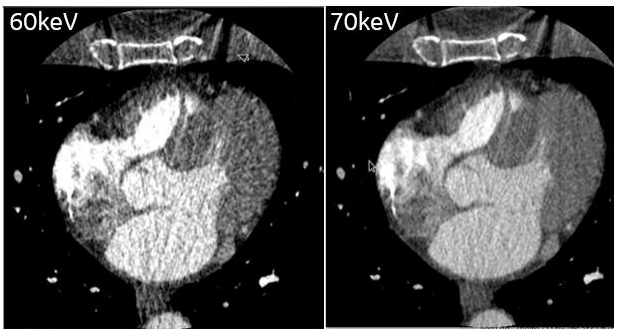


Figure 1 Monochromatic 60-keV image (left) and monochromatic 70-keV image (right). Contrast between ventricular chambers and surrounding myocardium is enhanced in lower-energy monochromatic images. Images are displayed at [window width, window level] = [800 1200].

C. Quantification of motion artifact reduction

We have previously reported quantitative results of our motion-correction techniques applied to both simulation and phantom data [8]. However, since we now want to assess the benefit of the motion correction techniques – both qualitatively and quantitatively – applied to clinical images acquired during the DECIDE-Gold clinical study, ground-truth motion-free images are not available as in the previous studies. As such, we postulated techniques to estimate the reduction of motion artifact.

Due to motion of the heart during data acquisition, shading artifacts in reconstructed CT images result, as well as blurring or doubling of chamber wall

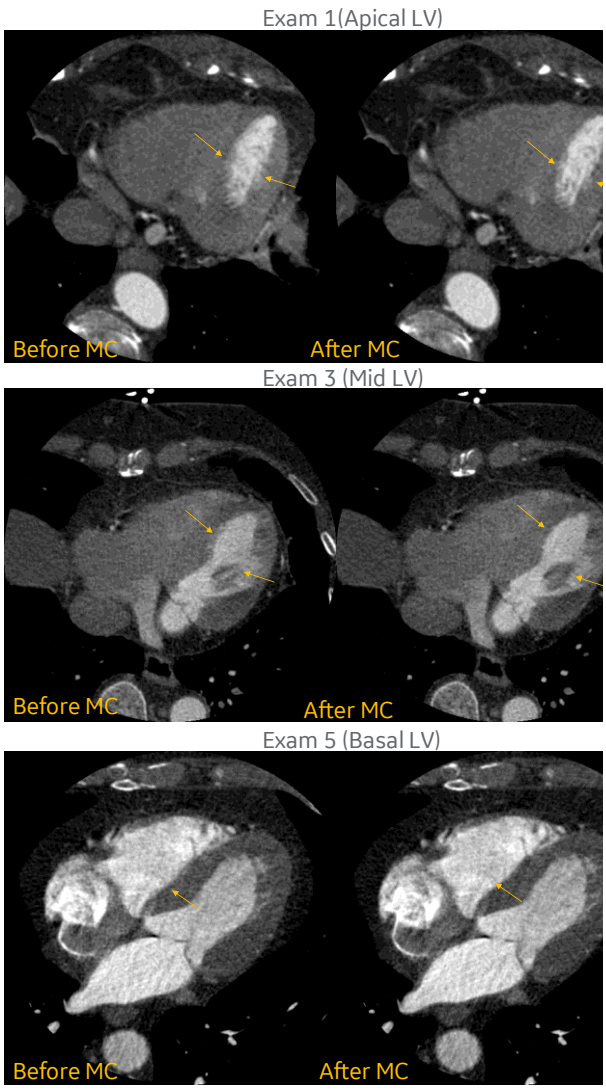


Figure 2 Axial CT images of basal (through the location of the mitral valve), mid (through the mid left ventricular myocardium) and apical (through the apex of the heart) regions of heart from a subset of exams from the DECIDE-Gold study. Left images and right images were generated before and after applying motion-correction (MC) algorithms, respectively. Arrows indicate regions in the myocardium impacted by motion artifacts and improved by correction. Images are displayed at [window width, window level] = [300 1000].

boundaries. When considering the impact on image quality, these effects will cause a broadening of the histogram of CT numbers (as represented by Hounsfield Units – HU) within ROIs about the motion control points. Subsequently, after successful application of motion estimation/correction, image quality improves, resulting in a reduction of shading artifacts and better delineation of chamber boundaries. Both effects should narrow the histogram of reconstructed CT numbers. When considering these observations – broadening and narrowing of CT number histograms for motion-corrupted and motion-corrected images, respectively – entropy is a reasonable choice for comparatively characterizing the CT number distribution in ROIs about the motion control points. As such, we computed the histogram of CT numbers within the ROIs before and after motion correction, and used a standard entropy functional, which computes the sum of the probability-weighted logarithm of the probability density distribution. One confounding factor will be image noise that contributes a "background" entropy that reduces the entropy contrast between motion-corrupted and motion-corrected images.

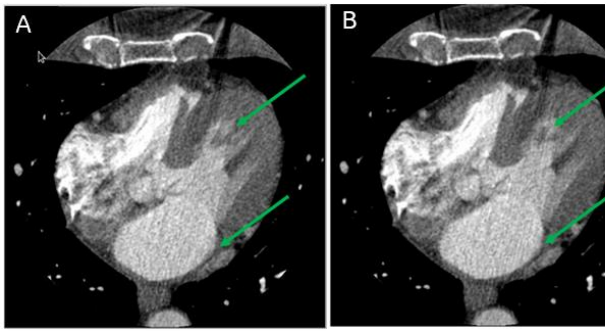


Figure 3 Motion-corrected monochromatic 70-keV images (A) using motion vectors estimated from monochromatic 70-keV images and (B) using motion vectors estimated from 60-keV images. Green arrows indicate locations of motion artifact reduction. Images are displayed at [window width, window level] = [800 1200].

III. Results

Using the subset of exams from the DECIDE-Gold study, we demonstrated reduction in motion artifacts, as shown in Figure 2. Three axial locations are used to assess improvement: (1) basal: an axial slice through the myocardium at the location of the mitral valve, (2) mid: an axial slice through the mid left ventricle, and (3) apical: a slice through the apex of the heart. Untreated images are shown on the left in Figure 2 and motion-corrected images are shown on the right for the representative patient exams. Arrows indicate regions where motion artifact is reduced, providing clearer delineation of the myocardial/chamber boundary.

We also identified clinical exams with suboptimal CNR, resulting in marginal improvement in image quality using the proposed algorithm. For those cases, we used monochromatic 60-keV images to extract motion vectors instead of monochromatic 70-keV

images, which are typically used. Motion correction in the myocardium region improved with motion vectors generated from the 60-keV images, as shown in Figure 3. The boundaries of the myocardium are clearer and doubling of chamber walls is reduced, shown at locations indicated by green arrows in Figure 3.

To quantitatively assess image quality improvement resulting from motion correction, we explored the entropy-based image quality metric. Figure 4 shows simple noise-free images mimicking the CT number contrast (400 HU) between a heart chamber and the surrounding myocardium: (left) original, resolution-blurred chamber; (center) motion-blurred chamber – sum of original image and a diagonally-shifted version; and (right) differential image. The location of a motion control point (location of maximum differential magnitude) for this image would be at the position of the red dot in Figure 4. The entropy computed from the normalized histogram of CT numbers, i.e., the probability density function of CT numbers, within an ROI centered on the motion control point for the original and motion-blurred images was 6.48 and 7.28, respectively. When uncorrelated white noise, with statistics comparable to noise observed in cardiac CT images, is added to the images, the computed entropy values are 7.49 and 7.60, respectively. As such, we are currently investigating bandpass filtering techniques to reduce the influence of noise in the entropy estimates while still capturing object blur, with the goal of increasing the entropy contrast.

IV. CONCLUSIONS

We identified a subset of exams from the DECIDE-Gold study where motion artifact and suboptimal CNR reduced image quality, impacting diagnostic capability, and applied the proposed motion-correction techniques to these exams. We demonstrated that the enhanced motion-correction methods reduce artifacts that may confound the clinical usefulness of CT perfusion information for assessing territorial perfusion deficits resulting from coronary artery disease. We postulated an entropy-based metric to quantitatively assess the improvement in image quality resulting from application of the proposed approaches, with an ultimate goal of assessing the impact of this improvement on diagnostic performance.

V. ACKNOWLEDGEMENT

Research reported in this publication was supported by the National Heart, Lung, And Blood Institute of the National Institutes of Health under Award Number R01HL111141. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Figure 4 Images of a simplified heart chamber and surrounding myocardium: original, resolution-blurred chamber (left), motion-blurred chamber (center), and differential image (right). The location of a motion control point is shown by the red dot in the differential image.

REFERENCES

[1] Pack, J. D., Yin, Z., Xiong, G., Mittal, P., Dunham, S. et al. Motion correction for improving the accuracy of dual-energy myocardial perfusion CT imaging. Proc. SPIE 9788, Medical Imaging 2016, doi:10.1117/12.2216986

[2] Yin, Z., Pack, J. D., Xiong, G., Dunham, S, Elmore, K., Mittal, P., Edic, P. M., Min, J. K., et al. Low-frequency Motion Artifact Correction for Myocardial Dual-Energy CT Perfusion Imaging. 2016 IEEE Nuclear Science Symposium and Medical Imaging Conference

[3] Truong, Q. A., Knaapen, P., Pontone, G., Andreini, D., Leipsic, J., Carrascosa, P., ... & Min, J. K. (2015). Rationale and design of the dual-energy computed tomography for ischemia determination compared to "gold standard" non-invasive and invasive techniques (DECIDE-Gold): A multicenter international efficacy diagnostic study of rest-stress dual-energy computed tomography angiography with perfusion. Journal of Nuclear Cardiology, 22(5), 1031-1040

[4] De Bruyne B, Sarma J. Fractional flow reserve: a review: invasive imaging. Heart. 2008;94:949–959

[5] Min, J. K., Arsanjani, R., Kurabayashi, S., Andreini, D., Pontone, G., Choi, B. W., … Leipsic, J. (2013). Rationale and design of the ViCTORY (Validation of an Intracycle CT Motion CORrection Algorithm for Diagnostic AccuracY) trial. Journal of Cardiovascular Computed Tomography, 7, 200–206.

[6] Fan, L., Zhang, J., Xu, D., & Dong, Z. (2015). CTCA Image Quality Improvement by Using Snapshot Freeze Technique Under Prospective and Retrospective Electrocardiographic Gating. Journal of Computer Assisted Tomography, 39 (2), p 202-206.

[7] https://clinicaltrials.gov/ct2/show/NCT02178904?term=DECIDE-Gold&rank=1

[8] Yin, Z., Pack, J. D., Edic, M. E., Elmore, K, Min, J. K. (2017). Low-frequency Motion Artifact Correction for Myocardial Dual-Energy CT Perfusion Imaging, Medical Physics (Submitted)

Papoulis-Gerchberg Algorithms for Limited Angle Tomography Using Data Consistency Conditions

Yixing Huang, Oliver Taubmann, Xiaolin Huang, Guenter Lauritsch and Andreas Maier

***Abstract*—The Papoulis-Gerchberg (P-G) algorithm is widely used for extrapolation of band-limited signals. It is applicable to limited angle tomography as well since typical imaged objects in computed tomography have a limited spatial extent, which means that the Fourier transforms of the objects can be considered band-limited signals. In computed tomography, some other band-limitation properties have been discovered as well, which are referred to as data consistency conditions. For example, the Fourier transform of a parallel-beam sinogram has an empty double-wedge region. The Chebyshev-Fourier transform of a parallel-beam sinogram only has nonzero values inside a wedge region and these values form a checkerboard pattern, which is Helgason-Ludwig consistency condition. In this paper, we propose two P-G algorithms to restore missing data in limited angle tomography using the above two consistency conditions. Numerical experiments on the Shepp-Logan phantom demonstrate that they can reduce streaks better than the conventional P-G algorithm.**

I. Introduction

In computed tomography (CT), image reconstruction from data acquired in an insufficient angular range is called limited angle tomography [1]–[3]. It arises when the gantry rotation of a CT system is restricted by other system parts or external obstacles. Because of missing data, artifacts, typically in the form of streaks, will occur in the reconstructed images.

Iterative algorithms with total variation are popular in limited angle tomography [4]–[7]. They incorporate sparsity at the gradient domain of medical images into the reconstruction and hence reduce streak artifacts. However, they are computationally expensive. Recently, deep learning has obtained impressive achivement for streak reduction in limited angle tomography [8]–[10], but it depends heavily on the availability and quality of training data.

Extrapolation/interpolation of the missing data is a common way to deal with data insufficiency. The Papoulis-Gerchberg (P-G) algorithm [11], [12] is well known for extrapolating band-limited signals. In limited angle tomography, according to the central slice theorem, some frequency components of an imaged object are missing. Also, the imaged object has a limited spatial extent, which means that the Fourier transform of the object can be considered a band-limited signal. Therefore, to extrapolate the missing frequency components from the measured ones, the P-G algorithm can be applied [13]–[15]. However, it requires strong *a priori* knowledge in that the object support should be known.

In computed tomography, many data consistency conditions have been discovered. For example, the Fourier transform of a parallel-beam sinogram has an empty double-wedge region [16]. The Helgason-Ludwig consistency condition [17], [18] states that Chebyshev-Fourier transform of a parallel-beam sinogram only has nonzero values inside a wedge region and these values form a checkerboard pattern. These consistency conditions are also band-limitation properties of CT data. Many experiments have demonstrated that consistency conditions are beneficial in restoring missing data [19]–[22]. Therefore, in this paper, we propose two P-G algorithms to restore missing data in limited angle tomography using the above two consistency conditions.

II. Methods

A. General P-G Algorithm

Given a measured segment $g_0(t)$ from a band-limited signal $g(t)$,

$$g_0(t) = \begin{cases} g(t) & |t| \le t_0, \\ 0 & |t| > t_0, \end{cases}$$

a characteristic function of the measurement interval $[-t_0, t_0]$,

$$M_g(t) = \begin{cases} 1 & |t| \le t_0, \\ 0 & |t| > t_0, \end{cases}$$

and a perfect low-pass filter $L(w)$,

$$L(w) = \begin{cases} 1 & |w| < w_c, \\ 0 & |w| > w_c, \end{cases}$$

where w_c is the band limit of $g(t)$, $g(t)$ can be estimated by the following algorithm [11], [12],

$$g^l(t) = g_0(t) + (1 - M_g(t)) \cdot \mathcal{F}_1^{-1}\left(L(w) \cdot \mathcal{F}_1\left(g^{l-1}(t)\right)\right),$$

where $\mathcal{F}_d$ and $\mathcal{F}_d^{-1}$ are the d-dimensional Fourier transform and inverse Fourier transform operators, respectively, $g^l(t)$ is an estimation of $g(t)$ at the l-th iteration, and $g^0(t) = g_0(t)$.

B. Conventional P-G Algorithm Using Object Support

We denote an imaged object by $\boldsymbol{f}(\boldsymbol{x})$. Typically the object has a compact support in the spatial domain, denoted by $\boldsymbol{S}$. A characteristic function for the support $\boldsymbol{S}$ is defined as,

$$L_{\boldsymbol{S}}(\boldsymbol{x}) = \begin{cases} 1 & \boldsymbol{x} \in \boldsymbol{S}, \\ 0 & \text{otherwise.} \end{cases}$$

Y. Huang is with Pattern Recognition Lab, Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen, Germany (e-mail: yixing.yh.huang@fau.de).

O. Taubmann, and A. Maier are with Pattern Recognition Lab, Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen, Germany, and also with Erlangen Graduate School in Advanced Optical Technologies (SAOT), Erlangen, Germany.

X. Huang was with Pattern Recognition Lab, Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen, Germany and now is with Institute of Image Processing and Pattern Recognition, Shanghai Jiao Tong University, Shanghai, China.

G. Lauritsch and O. Taubmann are with Siemens Healthcare GmbH, Forchheim, Germany.

The Fourier transform of the object $\boldsymbol{f}(\boldsymbol{x})$ is denoted by $\boldsymbol{F}(w,\theta)$ in polar coordinates, $\boldsymbol{F}(w,\theta) = \mathcal{F}_2\boldsymbol{f}(\boldsymbol{x})$. According to the symmetry property of the Fourier transform, we have $\mathcal{F}_2\boldsymbol{F}(w,\theta) = 2\pi\boldsymbol{f}(\boldsymbol{x})$. Therefore, $\boldsymbol{F}(w,\theta)$ is a band-limited function as, after applying a Fourier transform to it, its components are nonzero only in S.

In computed tomography, the parallel-beam sinogram of the object $\boldsymbol{f}(\boldsymbol{x})$ is denoted by,

$$p(s,\theta) = \int_{\boldsymbol{x}\cdot\boldsymbol{\theta}=s} \boldsymbol{f}(\boldsymbol{x})\mathrm{d}\boldsymbol{x},$$

where $\boldsymbol{\theta} = (\cos\theta, \sin\theta)$, $\theta \in [0, 2\pi)$ is the direction orthogonal to the X-rays, and $s \in (-\infty, \infty)$ is the detector index. The central slice theorem expresses that,

$$\mathcal{F}_1 p(s,\theta) = \boldsymbol{F}(w,\theta).$$

When the sinogram is measured for the whole angular range $[0,\pi)$, the object $\boldsymbol{f}$ can be reconstructed using the standard filtered back-projection (FBP) reconstruction algorithms. However, in limited angle tomography, the sinogram is measured only in a limited angular range, denoted by $[0,\theta_{\max})$ where $\theta_{\max}$ is the maximum scanned angle, $\theta_{\max} < \pi$. In this case, a double wedge region is missing in $\boldsymbol{F}$, i. e.,

$$\boldsymbol{F}_{\text{limited}}(w,\theta)|_{\theta_{\max}\le\theta<\pi,-\infty<w<\infty} = 0,$$

where $\boldsymbol{F}_{\text{limited}}$ is the measured frequency components of $\boldsymbol{f}$ in limited angle tomography. A characteristic function for the measured region is defined as,

$$M_{\boldsymbol{F}}(w,\theta) = \begin{cases} 1 & \theta \in [0,\theta_{\max}), \\ 0 & \text{otherwise.} \end{cases}$$

Restoring the complete $\boldsymbol{F}(w,\theta)$ from $\boldsymbol{F}_{\text{limit}}(w,\theta)$ can thus be considered as an extrapolation problem of a band-limited function. Therefore, the P-G algorithm can be applied,

$$\begin{aligned}\boldsymbol{F}^l(w,\theta) = \boldsymbol{F}_{\text{limited}}(w,\theta) + (1 - M_{\boldsymbol{F}}(w,\theta))\cdot \\ \mathcal{F}_2^{-1}(L_{\boldsymbol{S}}(\boldsymbol{x})\cdot\mathcal{F}_2\boldsymbol{F}^{l-1}(w,\theta)),\end{aligned} \tag{1}$$

where $\boldsymbol{F}^l(w,\theta)$ is an estimation of $\boldsymbol{F}(w,\theta)$ at the l-th iteration and $\boldsymbol{F}^0(w,\theta) = \boldsymbol{F}_{\text{limited}}$. Note that $\mathcal{F}_2$ and $\mathcal{F}_2^{-1}$ can change position here due to the symmetry property. This conventional P-G algorithm is denoted by P-G_{OS}.

C. P-G Algorithm Using Fourier Property Of Sinograms

The 2-D Fourier transform of a complete parallel-beam sinogram is as follows,

$$\boldsymbol{P}(w,k) = \mathcal{F}_2 p(s,\theta) = \frac{1}{2\pi}\int_0^{2\pi}\int_{-\infty}^{\infty} p(s,\theta)e^{-i(ws+k\theta)}\mathrm{d}s\mathrm{d}\theta.$$

A consistency condition in the sinogram's Fourier space is represented as follows [16],

$$\boldsymbol{P}(w,k) \approx 0 \text{ when } \left|\frac{k}{w}\right| > r_p,$$

where r_p is distance of the farthest point on the object to the isocenter. It means that a double-wedge region of $\boldsymbol{P}(w,k)$ is zero. Hence, $p(s,\theta)$ is a band-limited function. A characteristic function for the double-wedge region is defined as,

$$L_{\boldsymbol{P}}(w,k) = \begin{cases} 0 & \left|\frac{k}{w}\right| > r_p, \\ 1 & \text{otherwise.} \end{cases}$$

In limited angle tomography, we denote the measured sinogram by $p_{\text{limited}}(s,\theta)$, $\theta \in [0,\theta_{\max})$. According to $p(s,\theta) = p(-s,\theta+\pi)$, the limited angle sinogram is extended to a 2π angular range, denoted by $p'_{\text{limited}}(s,\theta)$. A characteristic function for the available part of the sinogram is defined as,

$$M_p(s,\theta) = \begin{cases} 1 & \theta \in [0,\theta_{\max}) \cup \theta \in [\pi,\pi+\theta_{\max}), \\ 0 & \text{otherwise.} \end{cases}$$

To extrapolate/interpolate the missing sinogram, the P-G algorithm can be applied,

$$\begin{aligned}p^l(s,\theta) = p'_{\text{limited}}(s,\theta) + (1 - M_p(s,\theta))\cdot \\ \mathcal{F}_2^{-1}\left(L_{\boldsymbol{P}}(w,k)\cdot\mathcal{F}_2 p^{l-1}(s,\theta)\right),\end{aligned} \tag{2}$$

where $p^l(s,\theta)$ is an estimation of $p(s,\theta)$ at the l-th iteration and $p^0(s,\theta) = p'_{\text{limited}}$. The proposed P-G algorithm using the double-wedge property of sinograms' frequency domain is denoted by P-G_{DW}. It was proposed for defect detector gap compensation of emission CT in [23] and sparse-view CT reconstruction in [19].

D. P-G Algorithm Using HLCC

We define the n-th order moment curve of a parallel-beam sinogram $p(s,\theta)$ as,

$$a_n(\theta) = \int_{-\infty}^{\infty} p(s,\theta)T_n(s)\mathrm{d}s,$$

where $T_n(s) = s^n$ and n is the order of the monomial. The Fourier transform of the moment curve is,

$$b_n(m) = \frac{1}{2\pi}\int_0^{2\pi} a_n(\theta)e^{-im\theta}\mathrm{d}\theta.$$

HLCC [17], [18] tells that,

$$b_n(m) = 0, \quad |m| > n \text{ or } m+n \text{ is odd.} \tag{3}$$

Therefore, the moment curves are band-limited functions. A characteristic function for HLCC is defined as,

$$L_{\text{HLCC}}(n,m) = \begin{cases} 1 & \text{if } |m| \le n \text{ and } m+n \text{ is even,} \\ 0 & \text{otherwise.} \end{cases}$$

When $T_n(s)$ is replaced by orthogonal polynomials, e.g., Chebyshev polynomials or Gegenbauer polynomials, $p(s,\theta)$ can be conveniently restored from $a_n(\theta)$ while $L_{\text{HLCC}}(n,m)$ remains the same. In this paper, we use the Chebyshev polynomial of the second kind,

$$U_n(s) = \frac{\sin\left((n+1)\arccos(s)\right)}{\sqrt{1-s^2}}.$$

$U_n(s)$ is a family of orthogonal polynomials at domain [-1, 1] with the scalar weight $W(s) = (1-s^2)^{1/2}$, i.e.,

$$\int_{-1}^{1} W(s)\cdot U_n(s)\cdot U_{n'}(s)\mathrm{d}s = \begin{cases} 0, & n \ne n' \\ \pi/2, & n = n'. \end{cases}$$

Note that here we normalize the detector index s to a range of [-1, 1]. Thus, an approximate sinogram can be restored by the inverse Chebyshev transform from the moment curves,

$$p_{n_r}(s,\theta) = \frac{2}{\pi}\sum_{n=0}^{n_r} a_n(\theta)\left(W(s)\cdot U_n(s)\right),$$

where n_r is the number of orders used.

In limited angle tomography, the moment curves $a_n(\theta)$ are available only at the angular range of $[0,\theta_{\max})$ and $[\pi,\pi+\theta_{\max})$. The available moment curve is denoted by $a_{n,\text{limited}}(\theta)$. An characteristic function for the available parts is defined as,

$$M_{a_n}(\theta) = \begin{cases} 1 & \theta\in[0,\theta_{\max}) \text{ or } \theta\in[\pi,\pi+\theta_{\max}),\\ 0 & \text{otherwise.}\end{cases}$$

To get $p_{n_r}(s,\theta)$, we need to extrapolate/interpolate the missing parts of $a_n(\theta), n=0,1,2,\ldots,n_r$. Hence, the P-G algorithm can be applied,

$$\begin{aligned} a_n^l(\theta) = a_{n,\text{limited}}(\theta) + (1-M_{a_n}(\theta))\cdot \\ \mathcal{F}_1^{-1}\left(L_{\text{HLCC}}(n,m)\cdot\mathcal{F}_1 a_n^{l-1}(\theta)\right),\end{aligned} \tag{4}$$

where $a_n^l(\theta)$ is an estimation of $a_n\theta$ at the l-th iteration. The proposed P-G algorithm using HLCC is denoted by P-G_{HLCC}.

Papoulis and Gerchberg have shown the convergence of P-G algorithms in the noise-free case [11], [12]. When the frequency band is known accurately, the missing signal can be extrapolated exactly with infinite iterations. However, in the presence of noise or discretization error, the missing signal typically cannot be recovered exactly. In [21], we find that the restoration of high order moment curves is severely ill-posed. The observation that the Fourier coefficients of the moment curves are sparse can be used to overcome the ill-posedness. Therefore, we define a soft-thresholding operator $\mathcal{S}_\tau$,

$$\mathcal{S}_\tau(v) = \begin{cases} v-\tau & v>\tau,\\ 0 & -\tau\le v\le\tau,\\ v+\tau & v<-\tau,\end{cases}$$

where v is the value to be soft-thresholded and τ is a threshold. Eq. (4) is then modified as follows,

$$\begin{aligned} a_n^l(\theta) = a_{n,\text{limited}}(\theta) + (1-M_{a_n}(\theta))\cdot \\ \mathcal{F}_1^{-1}\left(\mathcal{S}_\tau\left(L_{\text{HLCC}}(n,m)\cdot\mathcal{F}_1 a_n^{l-1}(\theta)\right)\right).\end{aligned} \tag{5}$$

Here $\mathcal{S}_\tau$ is applied to the imaginary and real Fourier coefficients element-wise. The proposed P-G algorithm using HLCC and soft-thresholding is denoted by P-$G_{\text{HLCC,ST}}$.

E. Simulation Experiments

To evaluate the performance of the proposed algorithms, experiments on the standard high-contrast Shepp-Logan phantom (Fig. 1) are conducted. The major and minor semi-axes of the outer ellipse of the phantom are 94.2 mm and 70.6 mm, respectively. The attenuation coefficients are converted to Hounsfield scale between [-1000, 3000] HU. A limited angle sinogram is computed analytically in a parallel-beam trajectory. The total scanned angular range is 160° and the angular step is 0.5°. The number of the equal-space detector pixels is 1537 and the detector element size is 0.2 mm. No noise is simulated but discretization error exists. The images are reconstructed using FBP with the Ram-Lak filter. The size of the reconstructed images is 512×512 with an isotropic pixel size of 0.4 mm.

Fig. 1. The Shepp-Logan phantom, window: [-1000, 3400] HU.

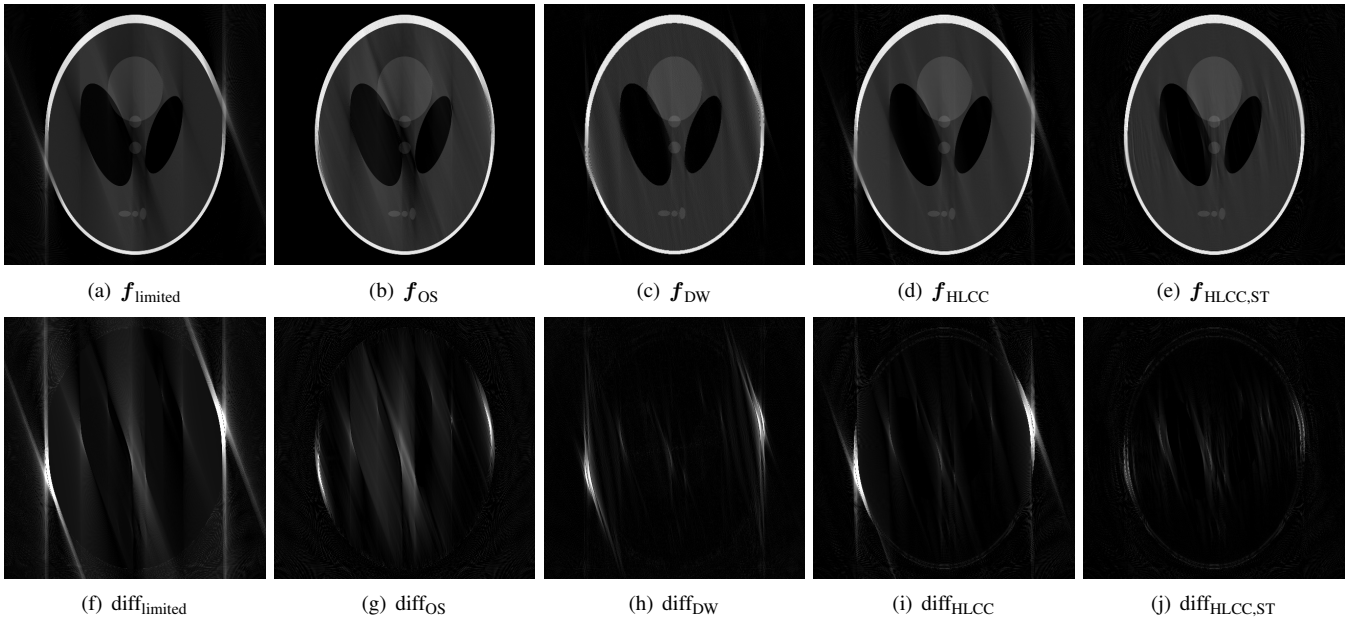
For P-G_{OS}, we assume that the support of the ground truth Shepp-Logan phantom is exactly known. For P-G_{DW}, we choose $r_p = 94$ mm, which is the top point of the phantom and can be accurately obtained from the limited angle reconstruction result using FBP. For P-G_{HLCC} and P-$G_{\text{HLCC,ST}}$, the number of used orders n_r is set to 2414. Empirically, the threshold τ is set to $\tau = 0.5\cdot(1-n/2500)$ for n-th moment curve. For all three algorithms, 1000 iterations are performed.

III. Results And Discussion

The images reconstructed from different algorithms and their absolute difference w. r. t. the image reconstructed from the full data are displayed in Fig. 2. The image reconstructed from the limited angle sinogram using FBP, denoted by $\boldsymbol{f}_{\text{limited}}$, is shown in Fig. 2(a). It suffers from streak artifacts. Especially, the outer boundary is severely distorted on the left and right sides. The image reconstructed from P-G_{OS}, denoted by $\boldsymbol{f}_{\text{OS}}$, is shown in Fig. 2(b). The streaks outside the boundary are totally removed, since strong prior knowledge of the ground truth object support is applied. However, streaks inside the support remain. The image reconstructed from P-G_{DW}, denoted by $\boldsymbol{f}_{\text{DW}}$, is shown in Fig. 2(c). Most streaks outside the boundary are reduced, although the boundary is still a little distorted. Its absolute difference image diff_{DW} displayed in Fig. 2(h) demonstrates that the streaks inside the boundary are also reduced. Figs. 2(d) and (e) are images reconstructed from P-G_{HLCC} and P-$G_{\text{HLCC,ST}}$, denoted by $\boldsymbol{f}_{\text{HLCC}}$ and $\boldsymbol{f}_{\text{HLCC,ST}}$ respectively. Most streaks remain in $\boldsymbol{f}_{\text{HLCC}}$ (Fig. 2(d)) since only low order moment curves are restored. On the contrary, most streaks at the boundary in $\boldsymbol{f}_{\text{HLCC,ST}}$ (Fig. 2(e)) are reduced and thus the boundary is reconstructed very well. Streaks inside the boundary of $\boldsymbol{f}_{\text{HLCC,ST}}$ are also reduced, although still some small streaks remain. The root-mean-square errors (RMSEs) of images reconstructed from different algorithm indicate that P-$G_{\text{HLCC,ST}}$ has the best performance with the lowest RMSE of 75 HU.

IV. Conclusion

In this paper, we propose two new P-G algorithms based on consistency conditions, P-G_{DW} and P-G_{HLCC}/P-$G_{\text{HLCC,ST}}$. P-G_{DW} uses the band-limitation property of sinograms, i. e., that the 2-D Fourier transform of a sinogram has a double-wedge zero region. P-$G_{\text{HLCC,ST}}$ uses the band-limitation property of moment curves according to HLCC. The conventional



(a) $\boldsymbol{f}_{\text{limited}}$ (b) $\boldsymbol{f}_{\text{OS}}$ (c) $\boldsymbol{f}_{\text{DW}}$ (d) $\boldsymbol{f}_{\text{HLCC}}$ (e) $\boldsymbol{f}_{\text{HLCC,ST}}$

(f) $\text{diff}_{\text{limited}}$ (g) diff_{OS} (h) diff_{DW} (i) $\text{diff}_{\text{HLCC}}$ (j) $\text{diff}_{\text{HLCC,ST}}$

Fig. 2. Reconstructions of the Shepp-Logan phantom using different algorithms and their absolute difference w. r. t. the full data reconstruction. The root-mean square errors for $\boldsymbol{f}_{\text{limited}}$, $\boldsymbol{f}_{\text{OS}}$, $\boldsymbol{f}_{\text{DW}}$, $\boldsymbol{f}_{\text{HLCC}}$, and $\boldsymbol{f}_{\text{HLCC,ST}}$ are 302 HU, 172 HU, 150 HU, 214 HU, and 75 HU, respectively. Window: [-1000, 3400] HU and [-1000, 1000] HU for the top and bottom images, respectively.

P-G algorithm P-G_{OS} requires strong *a priori* knowledge, the exact object support. In contrast, P-G_{DW} only needs the distance of the farthest point, which can be accurately estimated from a limited angle reconstruction. It reduces small streaks inside the boundary better than P-G_{OS}. However, it is unable to reconstruct the boundary well. P-G_{HLCC} only is not sufficient to reduce streaks due to the ill-posedness of high order moment curve extrapolation/interpolation. P-$G_{\text{HLCC,ST}}$ takes the advantage of the sparsity of the Fourier coefficients of the moment curves. It performs the best on streak reduction among the proposed algorithms.

Disclaimer: The concepts and information presented in this paper are based on research and are not commercially available.

References

[1] F. A. Grünbaum, "A study of fourier space methods for limited angle image reconstruction," *Numer Funct Anal Optim*, vol. 2, no. 1, pp. 31–42, 1980.

[2] M. E. Davison, "The ill-conditioned nature of the limited angle tomography problem," *SIAM J Appl Math*, vol. 43, no. 2, pp. 428–448, 1983.

[3] E. T. Quinto, "An introduction to x-ray tomography and radon transforms," vol. 63, p. 1, 2006.

[4] E. Sidky, C. Kao, and X. Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT," *J Xray Sci Technol*, vol. 14, pp. 119–139, 2006.

[5] Z. Chen, X. Jin, L. Li, and G. Wang, "A limited-angle CT reconstruction method based on anisotropic TV minimization," *Phys Med Biol*, vol. 58, no. 7, pp. 2119–2141, 2013.

[6] Y. Huang, O. Taubmann, X. Huang, V. Haase, G. Lauritsch, and A. Maier, "A new weighted anisotropic total variation algorithm for limited angle tomography," *Procs IEEE ISBI*, pp. 585–588, 2016.

[7] ——, "A new scale space total variation algorithm for limited angle tomography," *Procs CT-Meeting*, pp. 149–152, 2016.

[8] T. Würfl, F. C. Ghesu, V. Christlein, and A. Maier, "Deep learning computed tomography," *Procs MICCAI*, vol. 3, pp. 432–440, 2016.

[9] K. Hammernik, T. Würfl, T. Pock, and A. Maier, "A deep learning architecture for limited-angle computed tomography reconstruction," *Procs BVM*, pp. 92–97, 2017.

[10] J. Gu and J. C. Ye, "Multi-scale wavelet domain residual learning for limited-angle CT reconstruction," *Procs Fully3D*, pp. 443–447, 2017.

[11] R. Gerchberg, "Super-resolution through error energy reduction," *J Mod Opt*, vol. 21, no. 9, pp. 709–720, 1974.

[12] A. Papoulis, "A new algorithm in spectral analysis and band-limited extrapolation," *IEEE Trans Circuits Syst*, vol. 22, no. 9, pp. 735–742, 1975.

[13] G.-r. Qu, Y.-s. Lan, and M. Jiang, "An iterative algorithm for angle-limited three-dimensional image reconstruction," *Acta Math Appl Sin*, vol. 24, no. 1, pp. 157–166, 2008.

[14] G.-r. Qu and M. Jiang, "Landweber iterative methods for angle-limited image reconstruction," *Acta Math Appl Sin*, vol. 25, no. 2, pp. 327–334, 2009.

[15] M. Defrise and C. De Mol, "A regularized iterative algorithm for limited-angle inverse radon transform," *Opt Acta: Int J Opt*, vol. 30, no. 4, pp. 403–408, 1983.

[16] P. R. Edholm, R. M. Lewitt, and B. Lindholm, "Novel properties of the fourier decomposition of the sinogram," *Phys Eng Comput Multidimen Imag Process*, pp. 8–18, 1986.

[17] D. Ludwig, "The Radon transform on Euclidean space," *Comm Pure Appl Math*, vol. 19, no. 1, pp. 49–81, 1966.

[18] S. Helgason, *The Radon transform*. Boston, MA: Birkhauser, 1980.

[19] M. Pohlmann, M. Berger, A. Maier, J. Hornegger, and R. Fahrig, "Estimation of missing fan-beam projections using frequency consistency conditions," *Procs CT Meeting*, pp. 203–207, 2014.

[20] M. Berger, Y. Xia, W. Aichinger, K. Mentl, M. Unberath, A. Aichert, C. Riess, J. Hornegger, R. Fahrig, and A. Maier, "Motion compensation for cone-beam CT using Fourier consistency conditions," *Phys Med Biol*, vol. 62, no. 17, pp. 7181–7215, 2017.

[21] Y. Huang, X. Huang, O. Taubmann, Y. Xia, V. Haase, J. Hornegger, G. Lauritsch, and A. Maier, "Restoration of missing data in limited angle tomography based on Helgason-Ludwig consistency conditions," *Biomed Phys & Eng Express*, vol. 3, no. 3, p. 035015, 2017.

[22] Y. Xia, M. Berger, S. Bauer, S. Hu, A. Aichert, and A. Maier, "An improved extrapolation scheme for truncated CT data using 2D Fourier-based Helgason-Ludwig consistency conditions," *Int J Biomed Imaging*, vol. 2017, 2017.

[23] J. Karp, G. Muehllehner, and R. Lewitt, "Constrained fourier space method for compensation of missing data in emission computed tomography," *IEEE Trans Med Imaging*, vol. 7, no. 1, pp. 21–25, 1988.

Empirical Scatter Correction using the Epipolar Consistency Condition

Mathis Hoffmann, Tobias Würfl, Nicole Maaß, Frank Dennerlein, Anré Aichert and Andreas K. Maier

***Abstract*—Scatter affects every computed tomography (CT) image. Calibration-free software scatter reduction methods have not been used extensively in practice. Recently, consistency conditions have been applied successfully to other artifact reduction problems in CT imaging. We propose a scatter reduction method, that uses an epipolar consistency condition (ECC) to estimate parameters of an additive scatter model. We evaluate our approach by comparing it with an image-based empirical scatter correction method (ESC) that uses the same scatter model. We show that it performs equally well on simulated data. Further, ECC outperforms ESC regarding the computational load for the determination of the parameter models, because ECC is formulated in projection domain such that no image reconstruction is necessary. While some restrictions might apply for the stability of ECC on measured data, no prior information needs to be formulated regarding the reconstructed image, like it is required with ESC.**

I. Introduction

Scatter is caused by two different physical effects [1]. On the one hand, photons are scattered (change their direction) at particles that are much smaller than the wavelength of the radiation. This is known as Rayleigh scattering. On the other hand, the Compton effect describes incoherent scattering, where a scattered photon looses energy, increasing its wavelength. These effects contradict the assumption of reconstruction algorithms that the measured radiation solely consists of primary radiation. Therefore, the image quality is reduced. To weaken the effect, hardware- and software-based scatter reduction methods are applied [2].

Reference-less software methods aim to estimate the scatter radiation using a suitable scatter model and a cost function to estimate its parameters. Monte Carlo simulations are the well known gold standard to estimate the scatter component precisely. However, the simulation is computationally demanding. The situation can be improved by performing a very coarse Monte Carlo simulation and using the result to fit a scatter model [3]. Alternatively, a scatter model can be fitted by optimization of a cost function in image domain, for example by the minimization of the total variation (TV). Meyer et. al [4] show that this is possible without successive reconstruction steps by exploiting the linearity of the Radon transform.

M. Hoffmann, T. Würfl and A. K. Maier are with the Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany.
N. Maaß and F. Dennerlein are with Siemens Healthcare GmbH, Erlangen, Germany.

Recently, a new consistency condition based on redundant plane integrals was presented [5], [6] and has been successfully applied for geometry calibration [6], [7], motion compensation [8] and beam-hardening reduction [9].

We derive an algorithm using the epipolar consistency condition to estimate the parameters of a scatter model. The model we use was introduced by Ohnesorge et. al [10] and modified by Meyer et. al [4] to take the form of a linear combination. We exploit the results of Würfl et. al [9] to derive a closed-form solution that can be computed efficiently.

II. Methodology

In section II-A, we detail the model that we use to estimate the scatter component. In section II-B, we derive a closed-form solution to the estimation problem that uses the ECC as a cost function. In section II-C, we point at some insights that we gained from the model.

A. Model

Throughout this paper, we assume that scatter can be reduced using the additive model in line integral domain, that has been used by Meyer et. al [4]. The scatter-reduced raw-data $\boldsymbol{p}$ is given by

$$\boldsymbol{p} = \boldsymbol{q} - \sum_{n=1}^{N} w_n \boldsymbol{s}_n, \tag{1}$$

where $\boldsymbol{q}$ denotes the measured raw data. We need to estimate the raw-data scatter components $\boldsymbol{s}_n$ along with the weights w_n.

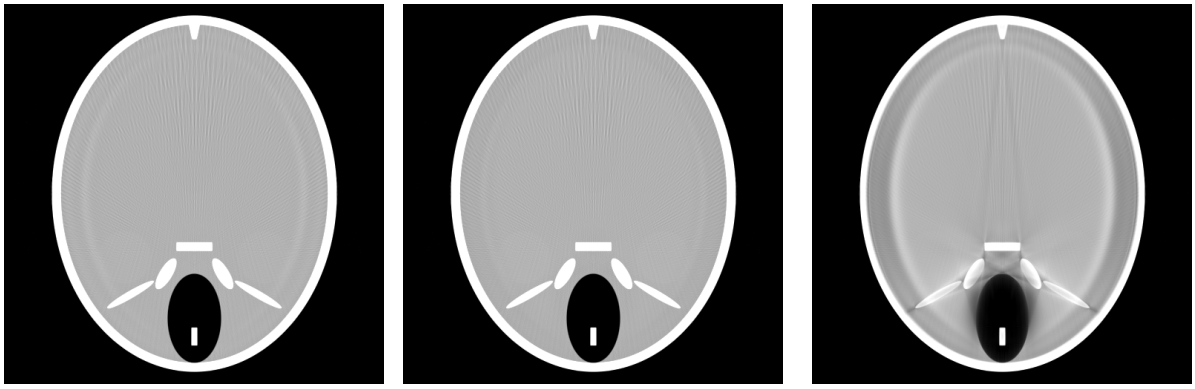
As the source for scatter resides in intesity domain, we need to estimate the scatter intensity $I_{\boldsymbol{s}}$ to obtain the scatter components $\boldsymbol{s}_n$. The scatter intensity consists of an intensity model $I_{\boldsymbol{f}}$ that is convolved with a model for the scatter point spread function. The intensity model is given by

$$I_{\boldsymbol{f}} = \alpha \boldsymbol{p} I_{\boldsymbol{p}}, \tag{2}$$

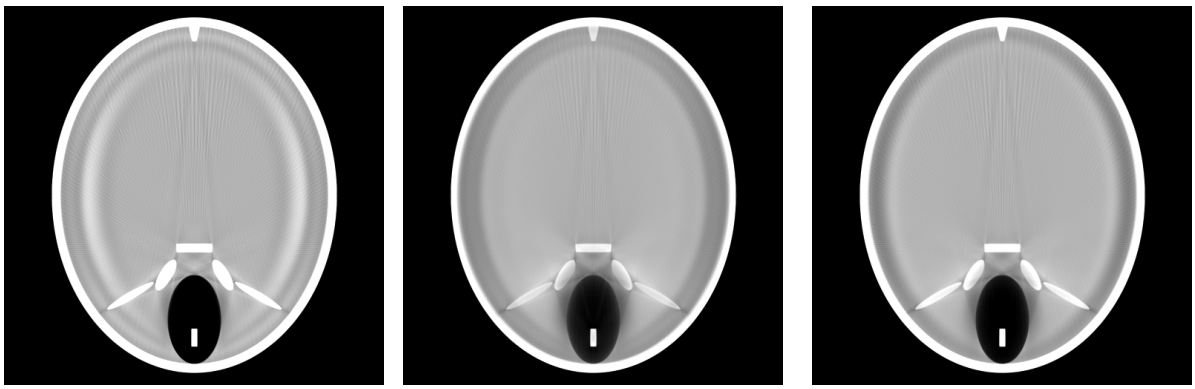
where $I_{\boldsymbol{p}} = e^{-\boldsymbol{p}}$ and α is a scale parameter that depends on the scatter cross section. From $I_{\boldsymbol{f}}$, the scatter component in intensity domain can be calculated as

$$I_{\boldsymbol{s}} = I_{\boldsymbol{f}} * \left[\exp\left(\frac{x+\beta}{\gamma}\right) + \exp\left(\frac{x-\beta}{\gamma}\right) \right], \tag{3}$$

where β and γ determine the shape of the scatter point spread function and x denotes the spatial location on the detector.



(a) ECC $[I_{\boldsymbol{p}}]$ (b) TV $[I_{\boldsymbol{p}}]$ (c) TV without mask $[I_{\boldsymbol{p}}]$



(d) ECC $[I_{\boldsymbol{q}}]$ (e) TV $[I_{\boldsymbol{q}}]$ (f) $I_{\boldsymbol{q}}$ - Uncorrected

Fig. 1: Results for the inverse crime scenario using the the proposed ECC-based method and the TV-based approach. The notation $[I_{\boldsymbol{p}}]$ denotes that the ideal projection data has been used to generate the scatter components (see Eq. (2)), whereas $[I_{\boldsymbol{q}}]$ denotes that the projection with simulated scatter has been used instead.

To enable a linear estimation using the ECC, we obtain the scatter component in line integral domain by

$$\begin{aligned} \boldsymbol{s} &= -\ln\left(I_{\boldsymbol{s}} + I_{\boldsymbol{p}}\right) + \ln\left(I_{\boldsymbol{p}}\right) \\ &= -\ln\left(\frac{I_{\boldsymbol{s}}}{I_{\boldsymbol{p}}} + 1\right). \end{aligned} \tag{4}$$

Eq. (1) already indicates that we use N of those scatter components to find the scatter-free image $\boldsymbol{p}$. We denote a specific scatter component that is obtained from parameters $(\alpha_n,\, \beta_n,\, \gamma_n)$ as $\boldsymbol{s}_n$. Note that this model assumes that the ideal, scatter-free data $I_{\boldsymbol{p}}$ is known to create the scatter components. This is generally not the case. A common solution is to apply any algorithm relying on this model in a fixed-point iteration scheme. Throughout the following, we only consider a single iteration of this algorithm.

B. Estimation

Given two X-ray projections, any plane which contains both source positions intersects the detectors in so-called corresponding epipolar lines. Epipolar consistency states that the derivative orthogonal to their common plane is identical in either image:

$$\frac{\partial}{\partial \boldsymbol{t}} \rho_{\boldsymbol{p}_0}\left(\boldsymbol{l}_0\right) = \frac{\partial}{\partial \boldsymbol{t}} \rho_{\boldsymbol{p}_1}\left(\boldsymbol{l}_1\right). \tag{5}$$

Here, $(\rho_{\boldsymbol{p}_0},\, \rho_{\boldsymbol{p}_1})$ denotes the Radon transform of a pair of projection images $(\boldsymbol{p}_0,\, \boldsymbol{p}_1)$, $(\boldsymbol{l}_0,\, \boldsymbol{l}_1)$ denote intersection lines of an epipolar plane with the corresponding detector planes and $\frac{\partial}{\partial \boldsymbol{t}}$ denotes the derivative of the Radon transformed projections perpendicular to the lines.

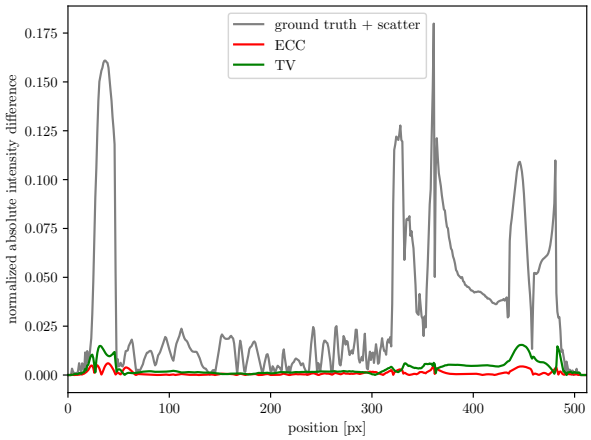
We combine the additive scatter model (Eq. (1)) with Eq. (5) to receive

$$\frac{\partial}{\partial \boldsymbol{t}} \rho_{\left(\boldsymbol{q}_0 - \sum_{n=1}^{N} w_n \boldsymbol{s}_{n,0}\right)}\left(\boldsymbol{l}_0\right) = \frac{\partial}{\partial \boldsymbol{t}} \rho_{\left(\boldsymbol{q}_1 - \sum_{n=1}^{N} w_n \boldsymbol{s}_{n,1}\right)}\left(\boldsymbol{l}_1\right), \tag{6}$$

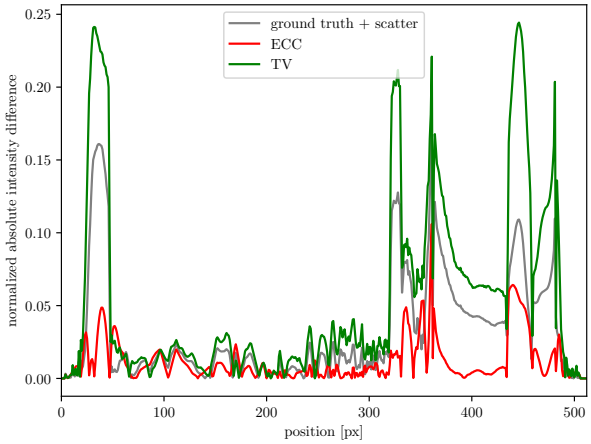
where $(\boldsymbol{s}_{n,0},\, \boldsymbol{s}_{n,1})$ are the n'th scatter components of the projection image pair $(\boldsymbol{q}_0,\, \boldsymbol{q}_1)$. Because the Radon transform and the derivative operator are linear, we reuse the idea of epipolar consistency guided beam hardening reduction [9] and rewrite Eq. (6) to

$$\boldsymbol{q}_0'\left(\boldsymbol{l}_0\right) - \sum_{n=1}^{N} w_n \boldsymbol{s}_{n,0}'\left(\boldsymbol{l}_0\right) = \boldsymbol{q}_1'\left(\boldsymbol{l}_1\right) - \sum_{n=1}^{N} w_n \boldsymbol{s}_{n,1}'\left(\boldsymbol{l}_1\right), \tag{7}$$

where we substituted $\boldsymbol{q}_i' = \frac{\partial}{\partial \boldsymbol{t}} \rho_{\boldsymbol{q}_i}$ and $\boldsymbol{s}_{n,i}' = \frac{\partial}{\partial \boldsymbol{t}} \rho_{\boldsymbol{s}_{n,i}}$, $i \in [0, 1]$. Unfortunately, the scatter components $\boldsymbol{s}_n$ are not linear in the parameters (see Eqs. (2), (3) and (4)). Therefore, we follow the idea from empirical scatter correction [4], generate N different estimates of the scatter components and



(a) $[I_{\boldsymbol{p}}]$



(b) $[I_{\boldsymbol{q}}]$

Fig. 2: Normalized absolute intensity difference of line profiles to the ground truth.

assume that they behave as a basis for the space of scatter components. In section II-C, we point out, that this approach has theoretical drawbacks. However, our experiments (section III) reveal that it performs well in practice.

In order to estimate the coefficients w_n, we optimize for consistency, such that the scatter components remain constant:

$$\min \left[\left(\sum_{n=1}^{N} w_n \boldsymbol{s}'_{n,0}\left(\boldsymbol{l}_0\right) - \boldsymbol{s}'_{n,1}\left(\boldsymbol{l}_1\right) \right) - \boldsymbol{q}'_0\left(\boldsymbol{l}_0\right) - \boldsymbol{q}'_1\left(\boldsymbol{l}_1\right) \right]^2 . \tag{8}$$

So far, we stated the problem for $(\boldsymbol{l}_0, \boldsymbol{l}_1)$, which corresponds to one epipolar plane and one pair of projection images. We want to solve this problem for K planes $\boldsymbol{l}^k_{p,i}$ across P pairs of projection images. Therefore, we have $M = PK$ measurements, producing an overdetermined system of linear equations:

$$\hat{\boldsymbol{w}} = \arg\min_{\boldsymbol{w}} \|\boldsymbol{A}\boldsymbol{w} - \boldsymbol{b}\|_2^2 \tag{9}$$

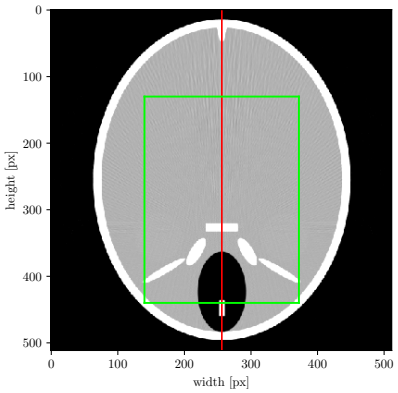


Fig. 3: Region (green) that has been used to estimate SSIM and PSNR (see Fig. 4) and line (red) along which the profiles have been calculated.

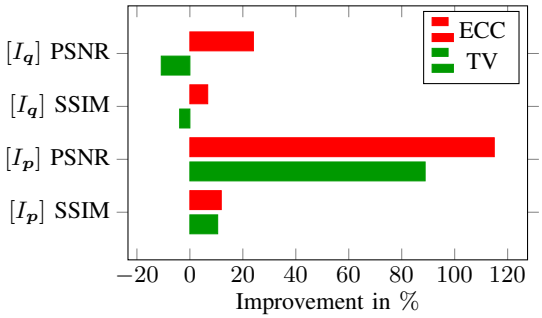


Fig. 4: Performance of our (ECC) and the TV-based method in terms of PSNR and SSIM. 0% improvement means that there is no improvement over the volume with simulated scatter. The measures have been computed in 2D within the region depicted in Fig. 3.

where

$$\boldsymbol{A} = \begin{bmatrix} a_{1,1,1} & \dots & a_{1,1,N} \\ \vdots & \ddots & \vdots \\ a_{1,K,1} & \dots & a_{1,K,N} \\ a_{2,1,1} & \dots & a_{2,1,N} \\ \vdots & \ddots & \vdots \\ a_{P,K,1} & \dots & a_{P,K,N} \end{bmatrix}$$

with

$$a_{p,k,n} = \boldsymbol{s}'_{n,p,0}\left(\boldsymbol{l}^k_{p,0}\right) - \boldsymbol{s}'_{n,p,1}\left(\boldsymbol{l}^k_{p,1}\right)$$

and

$$\boldsymbol{b} = \left[b_{1,1}, \dots, b_{1,K}, b_{2,1}, \dots, b_{P,K}\right]^\top$$

with

$$b_{p,k} = \boldsymbol{q}'_{p,0}\left(\boldsymbol{l}^k_{p,0}\right) - \boldsymbol{q}'_{p,1}\left(\boldsymbol{l}^k_{p,1}\right).$$

Note that the columns of $\boldsymbol{A}$ are linearly independent. Therefore, the solution to Eq. (9) is given by

$$\hat{\boldsymbol{w}} = (\boldsymbol{A}^\top \boldsymbol{A})^{-1} \boldsymbol{A}^\top \boldsymbol{b} = \boldsymbol{A}^+ \boldsymbol{b}. \tag{10}$$

C. Discussion

Typically, such linear combinations of basis images are applied in the domain where the artifact arises. However, this model using scatter basis functions in line integral domain is not equal to a linear combination of scatter basis functions in intensity domain (see Eq. (1)). To investigate this, we transform the linear combination $\sum_n w_n \boldsymbol{s}_n$ to intensity domain:

$$\sum_n w_n \ln\left(I_{\boldsymbol{s},n}\right) = \sum_n \ln\left(I_{\boldsymbol{s},n}^{w_n}\right) = \ln\left(\prod_n I_{\boldsymbol{s},n}^{w_n}\right). \quad (11)$$

This result shows that the linear combination in raw data domain does not correspond to a linear combination in intensity domain.

III. Experiments

We present two simulation experiments using the Forbild head phantom. We show the performance of the proposed algorithm in two inverse crime scenarios and compare the results to the method proposed by Meyer et. al that minimizes the total variation in volume domain, instead of the epipolar consistency in projection domain.

A. Setup

For the experiments, we define a feasible parameter range $\beta \in \left[\beta_l = 1\mathrm{e}^{-2},\, \beta_u = 1.5\mathrm{e}^{-2}\right]$, $\gamma \in \left[\gamma_l = \mathrm{e}^{-3},\, \gamma_u = 1\mathrm{e}^{-2}\right]$. The parameter α remains fixed at $\alpha = 4\mathrm{e}^{-10}$. We use $N = 4$ scatter basis images such that $(\beta_n,\, \gamma_n) \in (\beta_l,\, \beta_u) \times (\gamma_l,\, \gamma_u)$. The simulated scatter image is generated with $\beta_{\text{sim}} = 1.2\mathrm{e}^{-2}$, $\gamma_{\text{sim}} = 8\mathrm{e}^{-3}$. For the first simulation experiment, we generate the scatter components using the ideal projection intensity $I_{\boldsymbol{p}}$ (see Eq. (2)). The results of this experiment are annotated with $[I_{\boldsymbol{p}}]$. For the second simulation experiment, we use the projection intensity with simulated scatter $I_{\boldsymbol{q}}$ for the scatter components and annotate the results with $[I_{\boldsymbol{q}}]$.

B. Results

The results of our experiments are summarized in Fig. 1. We find, that both methods yield visually comparable results when $I_{\boldsymbol{p}}$ is used to generate the scatter components (Figs. 1a and 1b). However, both methods produce worse results, when $I_{\boldsymbol{q}}$ is used instead (Figs. 1c and 1d).

Fig. 2 depicts, how a line profile through the scatter-reduced volumes differs from the line profile through the ground-truth volume. For the case where the scatter components are based on $I_{\boldsymbol{p}}$, the results for both methods only differ slightly. Notably, there is an advantage for the ECC-based parameter estimation in the frontal sinus area within the phantom. This is due to a mandatory masking of soft- and hard-tissue regions in the TV-based approach (see also Fig. 1c) that eliminates contributions from regions with low attenuation to the loss. For the case where the scatter components are based on $I_{\boldsymbol{q}}$, worse results can be seen in both methods.

Finally, we compare the structure similarity (SSIM) and peak signal noise ratio (PSNR) of the scatter-reduced slices. The results are shown in Fig. 4. Regarding these measures, the ECC-based parameter estimation is superior to the TV-based approach in all cases.

IV. Conclusion

We have presented a novel reference-free scatter reduction algorithm and showed that it improves over a similar algorithm in terms of achievable image quality and computational efficiency. Overall, the PSNR is about 35% better, whereas the SSIM is also slightly better, compared to the reference algorithm. The estimation of the weights with our method happens to be 52 times faster using an unoptimized numerical Python implementation.

A key advantage of using a consistency condition is, that it does not impose any assumption about the object. We reuse the scatter reduction model from Meyer et. al to obtain a computationally efficient formulation of the optimization problem on intermediate functions in line integral domain with a closed-form solution. However, we show that this mathematical model has disadvantageous physical properties. Still we found the algorithm works well in simulation experiments, even if we restrict ourselves to a single iteration of the algorithm. In future work we want to evaluate the algorithm extensively on measured data. Additionally we want to explore different scatter reduction models better modeling the physical properties while preserving computational efficiency.

Acknowledgment

A. Aichert is supported by the German Research Foundation; DFG MA 4898/3-1.

Disclaimer

The concepts and information presented in this paper are based on research and are not commercially available.

References

[1] P. M. Joseph and R. D. Spital, "The effects of scatter in x-ray computed tomography," *Med. Phys.*, vol. 9, no. 4, pp. 464–472, 1982.

[2] E.-P. Rührnschopf and K. Klingenbeck, "A general framework and review of scatter correction methods in x-ray cone-beam computerized tomography. part 1: Scatter compensation approaches," *Med. Phys.*, vol. 38, no. 7, pp. 4296–4311, 2011.

[3] M. Baer and M. Kachelrieß, "Hybrid scatter correction for ct imaging," *Phys. in Med. and Bio.*, vol. 57, no. 21, p. 6849, 2012.

[4] E. Meyer, C. Maaß, M. Baer, R. Raupach, B. Schmidt, and M. Kachelrieß, "Empirical scatter correction (esc): A new ct scatter correction method and its application to metal artifact reduction," in *Nuclear Science Symposium Conference Record (NSS/MIC), 2010 IEEE*. IEEE, 2010, pp. 2036–2041.

[5] C. Debbeler, N. Maass, M. Elter, F. Dennerlein, and T. M. Buzug, "A new CT rawdata redundancy measure applied to automated misalignment correction," in *12th Int. Meet. Fully 3D Reconstr. Radiol. Nucl. Med*, 2013, pp. 264–267.

[6] A. Aichert, M. Berger, J. Wang, N. Maass, A. Doerfler, J. Hornegger, and A. K. Maier, "Epipolar consistency in transmission imaging," *IEEE Trans. Med. Imaging*, vol. 34, no. 11, pp. 2205–2219, 2015.

[7] N. Maass, F. Dennerlein, A. Aichert, and A. Maier, "Geometrical jitter correction in computed tomography," in *3rd Int. Conf. on image formation in X-ray computed tomography, edited by Frederic Noo, Salt Lake City*, 2014, pp. 338–342.

[8] M. Unberath, A. Aichert, S. Achenbach, and A. Maier, "Consistency-based respiratory motion estimation in rotational angiography," *Med. Phys.*, vol. 44, no. 9, 2017.
[9] T. Würfl, N. Maass, F. Dennerlein, X. Huang, and A. Maier, "Epipolar consistency guided beam hardening reduction - ECC^2," in *14th Int. Meet. Fully 3D Reconstr. Radiol. Nucl. Med*, 2017, pp. 181–185.
[10] B. Ohnesorge, T. Flohr, and K. Klingenbeck-Regn, "Efficient object scatter correction algorithm for third and fourth generation ct scanners," *European radiology*, vol. 9, no. 3, pp. 563–569, 1999.

Stereo Rectification for X-ray Data Consistency Conditions

André Aichert, Jérôme Lesaint, Tobias Würfl, Rolf Clackdoyle, Laurent Desbat and Andreas K. Maier

***Abstract*—Data consistency conditions (DCC) allow for the compensation of measurement errors and geometric inaccuracies of X-ray computed tomography (CT). Epipolar Consistency (EC), in particular, states that any two X-ray projections contain redundant information on epipolar lines. Recently, a fan-beam DCC was applied to cone-beam projections by re-projection and weighted integration in a virtual detector. It is related to EC, since the redundant fans lie in a common epipolar plane. This paper observes that the re-projection to a common virtual detector is merely an instance of stereo rectfication known from computer vision and shows how rectifying homographies can be used to generally express these DCCs. We derive an algorithm which allows for the evaluation directly in the original detector planes, without the need for memory- and time-intensive warping of all image pairs. Since the presented algorithm does not require a derivative, it may prove beneficial in applications, which aim at the estimation of absolute intensities, rather than geometry or motion. We have validated the equations in this paper numerically with phantom data.**

I. Introduction

Epipolar Consistency (EC) is a type of data consistency condition in X-ray computed tomography, which exploits redundancies along epipolar lines in two X-ray projection images [1]. They have been used for the correction of various artifacts in computed tomography. EC states that the orthogonal derivative of integrals over two corresponding epipolar lines are identical. An alternative formulation has been presented, which uses the ramp-filter instead of a derivative [3]. Recently, a similar set of conditions has been applied to calibration correction in rotational CT [4]. The authors show that by means of a reprojection to a virtual detector plane, pair-wise fan-beam consistency conditions (FBCC) can be applied [5]. The main difference between these conditions is (1) integration along epipolar lines followed the derivative or ramp filter orthogonal to epipolar lines versus (2) an integral along epipolar lines weighted by the distance of each pixel to the baseline. The derivative emphasizes high-frequency information, which improves alignment of edges in motion correction application. However, artifact correction methods which correct for intensity changes (e.g. beam-hardening, scatter or extinction) may benefit from the absolute intensity information still available through the latter FBCC. A major drawback of the algorithm is that re-projection to a common virtual detector is different for any two projections. Consequently, n projection images would produce $n \cdot (n-1)$ pairs of rectified 2D images.

André Aichert, Tobias Würfl and Andreas K. Maier are with the Pattern Recognition Lab, Friedrich-Alexander Universität Erlangen-Nürnberg, Germany.

Jérôme Lesaint, Rolf Clackdoyle and Laurent Desbat are with the TIMC-IMAG laboratory, CNRS UMR 5525 and Université Grenoble Alpes, Grenoble, France

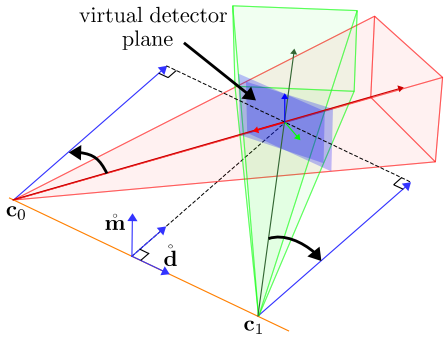


Figure 1. The stereo baseline connecting $\mathbf{c}_0$ and $\mathbf{c}_1$ (orange) and its relation to the virtual detector plane. The direction of the line $\mathring{\mathbf{d}}$ points from $\mathbf{c}_0$ to $\mathbf{c}_1$. The Plücker line moment $\mathring{\mathbf{m}}$ is the normal of the plane containing the origin and both source positions (i.e. the plane of rotation in case of circular trajectories). We choose the orthogonal vector $\mathring{\mathbf{d}} \times \mathring{\mathbf{m}}$ to be the normal to the virtual detector plane (blue).

This paper observes that re-projection of two central projections to a common plane is an instance of the stereo rectification problem in computer vision [2]. Stereo rectification is the process by which two rectifying homographies are applied to a pair of images to produce identical corresponding epipolar lines which are parallel to the horizontal image axis. It is the well-understood first step for the computation of a depth-map from two visible light images.

By application of standard computer vision techniques, Section II extends the FBCC algorithm by Lesaint et al. [4] to general projection matrices and thus non-circular trajectories. In Section III we show how to evaluate the presented DCCs without explicit projection of all pairs of images, thus rendering practical application feasible. We validated this work on real data of a pumpkin phantom and provide the full source code of our GPU implementation https://github.com/aaichert/EpipolarConsistency.

II. Rectification by Projection onto a Plane

A. Rectification

Consider two X-ray source-detector geometries defined by the projection matrices $\mathbf{P}_0$, $\mathbf{P}_1 \in \mathbb{R}^{3\times4}$. This section discusses how to find two rectifying homographies $\mathbf{H}_0$, $\mathbf{H}_1 \in \mathbb{R}^{3\times3}$, which transform projection images such that the respective

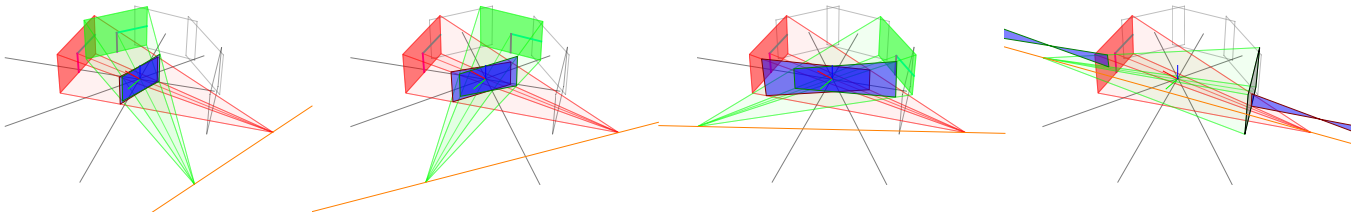


Figure 2. Rectified projection data is subject to increasingly strong distortion as the primary angle difference between the two projections increases. Example of a typical circular CT trajectory of a Siemens Artis Zeego: (1) little distortion up to $\approx 45°$, (2) considerable distortion at $\approx 90°$, (3) method is beginning to break down after $\approx 135°$, and (4) rectification not sensible for opposing views.

image planes of $\mathbf{H}_0\mathbf{P}_0$ and $\mathbf{H}_1\mathbf{P}_1$ are parallel to the baseline, while corresponding epipolar lines are equal and parallel to the u-axis. We will do this by re-projecting the image data to a common virtual detector plane. The source positions in homogeneous coordinates $\mathbf{c}_0 \cong (C_0^0,\, C_0^1,\, C_0^2,\, 1)^\top \in \mathrm{null}(\mathbf{P}_0)$ and $\mathbf{c}_1$, accordingly, define the stereo baseline, where $\cong$ denotes equality up to scale, define the stereo baseline. Its direction $\mathring{\mathbf{d}} \in \mathbb{S}^2$ is the unit vector

$$\mathring{\mathbf{d}} \cong \begin{pmatrix} C_2^0 - C_0^0 \\ C_1^1 - C_0^1 \\ C_1^2 - C_0^2 \end{pmatrix}, \tag{1}$$

where $\cong$ denotes equality up to positive scalar multiplication. Its moment shall be the unit vector $\mathring{\mathbf{m}} \in \mathbb{S}^2$

$$\mathring{\mathbf{m}} \cong \begin{pmatrix} C_2^0 \\ C_1^1 \\ C_1^2 \end{pmatrix} \times \begin{pmatrix} C_0^0 \\ C_0^1 \\ C_0^2 \end{pmatrix}, \tag{2}$$

which is the normal to the plane which contains both the baseline and the origin.

B. Rectification by Re-Projection

Our goal is to project both images to a virtual detector with an image u-axis pointing in the direction of the stereo baseline. Assuming that the iso-center is approximately in the coordinate origin, a good choice of virtual detector plane is

$$\mathbf{e} \cong \begin{pmatrix} \mathring{\mathbf{d}} \times \mathring{\mathbf{m}} \\ 0 \end{pmatrix} \in \mathbb{P}^3. \tag{3}$$

Since the u-axis must point in direction of the baseline, we establish a 2D coordinate frame for points in the plane $\mathbf{e}$ with 3D coordinate axis directions $\mathring{\mathbf{d}}$ and $\mathring{\mathbf{m}}$ according to

$$\mathbf{S} = \begin{pmatrix} \mathring{\mathbf{d}}^\top & 0 \\ \mathring{\mathbf{m}}^\top & 0 \\ \mathbf{0} & m \end{pmatrix} \in \mathbb{R}^{3\times 4}. \tag{4}$$

with arbitrary pixel size m in mm. Given an original image with projection matrix $\mathbf{P}$ and source position $\mathbf{c} \in \mathrm{null}(\mathbf{P})$, let $\mathbf{T}_\mathbf{c}^\mathbf{e} \in \mathbb{R}^{4\times 4}$ denote the central projection through $\mathbf{c}$ to 3D points on the virtual detector plane $\mathbf{e}$ (compare Section II-C) and $\mathbf{S}$ defined its 2D coordinate system. Then, the homography

$$\mathbf{H} = \mathbf{S}\mathbf{T}_\mathbf{c}^\mathbf{e}\mathbf{P}^+ \in \mathbb{R}^{3\times 3}, \tag{5}$$

directly relates pixels from the original image with pixels in the virtual detector plane $\mathbf{e}$. This situation is visualized in Figure 1.

C. Central Projection onto a Plane

Central projection from a point $\mathbf{c} \in \mathbb{P}^3$ to a plane $\mathbf{e} \in \mathbb{P}^3$ can be written in a single matrix $\mathbf{T}_\mathbf{c}^\mathbf{e} \in \mathbb{R}^{4\times 4}$. It projects a point $\mathbf{x} \in \mathbb{P}^3$ to a point $\mathbf{T}_\mathbf{c}^\mathbf{e}\mathbf{x} \in \mathbb{P}^3$ by intersection of the plane $\mathbf{e}$ with the line $\mathbf{r}$ through the points $\mathbf{c}$ and $\mathbf{x}$. Using the Plücker matrix [2] of the line $\mathbf{c}\mathbf{x}^\top - \mathbf{x}\mathbf{c}^\top$ we express the intersection as

$$\mathbf{T}_\mathbf{c}^\mathbf{e}\mathbf{x} = \mathbf{c}\mathbf{x}^\top\mathbf{e} - \mathbf{x}\mathbf{c}^\top\mathbf{e}. \tag{6}$$

By factoring out $\mathbf{x}$ we obtain the projection

$$\mathbf{T}_\mathbf{c}^\mathbf{e} = \mathbf{I}_4\mathbf{c}^\top\mathbf{e} - \mathbf{c}\mathbf{e}^\top, \tag{7}$$

We can rectify two projections with the homographies

$$\begin{aligned} \mathbf{H}_0 &= \mathbf{S}\mathbf{T}_{\mathbf{c}_0}^\mathbf{e}\mathbf{P}_0^+ \\ \text{and } \mathbf{H}_1 &= \mathbf{S}\mathbf{T}_{\mathbf{c}_1}^\mathbf{e}\mathbf{P}_1^+. \end{aligned} \tag{8}$$

III. The Radon Transform of Perspectively Distorted Functions

A. Line Integrals in Perspectivley Transformed Images

This section computes an integral along a line $\mathbf{l}'$ over an image which has been projectively transformed by a homography $\mathbf{H}$. We will express the integration over a line $\mathbf{l} = \mathbf{H}^\top\mathbf{l}'$ (see in Eq. 12) in the original image, weighted by the Jacobian. Let $\mathbf{l}$ be the line with angle to the u-axis α and signed distance to the origin τ

$$\mathbf{l} \cong \begin{pmatrix} l_0 \\ l_1 \\ l_2 \end{pmatrix} = \begin{pmatrix} \sin(\alpha) \\ -\cos(\alpha) \\ -\tau \end{pmatrix}. \tag{9}$$

We append a zero to the 2D vectors of its normal $\mathbf{n} = (l_0,\, l_1\,, 0)^\top$ and direction

$$\mathbf{d} = \begin{pmatrix} d_0 \\ d_1 \\ 0 \end{pmatrix} = \begin{pmatrix} l_1 \\ -l_0 \\ 0 \end{pmatrix}, \tag{10}$$

with unit length $\|\mathbf{n}\| = \|\mathbf{d}\| = 1$. The closest point to the origin is given by the homogeneous coordinates

$$\mathbf{o} \cong \begin{pmatrix} o_0 \\ o_1 \\ 1 \end{pmatrix} = \begin{pmatrix} -l_2 l_0 \\ -l_2 l_1 \\ 1 \end{pmatrix}. \tag{11}$$

Lines transform covariantly, therefore, the transformed line is given by

$$\mathbf{H}^{-\top}\mathbf{l} \cong \mathbf{l}' = \begin{pmatrix} l_0' \\ l_1' \\ l_2' \end{pmatrix} = \begin{pmatrix} \sin(\alpha') \\ -\cos(\alpha') \\ -\tau' \end{pmatrix}, \tag{12}$$

with direction

$$\mathbf{d}' = \begin{pmatrix} l'_1 \\ -l'_0 \\ 0 \end{pmatrix}. \tag{13}$$

The Radon transform is the integral over a line in an image

$$\begin{aligned} \mathcal{R}f(\alpha,\, \tau) = \mathcal{R}f(\mathbf{l}) &= \int f(o_0 + d_0 \cdot t,\, o_1 + d_1 \cdot t)\, \mathrm{d}t \\ &\stackrel{\text{def}}{=} \int f_{\mathbf{l}}(t)\, \mathrm{d}t, \end{aligned} \tag{14}$$

where $f_{\mathbf{l}}(t)$ is the image defined over the 1D sub-space parametrized by the line-coordinate t. Since a projective transformation $\mathbf{H}$ on the image is collinearity preserving, a mapping $\varphi(t) :\, t \mapsto t'$ must exist which fulfills

$$\mathcal{R}\mathbf{H}f(\alpha', \tau') = \mathcal{R}\mathbf{H}f(\mathbf{l}') = \int f_{\mathbf{l}}\left(\varphi^{-1}(t')\right)\, \mathrm{d}t', \tag{15}$$

where $\mathbf{H}f$ denotes the transformed image and t' is the line coordinate along the transformed line $\mathbf{l}'$. Equation 15 shows that only information on the original line $f_{\mathbf{l}}$ is required for this computation of $\mathcal{R}\mathbf{H}f(\mathbf{l}')$. A change of variables with $t' = \varphi(t)$ yields

$$\mathcal{R}\mathbf{H}f(\mathbf{l}') = \int f_{\mathbf{l}}\,(t)\, \frac{\partial}{\partial t}\varphi(t)\, \mathrm{d}t. \tag{16}$$

B. Projection to Line Coordinates

The goal of this section is to find a closed form for $\varphi(t)$ in Equation 15, by considering only the line coordinate t on the line $\mathbf{l}$ and t' on the line $\mathbf{l}'$. Because the Jacobian of a rigid transformation is one, we can consider w.l.o.g. the case of $\mathbf{l}_u = (0, 1, 0)^\top$, which is related to the line $\mathbf{l}$ with the rigid transformation

$$\mathbf{H}_{\mathbf{l}u} = \begin{pmatrix} l_1 & -l_0 & 0 \\ l_0 & l_1 & l_2 \\ 0 & 0 & 1 \end{pmatrix} \text{ with } \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \mathbf{H}_{\mathbf{l}u}^{-\top}\mathbf{l}. \tag{17}$$

Here, $\mathbf{H}_{\mathbf{l}u}$ takes image points $\mathbf{x} = (u,\, v,\, 1)^\top$ on the line $\mathbf{l}$ in the domain of the image $f(u, v)$ to line coordinates in the domain of $f_{\mathbf{l}}(t)$

$$\mathbf{x} = \mathbf{o} + \mathbf{d}t = \mathbf{H}_{\mathbf{l}u}^{-1} \begin{pmatrix} t \\ 0 \\ 1 \end{pmatrix} \text{ with } \mathbf{x}^\top\mathbf{l} = 0. \tag{18}$$

Analogously, we can transform line coordinates on $\mathbf{l}'$ with $\mathbf{H}_{\mathbf{l}'u}$ according to

$$\begin{pmatrix} t' \\ 0 \\ 1 \end{pmatrix} = \mathbf{H}_{\mathbf{l}'u}\mathbf{x}' \text{ iff. } \mathbf{x}'^\top\mathbf{l}' = 0. \tag{19}$$

However, if $\mathbf{H}$ is a rectifying homography according to Equation 8 and if $\mathbf{l}$ is an epipolar line, then $\mathbf{d}' = (1,\, 0,\, 0)^\top$ by definition and $\mathbf{H}_{\mathbf{l}'u}$ is only a translation.

C. Line-Induced Perspectivity

Using Equations 18 and 19 we have

$$\begin{aligned} & & \mathbf{x}' &\cong \mathbf{H}\mathbf{x} \\ \iff & & \mathbf{H}_{\mathbf{l}'u}^{-1} \begin{pmatrix} t' \\ 0 \\ 1 \end{pmatrix} &\cong \mathbf{H}\mathbf{H}_{\mathbf{l}u}^{-1} \begin{pmatrix} t \\ 0 \\ 1 \end{pmatrix} \\ \iff & & \begin{pmatrix} t' \\ 0 \\ 1 \end{pmatrix} &\cong \underbrace{\mathbf{H}_{\mathbf{l}'u}\mathbf{H}\mathbf{H}_{\mathbf{l}u}^{-1}}_{\mathbf{J}} \begin{pmatrix} t \\ 0 \\ 1 \end{pmatrix}. \end{aligned} \tag{20}$$

Since both sides of the equation contain a zero-element, only the four corner elements of the the matrix $\mathbf{J}$ are relevant. Note that $\mathbf{H}_{\mathbf{l}u}^{-1}$ and $\mathbf{H}_{\mathbf{l}'u}$ can be elegantly written by column, respectively, row vectors

$$\mathbf{H}_{\mathbf{l}u}^{-1} = \begin{pmatrix} | & | & | \\ \mathbf{d} & \mathbf{n} & \mathbf{o} \\ | & | & | \end{pmatrix}, \tag{21}$$

and

$$\mathbf{H}_{\mathbf{l}'u} = \begin{pmatrix} - & \mathbf{d}'^\top & - \\ - & \mathbf{l}'^\top & - \\ 0 & 0 & 1 \end{pmatrix}. \tag{22}$$

The entries of $\mathbf{l}$ can be computed via Equation 20, while we are only interested in the case where $\mathbf{H}$ is a rectifying homography and $\mathbf{l}$ is an epipolar line, we have more compactly, $\mathbf{d}' = (1,\, 0,\, 0)^\top$ and

$$J_{00} = \mathbf{d}'^\top\mathbf{H}\mathbf{d} \qquad = h_{00}l_1 - h_{10}l_0, \tag{23}$$

$$J_{02} = \mathbf{d}'^\top\mathbf{H}\mathbf{o} \qquad = h_{02} \quad - h_{00}l_0l_2, \tag{24}$$

$$J_{20} = (0,\, 0,\, 1) \cdot \mathbf{H}\mathbf{d} \quad = h_{20}l_1 - h_{21}l_0, \tag{25}$$

$$J_{22} = (0,\, 0,\, 1) \cdot \mathbf{H}\mathbf{o} \quad = h_{22} \quad - h_{20}l_0l_2 - h_{21}l_1l_2. \tag{26}$$

Assuming that the horizon is not contained in the support of f, de-homogenization of the mapping from $\mathbb{P}^1$ coordinates on the line $\mathbf{l}$ to $\mathbb{P}^1$ coordinates on the line $\mathbf{l}'$ is written

$$t' = \varphi(t) = \frac{J_{00} \cdot t + J_{02}}{J_{20} \cdot t + J_{22}}, \tag{27}$$

and

$$\frac{\partial}{\partial t}\varphi(t) = \frac{J_{00}J_{22} - J_{02}J_{10}}{J_{20}^2t^2 + 2 \cdot J_{20}J_{22} \cdot t + J_{22}^2}. \tag{28}$$

The Radon transform for a line $\mathbf{l}'$ can thus be computed using regular ray-casting in f by weighting each sample $\mathbf{x} = \mathbf{o} + t \cdot \mathbf{d}$ with $\frac{\partial}{\partial t}\varphi(t)$. For some regular sampling distance Δt in the original image (typically about 0.66 px), we have

$$\mathcal{R}\mathbf{H}f(\mathbf{l}') \approx \Delta t \sum_i \left.\frac{f(\mathbf{o} + t \cdot \mathbf{d}) \cdot (J_{00}J_{22} - J_{02}J_{20})}{J_{20}^2t^2 + 2 \cdot J_{20}J_{22} \cdot t + J_{22}^2}\right|_{t = i \cdot \Delta t}, \tag{29}$$

with the original line

$$\mathbf{l} \cong \mathbf{H}^\top\mathbf{l}' = \mathbf{H}^\top \begin{pmatrix} \sin(\alpha') \\ -\cos(\alpha') \\ -\tau' \end{pmatrix}. \tag{30}$$

IV. Discussion

We have presented a general and fast algorithm for evaluation of fan-beam DCCs on a virtual detector plane based on Lesaint et al. [4], which also handles non-circular trajectories. We have shown that reprojection to a common virtual detector is an instance of stereo rectification. However, rectifying each pair of projections, requires $n \cdot (n-1)$ rectified 2D images. In Figure 2 we show that rectification increasingly distorts the images, as the angle between the primary rays increases. For proper sampling rectified images would have to be considerably larger than the original projections, rendering the original algorithm impractical. We show how to compute line integrals over projectively transformed images on the fly, i.e. directly from the original projection image data, which is considerably faster and requires no additional memory. This

Algorithm 1 Radon transform of a projectively distorted image $\mathbf{H}f$ for a line $\mathbf{l}' = \mathbf{H}^{-\top}\mathbf{l}$ from original image f.

- **input** – Line $\mathbf{l}' = (\sin(\alpha'), -\cos(\alpha'), -\tau)^\top$
  – Homography $\mathbf{H} \in \mathbb{R}^{3\times3}$ (as in Equation 8).

1) Compute $\mathbf{l}$ according to Equation 30, its closest point $\mathbf{o}$ (Equation 11), and direction $\mathbf{d}$ (Equation 10).
2) Normalize $\|\mathbf{n}\| = \|\mathbf{d}\| = 1$ and dehomogenize $o_2 = 1$.
3) Compute J_{00} to J_{22} according to Equations 23-26.
4) Compute values t_{min} and t_{max}, such that $\mathbf{x}_{min} = \mathbf{o} + t_{min} \cdot \mathbf{d}$ and $\mathbf{x}_{max} = \mathbf{o} + t_{max} \cdot \mathbf{d}$ are the intersections of the line with the image border with $t_{min} < t_{max}$.
5) Initialize summation variable s.
6) **for** ($t = t_{min}$; $t \leq t_{max}$; $t \leftarrow t + \Delta t$)
   a) Compute pixel location $\mathbf{x} \leftarrow \mathbf{o} + t \cdot \mathbf{d}$
   b) Compute $j \leftarrow \frac{\partial}{\partial t}\varphi(t)$ according to Equation 28.
   c) Compute distance w of 3D location of $\mathbf{x}$ (on the virtual detector) to the source position [4], [5].
   d) Increment $s \leftarrow s + f(\mathbf{x}) \cdot j \cdot w \cdot \Delta t$.
7) **return** Variable s, which is the Radon transform of projectively distorted image $\mathbf{H}f$ for a line $\mathbf{l}'$.

paper gives a general implementation for this situation. For validation, we present in Figure 4, top, the redundant signals extracted from two X-ray shots of a pumpkin, along with the warped images in Figure 3. The fan-beam consistency condition can be computed explicitly as a sum over these rectified images weighted by the distance of each point to the stereo baseline. Integration in the original images, while taking into account the Jacobian according to Algorithm 1 is identical. The method breaks down for large primary angle difference, once the epipoles lie within the images, compare Figure 2, right. To demonstrate the fundamentally different nature of this consistency condition by comparison with epipolar consistency for the same epipolar planes in Figure 4, bottom. Since EC is based on a derivative orthogonal to the lines, we observe a roughly zero-mean curve and more detail for smaller structures. Generally speaking, rectification-based methods will be slower than their deriative-based counterparts for motion- and calibration-correction, because the Radon intermediate functions cannot be pre-computed. Future work will investigate, if avoidance of the derivative has potential benefits in applications, where absolute intensity values need to be estimated, such as beam-hardening, scatter correction or metal-artifact correction. Future work will investigate how the presented epipolar DCCs are affected by truncation and we will compare other DCCs.

Acknowledgments: This work is supported by the German Research Foundation; DFG MA 4898/3-1 "Consistency Conditions for Artifact Reduction in Cone-beam CT" and partially supported by the Agence Nationale de la Recherche (France), Labex CAMI, number ANR-11-LABX-0004-01 and project "DROITE", number ANR-12-BS01-0018. We thank NVIDIA Corporation for the donation of a Titan Xp GPU.

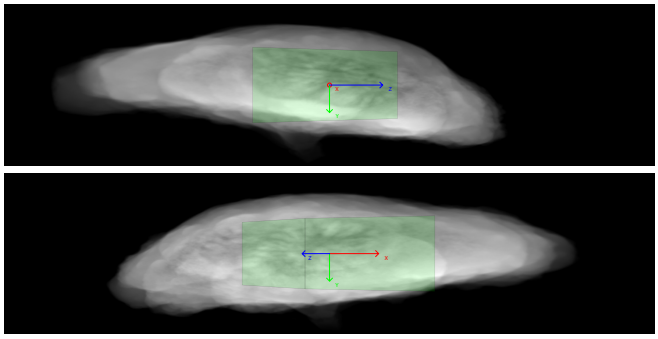


Figure 3. Rectified projection images of a pumpkin phantom in case of a primary angle difference of 120°. The green cube is intended to visualize the distortion introduced by rectification.

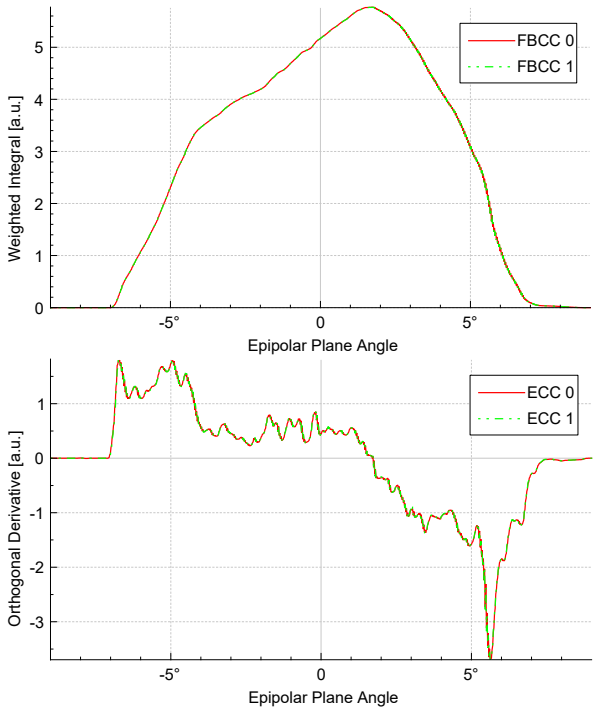


Figure 4. Top: Fan-beam consistency condition (FBCC) [5] as implemented in the original images. The computation of the rectified images in Figure 3 is avoided, but the curves are identical to a weighted summation over the intensity in these images. Bottom: Comparison to epipolar consistency conditions (ECC) for the same images (i.e. orthogonal derivative instead of rectification and weighting) is shown for corresponding epipolar plane angles.

References

[1] A. Aichert, K. Breininger, T. Köhler, and A. Maier. Efficient epipolar consistency. In *4th International Conference on Image Formation in X-Ray Computed Tomography*, pages 383–386, 2016.

[2] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521623049, 2000.

[3] J. Lesaint, R. Clackdoyle, S. Rit, and L. Desbat. Two cone-beam consistency conditions for a circular trajectory. In *Proc. 4th Int. Conf. Image Formation X-Ray Comput. Tomography*, pages 431–434, 2016.

[4] J. Lesaint, S. Rit, R. Clackdoyle, and L. Desbat. Calibration for circular cone-beam ct based on consistency conditions. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 1(6):517–526, 2017.

[5] M. S. Levine, E. Y. Sidky, and X. Pan. Consistency conditions for cone-beam ct data acquired with a straight-line source trajectory. *Tsinghua Sci Technol.*, 15(1):56–61, Feb 2010.

Fast Epipolar Consistency without the Need for Pseudo Matrix Inverses

Alexander Preuhs, Michael Manhart and Andreas Maier

***Abstract*—Interventional C-arm systems allow flexible 2-D imaging of a 3-D scene while being capable of cone beam computed tomography. Due to the flexible structure of the C-arm, the rotation speed is limited, increasing the acquisition time compared to conventional computed tomography. Therefore, patient motion frequently occurs during data acquisition inducing inconsistencies in the projection raw data. A framework using Grangeat's theorem and epipolar consistency was successfully applied for compensating rigid motion. This algorithm was efficiently parallelized, however, before each iteration, the pseudo-inverse of each projection matrix must be calculated. We present a geometric modification of the presented algorithm which can be used without a pseudo-inverse. As such, the complete algorithm can be implemented for low-level hardware without the need of a linear algebra package that supports the calculation of matrix inverse. Both algorithms are applied for head motion compensation and the runtime of both is compared.**

I. Introduction

A fundamental assumption in computed tomography (CT) is that the scanned object remains static during the acquisition process. If this assumption cannot be fulfilled, images produced with conventional reconstruction algorithms will suffer from artifacts. Current C-arm CT acquisitions last about 20 seconds. During the acquisition time, involuntary patient motion is often inevitable without patient fixation. However, if the motion can be assumed to be rigid and smooth, a motion compensated reconstruction can be computed by finding the correct geometric correspondence between the motion affected projections and the calibration data. Four categories of compensating motion artifacts have emerged in literature and they can be grouped into approaches using external markers [1], image metrics on the reconstruction volume [2], 3-D/2-D registration of the projection data to digitally rendered radiographs from the reconstruction volume [3], [4] and projection data consistency based metrics [5]–[8].

In this work, we focus on a consistency method based on the 3-D radon transform. The method exploits epipolar geometry to find lines on two detectors corresponding to an epipolar plane. Grangeat's theorem can be used to find a mapping between each epipolar line pair and the 3-D radon value corresponding to the epipolar plane [9]. This algorithm is denoted as epipolar consistency and was presented by Aichert et al. [5]. As the algorithm directly works on the projection domain without the need of a reconstruction, the computational cost is low. It basically consists of comparing corresponding line configurations. This can be accelerated by parallelizing the algorithm using graphics processing units (GPU) [10].

To apply the algorithm for rigid motion compensation, the consistency between all possible line pairs is evaluated in an iterative optimization process in order to find the set of parameters describing the motion within the scan [11], [12]. In [10] before each iteration the pseudo-inverse of the projection matrices must be calculated on the CPU. We propose a geometric modification that allows to calculate corresponding epipolar lines without the need of a pseudo-inverse.

A. Preuhs and A. Maier are with the Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany.

M. Manhart is with Siemens Healthcare GmbH, Forchheim, Germany.

Email: alexander.preuhs@fau.de

II. Methods

A. Grangeat's Theorem

In cone-beam CT an X-ray source radially emits photons, that — after attenuation — are measured at a detector. The attenuation process for a ray can be described by an integral. However, due to the radial structure of the rays, integrating along a detector line does not result in a plane integral of the underlying object f, instead it differs by a radial weighting.

Grangeat's theorem describes the connection between this weighted integral and a plane integral — i.e. the 3-D radon value $\mathcal{R}f(\mathbf{n}, d)$ describing the integral along a plane with normal $\mathbf{n} \in \mathcal{S}^2$ at distance d. Using a derivative operation the radial weighting can be canceled out. Grangeat defined an intermediate function $S_\lambda(\mathbf{n})$ that is calculated from the projection data which can be related to the derivative of the 3-D radon transform

$$S_\lambda(\mathbf{n}) = \int_{\mathcal{S}^2} \delta'(\mathbf{x}^\top \mathbf{n}) g_\lambda(\mathbf{x}) d\mathbf{x} = \frac{\partial}{\partial d} \mathcal{R}f(\mathbf{n}, d)|_{d=\mathbf{c}_\lambda^\top \mathbf{n}} \ , \quad (1)$$

where $g_\lambda(\mathbf{x})$ describes a single value on the detector with λ describing the projection index, $\mathbf{c}_\lambda$ the source position and $\mathbf{x}$ a vector from the source to a detector pixel. The geometry for two projections $\lambda = a$ and $\lambda = b$ is visualized in Fig. 1. Here $\delta'(\cdot)$ describes the derivative of the Dirac delta distribution. A detailed evaluation of Eq. (1) can be found in [9], and some simplifications are discussed in [5].

B. Epipolar Consistency

It directly follows from Eq. (1) that two projections a, b must satisfy

$$S_a(\mathbf{n}) = S_b(\mathbf{n}) \qquad \forall \mathbf{n} \in \mathcal{S}^2 : \mathbf{c}_b^\top \mathbf{n} = \mathbf{c}_a^\top \mathbf{n} \ . \quad (2)$$

If the geometry information is wrong, e.g. due to rigid object motion, then Eq. (2) will not hold. Thus, we can use it as a measure of inconsistency. Below we summarize the framework proposed by Aichert et al. [5], [10], which is used to evaluate the consistency of two views.

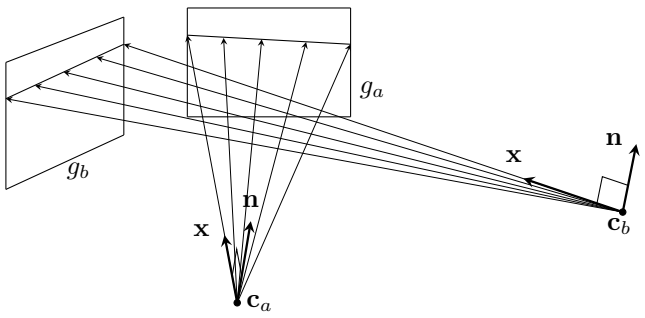


Fig. 1. Schematic drawing of a scene including two projections. The vector $\mathbf{n}$ describes the normal of a radon plane. Several realizations of the vectors $\mathbf{x}$ are drawn that are perpendicular to $\mathbf{n}$. The pixel intensity measured at the detector along ray $\mathbf{x}$ is described by $g_a(\mathbf{x})$ or $g_b(\mathbf{x})$, respectively.

The intermediate function $S_\lambda(\mathbf{n})$ can be precomputed for each projection. Then, the global indexing by the plane normal $\mathbf{n}$ can be replaced by a local projection-pair-dependent indexing using a line $\mathbf{l}_\lambda^\kappa$ defined on the detector described by g_λ. By epipolar geometry two epipolar lines $\mathbf{l}_a^\kappa$ and $\mathbf{l}_b^\kappa$ are found that belong to the same epipolar plane $\mathbf{E}^\kappa$ — i.e. the radon plane. The algorithm starts with a configuration of two projections described by their projection matrix $\mathbf{P}_a$ and $\mathbf{P}_b$, respectively. Using these two projection matrices a mapping matrix is derived that maps an angle κ to an epipolar plane $\mathbf{E}^\kappa$. Using the pseudo-inverse the respective epipolar lines $\mathbf{l}_a^\kappa$ and $\mathbf{l}_b^\kappa$ are computed. The respective values are then used to look up the values at the precomputed intermediate function S_a and S_b. This allows the indexing of Eq. (1) using an angle κ and two projection matrices

$$S_a(\kappa, \mathbf{P}_a, \mathbf{P}_b) = S_a(\mathbf{n}) \quad \forall \mathbf{n} \in \mathcal{S}^2 : \mathbf{c}_b^\top \mathbf{n} = \mathbf{c}_a^\top \mathbf{n} \ . \tag{3}$$

To evaluate the consistency of a whole scan, many different views must be compared to each other, while in each two-view comparison a multitude of line pairs are evaluated. As the operations are independent from each other, this can be evaluated in parallel allowing the efficient parallelization of the algorithm using GPUs.

C. Projective Geometry

Projective geometry can be seen as an extension to the common Euclidean geometry. In the context of image reconstructions, projective geometry is mostly used to describe the projection of a world point to a detector. Therefore, a projection matrix is created that performs a projective transformation on a world point. In this context, the world point must be converted to homogeneous coordinates first.

Homogeneous coordinates are the representation of n-dimensional points in the projective space and are written as $(n+1)$-component vectors. In $\mathbb{P}^3$ a point is described by $(x, y, z, w)^\top$, and we can obtain the euclidean representation by dividing with the last component $(x/w, y/w, z/w)^\top$. Similarly, a plane is described by $(a, b, c, d)^\top$. The vector can be understood as the parameters of a Hessian normal form, where the first three components describe the normal of the plane, and d is the scaled distance to the origin. If $a^2 + b^2 + c^2 = 1$ then d is exactly the signed distance to the origin. The concept that a four-component vector can either be interpreted as a point or a plane is called duality, where we refer to the point interpretation as primal form and the plane interpretation as dual form.

A special case is the representation of a line in $\mathbb{P}^3$. There is no direct description but we can construct the line as the connection of two points or the intersection of two planes. An intuitive derivation can be found in [13], we only state the relevant result of this derivation. The creation of a line as the incident of two planes $\mathbf{a}, \mathbf{b} \in \mathbb{P}^3$ is obtained by

$$\text{meet}(\mathbf{a}, \mathbf{b}) = \mathbf{L} = \begin{pmatrix} p \\ q \\ r \\ s \\ t \\ u \end{pmatrix} = \begin{pmatrix} a_z b_w - a_w b_z \\ a_y b_w - a_w b_y \\ a_y b_z - a_z b_y \\ a_x b_w - a_w b_x \\ a_x b_z - a_z b_x \\ a_x b_y - a_y b_x \end{pmatrix} \ , \tag{4}$$

where the six components of $\mathbf{L}$ are often referred to as Plücker coordinates. We can build an anti-symmetric matrix $\mathbf{L}_K$ from the Plücker coordinates that represents a line as the intersection of two planes — i.e. the dual representation of a line. A point $\mathbf{x}$ common to a plane $\mathbf{p}$ and the line $\mathbf{L}$ can be found by right-multiplication of $\mathbf{p}$ to $\mathbf{L}_K$

$$\mathbf{x} = \text{meet}(\mathbf{L}, \mathbf{p}) = \mathbf{L}_K \, \mathbf{p} = \begin{pmatrix} 0 & -p & -q & r \\ p & 0 & s & -t \\ q & -s & 0 & u \\ -r & t & -u & 0 \end{pmatrix} \mathbf{p} \ . \tag{5}$$

Note that there is also a primal representation of $\mathbf{L}$ which will not be discussed in this paper.

An extension in the projective geometry is the concept of geometric primitives at infinity. They are regular objects and thus can be handled as any other objects. A point at infinity is defined by a homogeneous coordinate $w = 0$. In $\mathbb{P}^3$ the plane at infinity is defined by $\boldsymbol{\pi}_\infty = (0, 0, 0, 1)^\top$. All previously introduced equations are also valid for objects at infinity. We could for example use Eq. (5) to find the incident of a line $\mathbf{L}$ with $\boldsymbol{\pi}_\infty$, which will be a point at infinity, where the first three component of that point are the direction of the line.

An advantage of using projective geometry is the representation of transformations based on matrix multiplication. A point $\mathbf{x}'$ which is the transformation of the point $\mathbf{x}$ under $\mathbf{T}$ is simply found by

$$\mathbf{x}' = \mathbf{T}\,\mathbf{x} \ . \tag{6}$$

The transformation rule for planes can be derived from the property that the distance from a point $\mathbf{x}$ incident to a plane $\mathbf{p}$ is zero. The distance between the transformed point $\mathbf{x}'$ and the plane $\mathbf{p}'$ will remain zero if they have been transformed under the same transformation $\mathbf{T}$. It therefore holds that

$$\mathbf{p}'^\top \mathbf{x}' = \mathbf{p}^\top \mathbf{x} = 0 \ . \tag{7}$$

Solving Eq. (6) for $\mathbf{x}$ and plugging that in Eq. (7) gives

$$\mathbf{p}'^\top \mathbf{x}' = \mathbf{p}^\top \mathbf{T}^{-1} \mathbf{x}' = \left((\mathbf{T}^{-1})^\top \mathbf{p} \right)^\top \mathbf{x}' \ , \tag{8}$$

it directly follows that

$$\mathbf{p}' = (\mathbf{T}^{-1})^\top \mathbf{p} \tag{9}$$

which describes the transformation of planes. The point $\mathbf{x}$ incident to a plane $\mathbf{p}$ and a line $\mathbf{L}$ can be found by right-multiplying the plane to the dual representation of $\mathbf{L}$ (cf. Eq. (5)). Further, a transformed point $\mathbf{x}'$ will be incident to the plane $\mathbf{p}'$ and line $\mathbf{L}'$ if both are transformed under a transformation $\mathbf{T}$, thus,

$$\mathbf{L}_K \, \mathbf{p} \; = \mathbf{x}, \qquad\qquad \mathbf{L}'_K \, \mathbf{p}' \; = \mathbf{x}' \;\; . \tag{10}$$

When we solve Eq. (6) and (9) for $\mathbf{x}$ and $\mathbf{p}$, respectively, we can plug the result in the left part of Eq. (10) which results in

$$\mathbf{L}_K \mathbf{T}^\top \, \mathbf{p}' = \; \mathbf{T}^{-1} \, \mathbf{x}' \; \Leftrightarrow \; \mathbf{T} \, \mathbf{L}_K \mathbf{T}^\top \, \mathbf{p}' = \; \mathbf{x}' \;\; . \tag{11}$$

Comparing the result with the right side of Eq. (10) it immediately emerges that the line $\mathbf{L}'$ which is the transformation of $\mathbf{L}$ under $\mathbf{T}$ can be calculated by

$$\mathbf{L}'_K \; = \; \mathbf{T} \, \mathbf{L}_K \, \mathbf{T}^\top \;\; . \tag{12}$$

D. Optimized Algorithm

The main purpose of the algorithm presented in Section II-B is to find the mapping between two lines $\mathbf{l}_a^\kappa$ and $\mathbf{l}_b^\kappa$ that can be used to look up the corresponding precomputed values S_a and S_b, respectively. This is achieved by first finding epipolar planes $\mathbf{E}^\kappa$ which are then mapped to the corresponding epipolar lines. The algorithm presented in [5] makes use of the pseudo-inverse to compute that mapping. However, the calculation of a pseudo-inverse is not supported on many GPUs, and must therefore be done on the CPU beforehand, whereas the rest of the framework is parallelizable. In addition a linear algebra library must be included to support the calculation of pseudo matrix inverses.

We propose a geometric modification that creates the mapping without the need of a pseudo-inverse. As shown in Section II-C, the transformation rule depends on the object that is to be transformed. It can be seen from Eq. (12) that lines are transformed using the transformation matrix and its transpose. Thus, transforming the plane to a line while preserving the relevant information will make the pseudo-inverse dispensable.

We can achieve this using the concept of infinity. The projective three-space is covered by the infinity plane $\boldsymbol{\pi}_\infty = (0, 0, 0, 1)$. Any plane intersects the infinity plane in a line incident to $\boldsymbol{\pi}_\infty$ and the plane itself, i.e. a line at infinity. The orientation of the plane is persevered by the direction of the line. In a last step, we can simply use Eq. (12) to project the line at infinity, resulting in the desired epipolar lines.

Therefore, we start with the epipolar plane $\mathbf{E}^\kappa$. Using Eq. (4) we can compute the line at infinity $\mathbf{L}^\kappa$ as the intersection of the epipolar plane with $\boldsymbol{\pi}_\infty$

$$\mathbf{L}^\kappa = \text{meet}(\mathbf{E}^\kappa, \boldsymbol{\pi}_\infty) \;\; . \tag{13}$$

Using the representation of the line at infinity now allows us to use the transformation rule as described by Eq. (12) to obtain the epipolar line $\mathbf{l}_\lambda^\kappa$

$$[\mathbf{l}_\lambda^\kappa]_\times \; = \; \mathbf{P}_\lambda \, \mathbf{L}_K^\kappa \, {\mathbf{P}_\lambda}^T \;\; . \tag{14}$$

The parameters of $\mathbf{l}_\lambda^\kappa$ are available from the 3×3 skew matrix $S = [\mathbf{l}_\lambda^\kappa]_\times$ as $\mathbf{l}_\lambda^\kappa = (S_{12}, S_{20}, S_{01})^\top$. As a result Eq. (13) and (14) replace the mapping from epipolar planes to lines presented in [5] and, therefore, makes the computation of pseudo-inverses unnecessary. The additional cost is the implementation of Eq. (4) on the GPU, however, this can be reused to simplify the calculation of source positions. As the three rows of the projection matrix can be interpreted as planes all passing the source, the incident of two of these planes will create a line. Using matrix multiplication (cf. Eq. 5) the source position is then found by the incident of that line with the third plane.

E. Optimization

If rigid motion occurs during the scan, the calibrated trajectory does not represent the true geometry of the acquired data. In order to restore the true geometry, a rigid transformation $\mathbf{T}_\lambda$ for each projection matrix $\mathbf{P}_\lambda$ must be found. The true geometry is expected to have minimal inconsistency. We therefore define the inconsistency between two projections a and b in dependence of the respective rigid transformations $\mathbf{T}_a$ and $\mathbf{T}_b$ by

$$d(\mathbf{P}_a\mathbf{T}_a, \mathbf{P}_b\mathbf{T}_b) = \frac{1}{N_\kappa} \sum_{k=0}^{K} \left[S_a(k\Delta\kappa, \mathbf{P}_a\mathbf{T}_a, \mathbf{P}_b\mathbf{T}_b) - S_b(k\Delta\kappa, \mathbf{P}_b\mathbf{T}_b, \mathbf{P}_a\mathbf{T}_a) \right]^2 \;\; , \tag{15}$$

where N_κ is the number of epipolar planes that hit both detectors and K is the total number of sampled epipolar planes. The angular step-size is denoted by $\Delta\kappa$. To be more robust for outliers we use the robust Cauchy norm and define the inconsistency of two views by

$$e_{a,b} = \frac{d(\kappa, \mathbf{P}_a\mathbf{T}_a, \mathbf{P}_b\mathbf{T}_b)}{1 + \frac{1}{c}\, d(\kappa, \mathbf{P}_a\mathbf{T}_a, \mathbf{P}_b\mathbf{T}_b)} \;\; . \tag{16}$$

The parameter c controls the penalty and should be selected according to the intensity of the projection images. We denote the vector of rigid transformations $\mathbf{T} = [T_1, ...T_N]$, with N being the number of projections of the trajectory. The corrected geometry is denoted by $\hat{\mathbf{T}}$ and found by solving

$$\hat{\mathbf{T}} = \arg\min_{\mathbf{T}} \sum_{a,b=1}^{N} e_{a,b} \;\; . \tag{17}$$

Since motion is expected to be smooth we model each rigid motion parameter in $\mathbf{T}$ by an Akima spline [14]. This also allows the reduction of the search space, as we must not find a transformation for each λ, but only for the nodes of the spline. The optimum is then found using the open source non-linear optimizer JPOP[1] in CONRAD [15].

III. Experiments

To evaluate the proposed method, we have acquired a 200° short scan (496 projections) of a head phantom using a robotic

[1] https://www5.cs.fau.de/research/software/java-parallel-optimization-package/

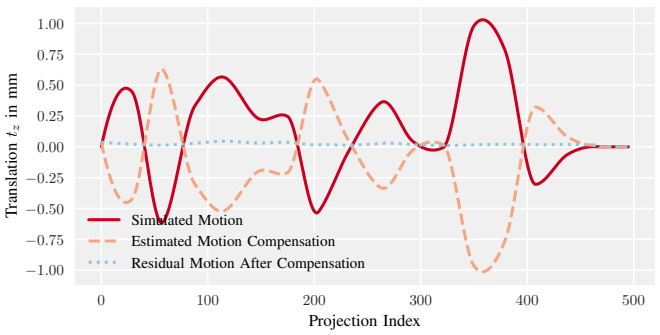


Fig. 2. Simulated, estimated and residual motion t_z for each projection.

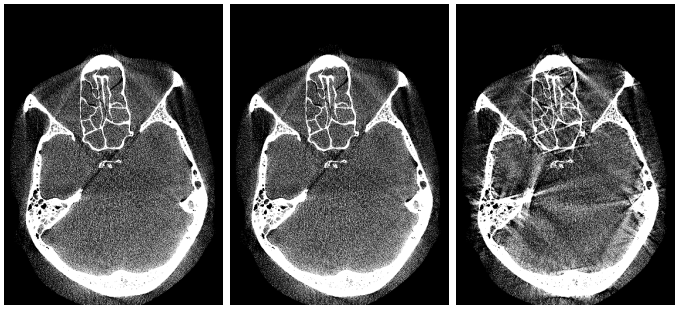


Fig. 3. Central slices of the reconstructed volume HU [-100, 100]. Left: ground truth, mid: with simulated motion after compensation, right: with simulated motion.

C-arm system (Artis zeego, Siemens Healthcare GmbH, Germany). Thereafter, we simulate rigid motion, which is directly incorporated in the projection matrices. This is done using a rigid motion creator[2].

Epipolar consistency is known to produce mostly horizontal epipolar lines in a majority of the projection pairs within a short scan. Only view pairs that are almost opposed to each other present diverging epipolar lines. Motion that is parallel to the epipolar lines is not detectable by the presented consistency measure. Thus, we only concentrate on motion orthogonal to the epipolar lines in all pairs, which is typically denoted as out-plane motion. Defining the rotation axis of the short scan as the z-axis, we only simulate translations in z-direction. The simulated motion pattern consisting of 17 spline nodes is shown in Fig. 2.

IV. Results

The reconstructions of the acquired head phantom is shown in Fig. 3 for the motion corrupted case (right), the motion compensated case (mid) and the ground truth (left). The corresponding motion is depicted in Fig. 2. Both algorithms produce the same results, only the runtime is expected to change. By skipping the sequential calculation of pseudo-inverses the runtime could be reduced by 1.29% using a standard computer with an Intel Core i7-4910MQ and a NVIDIA Quadro K2100M. The overall runtime for the motion parameter estimation was 841.7 seconds using the proposed modifications and 852.8 seconds if the inverse is pre-calculated before each optimization step.

V. Conclusion and Discussion

We presented a modification to the algorithm presented in [10] which avoids the calculation of inverse projection matrices. This is achieved by transforming the respective epipolar planes to lines at infinity. Lines are transformed — in contrast to planes — using only the transformation matrix and its transposed. Thus, only the projection matrix and its transposed must be available.

The runtime could be improved by 1.29% using a Java environment. Using more high-level programming languages, e.g. python, the runtime advantage could eventually increase, as more computations can be performed on a dedicated GPU. Furthermore, when implementing the algorithm in low-level programming languages, e.g. C++/CUDA, no linear algebra libraries need to be included.

Disclaimer: The concepts and information presented in this paper are based on research and are not commercially available.

[2]https://github.com/alPreuhs/MotionCreator

References

[1] J.-H. Kim, J. Nuyts, Z. Kuncic, and R. Fulton, "The feasibility of head motion tracking in helical CT," *Medical physics*, vol. 40, no. 4, 2013.
[2] A. Sisniega, J. W. Stayman, Q. Cao, J. Yorkston, J. H. Siewerdsen, and W. Zbijewski, "Motion estimation using a penalized image sharpness criterion for resolution recovery in extremities cone-beam CT," *Proceedings of the 3rd CT-meeting*, 2016.
[3] W. Wein and A. Ladikos, "Towards General Motion Recovery in Cone-Beam Computed Tomography," *Proceedings of the 12th Fully 3D*, 2013.
[4] M. Berger, K. Müller, M. Unberath, J. Thies, J.-H. Choi, R. Fahrig, and A. Maier, "Marker-free motion correction in weight-bearing cone-beam CT of the knee joint," *Medical Physics*, vol. 43, no. 3, 2016.
[5] A. Aichert, M. Berger, J. Wang, N. Maass, A. Doerfler, J. Hornegger, and A. K. Maier, "Epipolar Consistency in Transmission Imaging," *IEEE Transactions on Medical Imaging*, vol. 34, no. 11, 2015.
[6] C. Debbeler, N. , M. Elter, F. Dennerlein, and T. M. Buzug, "A New CT Rawdata Redundancy Measure applied to Automated Misalignment Correction," *Proceedings of the 12th Fully 3D*, 2013.
[7] N. Maass, F. Dennerlein, A. Aichert, and A. Maier, "Geometrical Jitter Correction in CT," in *Proceedings of the 3rd CT-meeting*, 2014.
[8] M. Berger, Y. Xia, W. Aichinger, K. Mentl, M. Unberath, A. Aichert, C. Riess, J. Hornegger, R. Fahrig, and A. Maier, "Motion Compensation for Cone-Beam CT Using Fourier Consistency Conditions," *Physics in Medicine and Biology*, vol. 62, no. 17, 2017.
[9] M. Defrise and R. Clack, "A Cone-Beam Reconstruction Algorithm Using Shift-Variant Filtering and Cone-Beam Backprojection," *IEEE Transactions on Medical Imaging*, vol. 13, no. 1, 1994.
[10] A. Aichert, K. Breininger, T. Köhler, and A. Maier, "Efficient epipolar consistency," *Proceedings of the 4rd CT-meeting*, 2016.
[11] R. Frysch and G. Rose, "Rigid motion compensation in C-arm CT using consistency measure on projection data," in *MICCAI*, 2015.
[12] M. Unberath, A. Aichert, and A. Maier, "Consistency-based Respiratory Motion Estimation in Rotational Angiography," *Medical Physics*, 2017.
[13] J. F. Blinn, "A Homogeneous Formulation for Lines in 3 Space," *Computer Graphics (SIGGRAPH '77 Proceedings)*, vol. 11, 1977.
[14] H. Akima, "A New Method of Interpolation and Smooth Curve Fitting Based on Local Procedures," *Journal of the ACM*, vol. 17, no. 4, 1970.
[15] A. Maier, H. Hofmann, M. Berger, P. Fischer, C. Schwemmer, H. Wu, K. Müller, J. Hornegger, J.-H. Choi, C. Riess, A. Keil, and R. Fahrig, "CONRAD - A software framework for cone-beam imaging in radiology," *Medical Physics*, vol. 40, no. 11, 2013.

A sinogram extrapolation method for CT field of view extension

Qiulin Tang, Masakazu Matsuura and Zhou Yu

***Abstract*– Truncation is a common issue encountered in computed tomography. It degrades the reconstructed image with severe artifacts. These artifacts can lead to misdiagnosis. Artifacts include biased HU values which will cause inaccurate treatment planning in radiotherapy, as well as inaccurate attention correction in PET reconstruction. In this work, we propose a sinogram based extrapolation method to complete the missing data in truncation cases. The proposed method not only eliminates truncation artifacts, but also extends the reconstruction field of view to a larger area, providing important information for radiation therapy planning and PET attenuation correction. The proposed method has two steps: first, the support of the sinogram is estimated; second, fill the sinogram from last measured pixel to the sinogram boundary using a hybrid polynomial fitting algorithm. The proposed method has been evaluated by physical phantoms and clinical data.**

I. INTRODUCTION

The reconstruction field of view (RFOV) of general CT is usually smaller than the full size of the bore. For example a typical scanner has an RFOV of 500mm in diameter, where every pixel has 360 degrees worth of samples. However, the scanner bore size can be as high as 700-800 mm. Here an object would fall outside RFOV when the patient size is large, the patient is positioned off-center, or there is medical equipment inside the bore, etc. When an object falls outside the RFOV it does not have enough samples to be correctly reconstructed. Moreover, it causes discontinuity in the sinogram and thus results in truncation artifacts (e.g. bright shading) inside RFOV. Consequently, it may lead to false diagnosis in routine CT examinations; inaccurate treatment planning for image guided radiotherapy; and inaccurate attenuation correction for PET imaging.

Several algorithms have been developed to compensate for truncation artifacts [1-6]. One approach uses data consistency to reduce the truncation artifacts and is only used for 2-dimentional parallel-beam or fan-beam scans [1-2]. Hsieh, et al developed a water cylinder fitting method to fill the region outside of RFOV. In Hsieh's method, the truncated sinogram is extrapolated assuming that the truncated region is part of a cylinder filled with water. For each view and each detector row, the attenuation of the water cylinder F_w is determined by water attenuation $\mu_w$and radius of fitted cylinder R [5]:

$$F_w = 2\mu_w\sqrt{R^2 - x^2} \quad (1)$$

where R is the radius of the water cylinder and x is the distance to origin of the cylinder. Figure 1 illustrates this method, where the green curve is the extrapolated sinogram using equation (1) and the black curve is the actual measurement. The parameter R ensures a smooth transition between the measurement and the extrapolated signal. This method is widely used due to its simplicity. The water cylinder assumption works reasonably well since soft tissue can be well approximated by water. However, the method relies on only a few edge detected pixels for the fitting, the fitting might become less accurate when noise is high. Moreover, in severe truncation cases, the water cylinder assumption is no longer accurate. The truncated object is no longer round, and it might contain materials other than soft tissue (e.g. bone, etc.). Figure 2 shows an example of water cylinder extrapolation on severely truncated data. The

Qiulin Tang, Masakazu Matsuura, Zhou Yu are with Canon Medical Research Institute, USA. 706 N. Deerpath Dr., Vernon Hills, IL, 60061, USA. (e-mail: qtang@mru.medical.canon, mmatsuura@mru.medical.canon, zyu@mru.medical.canon).

extrapolated sinogram is very irregular, and therefore cannot provide an accurate reconstruction outside the FOV. To address these limitations, Bruder et al developed an image and projection domain based method to extend the reconstruction FOV [7]. However, Bruder's method requires images to be reconstructed twice plus an extra forward projection; therefore it is computationally expensive.

In this work, we propose a sinogram domain only extrapolation method to extend the reconstruction field of view outside the fully sampled region. The proposed method can address the limitation of Hsieh's method without extra reconstructions. The proposed method is composed of two steps. First, it estimates a smooth boundary of scanned objects in a sinogram. Second, we use a polynomial model to complete the missing data between the boundary and real detector edge pixels, which allows us to handle more complicated truncation objects. Figure 2 shows the estimated sinogram using our proposed method. Both a physical phantom and clinical data were used to evaluate the proposed method.

II. ALGORITHMS

The proposed method has two main steps.

Step 1: Estimate sinogram support

For each row and each view, the sinogram is a truncated 1D signal, as shown in Figure 1. The first step is to estimate the support of the 1D signal, that is, where the signal approaches zero (see Fig. 1). This estimation can be done in two steps. First, we extrapolate the existing signal using water cylinder fitting described in [5]. This gives us the first pass estimate of the signal support for each row and view.

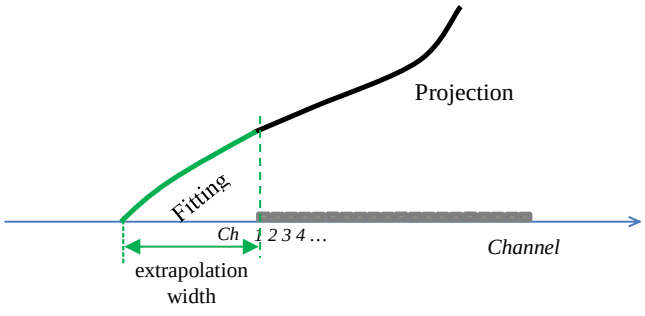


Fig. 1 Illustration of extrapolation width by water cylinder fitting.

The combination of extrapolation widths for all of the rows and projection views makes up a 2D function, referred to as extrapolation width map $m(i,j)$ and shown in Figure 2, where i, and j are row view indices. The extrapolation width map can have a jagged profile in both the row and view directions due to noise (see Fig. 3). Therefore, the extrapolation width map m cannot be used directly.

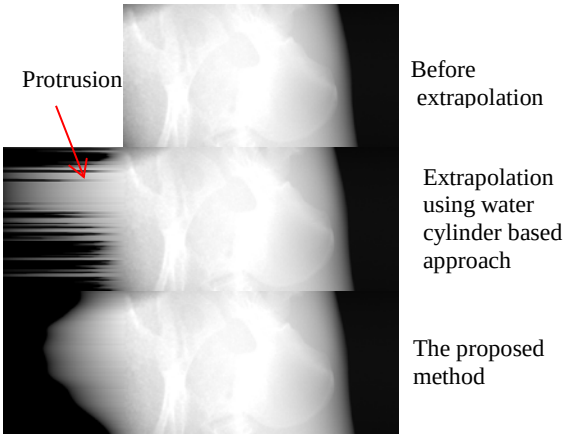


Fig. 2. Extrapolation width along rows in one projection view. It is jagged and has protrusion.

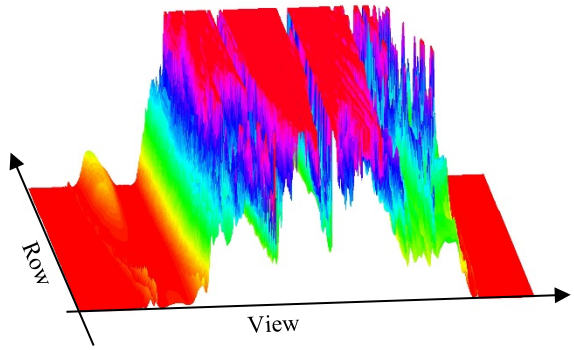


Fig. 3. Illustration of 2D extrapolation width map.

Secondly, a weighted Gaussian filter is applied on the extrapolation width map to obtain a smooth and more reliable map m' as follows:

$$m' = \frac{(wm * f)}{(w * f)}, \tag{2}$$

where f is a 2D Gaussian filter, $*$ is 2D dimensional convolution and w is a function of m:

$$w(i,j) = \exp(-m(i,j)^2/\sigma^2), \tag{3}$$

where σ is an empirical constant. The design of such a weighting function w is based on empirical observations that water cylinder based estimation is more accurate when truncation is small. Therefore,

smaller extrapolation widths are considered more reliable, and are given more weight in determining a moving average (i.e., smaller extrapolation widths have more contribution to filter).

The profiles along rows and views of the optimized and smoothed extrapolation width map are shown in Fig. 4.

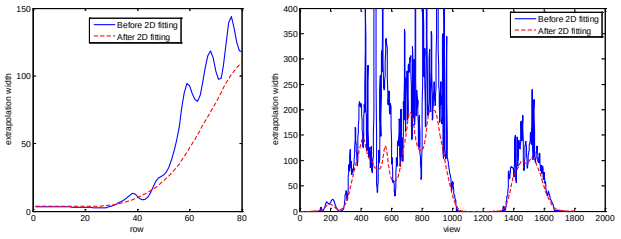


Fig. 4. The profile along rows and views of extrapolation width map before and after smoothing and optimization.

Step 2. Extrapolation by polynomial fitting

As discussed above, the water-cylinder approximation is based on the hypothesis that truncation regions are predominantly made up of materials consisting of water or exhibiting X-ray attenuation similar to water and having a shape that is approximately cylindrical. This hypothesis begins to fail when the truncation region includes bone or other materials that differ from water, are not cylindrical in shape or are inhomogeneous. Therefore, a different extrapolation function is desired especially when bone is in the truncated regions, which is common in cases of severe truncation. The bone will contribute a local peak on projection along channels, and the function F_w does not fit the profile well. In this work, a second order polynomial F_p is added to the fitting function to fit the bone and thus improve the fitting. This combination of fitting functions is called a hybrid polynomial fitting:

$$F = aF_w + bF_p, \tag{3}$$

where a, b are coefficients and F_p is given by

$$F_p = c_0 + c_1 ch + c_2 ch^2, \tag{4}$$

c_0, c_1, and c_2 are coefficients which can be optimized as well as a and b during fitting. The fitting is shown in Fig. 5.

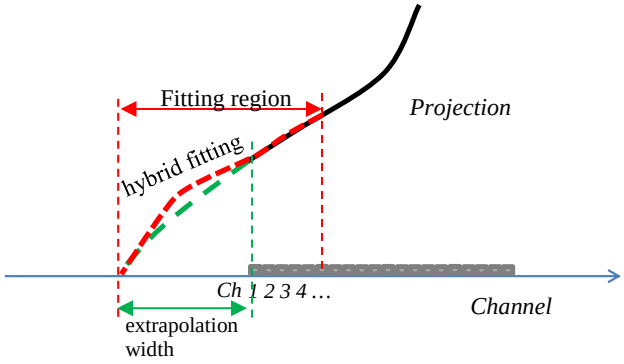


Fig. 5 Illustration of hybrid local multi-material fitting.

III. Experiment and Results

In this study, physical phantom and clinical data were acquired by a single source, 320-row CT scanner (Aquilion ONE^{TM}, Toshiba, Otawara, Japan) with gantry rotation speeds of 350 ms/rotation. The performance of the proposed method was compared with the zero extrapolation method which pads zeros to all truncated projection.

Figures 6-7 show results with the physical phantom.

Figure 6 shows the axial view and coronary view of reconstruction with the zeros extrapolation method and the proposed method. It shows that the proposed method reduces the truncation artifacts significantly and recovers most of the truncated region.

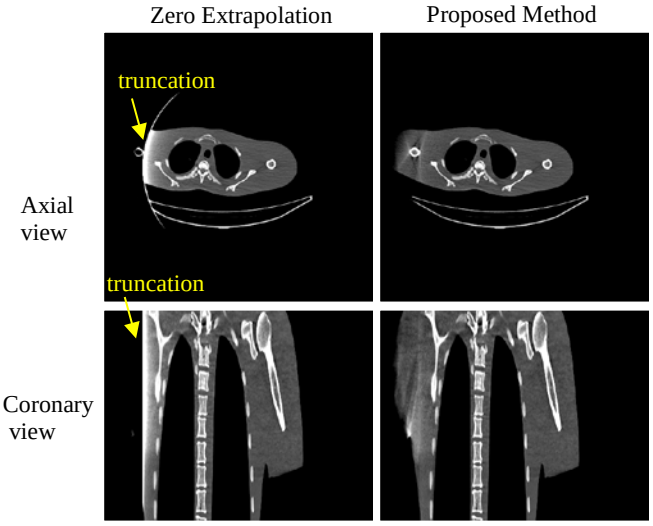


Fig. 6. Reconstructed images: top row is from axial view and bottom is from coronary view. From left to right columns are from zero extrapolation method and the proposed method, respectively. Window width and level are 400 and 50 HU, respectively.

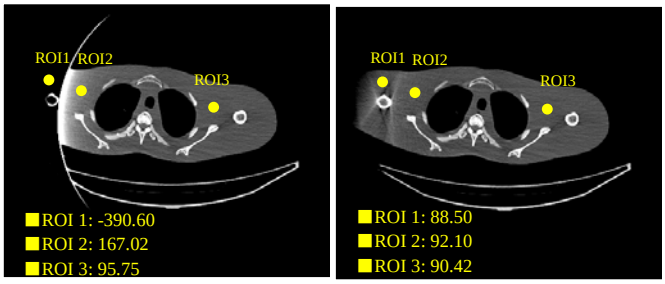


Fig. 7. HU value of three ROIs.

Figure 7 shows HU of three selected regions of interest. It shows that the zero extrapolation method has a huge bias in and near the truncation region (ROI1 and ROI2), while the proposed method produces more accurate and consistent HU values in the whole phantom.

Figures 8 shows results with clinical data. It can be observed that the proposed method recovers the truncation region with more accurate HU values especially when truncation is severe.

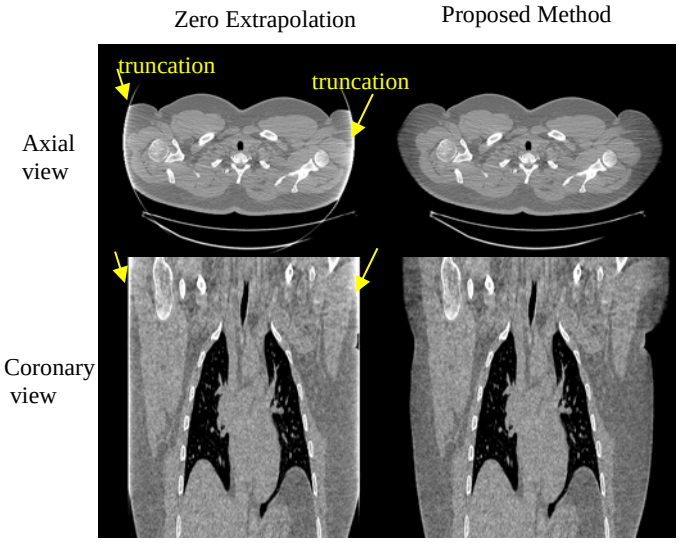


Fig. 8. Reconstructed images: top row is from axial view and bottom is from coronary view. From left to right columns are from zero extrapolation method and the proposed method, respectively. Window width and level are 400 and 50 HU, respectively.

IV. Conclusions

We have developed a new projection based data extrapolation method to reduce truncation artifacts. The proposed method includes a novel projection based edge detection scheme and a hybrid local polynomial fitting to fill the missing data. This proposed method has been tested with physical phantom and clinical data. Results shows that the proposed method can reduce truncation artifacts significantly and achieve overall consistent HU value even in severely truncated regions.

References

[1] G. Chen, “Correction of CT images for truncated or incomplete projections,” U. S. patent 7 050 528 B2

[2] R. Clackdoyle and L. Desbat,“Data consistency conditions for truncated fanbeam and parallel projections,” Med Phys, 42(2), pp 831-45 (2015)

[3] J. Boese and B. Scholz, “Method for correction of truncation artifacts in a reconstruction method for tomographic imaging with truncated projection data,” U.S. patent 7 660 380 B2 (2010).

[4] T. Flohr, B. Ohnesorge, and K. Schwarz,“Method for image reconstruction in a computed tomography device,” U.S. patent 6 307 909 (2001).

[5] J. Hsieh, et al,“A novel reconstruction algorithm to extend the CT scan field-of-view”, Medical Physics 31, 2385 (2004);

[6] J. Boese and B. Scholz, “Method for correction of truncation artifacts in a reconstruction method for tomographic imaging with truncated projection data,” U.S. patent 7 660 380 B2 (2010).

[7] H. Bruder, M. Sedlmair, K. Stierstorfer, T. Flohr, “High Definition Extended Field of View (HD FOV) Reconstruction in CT,” Radiological Society of North America 2011 Scientific Assembly and Annual Meeting, November 26 - December 2, 2011 ,Chicago IL.

Early investigation of CT thermometry as a tool to monitor the ablation zone during thermal ablation therapy

April M. Griffith, Frederic Noo, and Gabriel C. Fine

***Abstract*— In this paper, we investigate the use of CT as a reliable predictor of changes in temperature. This becomes clinically relevant when considering thermal ablation, a treatment which uses directed heating to induce cell death and has applications in a multitude of pathologies, including treating hepatocellular carcinoma. Utilizing the fundamental decrease in mass density associated with increased temperature, we hypothesize that we can use CT to reliably correlate changes in the temperature of water with changes in mass density.**

We demonstrate that changes in the density of water can be reliably correlated to changes in temperature. Furthermore, the difference in attenuation coefficient between body temperature and the temperature at which cell death occurs is 11.52 HU, which is a difference that can be clinically differentiated by diagnostic radiologists. While this data is preliminary, it implies that there is a fundamental density change that occurs in tissue during heating which can be measured using CT thermometry.

***Index Terms*—Computed tomography, Thermal Ablation, CT Thermometry**

I. Introduction

Percutaneous thermal ablations have become increasingly popular interventions to treat a wide range of pathologies, including liver, renal, lung, and soft tissue/bone disease [1]. There are multiple heat-based modalities in use, including microwave ablation (MW), radiofrequency ablation (RFA), laser ablation, and high frequency ultrasound (HIFU). These methods take advantage of coagulative necrosis, which occurs at approximately 55°C. Imaging guidance is critical in the delivery of thermal ablation treatments. Among the many imaging modalities, X-ray computed tomography (CT) is often preferred.

Current assessment of the ablation zone is based on manufacturer guidelines and differs depending on modality, manufacturer, specific probe being used, number of probes, and duration of burn. In the majority of cases, manufacturers predict ablation zones from *ex vivo* studies, which do not consider vascular perfusion or tissue type. Individual patient variables, such as the tissue type being ablated, tumor composition, and variances in clinician's technique, are not accounted for.

The inability to determine the true margin of the ablation zone is a major limitation of thermal ablation. For example, while MW ablation is up to 98% effective in the treatment of hepatocellular carcinoma (HCC) in tumors smaller than 2 cm, it is less effective in the treatment of larger tumors [2]. Incomplete ablation has been shown to precipitate local recurrence of cancer, resulting in the need for multiple procedures and increasing both the morbidity and mortality of the disease [2, 3]. Reliable evaluation of intraparenchymal temperature changes is critically needed to refine thermal ablation and decrease post-procedural morbidity.

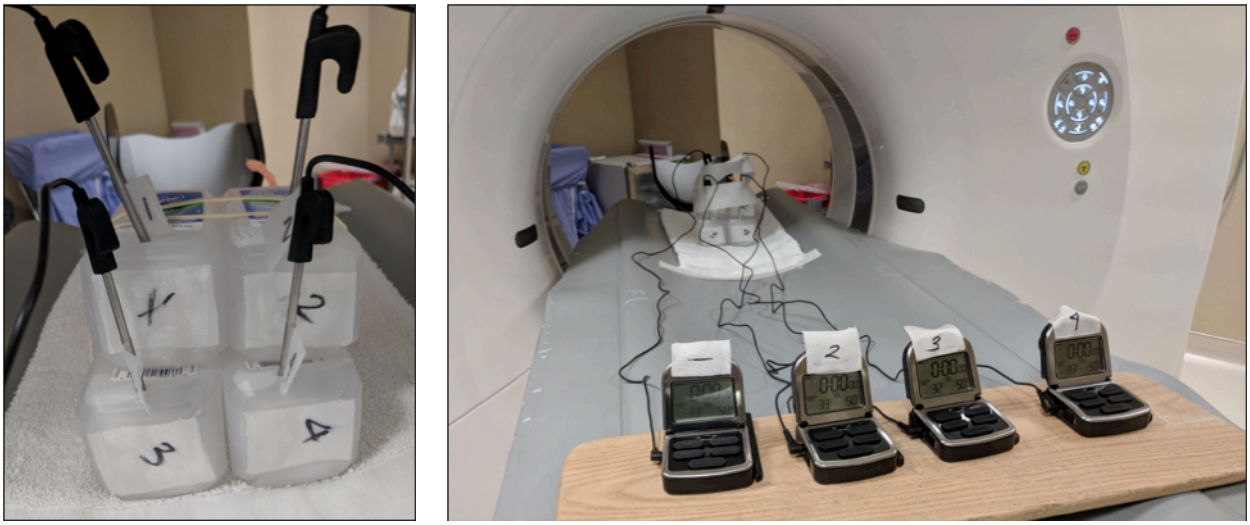
Researchers have investigated multiple approaches to better identify ablative zones, however none have provided a lasting solution. While heat-mapping has been explored in MRI [4] and quantitative ultrasound (US) [5], there are significant limitations associated with each modality which have prevented widespread clinical implementation. CT would be the preferred method for peri-procedural heat-mapping in thermal ablation as it is the modality already utilized for percutaneous probe placement. CT monitoring of temperature changes is viable because the linear attenuation coefficient of X-rays depends on mass density, which changes with temperature. For example, consider water: as it approaches boiling, it becomes less dense relative to colder water, and thus has a lower attenuation [6, 7]. This is in concert with the clinical CT findings seen in scans taken following thermal ablation, which show predictable differences in parenchymal attenuation [8].

Despite its potential to monitor temperature changes and thereby successful thermal ablation, CT thermometry is currently not a developed imaging tool. Our long-term goal is to remedy this important clinical deficiency. As part of this effort, in this work, we report on a study that tests the hypothesis that a reliable, clinically significant decrease in attenuation can be seen in water heated from body temperature (36°C) to the temperature at which cell death occurs (above 55°C). We hypothesize that the density changes that occur in water with increasing temperature will create reliable attenuation changes seen on CT.

II. Experimental Setup

Four bottles of water were stacked in a 2 x 2 arrangement as shown in Figure 1a. The water inside the bottles was initially heated to 75°C and then passively cooled to 30°C. This temperature range was chosen as it spans from the temperature of the human body, 36°C, to the excess temperature of 55°C, above which coagulative necrosis occurs. The temperature in each bottle was measured using independent probes that each had a degree of accuracy of +/- 0.5°C, as seen in Figure 1b.

April M. Griffith, Frederic Noo, and Gabriel C. Fine are with the Department of Radiology and Imaging Sciences, University of Utah, Salt Lake City, UT. Contact email: April.Griffith@hsc.utah.edu



(a) (b)

Fig. 1. (a) Arrangement of bottles with heated water. (b) Positioning of independent thermometers used to monitor temperature changes within each bottle. To monitor changes in the attenuation coefficient as a function of temperature, the bottles were placed within the bore of the CT scanner.

To assess the changes in the attenuation value of water over time as it cooled, the arrangement of bottles was placed within the bore of a diagnostic CT scanner. CT scans were taken after every 2°C temperature decrease until our full temperature range was obtained. The settings for the CT scans were 120 kV and a routine abdominal protocol.

DICOM files corresponding to 5 mm thick reconstructed slices with smooth ramp filter kernel typically used for soft tissue were processed using MatLab (MathWorks). A representative central slice passing through the four bottles was chosen. Within this slice, we selected an 81 x 81 region of interest within each bottle for analysis. This is demonstrated in Figure 2.

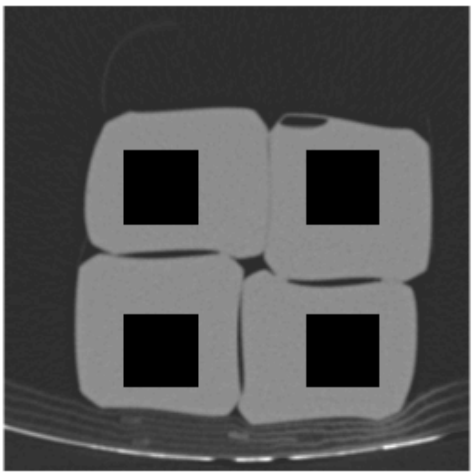


Fig. 2. CT slice selected through the four bottles with an 81 x 81 region of interest marked as a black square in each bottle. These regions of interest were used to assess the mean attenuation coefficient as a function of temperature.

III. Experimental Results

Figure 3 shows the evolution of the CT slice through the four bottles over time as observed in Hounsfield units (HU). Figure 4 shows the result of our analysis of the observed attenuation values as a function of temperature.

As can be seen in Figure 4, at 73°C (the temperature measured during the initial CT scan), the attenuation of water was -24.6 ± 0.4 HU. As the water cooled to 33°C (the temperature measured during the last scan), the attenuation increased to -5.0 ± 0.6 HU. The measurements were consistent across all 4 bottles. A linear model was fitted to all measurements; the R^2 value for this fit was 0.98, implying strong linear dependence. Using the model, we assess an 11.52 HU change is measurable between 36°C and 60°C. More precisely, using the Student distribution with 88 degrees of freedom, we obtain the following 95% confidence interval for the change in HU: 11.21-11.84.

IV. Conclusion

In this study, we showed that there is a measurable decrease in CT attenuation between body temperature and 60°C, the temperature at which cell death is guaranteed. Moreover, the observed 11.52 HU change in attenuation can be clinically differentiated by radiologists as demonstrated in Figure 3, indicating that parenchymal changes due to heating could be measured using CT.

Our current analysis has not assessed the radiation exposure required to successfully assess CT attenuation changes with temperature. Future work is needed to determine the minimum amount of radiation exposure required to detect changes of 10-12 HU for lesion sizes as small as 10 mm. Although radiation exposure is always a clinical concern, we note that the potential stochastic risks are outweighed in the targeted patient population, which primarily consists of patients receiving thermal ablation treatment for malignancies without other curative or palliative treatment options.

Future investigations into CT thermometry will also be aimed at exploring the effects that heating has on density changes in animal and human tissue. In particular, we plan to

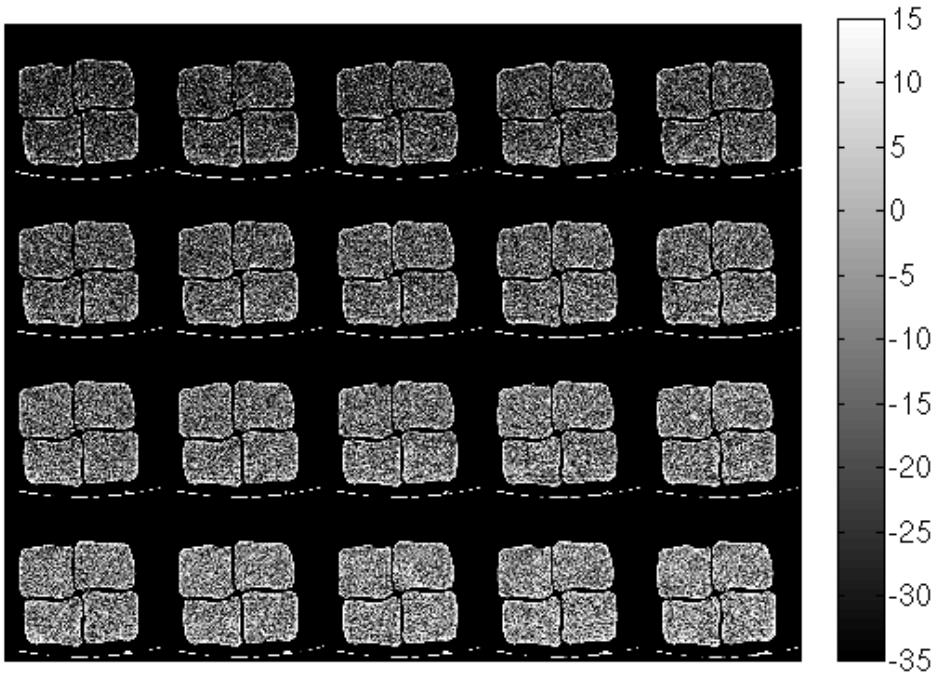


Fig. 3. CT reconstruction of the four bottles as a function of time. The time increases (and thus the temperature decreases) from row to row starting from the top row, and from left to right within each row. Grayscale window: [-35, 15] HU.

further explore the utility that CT has in evaluating temperature changes in *ex vivo* animal models, in *in vivo* animal models, and ultimately prospectively comparing proven ablation zones with intra-procedural findings in human subjects.

Reliable evaluation of the ablation zone using CT thermometry will allow for more accurate cancer therapy and decrease the increased morbidity and mortality associated with cancer recurrence and the need for multiple therapies.

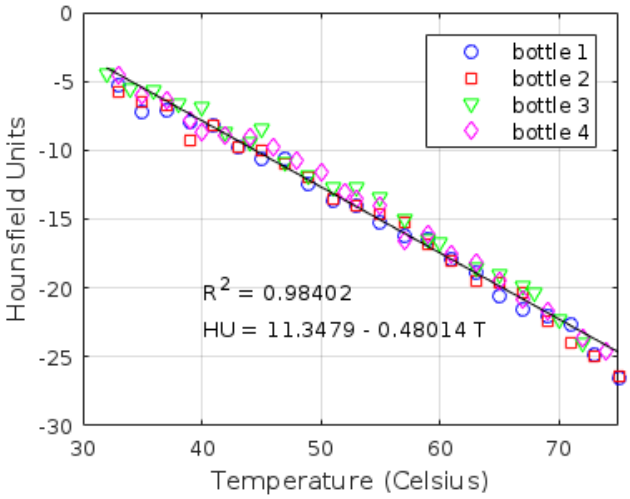


Fig. 4. Mean attenuation values as a function of temperature as measured within each bottle. Linear fitting was applied to all measurements as reported within the figure.

V. References

[1] Hinshaw JL, Lubner MG, Ziemlewicz TJ, Lee FT, Brace CL. Percutaneous tumor ablation tools: microwave, radiofrequency, or cryoablation--what should you use and why? Radiographics. 2014;**34**(5):1344-1362.

[2] Poggi G, Tosoratti N, Montagna B, Picchi C. Microwave ablation of hepatocellular carcinoma. World J Hepatol. 2015;**7**(25):2578-2589.

[3] Ferlay J, Soerjomataram I, Dikshit R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2015;**136**(5):E359-386.

[4] Zhu M, Sun Z, Ng CK. Image-guided thermal ablation with MR-based thermometry. Quant Imaging Med Surg. 2017;**7**(3):356-368.

[5] Samimi K, White JK, Brace CL, Varghese T. Monitoring Microwave Ablation of Ex Vivo Bovine Liver Using Ultrasonic Attenuation Imaging. Ultrasound Med Biol. 2017;**43**(7):1441-1451.

[6] Homolka, P., A. Gahleitner, and R. Nowotny, Temperature dependence of HU values for various water equivalent phantom materials. Phys Med Biol. 2002. **47**(16): p. 2917-23.

[7] Pandeya GD, Klaessens JH, Greuter MJ, et al. Feasibility of computed tomography based thermometry during interstitial laser heating in bovine liver. Eur Radiol. 2011;**21**(8):1733-1738.

[8] Park MH, Rhim H, Kim YS, Choi D, Lim HK, Lee WJ. Spectrum of CT findings after radiofrequency ablation of hepatic tumors. Radiographics. 2008;**28**(2):379-390; discussion 390-372.

Volume-of-interest CT imaging with dynamic beam filtering using multiple aperture devices

Wenying Wang, Grace J. Gang, Andrew Mao, Alejandro Sisniega, Jeffrey H. Siewerdsen, and J. Webster Stayman

***Abstract*—Interior tomography is promising approach for retaining high quality CT images within a volume-of-interest (VOI) while reducing the total patient dose. A static collimating filter can only image a centered symmetric VOI, which requires careful patient positioning and may be suboptimal for many clinical applications. Multiple aperture devices (MADs) are an emerging technology based on sequential binary filters that can provide a wide range of fluence patterns that may be adjusted dynamically with relatively small motions. In this work, we introduce a general approach for VOI imaging using MAD-based fluence field modulation (FFM). Physical experiments using a CT test bench are conducted illustrating off-center x-ray beam control for imaging the spine in an abdominal phantom. Image quality and dose metrics are computed for both standard full-field CT and VOI CT. We find that the image quality within the VOI can be preserved for VOI CT with a significant drop in integral dose as compared with a standard full-field protocol.**

I. INTRODUCTION

X-ray CT is widely used for diagnosis, disease monitoring, and in interventional procedures. However, increased usage has raised concern about excessive radiation exposure. In many CT applications, only a limited volume is required for assessment. For example, target positions are generally known in single-organ studies (e.g. cardiac imaging) and in interventional procedures focused on particular sites (e.g. spine imaging). Despite the need for only local structural information, CT scans routinely cover the entire lateral extent of the patient. Volume-of-interest (VOI) scanning gives the opportunity to spare dose to surrounding tissues, while providing the information required inside the VOI [1].

Implementing VOI imaging can be challenging for arbitrary VOIs. Many VOI studies have used static collimation to restrict the X-ray beam width [2], or static zonal filtering to decrease the fluence intensity outside the VOI [3]. However, these methods can only collect centered, cylindrical VOIs (using standard source-detector trajectories) limiting practical application. Several attempts have been made to increase versatility in VOI imaging. Kolditz *et al.* took advantage of mobility of C-arm CT system and realized non-centered VOI imaging [4]; and dynamic collimation has been applied to achieve adaptable fluence patterns with off-centered VOIs. [5]

In previous work, we introduced a novel filtering scheme using fine-scale binary filters called multiple aperture devices (MADs) [6]. In this scheme, two MADs are used in series. Translations of the dual-MAD filtering system as a whole shifts the peak of the incident beam and small relative displacements between two components changes the width of transverse fluence profile. This new filtering system provides the capability for sophisticated fluence-field modulation (FFM) based on the specific imaging task. In this work, we present a dual-MAD trajectory design for an arbitrary VOI imaging task. VOI CT acquisition is conducted on a CT bench and truncated projection data are reconstructed using an extrapolation-based method. Image quality and dose are assessed for both standard full-field CT and MAD-based VOI CT. Dose is estimated using Monte Carlo simulation, bare-beam fluence measurements, and an estimated phantom density map.

Department of Biomedical Engineering, Johns Hopkins University. email: web.stayman@jhu.edu.

II. MATERIALS AND METHODS

A. Filter Trajectory Design for arbitrary VOI

MADs-based FFM provides much flexibility for general VOI imaging. We will describe the desired VOI using a 3D binary mask, m_{VOI}, where voxels within the VOI have value 1, and outside the VOI are 0. Furthermore, let $m_{\text{VOI}} \in \mathbb{R}^N$ be a column vector with N denoting the total number of voxels in a full-field image volume. At each projection angle θ, a detector pixel mask associated with the VOI projection area is

$$g(\theta) = \Theta\{\mathbf{A}(\theta) m_{\text{VOI}} > 0\} \tag{1}$$

where $\mathbf{A}(\theta) \in \mathbb{R}^{P \times N}$ is the system matrix at projection angle θ (with P being the total number of pixels on detector.) The logical operator, $\Theta\{\cdot\}$, returns 1 if the argument is true and 0 if false. Thus, $g(\theta) \in \mathbb{R}^P$ is the binary mask describing the data required to reconstruct the VOI. (We note that such data is not generally mathematically complete [7].)

Current MAD-based FFM has focused only on modulation of the X-ray beam in the transverse direction (along the detector). We denote $g(\theta)$ reshaped to a 2D mask corresponding to the face of the detector as $G_\theta(u, v) \in \mathbb{R}^{U \times V}$ (where u and v are detector coordinates in the transverse and transaxial directions, respectively; U/V denote the number of detector column/rows on the detector panel; and $P = U \times V$). Thus all column indices required for obtaining a desired VOI (illustrated in Fig. 1) may be written as

$$\{u_{\text{VOI}}\}(\theta) = \{u | \textstyle\sum_v G_\theta(u, v) > 0\}. \tag{2}$$

To find the optimal actuation of MAD filters for VOI imaging, we consider two requirements: 1) Maximize the fraction of fluence transmitting through VOI region. And, 2) constrain the minimum filter transmissivity passing through the VOI region to be no less than some threshold f_{min}. (E.g., There is some minimal level of fluence required to obtain

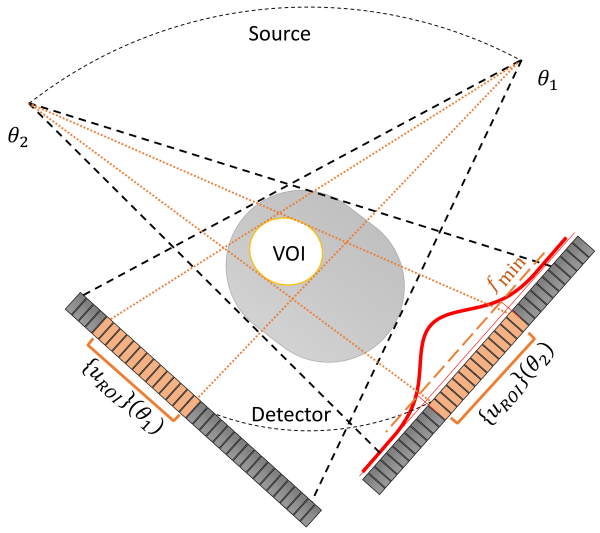


Fig. 1: MAD-based VOI scan diagram and notation: $\{u_{\text{VOI}}\}(\theta)$ is defined as a set of detector indices based on projection of the VOI at each projection angle θ; the incident fluence to $\{u_{\text{VOI}}\}(\theta)$ is constrained to be above $f_{\min}$.

useful images in the VOI.) Mathematically, we may then write full FFM objective function as

$$(\hat{t}_0, \hat{t}_1)(\theta) = \arg \max_{(t_0,t_1)\in\mathcal{T}_\theta} \frac{\sum_{u\in\{u_{\text{VOI}}\}(\theta)} f(u;t_0,t_1)}{\sum_u f(u;t_0,t_1)}, \quad (3)$$

$$\mathcal{T}_\theta = \{(t_0,t_1)| \min_{u\in\{u_{\text{VOI}}\}(\theta)} f(u;t_0,t_1) > f_{\min}\}. \quad (4)$$

where $f(u;t_0,t_1)$ denotes all fluence field patterns (a function of u) achievable from the dual-MAD system which is controlled through two actuation parameters: t_0, the translation of MAD0, and t_1 the translation of MAD1 relative to MAD0. That is, the parameter pair (t_0,t_1) defines the fluence profile $f(u;t_0,t_1)$ which is determined experimentally through a pre-calibration using an exhaustive sweep through varying (t_0,t_1).

The above optimization is complicated by the periodic nature of $f(u;t_0,t_1)$. That is, the MAD filters themselves are quasi-periodic which means that fluence patterns will repeat for motions on the order of one MAD period. To avoid difficulties in optimization and to enforce actuations with relatively small displacements between two acquisition frames, we make the following modifications. Given a circular orbit, the projection of VOI center may be approximated by a sinusoidal curve. Thus, we perform an initial optimization using (3) and (4) and then perform a sinusoidal fitting $\tilde{t}_0(\theta)$ to $\hat{t}_0(\theta)$. Subsequently, we perform a second optimization, constraining the final displacements $\hat{t}_0^{\text{fit}}$ to be close to the sinusoidal approximation, with

$$(\hat{t}_0^{\text{fit}}, \hat{t}_1^{\text{fit}})(\theta) = \arg \max_{(t_0,t_1)\in\mathcal{T}_\theta^{\text{fit}}} \frac{\sum_{u\in\{u_{\text{VOI}}\}(\theta)} f(u;t_0,t_1)}{\sum_u f(u;t_0,t_1)}, \quad (5)$$

$$\mathcal{T}_\theta^{\text{fit}} = \{(t_0,t_1)\in\mathcal{T}_\theta : |\hat{t}_0^{\text{fit}}(\theta) - \tilde{t}_0(\theta)| < \Delta t\}. \quad (6)$$

In our test-bench experiments, we select $f_{\min} = 0.2$ and $\Delta t = 0.1$ mm (this is small relative to the MAD period of 0.9 mm).

B. Truncated CT reconstruction

Projection data acquired using MAD filters require careful calibration for a number of physical effects. This includes sensitivity to focal spot changes and spectral effects. A modified forward model is employed to account for these effects. Specifically,

$$\bar{y}(\theta) = g_{I_0} g_D g_M(\theta) e^{-\alpha(\theta) l} \quad (7)$$

where g_{I_0} is the emitted X-ray fluence strength distribution, g_D is the detector sensitivity map, and $g_M(\theta)$ is the designed MAD modulation profile (including calibrations for focal spot positioning). To compensate for spectral effects, we estimate a first-order spectral correction term $\alpha(\theta)$ from variable thickness slabs, which may be applied as a data correction (much like ray-based beam hardening corrections). Inverting the above forward model [8], we compute a sinogram $l(u(\theta),v,\theta)$ from noisy measurements y.

Since VOI data is highly truncated, direct FDK reconstructions suffer severe artifacts within the ROI. Here we extrapolate the sinogram along the transverse direction using a quadratic model [9] to alleviate the influence of truncation within the VOI. Specifically, for each line integral profile on the v^{th} row at one projection angle θ, extrapolations are performed on each truncated side independently,

$$l_{\text{ext}}(u) = \alpha(u-u_0(\theta))^2 + \beta(u-u_0(\theta)) + \gamma \quad (8)$$

where u_0 is the index of the nearest truncation point. Assuming continuity at the truncation point ($u = u_0$) and $l = 0$ at the detector boundary $u = u_b$, and we derive

$$\beta(v,\theta) = \frac{\partial l(u,v,\theta)}{\partial u}|_{u=u_0(\theta)}, \quad (9)$$

$$\gamma(v,\theta) = l(u_0(\theta),v,\theta), \quad (10)$$

$$\alpha(v,\theta) = -\frac{\beta(v,\theta)(u_b-u_0(\theta)) + \gamma(v,\theta)}{(u_b-u_0(\theta))^2}. \quad (11)$$

With noisy line integral profiles, $\beta(v,\theta)$ and $\gamma(v,\theta)$ are estimated with a local linear regression. Extrapolated line integrals are enforced to be non-negative, and are used as a direct input into a standard FDK algorithm to reconstruct the VOI.

C. Experiment Setup

To investigate the MAD-based VOI imaging feasibility, experiments were conducted on a cone-beam CT test bench with dual MAD filters (Fig. 2a). A circular region around the spine in a QRM abdomen phantom (Fig. 2b) is chosen as the VOI. Note that MAD-based fluence field patterns allow for control of beam width, permitting noncircular VOI.

In addition to MAD-VOI scans, reference "high"-dose and low-dose scans without the MAD filters are acquired for comparison. In each scan, 720 frames are acquired in a single rotation with constant angular steps. Experiments were performed at 100 kVp, 35 mA. For the high-dose and low-dose scans, pulse widths of 18.2 ms and 3.2 ms were used, respectively. This corresponded to an incident fluence, I_0, of 8.1×10^5 photons per pixel and 1.58×10^5 photons per pixel for the high- and low-dose scans, respectively (estimated based on bare-beam variance measurements).

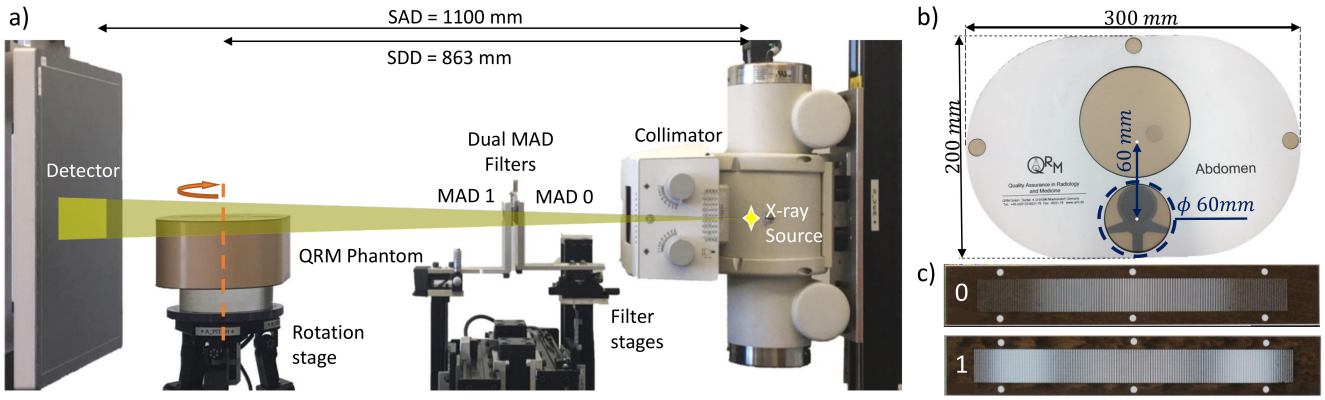


Fig. 2: a) Test bench experiment setup. b) Abdomen QRM phantom. c) Photograph of manufactured dual MADs. [10]

MAD-VOI acquisition used the same tube settings as the high-dose scan. After appropriate gain-correction and constant scatter correction (with single scatter-to-primary-ratio tuned to remove the cupping artifact), Feldkamp (FDK) reconstruction was performed with a Hamming window, 0.4 Nyquist frequency cutoff, and a $120 \times 120 \times 10$ volume with cubic 0.5 mm voxels.

For MAD-VOI and low-dose experiments, we used Monte-Carlo simulation to estimate the dose distribution. [11] The input fluence fields were acquired using bare-beam measurements scaled by the estimated I_0 values. Contrast-to-noise ratios and dose are reported for all three experiments.

III. Results

A. Filter Trajectory Design

The estimated filter trajectory based on the proposed optimization for the spine VOI is shown in Fig. 3a. The blue dashed line shows the MAD0 translations at each projection angle, and the red curve shows the MAD1 translation relative to the MAD0 position. This designed trajectory results in the bare-beam dynamic fluence-field modulation shown in Fig. 3b.

As the long axis of QRM phantom is 300 mm, the full projection data will occupy over 90% of the detector transversely. In contrast, the diameter of the VOI is only 60 mm, ∼1/5 of the long axis size. The designed modulation profile only illuminates ∼1/4 of the detector with filter transmissivity larger than 20% suggesting a good match with the design objective.

The VOI in this study is cylindrical and off-center. Thus, the x-ray beam center is shifted sinusoidally along with the axial center of the VOI. The width of beam is nearly constant over projection angle. Relative translation between the two MADs is required even for constant width designs due to obliquity effects that narrow beam profiles with an off-centered MAD0.

B. Image quality analysis and dose calculation

MAD-VOI was processed according to the modified forward model (7) with extrapolation to approximate a complete sinogram (Fig. 4a). FDK reconstruction of the VOI for full-field high-dose, full-field low-dose, and MAD-VOI scans are shown in Figures 4b, 4c, 4d, respectively. The noise distribution was estimated by computing the standard deviation over 10 (axial) slices and are shown in Figures 4e, 4f, 4g. We observe that the MAD-VOI reconstruction has lower noise than full-field low-dose result, and the CT number bias appears smaller. The MAD-VOI reconstruction has some mild structured noise and ring artifacts. We conjecture that

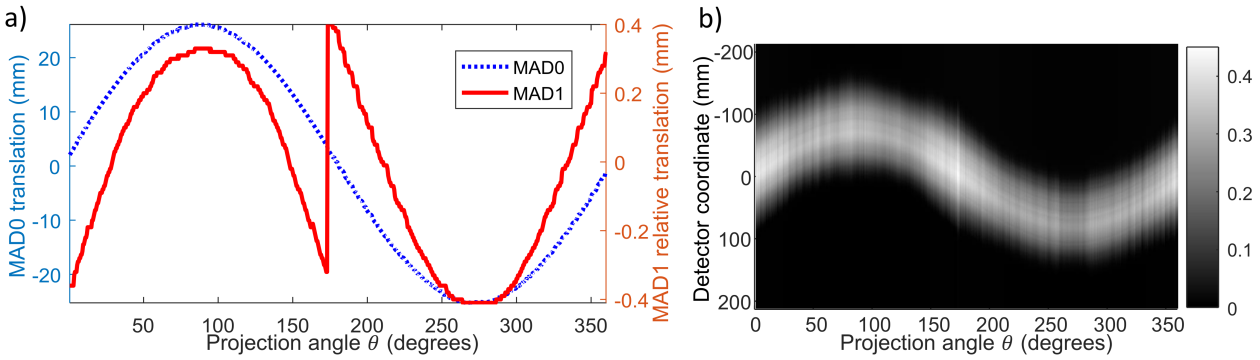


Fig. 3: a) Estimated actuation trajectories for MAD-VOI acquisition. b) Measured bare-beam fluence for the off-center VOI.

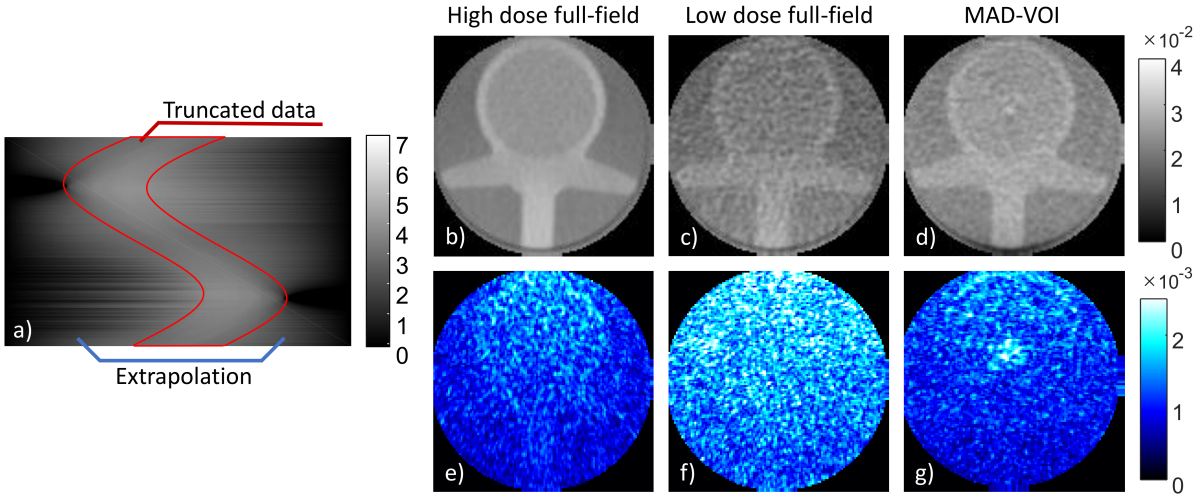


Fig. 4: a) Central slice extrapolated sinogram. The truncated data is marked in red, while the extrapolation region is marked in blue. Upper row: FDK reconstructions of the VOI in the: b) Full-field high-dose scan; c) Full-field low-dose scan; and d) MAD-VOI scan. Bottom row: e)-g): Reconstruction noise based on the standard deviation over 10 slices.

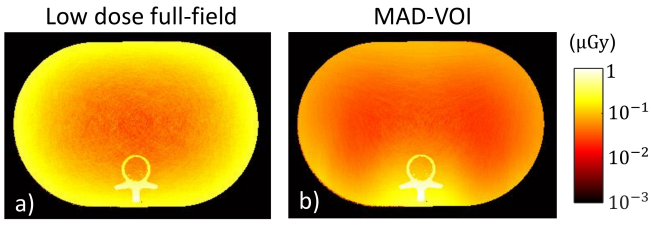


Fig. 5: Absorbed dose distribution in the a) full-field low-dose scan and the b) MAD-VOI scan. For optimal display effect, the colormap is shown in log-scale.

these residual imperfections are due to errors in the MAD gain correction process.

The QRM phantom provides uniform regions in "spinal bone" and "soft tissue" interior to the spine. We computed the contrast-to-noise ratio (CNR) between these two tissues divided by noise in the background.

$$\mathrm{CNR} = \frac{\mu_{\text{bone}} - \mu_{\text{soft-tissue}}}{\sigma_{\text{background}}} \tag{12}$$

The results of the Monte Carlo simulation of dose distribution maps in MAD-VOI and full-field low-dose scans are shown in Fig. 5. In the MAD-VOI scan, the dose is accumulated more highly in the VOI rather than the more even distribution along peripheral tissues in the full-field scan.

CNR, integral dose and dose accumulated exterior to the VOI estimates for all methods are summarized in Table I. Note that the integral dose in MAD-VOI scan is reduced by 40% while the CNR is not decreasing (as compared with the low-dose scan). Also note that while the (reference) high-dose scan and the MAD-VOI scan used the same x-ray technique, the dual-MAD filter transmissivity peaks around 40% and one should not expect to achieve the same CNR. The comparison between the full-field low-dose and MAD-VOI scans suggests that MAD-VOI has the potential to retain high quality VOI images while reducing integral dose.

TABLE I: CNR and Integral Absorbed Dose

Experiment	CNR	Dose (mGy)		
		Outside VOI	In VOI	Total
High-dose full-field	7.27	29.3	2.87	32.1
Low-dose full-field	2.73	5.71	0.556	6.27
FFM-VOI	2.90	2.72	0.784	3.51

IV. Discussion

In this work, we apply novel MAD-based FFM to VOI CT imaging. These dynamic filters can tailor the beam width and center for non-centered and irregular-shaped VOIs (to be explored additionally in ongoing studies). A preliminary bench-top study on a QRM phantom shows that MAD-VOI imaging retains good image quality while significantly decreasing the integral dose as compared with full-field exposures.

These preliminary studies focused on relatively simple extrapolation-based FDK for truncated data. Future work will employ more sophisticated (e.g., Hilbert-transform methods [7] [12] and model-based reconstruction [13]). Similarly, ongoing studies are considering combination of fluence-field modulation and tube-current modulation for additional control of dose and image quality. Other future work includes the development of new MAD filters for more varied fluence patterns including bimodal beam patterns and profiles that are peaked but non-zero outside a region of interest. Such patterns may facilitate more sophisticated scans including two disconnected VOIs (e.g. lungs, kidneys) and untruncated scans with two dose/image quality levels inside/outside a specified VOI [14].

ACKNOWLEDGMENT

This work was supported, in part, by NIH grant U01EB018758.

REFERENCES

[1] G. Wang, "The meaning of interior tomography," *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, pp. 5764–5767, 2011.

[2] L. Chen, C. C. Shaw, M. C. Altunbas, C. J. Lai, X. Liu, T. Han, T. Wang, W. T. Yang, and G. J. Whitman, "Feasibility of volume-of-interest (VOI) scanning technique in cone beam breast CT - A preliminary study," *Medical Physics*, vol. 35, no. 8, pp. 3482–3490, 2008.

[3] R. Chityala, K. R. Hoffmann, S. Rudin, and D. R. Bednarek, "Region of interest (ROI) computed tomography (CT): comparison with full field of view (FFOV) and truncated CT for a human head phantom," vol. 5745, no. 1, apr 2005, p. 583.

[4] D. Kolditz, Y. Kyriakou, and W. A. Kalender, "Volume-of-interest (VOI) imaging in C-arm flat-detector CT for high image quality at reduced dose," *Medical Physics*, vol. 37, no. 6, pp. 2719–2730, 2010.

[5] E. Pearson, X. Pan, and C. Pelizzari, "Dynamic intensity-weighted region of interest imaging for conebeam CT," *Journal of X-Ray Science and Technology*, vol. 24, no. 3, pp. 361–377, jun 2016.

[6] J. W. Stayman, A. Mathews, W. Zbijewski, G. Gang, J. Siewerdsen, S. Kawamoto, I. Blevis, and R. Levinson, "Fluence-field modulated x-ray CT using multiple aperture devices," in *Proc SPIE Int Soc Opt Eng*, vol. 28, no. 10, mar 2016, p. 97830X.

[7] M. Defrise, F. Noo, R. Clackdoyle, and H. Kudo, "Truncated Hilbert transform and image reconstruction from limited tomographic data," *Inverse Problems*, vol. 22, no. 3, pp. 1037–1053, 2006.

[8] G. J. Gang, A. Mao, J. H. Siewerdsen, and J. W. Stayman, "Dynamic Fluence Field Modulation with Multiple Aperture Devices: Initial Results from an Experimental Cone-Beam CT Bench," in *International Conference on Image Formation in X-Ray Computed Tomography*, 2018, p. (submitted).

[9] S. Zhao, K. Yang, and X. Yang, "Reconstruction from truncated projections using mixed extrapolations of exponential and quadratic functions," *Journal of X-Ray Science and Technology*, vol. 19, no. 2, pp. 155–172, 2011.

[10] A. J. Mathews, G. Gang, R. Levinson, W. Zbijewski, S. Kawamoto, J. H. Siewerdsen, and J. W. Stayman, "Experimental evaluation of dual Multiple Aperture Devices for Fluence Field Modulated X-Ray Computed Tomography," *SPIE Medical Imaging (accepted paper)*, no. March 2017, pp. MI101–361, 2017.

[11] A. Sisniega, W. Zbijewski, A. Badal, I. S. Kyprianou, J. W. Stayman, J. J. Vaquero, and J. H. Siewerdsen, "Monte Carlo study of the effects of system geometry and antiscatter grids on cone-beam CT scatter distributions," *Medical Physics*, vol. 40, no. 5, pp. 1–19, 2013.

[12] L. Li, K. Kang, Z. Chen, L. Zhang, and Y. Xing, "A general region-of-interest image reconstruction approach with truncated Hilbert transform," *Journal of X-Ray Science and Technology*, vol. 17, no. 2, pp. 135–152, 2009.

[13] H. Yu and G. Wang, "Compressed sensing based interior tomography," *Physics in Medicine and Biology*, vol. 54, no. 9, pp. 2791–2805, 2009.

[14] S. Bartolac, S. Graham, J. Siewerdsen, and D. Jaffray, "Fluence field optimization for noise and dose objectives in CT," *Medical Physics*, vol. 38, no. SUPPL.1, 2011.

Extrapolation of Truncated C-arm CT Data using Grangeat-based Consistency Measures

Daniel Punzet, Robert Frysch and Georg Rose

***Abstract*—A common key problem of cone-beam computed tomography (CBCT) is the reconstruction of tomographic images from incomplete projection data. The issue of acquired data sets containing only a part of the patient's projection is referred to as truncation. Truncation occurs when the irradiated scan field of view (SFOV) is constrained to not contain the whole patient but only a limited volume-of-interest (VOI). VOI imaging is a technique mainly used as a means of patient dose reduction, typically in an interventional scenario. The truncated data gives rise to image artifacts in the reconstructed images, though, decreasing the effectively usable VOI and degrading image quality in general. A typical strategy to counter these is the extrapolation of the measured region by some smooth function or by incorporating other a-priori information. In this work we propose an extrapolation algorithm based on optimization of data consistency conditions (DCCs) which can be obtained from Grangeat's fundamental relation. Our aim is to demonstrate the suitability of Grangeat-based consistency measures for truncation problems. To support the optimization we further incorporate available a-priori information into the choice of an appropriate extrapolation model. Our preliminary results show that an improvement in reconstructed image quality is achieved even for less-than-optimal extrapolations.**

I. Introduction

With the rising acceptance of C-arm-based cone-beam CT (CBCT) systems, a key problem of CBCT remains to be the reconstruction of tomographic images from incomplete projection data. The issue of acquired data sets containing only a part of the patient's projection is referred to as truncation. Truncation occurs when the scan field of view (SFOV) is constrained to not contain the whole patient, but only a limited volume-of-interest (VOI). This technique is called VOI imaging and is commonly used as a means of patient dose reduction. VOI imaging is typically applied for intraoperative 3D imaging during locally constrained interventions. Besides in VOI imaging, truncation can generally occur due to the limited detector size of C-arm based CT systems. The truncation of the projection data leads to ring-like cupping artifacts and degraded contrast in the reconstructed images, though, reducing the effectively usable VOI and image quality in general.

A typical strategy to counter these artifacts is the extrapolation of the truncated projection data. For truncation problems in conventional (2D) CT many different strategies have been proposed, ranging from simple data mirroring [1] and water cylinder extrapolation [2] to the use of the Helgason-Ludwig consistency conditions (HLCC) [3]. For cone-beam CT, there also exists a data consistency condition (DCC) based on the fundamental relation of Grangeat. So far, this DCC has mainly been used for motion compensation [4], beam-hardening correction [5] and detector calibration [6].

In this work we investigate the suitability of the Grangeat data consistency measure for use in truncation problems. First we introduce a consistency measure based on the Grangeat DCC and investigate the general effect of truncation on this measure. Furthermore we propose an algorithm that utilizes a-priori information about the intervention to extrapolate the truncated data with a simple extrapolation model which is then optimized regarding the Grangeat-based consistency measure. Finally, some resulting reconstructions are presented in order to evaluate the achievable image quality enhancement.

D. Punzet, R. Frysch and G. Rose are with the Institute for Medical Engineering, Otto-von-Guericke University, Magdeburg, Germany and also with the Forschungscampus STIMULATE, Otto-von-Guericke University, Magdeburg, Germany

send correspondence to: daniel.punzet@ovgu.de

II. Methods

A. Grangeat-based consistency measures

The 3D Radon transform is an extension of the 2D Radon transform which instead of rays integrates over planes. Given the volume $f(\boldsymbol{r})$ ($\boldsymbol{r} \in \mathbb{R}^3$) to be imaged, the 3D Radon transform is defined as

$$\mathcal{R}_3(\hat{\boldsymbol{n}}, \xi) \coloneqq \iiint_{\mathbb{R}^3} f(\boldsymbol{r})\delta(\boldsymbol{r}\cdot\hat{\boldsymbol{n}} - \xi)\ \mathrm{d}\boldsymbol{r} \coloneqq \Gamma(\hat{\boldsymbol{n}}, \xi)\,. \tag{1}$$

Here, $(\hat{\boldsymbol{n}}, \xi)$ parametrizes a plane by its normal unit vector $\hat{\boldsymbol{n}}$ and its distance ξ to the origin of the patient's world coordinate system. δ denotes the Dirac delta function.

What is measured in a CBCT acquisition is called the X-Ray transform. It is defined by

$$\mathcal{X}(\boldsymbol{C}_i, \hat{\boldsymbol{m}}) \coloneqq \int_0^\infty f(\boldsymbol{C}_i + \hat{\boldsymbol{m}}\lambda)\ \mathrm{d}\lambda \coloneqq g_i(\hat{\boldsymbol{m}})\,. \tag{2}$$

With $\boldsymbol{C}_i$ being the *i*'th position on the x-ray source trajectory lying in the plane given by $(\hat{\boldsymbol{n}}, \xi)$, which in the following will be assumed a circular trajectory. And $\hat{\boldsymbol{m}} \in S^2$ pointing in the direction of a ray starting from $\boldsymbol{C}_i$.

Grangeat's fundamental relation formulates a connection between the 3D Radon transform over planes and the X-Ray transform along rays via the intermediate function [7] [8]:

$$S_i(\hat{\boldsymbol{n}}) \coloneqq -\int_{S^2} \mathrm{d}\hat{\boldsymbol{m}}\delta'(\hat{\boldsymbol{m}}\cdot\hat{\boldsymbol{n}})g_i(\hat{\boldsymbol{m}}) = \frac{\partial}{\partial\xi}\Gamma(\hat{\boldsymbol{n}}, \xi)|_{\xi=\boldsymbol{C_i}\cdot\hat{\boldsymbol{n}}}\,. \tag{3}$$

In the context of projective geometry, this has also been derived under the name of epipolar consistency (ECC) in [9].

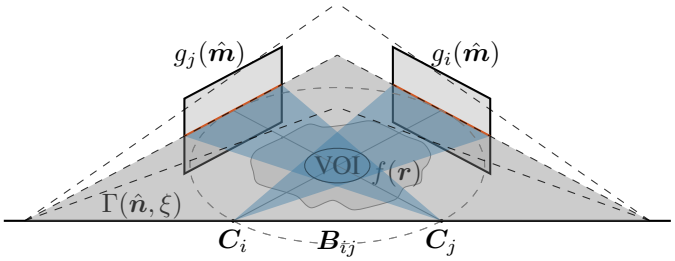


Fig. 1: Fan of Radon planes around the base line $\boldsymbol{B}_{ij}$ and corresponding cone-beam projections of 2 source positions $\boldsymbol{C}_i$ and $\boldsymbol{C}_j$ in a VOI imaging scenario with transversal truncation. In the case of transversal truncation, the support of $f(\boldsymbol{r})$ (dark gray) is still completely included in the Radon plane (gray), but not in the limited field of view of the line integrals of the cone-beam projections (blue). Therefore the line integrals on the intersection of the detector and the radon plane (orange) differ between the 2 projections.

Grangeat's relation shows that the intermediate function depends only on the Radon plane $(\hat{\boldsymbol{n}}, \xi)$ and that it can also be identified as a line integral on the cone-beam projections from a corresponding source position $\boldsymbol{C}_i$. Since we assumed circular source trajectories, there always exists another source position $\boldsymbol{C}_j$ in the same plane, which therefore, following Grangeat's relation, results in the same intermediate function

$$S_i(\hat{\boldsymbol{n}}) = S_j(\hat{\boldsymbol{n}})\,. \tag{4}$$

This redundancy, however, only holds as long as no transversal truncation occurs, since in the presence of transversal truncation the support of $f(\boldsymbol{r})$ is still included completely in the integral over the Radon plane, but not in the cone-beam projections (see Figure 1). For planes that are very far from the central plane (indicated as dashed in Fig. 1), also axial truncation becomes relevant, especially for increasingly orthogonal pairs of projections. With increasing truncation the mismatch $S_i(\hat{\boldsymbol{n}}) \neq S_j(\hat{\boldsymbol{n}})$ also increases. Therefore, the difference $|S_i(\hat{\boldsymbol{n}}) - S_j(\hat{\boldsymbol{n}})|$ or similar measures can be used as a consistency measure between a pair of projections. For the extrapolation algorithm proposed in the next section, we decided to use the cross-correlation of $S_i(\hat{\boldsymbol{n}})$ and $S_j(\hat{\boldsymbol{n}})$ as we experienced this measure to be more robust.

B. Extrapolation of truncated projection data

The extrapolation procedure we propose takes the set of truncated projections and tries to extrapolate them in the projection domain with projections of an appropriate extrapolation model (for the choice of the extrapolation model see section II-C). The parameter vector $\boldsymbol{\Xi}$ of the extrapolation model is chosen by optimization with respect to the average consistency measure between N pair-wise compared projections:

$$\begin{aligned} \max_{\boldsymbol{\Xi}} \quad & \frac{1}{N}\sum_{i=1}^{N}\sum_{j\neq i}^{N}\rho_{ij} \qquad j = 2,3,\ldots,N \\ \text{s.t.} \quad & \boldsymbol{\Xi} \in \boldsymbol{\Omega}\,. \end{aligned} \tag{5}$$

Where ρ_{ij} denotes the zero-normalized cross-correlation (ZNCC) defined as

$$\rho_{ij} \coloneqq \frac{1}{K-1}\sum_{k=1}^{K}\left(\frac{S_i^k - \overline{S_i}}{\sigma_{S_i^k}}\right)\left(\frac{S_j^k - \overline{S_j}}{\sigma_{S_j^k}}\right) \tag{6}$$

and $\boldsymbol{\Omega}$ being the space of all valid solutions, which is constrained to include only models which roughly approximate the human body parts they are supposed to model, e.g. a head.

Since the computation of the consistency measure is rather expensive, the choice of the optimization algorithm is crucial. We used an implementation of the differential evolution (DE) algorithm [10] as we experienced it to be capable of effectively searching the big space of possible solutions in this case, giving reliable results. DE is a stochastic global optimization based on evolutionary algorithms. However, more investigation in the behavior of the optimization for the given problem is needed and there is likely another algorithm to be found which is suited better for the task at hand.

C. Incorporating a-priori information

In order for the optimization to succeed in finding the global maximum, it relies on a good initial guess. To achieve this, we focussed on the internal reconstruction problem, where the VOI is completely embedded in the patient. This truncation scenario is common for procedures like stent placement and heart-catheter interventions. Typically, for this kind of intervention the location and surroundings of the VOI are roughly known.

We can now make use of this a-priori information to choose a simple but appropriate extrapolation model based on the location of the intervention, e.g. an ellipsoid to model the head for stroke interventions, or a cylinder to model the thorax for heart-catheter interventions.

For a simple, homogeneous ellipsoid extrapolation model the optimization vector $\boldsymbol{\Xi}$ consists of 7 parameters:

$$\boldsymbol{\Xi} \coloneqq [x_0, y_0, z_0, a, b, c, \phi]\,. \tag{7}$$

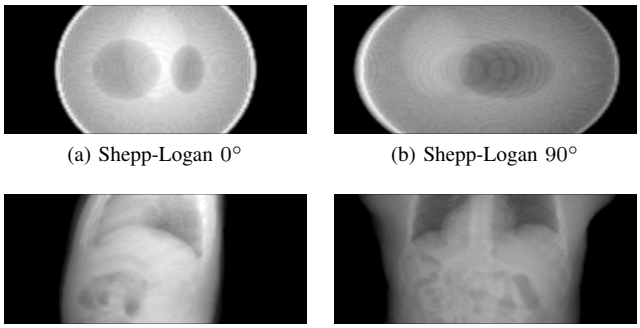
With x_0, y_0, z_0 specifying the position of the ellipsoid center, a, b, c the semiaxes and ϕ the rotation.

In order to reduce the space of possible solutions for the optimization, the absorption coefficient is not being optimized by the algorithm. Instead, we introduce a smoothness-constraint between the truncated projections and the extrapolation which has to be satisfied. For this, the absorption coefficient of the ellipsoid is chosen so that the average discontinuity in gray values between truncated projections and extrapolation from all evaluated projectional angles is minimal.

III. Results

A. Examination of the consistency measure

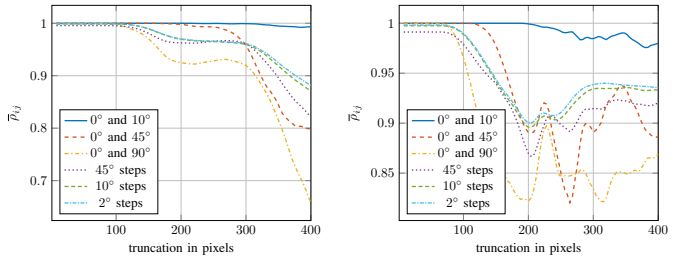
The relationship between increasing transversal truncation and the consistency measure is shown in Figure 3. For this, we considered 2 different software phantoms. The modified Shepp-Logan phantom and a thorax phantom which is part of the TIGRE toolbox [11]. Orthogonal projections of these phantoms are shown in Fig. 2. The curves in Fig. 3 are plotted



(a) Shepp-Logan 0° (b) Shepp-Logan 90°

(c) Thorax 0° (d) Thorax 90°

Fig. 2: Orthogonal projections of the phantoms used with constant axial truncation. Dimensions: 616 x 262 px. Gray value windowing: [0% 100%]



(a) Thorax phantom (b) Shepp-Logan phantom

Fig. 3: Examination of the consistency measure for increasing transversal truncation. The truncation is given in total amount of pixels of transversal truncation. Additionally, the phantoms possessed a constant axial truncation to model a more realistic scenario.

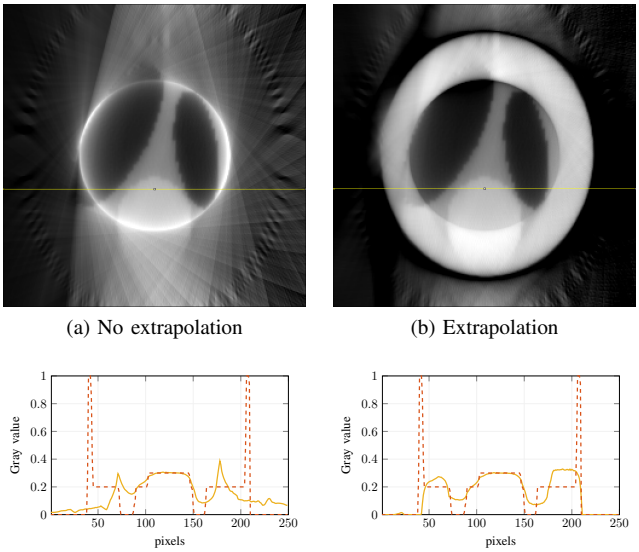
for the different projections which have been considered for the computation of the consistency measure.

The first observation that can be made is that the consistency measure decreases with increasing transversal truncation no matter which projections are considered for the consistency computation. This indicates that the Grangeat-based consistency measure is generally susceptible for truncation. However, the consistency measure derived from orthogonal projections shows the steepest decline in consistency for moderate truncation (note that the phantom has not really been truncated until roughly 150 px truncation for the thorax and about 80 px for the Shepp-Logan phantom because less truncation basically just removes the black surroundings of the phantom and not much of the phantom itself). Whereas the effect of truncation on the consistency calculated from projections from similar angles (0° and 10°) is largely negligible. This is not surprising since neighboring projections are redundant to a high degree as their SFOVs cover largely the same area of the phantom, whereas orthogonal projections generally contain the highest amount of additional information to be considered by the DCC since they share a smaller common SFOV (compare Fig. 1).

Observe also, that the consistency measure has several local maxima for strong truncation of the Shepp-Logan phantom. This can be explained by the symmetrical properties of the Shepp-Logan phantom as DCCs are generally known to not be suited well for highly symmetrical objects.

B. Performance of the extrapolation procedure

Figure 4 shows some results obtained with the proposed extrapolation procedure for the, regarding the consistency curves, demanding Shepp-Logan phantom (128 x 128 x 128 voxels). The results shown were obtained over 5 DE-iterations with an initial population of 70 parents. For the extrapolation model we assumed a homogenous ellipsoid as described in section II-C to approximate the overall shape of the Shepp-Logan phatom. The consistency was computed by pair-wise comparing projections from 0°, 45°, 90° and 135°. The resulting extrapolated projections were reconstructed iteratively via an Algebraic Reconstruction Technique (ART) since it is supposedly less susceptible to truncation than FDK.



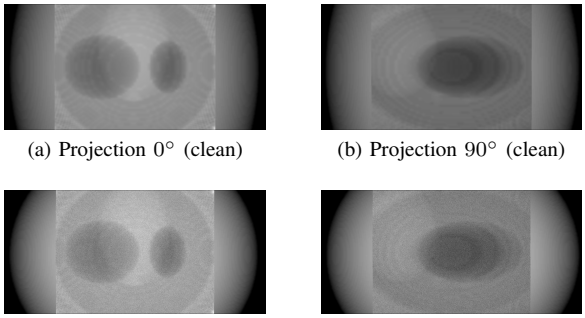
(a) No extrapolation (b) Extrapolation

(c) Cross-section profile of Fig. 4a (d) Cross-section profile of Fig. 4b

Fig. 4: Reconstructions from truncated projections with and without extrapolation including corresponding cross-section profile plots (yellow line in the reconstruction). The ground truth is shown in red for comparison. The central slice of the reconstruction is shown. Gray value windowing: [0% 100%]

Observe how the optimized ellipsoid does not exactly match the boundaries of the phantom which is barely visible in the background and therefore also appears a bit too bright. However, the improvement to image quality inside the VOI is clearly visible. Especially the bright ring-artifact surrounding the VOI is mostly reduced. This allows to identify the limit of the left ellipsoid inlet of the phantom near the border of the VOI in Fig. 4b which is not identifiable in Fig. 4a. Also the contrast of the different inlets of the phantom is increased (compare the profiles in Fig. 4c and 4d). Since the optimized ellipsoid is also visible in the reconstruction, the region outside of the VOI has to be considered with care to not mistaken the ellipsoid for the imaged phantom.

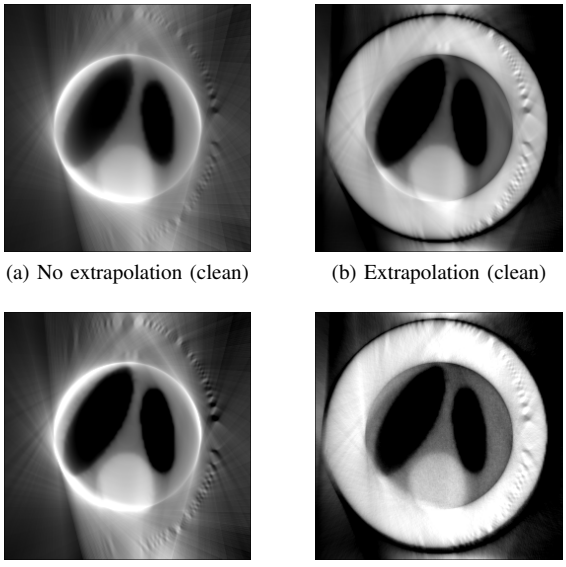
Figures 5 and 6 show the performance of the extrapolation with and without noise added to the truncated projection data.



(a) Projection 0° (clean) (b) Projection 90° (clean)

(c) Projection 0° (noisy) (d) Projection 90° (noisy)

Fig. 5: Exemplary extrapolated projections of clean and noisy projection data from orthogonal source positions. Gray value windowing: [0% 100%]



(a) No extrapolation (clean) (b) Extrapolation (clean)

(c) No Extrapolation (noisy) (d) Extrapolation (noisy)

Fig. 6: Reconstructions from truncated and extrapolated projections with and without added noise. The associated clean/noisy projection data is shown in Figure 5. Gray value windowing: [0% 100%]

The noise is clearly visible in the projections in Figure 5c and 5d. Note that the extrapolations shown resulted from two independent optimization runs. Both extrapolations match well, which indicates the algorithm's robustness against noise. However, looking at Figures 5 and 6 the algorithm's tendency to favour ellipsoids with a low eccentricity becomes apparent.

IV. Discussion and Conclusion

In this paper a novel extrapolation method for truncation problems based on Grangeat-based DCCs is introduced. The behavior of the resulting consistency measure has been investigated and shown to be susceptible for transversal truncation even in the presence of constant axial truncation and also for phantoms with high symmetric properties like the Shepp-Logan phantom. The reconstructions indicate a slight increase in image contrast and a good cancellation of the ring-artifact typical for truncation. The quality of the approximation of the real phantom boundary by the extrapolation model still needs to be improved, though. This is likely due to the secondary maxima shown in Fig. 3.

Further enhancements can be made to the choice and weighting of projections for the computation of the consistency measure and to the optimization itself. Furthermore, more sophisticated extrapolation models based on empirical studies of body-part measurements are thinkable to overcome the limitations of purely homogenous extrapolation models.

V. Acknowledgement

This work was conducted within the context of the International Graduate School MEMoRIAL at Otto von Guericke University (OVGU) Magdeburg, Germany, kindly supported by the European Structural and Investment Funds (ESF) under the programme "Sachsen-Anhalt WISSENSCHAFT Internationalisierung" (project no. ZS/2016/08/80646).

The work of this paper is also partly funded by the German Ministry of Education and Research (BMBF) within the Forschungscampus STIMULATE under grant number 13GW0095A.

References

[1] B. Ohnesorge, T. Flohr, K. Schwarz, J. P. Heiken, and K. T. Bae, "Efficient correction for ct image artifacts caused by objects extending outside the scan field of view," *Medical physics*, vol. 27, no. 1, pp. 39–46, 2000.

[2] J. Hsieh, E. Chao, J. Thibault, B. Grekowicz, A. Horst, S. McOlash, and T. J. Myers, "A novel reconstruction algorithm to extend the ct scan field-of-view," *Medical physics*, vol. 31, no. 9, pp. 2385–2391, 2004.

[3] G. van Gompel, M. Defrise, and D. van Dyck, "Elliptical extrapolation of truncated 2d ct projections using helgason-ludwig consistency conditions," in *Medical Imaging*, ser. SPIE Proceedings, M. J. Flynn and J. Hsieh, Eds. SPIE, 2006, p. 61424B.

[4] R. Frysch and G. Rose, "Rigid motion compensation in interventional c-arm ct using consistency measure on projection data," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, ser. Lecture Notes in Computer Science, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Cham: Springer International Publishing, 2015, vol. 9349, pp. 298–306.

[5] S. Abdurahman, R. Frysch, R. Bismark, M. Friebe, and G. Rose, "Calibration free beam hardening correction using grangeat-based consistency measure," Strasburg, 2016.

[6] C. Debbeler, N. Maass, M. Elter, F. Dennerlein, and T. M. Buzug, "A new ct rawdata redundancy measure applied to automated misalignment correction," in *Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, 2013, pp. 264–267. [Online]. Available: http://www.fully3d.org/2013/Fully3D2013Proceedings.pdf

[7] P. Grangeat, "Mathematical framework of cone beam 3d reconstruction via the first derivative of the radon transform," in *Mathematical Methods in Tomography*, ser. Lecture Notes in Mathematics, G. T. Herman, A. K. Louis, and F. Natterer, Eds. Berlin and Heidelberg: Springer, 1991, vol. 1497, pp. 66–97.

[8] M. Defrise and R. Clack, "A cone-beam reconstruction algorithm using shift-variant filtering and cone-beam backprojection," *IEEE transactions on medical imaging*, vol. 13, no. 1, pp. 186–195, 1994.

[9] A. Aichert, M. Berger, J. Wang, N. Maass, A. Doerfler, J. Hornegger, and A. K. Maier, "Epipolar consistency in transmission imaging," *IEEE transactions on medical imaging*, vol. 34, no. 11, pp. 2205–2219, 2015.

[10] R. Storn and K. Price, "Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, no. 4, pp. 341–359, 1997.

[11] A. Biguri, M. Dosanjh, S. Hancock, and M. Soleimani, "Tigre: a matlab-gpu toolbox for cbct image reconstruction," *Biomedical Physics & Engineering Express*, vol. 2, no. 5, p. 055010, 2016. [Online]. Available: http://iopscience.iop.org/article/10.1088/2057-1976/2/5/055010/pdf

Analysis of Tomographic Influences on the Surface Definition in Industrial X-Ray Computed Tomography

Dierck Matern and Frank Herold

***Abstract*—In X-ray computed tomography (CT), the surface is defined by different grey values in the tomographic reconstruction [1]. This surface therefore is influenced by any disturbances that may alter the voxel data. However, many applications, especially metrology [2], rely heavily on the surface definition. Distances are measured between the edges or coordinates defined by the surface. If the voxel information changes, this surface may alter, too, leading to false proportions of the object. In this work, an analysis method for the edges is proposed as a qualification method.**

I. Introduction

Lately, dimensional metrology has been applied to industrial CT [3]. This is the study of the sizes of tested objects. For example, in quality assurance in manufacturing of objects, one may be interested if a drilled hole has the correct size, or if different elements have the correct distance to each other [4]. CT has become popular for this task, because it is a non-destructive method which can describe inner structures as well as the exterior [5].

However, this technology has its limits because of its physical nature. One shortcoming for this application can be geometrical misalignments of the CT system [6], [7]. An example of the projection geometry can be seen in Figure 1; a misalignment is some unknown difference from this perfect geometry. Other errors result from the cone-beam geometry often used for this task [8], influences of the alignment of the test object [9], or the drift of the X-ray focal spot [10]. Further, industrial X-ray sources usually do not provide a monochromatic spectrum [11], and the different wavelengths are absorbed and scattered differently [1], [12], [13]. This is a non-geometrical error, but it influences the grey values of the tomogram, meaning that it has influence on the surface definition, and therefore on the dimensional measurement.

Several correction methods try to reduce these effects. For example, in [9] and [13], the scattering kernel is estimated. In [14], the influence of beam hardening effects on the dimensional measurement are analysed, with inconclusive results. The assumptions of beam hardening corrections are usually applicable if there is only one material in the object, so this is not a conclusive method, though.

Furthermore, reconstruction methods are necessary to acquire the tomogram [15], [16]. This may also influence the measurement, since it processes the raw data, which can be lossy.

YXLON International GmbH, e-mail: {matern,herold}@hbg.yxlon.com

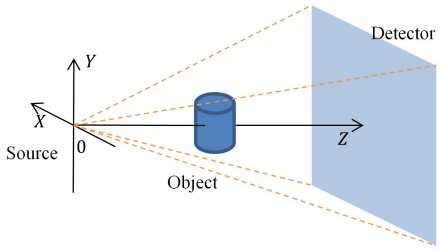


Fig. 1. Typical cone beam geometry. An object is rotated between the source and the detector. Misalignments in this geometry can result in errors in the dimensional measurements.

In Figure 2, results of the measurement of the sphere distance in simulated data are shown. The determination of the sphere distance error is a typical challenge in the dimensional metrology [17]. The task is to measure distances of spherical markers in a known geometry, see Figure 2, on the left side. The errors with a simulated CT are shown on the right hand side. Even though no scattering was included in this simulation, the errors are biased and larger than zero. Furthermore, the orientation of the gauge, which changes the beam hardening, has little influence on these results [18]. Hence, the reconstruction algorithm, or the digital sampling of the detector, seems to have influence on the dimensional measurement. But their influence on the voxels is relatively hard to measure, because of the lacking of a ground truth.

Recently, the surface has become of interest in the study of the dimensional measurement in CT. In [19], the surface is analysed in a perpendicular direction. However, this can be used when the surface is smooth, as this methods analyse relatively large chunks of the surface. In [20], the analysis is performed on image quality parameters of the edges, that is, the step from the material into the air. There, the error function is fitted to the surface; this function is related to the Gaussian distribution [21]. While this seems legit to smooth edges, it holds certain disadvantages, especially if the edge is not very smooth, and closer to an ideal step function.

One part of this work is introducing a generalization to the normal function, the generalized Gaussian, to the problem to describe the surface. Furthermore, a calculation of the difference to this assumption is performed, which is the main quality measure of this work. Hence, the measure predicts how certain the edge can be found.

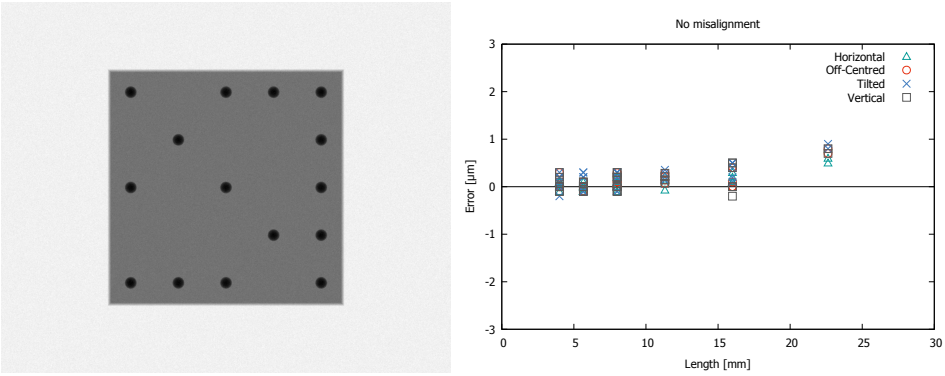


Fig. 2. Sphere distance: typical benchmark for the dimensional measurement in industrial CT. A gauge with known (that is, calibrated) dimensions is scanned. In this case, spherical markers of high absorbing material (ruby) are used (left). In the tomogram, the surface is extracted, and the distance between the centres of the spheres are measured, and compared to the known values. On the right hand side, the error to the known values in an ideal geometry, with different alignments of the gauge, are shown. Even though there are no misalignments in this simulation, the sphere distance errors are biased.

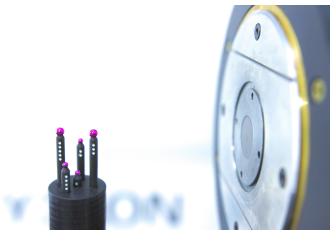


Fig. 3. Tested gauge in front of the used X-ray source. The gauge consists of several ruby spheres. The X-ray source is a microfocus tube with transmission target.

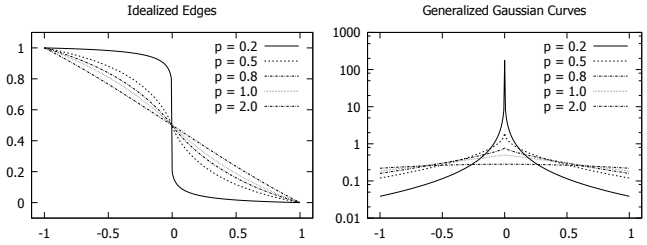


Fig. 4. Left: different ideal edges. The steeper the edge, the more stable is the turning point. Right: the differentiated edge on the left is similar to the generalized Gaussian. The steeper the grey value step, the "pointier" the curve becomes. This happens when the parameter p shrinks.

II. Materials and Methods

The test object analysed in this study is a probe consisting of several ruby spheres, see Figure 3. We scan this object in different magnifications with the same X-ray parameters. The part of the object we want to analyse are two spheres at the topmost edge of the gauge. In other scans, obstructing materials are included, either in the inside of the reconstructed volume, or outside of it. Then, tomographic slices are extracted, and the edge responses are analysed.

This is similar to the edge response function (ERF) as discussed in the ASTM E1695 standard [22]. However, in this standard, the ERF is used on the whole slice only, not in a specific direction, and is not applicable in arbitrary CT scans. The influence of the obstructing object can be very specific to the direction.

A. Generalized Gaussian for Edge Response

The generalized normal distribution or generalized Gaussian distribution is discussed lately as a generalization of the well known normal distribution [21]. Because its statistical properties are not discussed in this paper, the function is called here a *generalized Gaussian*, so it is not used as a distribution. In detail, let $\sigma, p > 0$, $\mu \in \mathbb{R}$. Then the generalized Gaussian g is defined as

$$A(p,\sigma) = \left(\frac{\sigma^2\Gamma(1/p)}{\Gamma(3/p)}\right)^{1/2},$$
$$g(x;\mu,\sigma,p) = \frac{p}{2\Gamma(1+1/p)A(p,\sigma)} \exp\left(-\left|\frac{x-\mu}{A(p,\sigma)}\right|^p\right) \quad (1)$$

for all $x \in \mathbb{R}$ [23]. Note that for $p = 2$, this is the standard normal distribution, hence the form of the edge is the same as using the error function as in [20].

This can be used to define a quality measure for the edges. Say a line profile, that is, a slice through the image of finite length in an arbitrary direction, is set over an edge of a mono material, from the inside to the outside. Then the negative gradient of this line profile forms a curve similar to a bell, so a generalized Gaussian is a suitable description. The steeper the transition from the material to the outer regions of the materials in grey values, the "pointier" is this gradient, so p becomes smaller. In the extreme, the ideal transition, this becomes a Dirac impulse.

Furthermore, surface definitions rely on the turning point of the curve [24], that is, the maximum of the gradient. Since this is inflicted by the symmetry of the flanks of this edge, a comparison to a symmetrical function provides a robust measure for the reliability.

In detail, this comparison is as follows. First, the line profile is extracted from a tomographic slice, at the very surface of

the sphere. This is long enough to include both material and air, and orthogonally to the surface, approximately ten pixels long. Then, the derivative is calculated, and a curve is fitted to this result. The fit then is evaluated to provide a measure of ambiguity.

Let the ith curve be defined as l_i with $l_i(-1)$ is the limit in the material, and $l_i(1)$ the limit on the outside. Then the bell shaped curve at $t \in \Omega = [-1, 1]$ is $-\delta l_i(t)$. In praxis, this can be calculated using a fitted polynomial [22]. Further, let ν be a normalization of this curve with

$$s(f) = \int_\Omega f(x) - \min_\Omega(f) dx,$$
$$\nu(f, x) = \frac{f(x) - \min_\Omega(f)}{s(f)}. \quad (2)$$

Then fitting a generalized Gaussian is a minimization problem

$$\hat{\theta}_i \leftarrow \min_\theta \int_\Omega |\nu(-\delta l_i, x) - \nu(g(\,\cdot\,; \mu, \sigma, p), x) - a| dx, \quad (3)$$

where $\theta = \{\mu, \sigma, p, a\}$, the offset a is a slack variable. Note that this optimization is performed for lines individually, hence a set for each analysed direction is computed.

Furthermore, we are interested in the suitability of the curve model, and therefore reliability of the surface definition. For this, the Kullback-Leibler divergence is used [25]. This is defined as

$$KL(a||b) = -\sum_{x \in \Omega} a(x) \cdot \log\left(\frac{a(x)}{b(x)}\right). \quad (4)$$

This measure is used to define the difference between the measured edge response and the best fitted generalized Gaussian, that is $KL(-\delta\tilde{l}_i||\tilde{g}(\,\cdot\,, \hat{\mu}, \hat{\sigma}, \hat{p}))$, where the curves are normalized to form a probability distribution on Ω, denoted by the tilde.

III. Experiments

The following analysis is performed with several CT scans of the gauge as seen in Figure 3. The scans are in two different magnifications, 13.3333 and 4.7059, and the distance between focus and detector is 800mm. The X-ray source is set to 100kV and 30μA. In our experience, those are fairly common parameters for industrial CT with the utilized microfocus tube, see Figure 3.

Several scans for each case are used, with and without obstructing objects, which obfuscate the gauge in several angles. Furthermore, the spheres obfuscate each other.

After the reconstruction, a slice is extracted, see Figure 5 for an example. At the edges, in 720 equally spaced directions, orthogonally to the edge, lines are extracted, the curves are fitted, and the Kullback-Leibler divergence is computed.

Examples for the curve fit, at two different angles, can be seen in Figure 6. Those are extracted from the right slice in Figure 5, upper left sphere. Even though it is the same tomogram and even the same sphere, the results differ.

In Figure 7, an overview of the Kullback-Leibler divergence for the whole edge can be seen. In the case without the obstructing object, the effect is rather low, but still measurable.

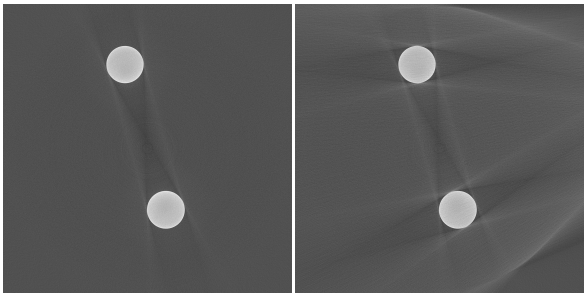


Fig. 5. Slices through the reconstruction of the gauge where the edge is evaluated, without (left) and with (right) obstructing object, located at the right side in this view. Magnification is 13.3333.

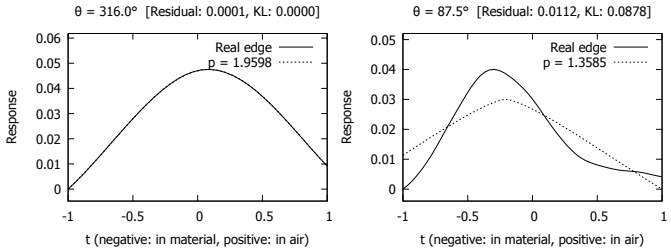


Fig. 6. Sample of a fitting edge (left) and an edge where the curve fitting results in a large difference (right). In the latter case, the Kullback-Leibler divergence is higher by several magnitudes. Curves from the same slice as in Figure 5, right side; $\theta = 0$ is oriented to the right, the angle is in positive direction.

That is, the effects of the neighbour sphere can be seen from -10 degrees (350 degrees) to 40 degrees and from 170 degrees to 220 degrees. Referring to the slice, as this represents the upper sphere, this is where the tangents from the lower sphere touch the upper one. The disturbances are still present at lower magnification, see Figure 8.

In the cases where the object is obstructed, the signal around the sphere is noisier. While the effects of the neighbour sphere are still present, the object has much higher influence, that is around 90 degrees and 270 degrees, respectively. This behaviour is consistent throughout all of the measurements performed in this study, also at different magnifications. Hence, this measure is sensitive to disturbances, and therefore is well suited to evaluate the ambiguity of the surfaces.

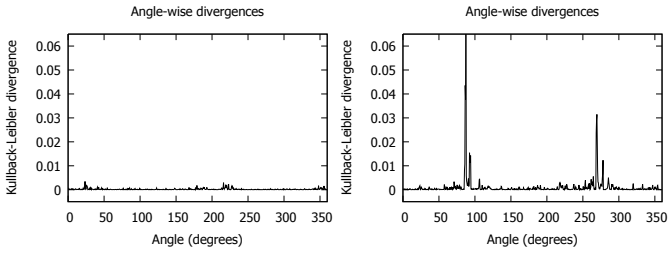


Fig. 7. Kullback-Leibler divergence over the full circle at magnification 13.3333. Without (left) and with (right) obstructing objects in the scan. The influences of the spheres onto each other are measurable, but low. The obstructing object has clear effects on the divergence.

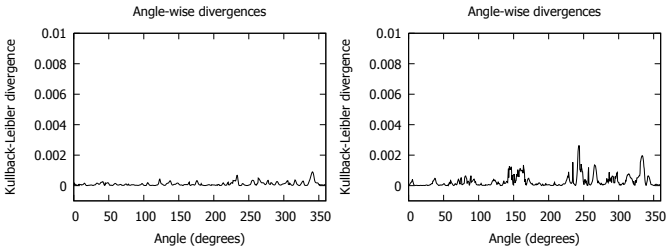


Fig. 8. Evaluation at magnification 4.7059. The influence of the obstructing object is much lower in this case (right), but still measurable. Left: measurement without obstructing object.

IV. Summary and Outlook

In this paper, we discussed a measure to evaluate edges in reconstructions in industrial CT. A model for edges has shown to be accurate for "good" edges. The experiments show that the Kullback-Leibler divergence is an adequate method to evaluate effects on the edge. Because this measure evaluates the difference to the assumption on which the surface estimation depends on, it is also a measure for the ambiguity in this procedure.

There are two major directions in which this information can be used. One is the evaluation of different models for edges to find the most accurate one. As the divergence is a measure of entropy [23], this follows the principle of maximum entropy [26].

The other route is to use this measure to predict the uncertainty of the real measurement [17], [27]. So far, the connection between the ambiguity as measured with the Kullback-Leibler divergence and this standardized information is not well understood, that is, we cannot predict confidence intervals from this measure.

In total, this measure provides a decent understanding of the surface, and therefore is useful for a general overview of the reliability of the length measurements. Hence, it is a solid basis for both, the evaluation of edge definitions, and the prediction of uncertainties. However, the surface of the gauge is analysed in a similar way as the MTF. In the future, we will study how the different measures for image quality rely to each other to improve the understanding of the uncertainty. The connection is especially interesting, because in this study, the starting point is a (simulated) probing test as defined in [17]. Hence, this might be a connection of those very different quality measures.

References

[1] T. M. Buzug, *Computed Tomography*, 1st ed. Springer-Verlag, 2008.

[2] J.-P. Kruth, M.Bartscher, S.Carmignato, R.Schmitt, L. Chiffre, and A.Weckenmann, "Computed Tomography for Dimensional Metrology," *CIRP Annals-Manufacturing Technology*, vol. 60, no. 2, pp. 821–842, 2011.

[3] J. Hiller, S. Kasperl, U. Hilpert, and M. Bartscher, "Koordinatenmessung mit industrieller Röntgen-Computertomografie (Coordinate Measuring with Industrial X-Ray Computed Tomography)," *tm - Technisches Messen*, vol. 74, pp. 553–564, 2007.

[4] J. A. B. Angel, "Quality Assurance of CT Scanning for Industrial Applications," Ph.D. dissertation, Technical University of Denmark, March 2014.

[5] F. Herold, P. Kramm, and U. Adamczak, "Automating the X-Ray Inspection," in *Proceedings of the 10th European conference and exhibition on non-destructive testing*, 2010.

[6] M. Ferrucci, R. K. Leach, C. Giusca, S. Carmignato, and W. Dewulf, "Towards Geometrical Calibration of X-Ray Computed Tomography Systems Review," *Measurement Science and Technology*, vol. 26, no. 9, p. 092003, 2015.

[7] D. Gross, U. Heil, R. Schulze, E. Schoemer, and U. Schwanecke, "Auto Calibration of a Cone-Beam-CT," *Med. Phys.*, vol. 39, no. 10, pp. 5959–5970, October 2012.

[8] W. C. Scarfe and A. G. Farman, "What is Cone-Beam CT and How Does it Work?" *The Dental Clinics of North America*, vol. 52, no. 4, pp. 707–730, 2008.

[9] J. Kumar, A. Attridge, P. K. C. Wood, and M. A. Williams, "Analysis of the Effect of Cone-Beam Geometry and Test Object Configuration on the Measurement Accuracy of a Computed Tomography Scanner used for Dimensional Measurement," *Measurement Science and Technology*, vol. 22, no. 3, p. 035105, 2011.

[10] N. Flay, W. Sun, S. Brown, R. Leach, and T. Blumensath, "Investigation of the Focal Spot Drift in Industrial Cone-beam X-ray Computed Tomography," in *Digital Industrial Radiology and Computed Tomography*, 2015.

[11] A. Deresch, G. R. Jaenisch, C. Bellon, and E. Warrikhoff, "Simulating X-ray Spectra: From Tube Parameters to Detector Output," *Proceedings of the 18th World Conference on Nondestructive Testing*, vol. 1, pp. 1–8, 2012.

[12] S. Schlesinger and H. Bomsdorf, "Non-destructive X-ray Diffraction Tomography with Cone-Beam CT Systems-Visualizing the Local Charge Distribution of Li-ion Batteries," *Proceedings of the 7th Conference on Industrial Computer Tomopgraphy*, pp. 1–8, 2017.

[13] H. Li, R. Mohan, and X. R. Zhu, "Scatter Kernel Estimation with an Edge-spread Function Method for Cone-beam Computed Tomography Imaging," *Physics in Medicine and Biology*, vol. 53, no. 23, p. 6729, 2008.

[14] W. Dewulf, Y. Tan, and K. Kiekens, "Sense and Non-sense of Beam Hardening Correction in CT Metrology," {*CIRP*} *Annals - Manufacturing Technology*, vol. 61, no. 1, pp. 495 – 498, 2012.

[15] M. Grass, T. Koehler, and R. Proksa, "3D cone-beam CT reconstruction for circular trajectories," *Physics in Medicine and Biology*, vol. 45, no. 2, p. 329, 2000.

[16] R. Grimmer, M. Oelhafen, U. Elstroem, and K. Marc, "Cone-Beam CT Image Reconstruction with Extended z Range," *Medical Physics*, vol. 36, no. 7, pp. 3363–3370, 2009.

[17] "VDI/VDE 2630 Computertomografie in der dimensionellen Messtechnik," Verein Deutscher Ingenieure.

[18] D. Matern and F. Herold, "On Gauge Orientation for the Qualification of Industrial Computed Tomography," *Proceedings of the 8th Conference on Industrial Computed Tomography*, 2018.

[19] E. H. M. Flessner, A. Müller and T.Hausotte, "Evaluating and Visualizing of the Quality of Surface Points Determined from Computed Tomography Volume Data," in *MacroScale*, 2014.

[20] T. Schönfeld and M. Barscher, "Verification and Application of Quality Measures in Dimensional Computed Tomography," in *DIR*, 2015.

[21] T. L. Toulias and C. P. Kitsos, "On The Properties of the Generalized Normal Distribution," *Discussiones Mathematicae - Probability and Statistics*, vol. 34, pp. 35–49, 2014.

[22] *Standard Test Method for Measurement of Computed Tomography (CT) System Performance*, ASTM International Std. E1695-95 (2013), 2013.

[23] S. Nadarajah, "A Generalized Normal Distribution," *Journal of Applied Statistics*, vol. 32–7, pp. 685–694, 2005.

[24] W. E. Lorensen and H. E. Cline, "Marching Cubes: A High Resolution 3D Surface Construction Algorithm," in *ACM siggraph computer graphics*, vol. 21, no. 4. ACM, 1987, pp. 163–169.

[25] J. C. Baez and T. Fritz, "A Bayesian Characterization of Relative Entropy," *CoRR*, vol. abs/1402.3067, 2014.

[26] A. Giffin and A. Caticha, "Updating Probabilities with Data and Moments," in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, ser. American Institute of Physics Conference Series, K. H. Knuth, A. Caticha, A. Giffin, and C. C. Rodríguez, Eds., vol. 954, Nov. 2007, pp. 74–84.

[27] *Geometrical Product Specifications (GPS) - Coordinate Measuring Machines (CMM): Technique for Determining the Uncertainty of Measurement - Use of Calibrated Workpieces or Measurement Standards*, DIN EN ISO Std. 15 530-3, 2011.

Dual-energy CT with Nanotube X-ray Source Array for Security Scanning

Buxin Chen, Dan Xia, Zheng Zhang, Emil Y. Sidky, and Xiaochuan Pan

***Abstract*—Dual-energy CT can reduce the false alarm rate of CT-based explosive detection systems, based on its ability to recover spectral response of different material types and thus differentiate explosives such as ANFO from water. On the other hand, CT systems based upon nanotube X-ray tubes have been considered for check-point bag screening. As the array of focal spots in nanotube X-ray rube is individually controllable, nanotube-based X-ray tube can be used for dual-energy CT imaging, with focal spots in one tube alternately emitting low- and high-kVp spectra. In this work, we investigate a dual-energy CT setup based on 3 arc-shaped nanotube X-ray source arrays with alternating low- and high-kVp focal spots. A simulation study has been carried out with a digital dual-energy phantom. The results suggest that the proposed dual-energy setup combined with the optimization-based reconstruction method can accurately recover the spectral responses of different material types and adequately differentiate them from each other.**

I. Introduction

Dual-energy CT may reduce the false alarm rate of CT-based explosive detection systems [1], [2]. In single-energy CT, explosives, such as ANFO (Ammonium Nitrate and Fuel Oil), can have similar physical density with water, and thus it can be challenging to adequatly differentiate ANFO from water. In dual-energy CT, the effective atomic number of the material can be estimated through the decomposition model and used to better discriminate ANFO from water [1]. As a result, reduced false alarm rate can lead to increased throughput of the checked luggage screening. On the other hand, CT systems based upon nanotube multi-beam X-ray tubes have been considered as a potential scan technique for possible check-point bag screening [3] and small animal imaging [4]. Such systems are usually stationary, without the need to rotate X-ray tube. Also, the nanotube-based X-ray tube contains array of focal spots, or X-ray firing positions, which are individually controllable. As a result, the focal spots in one nanotube X-ray source array can be programmed to generate alternating low- and high-kVp spectra, achieving dual-energy scanning. In this work, we investigate a dual-energy CT setup based on the nanotube X-ray source array with alternating low- and high-kVp-generating focal spots. A simulation study with a digital dual-energy phantom is carried out, and the evaluation focuses on the differentiation between two materials and the recoveries of spectral responses of these two materials. Such a study can reveal the potential of material differentiation and possibly explosive detection by use of the proposed dual-energy CT setup and the optimization-based reconstruction method.

B. Chen, D. Xia, Z. Zhang, and E. Y. Sidky are with the Department of Radiology, The University of Chicago, Chicago, IL 60637, USA

X. Pan is with the Departments of Radiology & Radiation and Cellular Oncology, The University of Chicago, Chicago, IL 60637, USA

II. Materials and Methods

A. Data Generation

In this preliminary investigation, a 2D CT geometry is simulated, as shown in Fig. 1. The setup consists of three arc-shaped X-ray nanotube source arrays. Each arc covers 90 degrees, with 30 degrees of separation between any two. On each array, there are 300 X-ray focal spots, 150 of which are set at low voltage, i.e., 80 kVp in this work and the other 150 are set at high voltage of 140 kVp. The low and high kVp focal spots are aligned in an alternating fashion, separated by 3.14 mm, on the panel. The curved detector has 896 detector bins and is subject to a fan angle of 49 degrees. In this simulation, the detector rotates as the firing position of the X-ray focal spot moves along the array, while a stationary detector panel corresponding to each nanotube source array is also feasible in reality. The detector is always at 1100 mm away from its corresponding X-ray focal spot, while the center of the rotation axis is at 600 mm away from the source array. As a result, a complete scan results in low- and high-kVp data sets, each of which has 450×896 data points.

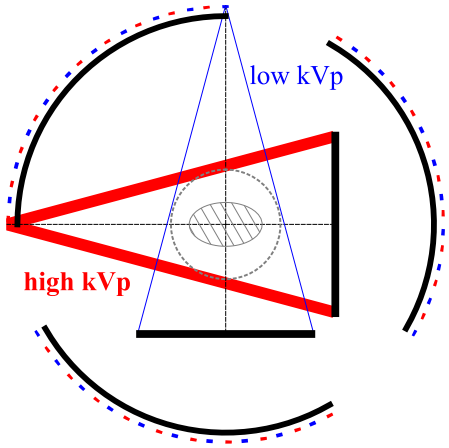


Fig. 1: Scan configuration consisting of 3 nanotube X-ray source array. Alternating dots of blue and red color on the panel indicate X-ray source firing positions of low and high kVp spectra.

The low and high X-ray spectra of 80 and 140 kVp are generated by use of the TASMICS worksheet (v1.0) [5], assuming a tungsten anode and a 5-mm Al filter. A digital GAMMEX DE-472 phantom [6] is used for data simulation. The digital phantom is sampled on a 512×512 image array with 0.68-mm square pixels. Each pixel is labeled with a material type from the spec sheet of the phantom. The polychromatic forward model, equation (5) in Ref. [7], is used to generate data, where the decomposition error could be present, based on the scanning geometry described above. Poisson noise is also added to the data, assuming a fluence of 2×10^6 photons for each ray in the air scan.

B. Optimization-Based Reconstruction

1) Discrete-to-discrete data model: In this work, a two-basis decomposition model with water and bone is used. Therefore, we can express the discrete-to-discrete data model as

$$g_j^{[s]}(\mathbf{b}_1,\mathbf{b}_2) = -\ln\sum_m q_{jm}^{[s]}\exp\left(-\sum_i a_{ji}^{[s]}\left(\mu_{1m}b_{1i}+\mu_{2m}b_{2i}\right)\right), \quad (1)$$

where $j\in\{0,\cdots,J^{[s]}-1\}$ is a ray index for either low ($s=1$) or high ($s=2$) kVp scan, $g_j^{[s]}(\mathbf{b}_1,\mathbf{b}_2)$ denotes the model data for the jth ray in scan s, $q_{jm}^{[s]}$ the ray-dependent, normalized X-ray spectrum, satisfying $\sum_m q_{jm}^{[s]}=1$, at energy m for the jth ray in scan s, and $a_{ji}^{[s]}$ the intersection length of the jth ray in scan s with the ith voxel. Vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ denote basis images of size I, μ_{1m} and μ_{2m} the decomposition coefficients. Clearly, the data model is a non-linear function of basis images $\mathbf{b}_1$ and $\mathbf{b}_2$. The data model in equation (1) can be re-written as

$$g_j^{[s]}(\mathbf{b}_1,\mathbf{b}_2) = \bar{g}_j^{[s]}(\mathbf{b}_1,\mathbf{b}_2)+\Delta g_j^{[s]}(\mathbf{b}_1,\mathbf{b}_2), \quad (2)$$

where

$$\bar{g}_j^{[s]}(\mathbf{b}_1,\mathbf{b}_2) = \sum_i a_{ji}^{[s]}\left(\bar{\mu}_{j1}^{[s]}b_{1i}+\bar{\mu}_{j2}^{[s]}b_{2i}\right), \quad (3)$$

$$\Delta g_j^{[s]}(\mathbf{b}_1,\mathbf{b}_2) = -\ln\sum_m q_{jm}^{[s]}\exp\left(-\sum_i a_{ji}^{[s]}\left(\Delta\mu_{j1m}^{[s]}b_{1i}+\Delta\mu_{j2m}^{[s]}b_{2i}\right)\right), \quad (4)$$

$\bar{\mu}_{jk}^{[s]}=\sum_m q_{jm}^{[s]}\mu_{km}$ is an energy-independent term and $\Delta\mu_{jkm}^{[s]}=\mu_{km}-\bar{\mu}_{jk}^{[s]}$ remains energy dependent [8], [9].

2) Non-convex optimization program: A constrained optimization program is formulated as

$$\begin{aligned}(\mathbf{b}_1^*,\mathbf{b}_2^*) = &\arg\min_{\mathbf{b}_1,\mathbf{b}_2}\left(||\mathbf{b}_1||_{\mathrm{TV}}+||\mathbf{b}_2||_{\mathrm{TV}}\right)\\ &\text{s.t.}\quad D\left(\bar{\mathbf{g}}(\mathbf{b}_1,\mathbf{b}_2),\,\mathbf{g}_{\mathcal{M}}-\Delta\mathbf{g}(\mathbf{b}_1,\mathbf{b}_2)\right)\le\epsilon,\\ &\quad\ \mathbf{b}_1\succeq0\ \&\ \mathbf{b}_2\succeq0,\end{aligned} \quad (5)$$

where $||\cdot||_{\mathrm{TV}}$ denotes the image total-variation (TV), $\bar{\mathbf{g}}(\mathbf{b}_1,\mathbf{b}_2)$ and $\Delta\mathbf{g}(\mathbf{b}_1,\mathbf{b}_2)$ are aggregate vectors formed by concatenating $\bar{g}_j^{[s]}(\mathbf{b}_1,\mathbf{b}_2)$ and $\Delta g_j^{[s]}(\mathbf{b}_1,\mathbf{b}_2)$, as defined in equations (3) and (4), respectively, and $\mathbf{g}_{\mathcal{M}}$ is the aggregate vector for measured data. The ℓ_2 norm is used for the data fidelity function as

$$\begin{aligned}&D(\bar{\mathbf{g}}(\mathbf{b}_1,\mathbf{b}_2),\mathbf{g}_{\mathcal{M}}-\Delta\mathbf{g}(\mathbf{b}_1,\mathbf{b}_2))\\ &\qquad=||\bar{\mathbf{g}}(\mathbf{b}_1,\mathbf{b}_2)-\mathbf{g}_{\mathcal{M}}+\Delta\mathbf{g}(\mathbf{b}_1,\mathbf{b}_2)||_2,\end{aligned} \quad (6)$$

and $\epsilon>0$ is the data constraint parameter that specifies the degree of data inconsistency.

3) Reconstruction algorithm: The ASD-NC-POCS algorithm [7] is used to numerically solve the non-convex optimization program in equation (5). A NC-POCS step is designed to lower the data fidelity function, in which a POCS step based on the linear term in the data model in equation (3) is followed by a non-linear term correction by estimating the non-linear term using equation (4) with the current iteration of basis images and subtracting the estimated non-linear term from the measured data. We then combine the NC-POCS step with the adaptive steepest descent (ASD) step for lowering the TV objective function.

4) Reconstruction parameter: Most of the reconstruction parameters are the same as those used in the data generation, including spectra, image pixels, and system matrices. The data fidelity parameter ϵ is selected as the one yielding the lowest value for the surrogate metric for the spectral recovery task, as defined in equation (9) below.

In addition, the practical convergence conditions, which are necessary conditions for the local optimality of the non-convex optimization program, are defined as

$$\begin{aligned}&\left|D\left(\bar{\mathbf{g}}(\mathbf{b}_1^{(n)},\mathbf{b}_2^{(n)}),\,\mathbf{g}_{\mathcal{M}}-\Delta\mathbf{g}(\mathbf{b}_1^{(n)},\mathbf{b}_2^{(n)})\right)-\epsilon\right|/\epsilon\le10^{-3},\\ &\frac{\left|\mathrm{Obj}^{(n+1)}-\mathrm{Obj}^{(n)}\right|}{\left|\mathrm{Obj}^{(n+1)}+\mathrm{Obj}^{(n)}\right|}\le10^{-3},\end{aligned} \quad (7)$$

where $\mathrm{Obj}^{(n)}=||\mathbf{b}_1^{(n)}||_{\mathrm{TV}}+||\mathbf{b}_2^{(n)}||_{\mathrm{TV}}$ is the objective function at iteration n.

C. Evaluation

The advantage of using dual-energy CT for explosive detection is its ability to differentiate explosive materials, such as ANFO (Ammonium Nitrate and fuel oil), from water, which has similar density with ANFO and thus display the similar gray value in a single energy CT scan [1]. Such differentiating power from dual energy is based on the decomposition of material interactions into Compton scattering and photoelectric effect and consequently the estimation of the effective atomic number of the material. Such decomposition is equivalent to the recovery of the spectral response of the material. As a result, in the absence of explosive ANFO in our digital phantom in this work, we evaluate the proposed dual-energy CT setup using nanotube X-ray source array and the reconstruction method for explosive detection by focusing on two inserts in the DE-472 phantom, 15-mg/ml iodine and 100-mg/ml calcium, that are difficult to be separated in the single-energy CT. Visual inspection is first made to assess if monochromatic images at different energy levels can better differentiate the two inserts. Then, quantitative analysis is

carried out to evaluate how accurate the spectral response of the two inserts are from the reconstructed basis images, as compared to the ground truth information. For each insert, an average relative error is calculated as

$$\text{relErr}_r = \frac{1}{N_m}\sum_m \frac{\frac{1}{N_r}\sum_{i\in I_r}(\mu_{1m}b_{1i}+\mu_{2m}b_{2i}) - f_{mr}}{f_{mr}}, \tag{8}$$

where I_r is the set containing all the pixel indices for ROI (insert) r, N_r the number of pixels in ROI r, f_{mr} the ground-truth linear attenuation value at energy m for ROI r, and N_m the number of energy levels to be averaged. In this work, $N_m = 7$ with energy levels at 40, 60, 80, 100, 120, 140, and 160 keV. A surrogate metric to characterize the spectral recovery of the two inserts under consideration is thus defined as

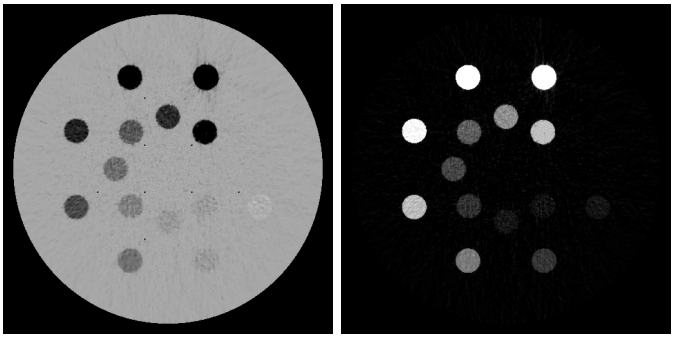
$$\text{relErr} = \frac{1}{2}(\text{relErr}_4 + \text{relErr}_{16}), \tag{9}$$

where $r = 4$ denotes the 15-mg/ml iodine insert and $r = 16$ the 100-mg/ml calcium one.

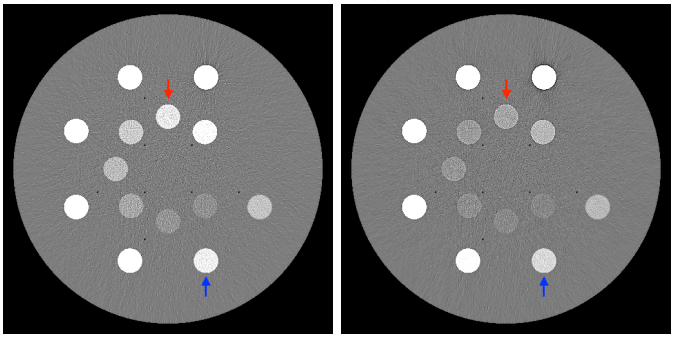
III. Results

We first show in Fig. 2 the water- and bone-basis images and monochromatic images at 80 and 140 keV reconstructed using the ASD-NC-POCS algorithm from the data collected with the nanotube X-ray source arrays as shown above. In the first row, Figs. 2a and 2b are the reconstructed water- and bone-basis images, respectively, with a displaying window of [0,1.5]. Good separation of basis materials can be observed. In the second row, Figs. 2c and 2d are the monochromatic images at 80 and 140 keV, respectively, with a displaying window of [-300, 300]HU. The two inserts under consideration in this work are pointed out by arrows in the figure, with ↓ indicating the 15-mg/ml iodine insert and ↑ the 100-mg/ml calcium one. It is difficult to differentiate the two inserts pointed by arrows in the 80-keV image in Fig. 2c, since their gray values in the image are very similar (as a matter of fact, their ground-truth HU values at 80 keV are 275.6 and 274.4 for 15-mg/ml iodine and 100-mg/ml calcium, respectively). Incorporating the dual-energy information and forming the monochromatic image at the 140 keV energy level, the contrast between these two inserts is increased (76.0 and 208.3 ground-truth HU values, respectively), as visually shown in Fig. 2d. Additionally, we show the ground-truth monochromatic images at 80 and 140 keV in the third row of Fig. 2. Similar observation regarding the two inserts at focus can be made from the ground-truth monochromatic images.

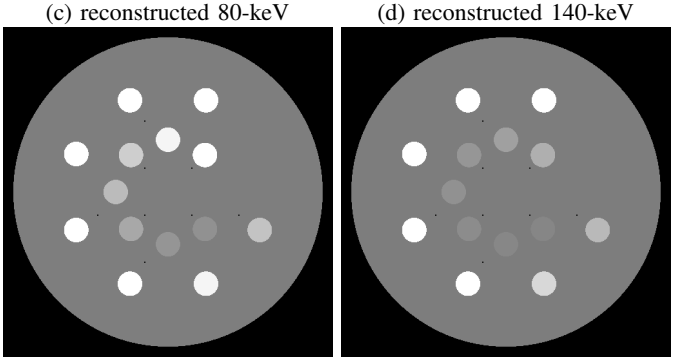
Next, the reconstructed basis images are used to form monochromatic images at 7 energy levels: 40, 60, 80, 100, 120, 140, and 160 keV. The linear attenuation coefficient values, estimated as the mean pixel value within the ROI of the insert at each energy level, are compared to the ground-truth values calculated from the manufacturer's spec sheet and the NIST XCOM data [10]. The results are shown in Fig. 3. The solid lines in the two plots denote the ground-truth values, where the circles (•) represent the linear attenuation coefficient calculated from the monochromatic images at the 7 energy levels. In both plots Figs. 3a and 3b, the



(a) reconstructed water basis (b) reconstructed bone basis



(c) reconstructed 80-keV (d) reconstructed 140-keV

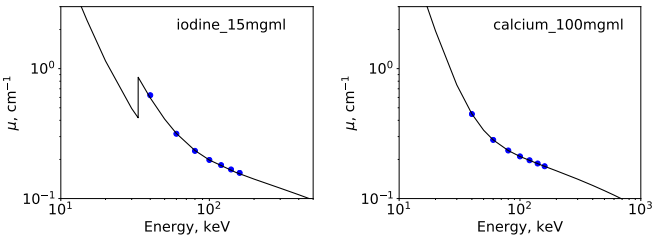


(e) ground-truth 80-keV (f) ground-truth 140-keV

Fig. 2: First row: reconstructed (a) water- and (b) bone-basis images with the ASD-NC-POCS algorithm from the data collected with the nanotube X-ray source array. Second row: reconstructed monochromatic images at (c) 80 and (d) 140 keV. Down and up arrows (↓ & ↑) point to two inserts of different materials (15-mg/ml iodine and 100-mg/ml calcium, respectively) that are difficult to separate in a single-energy CT, such as in the 80-keV monochromatic image, but are better differentiated in dual-energy CT with the help of 140-keV monochromatic image. Such two inserts at focus are to mimic the behavior of water and explosive ANFO in the scan of single- and dual-energy CT. Third row: ground-truth monochromatic images at (e) 80 and (f) 140 keV.

estimated values seem to overlap on the ground-truth curves, indicating accurate recoveries of the spectral responses of the corresponding materials. The relative errors averaged over the 7 energy levels for the two inserts, as calculated using equation (8), are $\text{relErr}_4 = 1.38\%$ and $\text{relErr}_{16} = 0.17\%$ for the 15-mg/ml iodine and 100-mg/ml calcium, respectively. The resulting surrogate metric as defined in equation (9) is

then $\mathrm{relErr} = 0.77\%$.



(a) 15-mg/ml iodine insert (b) 100-mg/ml calcium insert

Fig. 3: Ground-truth (solid-line) and estimated (•) linear attenuation coefficients at different energy levels for (a) 15-mg/ml iodine and (b) 100-mg/ml calcium inserts. The estimated values overlap on the ground-truth curve, indicating accurate recoveries of the spectral responses of the corresponding materials.

IV. Conclusion and discussion

In this work, we have proposed a dual-energy CT setup for explosive detection based on nanotube X-ray source array with alternating low and high X-ray firing positions. The scan configuration is enabled by the optimization-based reconstruction method. A simulation study with a digital version of the standard clinical dual-energy phantom, GAMMEX DE-472, was carried out, where decomposition error and Poisson noise were considered in the data. Two specific inserts of different materials, 15-mg/ml iodine and 100-mg/ml calcium, were at focus in the investigation to mimic the responses of water and ANFO in explosive detection. Visual inspection shows that dual-energy setup proposed in this work can better differentiate the two materials types, that are otherwise of similar gray value in a single-energy CT image. Further, quantitative analysis revealed accurate recoveries of the spectral responses of the two materials by use of the proposed dual-energy CT with nanotube X-ray source arrays and the ASD-NC-PCOS algorithm. The relative error of the estimated linear attenuation coefficients averaged over 7 energy levels and over 2 material types is 0.77%.

While the setup investigated in this work assumes arc-shaped nanotube X-ray source arrays with rotating detector panels, it is straightforward to apply the optimization-based reconstruction method to other designs of hardware and geometry configurations, such as straight nanotube X-ray source array and stationary detector panels at opposing position. Further investigation of such configuration is underway and will be reported at the conference. While a dual-energy phantom was used in this work for the differentiation between an iodine insert and a calcium one, the underlying advantage and power for recovering spectral responses of different material types from the dual-energy CT scanning is expected to remain unchanged for a more direct task involving water and explosive materials such as ANFO. Further investigation will focus on phantoms with water and explosive ANFO materials, when the method incorporates interaction type based decomposition model using Compton scattering and photoelectric effect, in order to calculate the Z-image, or the effective atomic number [1]. These results will be reported at the conference.

Acknowledgment

This work was supported in part by NIH R01 Grants Nos. CA182264 and EB018102. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official NIH views.

References

[1] Z. Ying, R. Naidu, and C. R. Crawford, "Dual energy computed tomography for explosive detection," *Journal of X-Ray Science and Technology*, vol. 14, no. 4, pp. 235–256, Jan. 2006. [Online]. Available: https://content.iospress.com/articles/journal-of-x-ray-science-and-technology/xst00163

[2] S. Singh and M. Singh, "Explosives detection systems (EDS) for aviation security," *Signal Processing*, vol. 83, no. 1, pp. 31–55, Jan. 2003. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165168402003912

[3] B. Gonzales, D. Spronk, Y. Cheng, A. W. Tucker, M. Beckman, O. Zhou, and J. Lu, "Rectangular Fixed-Gantry CT Prototype: Combining CNT X-Ray Sources and Accelerated Compressed Sensing-Based Reconstruction," *IEEE Access*, vol. 2, pp. 971–981, 2014.

[4] G. Cao, Y. Z. Lee, R. Peng, Z. Liu, R. Rajaram, X. Calderon-Colon, L. An, P. Wang, T. Phan, S. Sultana, D. S. Lalush, J. P. Lu, and O. Zhou, "A dynamic micro-CT scanner based on a carbon nanotube field emission x-ray source," *Physics in Medicine & Biology*, vol. 54, no. 8, p. 2323, 2009. [Online]. Available: http://stacks.iop.org/0031-9155/54/i=8/a=005

[5] A. M. Hernandez and J. M. Boone, "Tungsten anode spectral model using interpolating cubic splines: Unfiltered xray spectra from 20 kV to 640 kV," *Medical Physics*, vol. 41, no. 4, pp. 042 101–1–042 101–15, Apr. 2014. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.1118/1.4866216/abstract

[6] JRTAssociates, "Dual energy CT phantom - JRT Associates," 2015. [Online]. Available: http://jrtassociates.com/pdfs/472_DECTPhantom2.pdf

[7] B. Chen, Z. Zhang, E. Y. Sidky, D. Xia, and X. Pan, "Image reconstruction and scan configurations enabled by optimization-based algorithms in multispectral CT," *Physics in Medicine & Biology*, vol. 62, no. 22, p. 8763, 2017. [Online]. Available: http://stacks.iop.org/0031-9155/62/i=22/a=8763

[8] Y. Zou and M. D. Silver, "Analysis of fast kV-switching in dual energy CT using a pre-reconstruction decomposition technique," in *Proc. SPIE, Medical Imaging: Physics of Medical Imaging*, vol. 6913, 2008, pp. 691 313–691 313–12. [Online]. Available: http://dx.doi.org/10.1117/12.772826

[9] X. Pan, B. Chen, Z. Zhang, E. Pearson, E. Sidky, and X. Han, "Optimization-based Reconstruction Exploiting Spectral Information in CT," in *The Third International Conference on Image Formation in X-Ray Computed Tomography*, 2014, pp. 228–232.

[10] M. Berger, J. Hubbell, S. Seltzer, J. Chang, J. Coursey, R. Sukumar, D. Zucker, and K. Olsen, "XCOM: Photon cross section database (version 1.5)," 2010. [Online]. Available: http://physics.nist.gov/xcom

Edge information diffusion based reconstruction (EIDR) for cone beam computed laminography

Yunsong Zhao, Jinqiu Xu, Hongwei Li*, and Peng Zhang

***Abstract*—Computed laminography (CL) reconstruction is a challenging task because incomplete projection data is acquired from the CL scanning. Conventional computed tomography reconstruction would blur the vertical edges (singularity in the z-direction) in the reconstructed images, while the horizontal edges (singularity within slices) can be quite accurately reconstructed. Based on this observation, an edge information diffusion method is developed, which fixes the horizontal edges and propagates their values within slices. Experiments on simulatation data are performed to verify the effectiveness of the proposed method.**

***Index Terms*—computed laminography, image reconstruction, edge information diffusion.**

I. Introduction

Micro-CT systems usually need full angular scanning and small source-to-detector distance (for sufficient magnification ratio). This scanning mode could not be applied for inspecting flat objects with large length-thickness ratio, simply because of geometrical limitations. Another reason is that the X-rays would be much more attenuated or even completely absorbed when penetrating the long edges of the flat objects, which could lead to poor reconstructions.

Computed laminography (CL), or laminographic scanning [1], [2], as illustrated in Figure 1, provides a solution for the above mentioned difficulties. An advantage of this scanning geometry is that the X-ray source can be very close to the scanned object to achieve a high magnification ratio and hence high-resolution reconstructions, even for large objects. Note that there exists other laminographic scanning geometries, please refer to [2].

CL scan can be achieved by adapting Computed Tomography (CT) device by tilting its rotation axis. It also can be thought of as a special circular cone beam CT scan with the detector placed far away from the mid-plane [2]. So, CL scan produces incomplete data, and direct reconstruction will introduce heavy cone angle artifacts [3], [4]. To reconstruct high quality images from incomplete projection data, strong prior information needs to be incorporated into the reconstruction algorithm. Total Variation regularization (TV) is one of the most popular algorithms, which could produce high quality images with very limited sparse-view data [5], [6]. DART [7] is another regularization method dedicated to limited-angle reconstruction, which incorporates the sparsity assumption that the scanned object consists of only a few components, each corresponding to a constant gray value in the reconstructed images. This idea is further utilized by TVR-DART [8], which estimates the gray values automatically.

Ihe cone angle artifacts appear mainly as z-direction inter-slice aliasing. However, the in-slice edges would be reconstructed rather accurately, as we have projection data corresponding to the X-ray paths that are tangent to the in-slice edges [9].

Based on this observation, we will propose an edge information diffusion based reconstruction (EIDR) method for flat objects imaging from cone beam CL scanning. The basic idea is that since the in-slice edges can be reconstructed correctly, and their gray values are trustable, then we can fix them and propagate their values to other areas upon the assumption that the true images are piecewise linear. This edge diffusion process is then utilized as a regularizer and combined with the SART [10] method which lead to an iterative reconstruction algorithm.

II. The idea of edge information diffusion

As mentioned previously, the CL scan is actually a circular cone beam CT scan with a very large cone angle and a displaced detector. It is well known that the data of circular cone beam CT scan is incomplete, and the non-central slices of the reconstructed images suffer from the so-called cone angle artifacts [3], [4]. The larger is the cone angle, the more serious are the artifacts. However, the in-slice edges are much

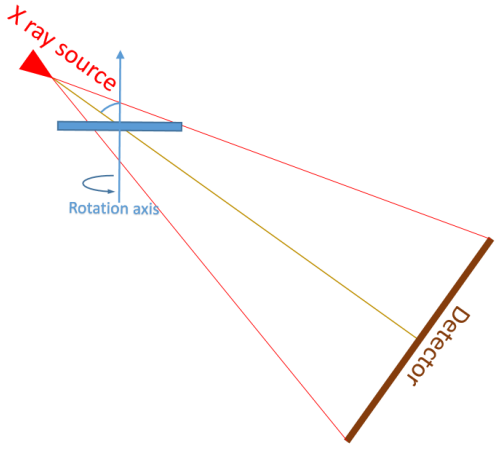


Fig. 1: Illustration of a CL scanning geometry.

Asterisk indicates corresponding author
Yunsong Zhao, Jinqiu Xu, Hongwei Li, and Peng Zhang are with School of Mathematical Sciences, Capital Normal University, Beijing, 100048, China, and are also with Beijing Advanced Innovation Center for Imaging Technology, Capital Normal University, Beijing, 100048, China.
e-mail: hongwei.li91@cnu.edu.cn

less affected by the cone angle artifacts, since we have X-ray paths that are tangent to these edges.

So in the reconstructed images by conventional algorithms like SART, the in-slice derivatives are actually rather accurate and trustable, such that the true image even could by recovered by just interpolating the values from the in-slice left and right edges. This is the key observation which motivates the proposed method. our method is an iterative one that consists of two main procedures. Let $u^{(n)}$ denotes the solution after the nth iteration. The first procedure of our method updates the image $u^{(n)}$ to obtain $u^{(n+1/2)}$ by the SART reconstruction algorithm from the projection data, then the second procedure corrects $u^{(n+1/2)}$ and obtain $u^{(n+1)}$ by propagating the in-slice edge information.

One difficulty of implementing the propagation procedure it that the in-slice image edges could not be identified accurately, i.e. the magnitude of the in-slice derivatives is not an trustable index for edge points. So a diffusion procedure will be introduced below to propagate the edge information in a trustable way. We propose the following minimization problem

$$\begin{aligned}&\min_u \Phi(u),\\ &\Phi(u) = \Phi(u;\phi) := \frac{1}{2}\int K(x,y,z)\left(u_x^2+u_y^2\right)dxdydz\\ &+\frac{\mu}{2}\int \lambda(x,y,z)\left(u(x,y,z)-\phi(x,y,z)\right)^2 dxdydz, \quad (1)\end{aligned}$$

where $\phi(x,y,z)$ refers to the current solution, i.e. $u^{(n+1/2)}$. The function $\lambda(x,y,z)$ is an edge indicator, whose value should be large for edge points while small for non-edge points. One could define $\lambda(x,y,z)$ to be a function of the gradient of $\phi(x,y,z)$, such as $\lambda = e^{c|\nabla\phi|^2}$ with scalar parameter $c>0$. It should be pointed out that the first term of $\Phi(u)$ involves only in-slice derivatives, i.e. u_x and u_y, while the inter-slice derivative u_z is missing. The proposed method can be summarized as the following **Algorithm 1**.

Algorithm 1 the pseudocode of the proposed method.

Start from some initial image $u^{(0)}$ and the projection data p
$n := 0$
while (stop criterion is not met) **do**
begin
Update the images with the acquired projections

$$u^{(n+1/2)} = \text{SART}\left(u^{(n)}, p\right)$$

Update the image by solving the optimization problem

$$u^{(n+1)} = \arg\min \Phi(u; u^{(n+1/2)})$$

$n := n+1$
end

Considering how to solve the optimization problem (1). As only derivatives with respect to the variable x and y are involved, the optimization problem is actually separable regarding to the variable z. So the original problem could be decoupled into a series of similar optimization problems defined for each fixed z, which could be formulated as

$$\begin{aligned}\Phi_z(u) = &\frac{1}{2}\int K(x,y)\left(u_x^2+u_y^2\right)dxdy\\ &+\frac{\mu}{2}\int \lambda(x,y)\left(u(x,y)-\phi(x,y)\right)^2 dxdy. \quad (2)\end{aligned}$$

The Euler-Lagrange equation of the above functional reads

$$\begin{aligned}&\mu\lambda(x,y)\left(u(x,y)-\phi(x,y)\right) - \frac{\partial}{\partial x}\left(K(x,y)u_x\right)\\ &-\frac{\partial}{\partial y}\left(K(x,y)u_y\right) = 0. \quad (3)\end{aligned}$$

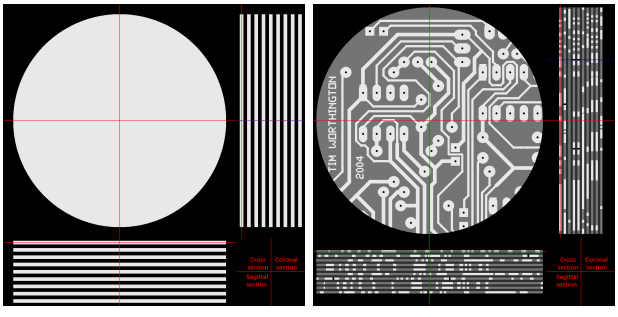
We adopt the well-known time-marching technique to solve the above problem:

$$\begin{cases} u_t(t;x,y) = -\mu\lambda(x,y)\left(u(t;x,y)-\phi(x,y)\right)\\ \qquad\qquad +\frac{\partial}{\partial x}\left(K(x,y)u_x\right)+\frac{\partial}{\partial y}\left(K(x,y)u_y\right)\\ u(0;x,y) = \phi(x,y)\\ u(t;x,y)\,|_{\partial\Omega} = 0\end{cases} \quad (4)$$

where $\partial\Omega$ denotes the boundary of the image. Then the Peaceman-Rachford scheme (PR) [11] is adopted, which is one of alternating direction implicit schems (ADI), or alternating direction multiplier methods (ADMM) [12], [13] in terms of the original energy functional.

III. Numerical experiments

Two phantoms, a disk phantom and a simulated piece of PCB, have been constructed for the simulations, as shown in Figure 2. The resolution of the two phantoms are both $1024\times1024\times300$. The disk phantom includes 9 disks stacked on each other and the thickness of each disk and the gaps between them are both 16 voxels. The PCB phantom includes 10 routing layers, each of which is of 10 voxels in thickness. The gap between the routing layers is also 10 voxels. Both noise-free and noisy projection data are simulated. The noisy data are obtained by adding Poisson noise to the noise-free data with mean value 10^6 for each detector cell. The SART method, DART method, and the proposed method are implemented to compare with each other.



(a) The disk phantom. (b) The PCB phantom.

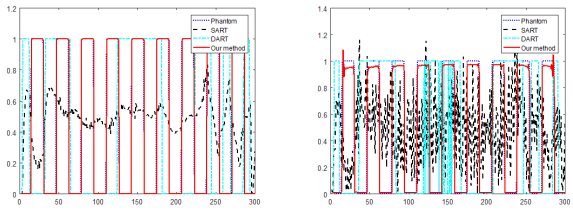
Fig. 2: Simulated disk phantom and PCB phantom.

A. Results with the disk phantom

The results for the disk phantom are shown in Figure 3, where the left column and the right column show the results

of the three methods for noise-free data and noisy data, respectively. Figure 3(a) and (b) demonstrate the results of the SART method after 20 iterations. Figure 3(c) and (d) show the results of the DART method after 100 iterations. For the proposed method, we also computed 100 iterations, and the results are shown in Figure 3(e) and (f). In each subfigure, three plane sections of the volume image, including the cross section, the coronal section, and the sagittal section of the top disk are shown for each method. The display window is set to $[0, 1.1]$.

For the results of the SART method, all the disks are mixed up and hardly distinguishable from each other. For the results of the DART method, some pseudo-structures are reconstructed, as shown in Figure 3(c) and (d). It seems that these structural artifacts can be attributed to its classification errors. DART method considers mainly the gray value information of the images, which encodes little structural information. When the reconstructed images by the SART method exhibit large distortions due to cone angle artifacts, the classification procedure of the DART method would result in large errors. When looking at the results of the proposed method shown in Figure 3(e) and (f), we can see that the disks are reconstructed correctly, and the inter-slice aliasing has been removed effectively. From Figure 3(f), one can also draw the conclusion that the proposed method is robust against noise.



(a) (b)

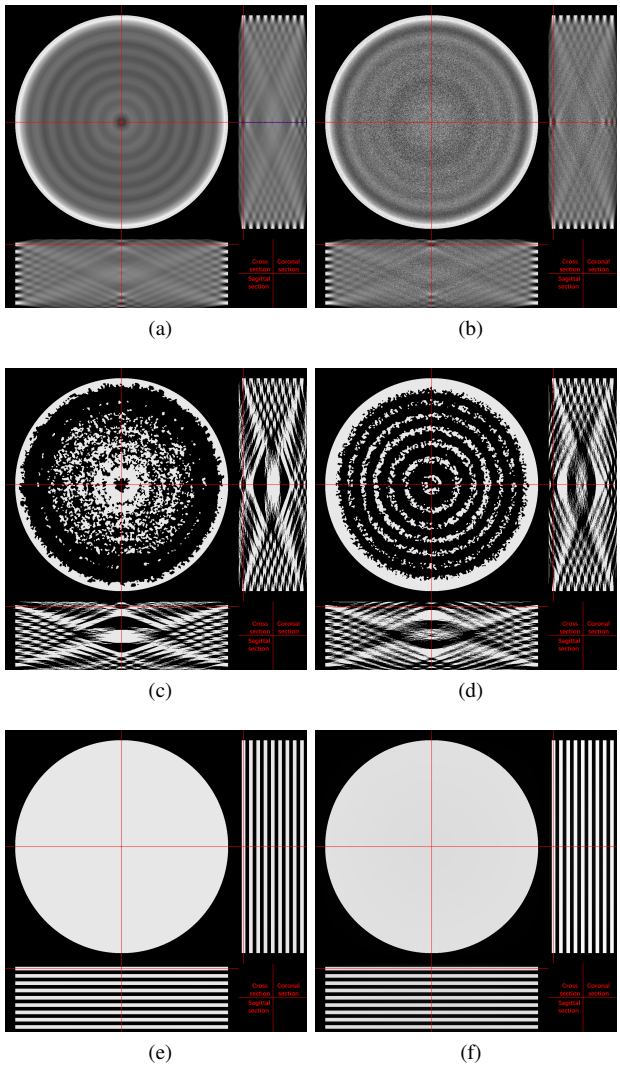
Fig. 4: The profiles of of the center rows of the sagittal sections of the reconstructed images shown in Figure 3. (a) is the profiles for noise-free data and (b) is the profiles for noisy data.

To examine the results in a quantitive way, we have plotted the profiles of the center rows of the sagittal sections of the reconstructed images by the three methods, which are shown in Figure 4. We can see that the reconstructed gray values with the proposed method are much more accurate than with the other two methods.

1) Results with the PCB phantom: Figure 5 shows the reconstruction results of the PCB phantom with the three methods, where the left column shows the results for noise-free data, while the right column shows the results for noisy data.

The strategies for the reconstructions of the disk phantom are again employed here for the reconstructions of the PCB phantom, except that the set of gray values for the DART method contains more elements here. The display window for the images is still $[0, 1.1]$.

[H]



(a) (b) (c) (d) (e) (f)

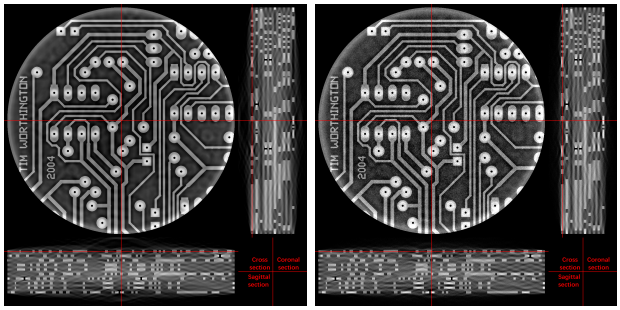
Fig. 3: The reconstruction results of the disk phatom with the SART method (a) and (b), the DART method (c) and (d), and the proposed method (e) and (f) from noise-free data (a), (c), and (e) and noisy data (b), (d), and (f). The display window is set to $[0, 1.1]$.

In all the three cross sections shown in Figure 5, the routings of the PCB can be seen. But the borders parallel to the PCB are blurred with the SART method. In the images reconstructed by the DART method, there are obvious classification errors, and wrong superfluous borders appear here and there. When inspecting the results with the proposed method, however, one can see that all the routings have been reconstructed correctly.

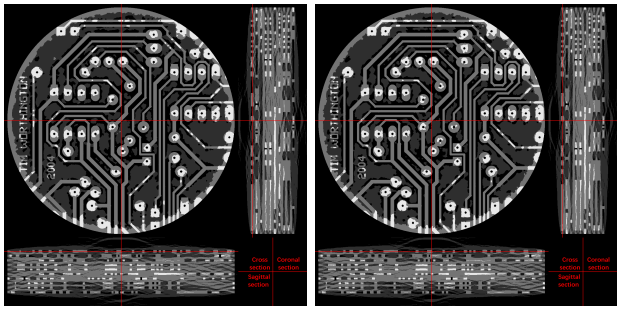
Figure 6 shows the profiles of the center rows of the sagittal sections of the images shown in Figure 5, which clearly shows that the proposed method constructs much more accurate gray values than the other two methods.

IV. Conclusion and Discussion

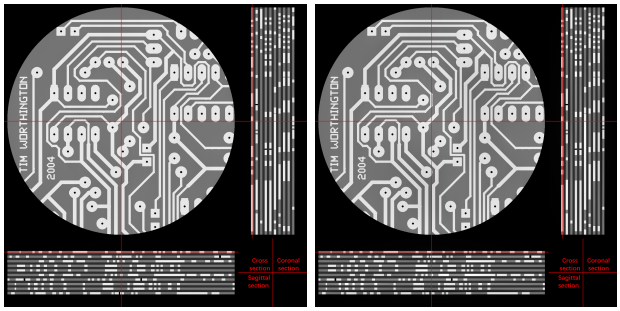
In this paper, a reconstruction method for CL system for inspecting flat objects is proposed, which can effectively



(a) (b)



(c) (d)



(e) (f)

Fig. 5: The reconstruction results of the PCB phantom with the SART method ((a) and (b)), the DART method ((c) and (d)), and the proposed method ((e) and (f)) from noise-free data ((a), (c), and (e)) and noisy data ((b), (d), and (f)). The display window is set to $[0, 1.1]$.

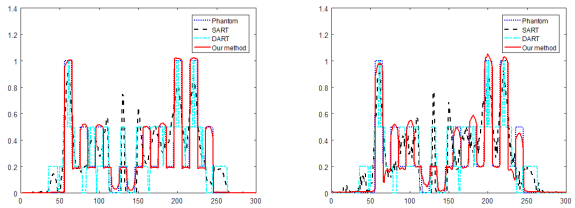
alleviate the inter-slice aliasing caused by the incompleteness of the scanned data. Experiments on simulated data suggest that the proposed method is effective, and outperforms the state-of-art methods.

The in-slice diffusion requires that the cross section of the image is parallel to the flat face. However, this is not exactly true in general situations. To overcome this limitation, a pre-reconstruction can be performed to obtain a initial image, from which we can estimate the geometric features of the objects.

REFERENCES

[1] S. Gondrom, J. Zhou, M. Maisl, H. Reiter, M. Kröning, and W. Arnold, "X-ray computed laminography: an approach of computed tomography for applications with limited access," *Nuclear engineering and design*, vol. 190, no. 1, pp. 141–147, 1999.

[2] N. S. OBrien, R. P. Boardman, I. Sinclair, and T. Blumensath, "Recent advances in x-ray cone-beam computed laminography," *Journal of X-ray science and technology*, vol. 24, no. 5, pp. 691–707, 2016.



(a) (b)

Fig. 6: The profiles of the center rows of the sagittal sections of the images shown in Figure 5. (a) is the profiles for noise-free data and (b) is the profiles for noisy data.

[3] J. F. Barrett and N. Keat, "Artifacts in CT: recognition and avoidance," *Radiographics*, vol. 24, no. 6, pp. 1679–1691, 2004.

[4] J. Hsieh, *Computed tomography: principles, design, artifacts, and recent advances*. SPIE press, 2003, vol. 114.

[5] E. Y. Sidky, C.-M. Kao, and X. Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT," *Journal of X-ray Science and Technology*, vol. 14, no. 2, pp. 119–139, 2006.

[6] H. Yu and G. Wang, "Sart-type image reconstruction from a limited number of projections with the sparsity constraint," *Journal of Biomedical Imaging*, vol. 2010, p. 3, 2010.

[7] K. J. Batenburg and J. Sijbers, "DART: a practical reconstruction algorithm for discrete tomography," *IEEE Transactions on Image Processing*, vol. 20, no. 9, pp. 2542–2553, 2011.

[8] X. Zhuge, W. J. Palenstijn, and K. J. Batenburg, "TVR-DART: a more robust algorithm for discrete tomography from limited projection data with automated gray value estimation," *IEEE Transactions on Image Processing*, vol. 25, no. 1, pp. 455–468, 2016.

[9] J. Frikel and E. T. Quinto, "Characterization and reduction of artifacts in limited angle tomography," *Inverse Problems*, vol. 29, no. 12, p. 125007, 2013.

[10] A. H. Andersen and A. C. Kak, "Simultaneous algebraic reconstruction technique (SART): a superior implementation of the ART algorithm," *Ultrasonic imaging*, vol. 6, no. 1, pp. 81–94, 1984.

[11] W. H. Hundsdorfer and J. G. Verwer, "Stability and convergence of the peaceman-rachford adi method for initial-boundary value problems," *Mathematics of Computation*, vol. 53, no. 187, pp. 81–101, 1989.

[12] D. Hawkins, "Constrained optimization and lagrange multiplier methods," *Athena Scientific, Belmont*, 1982.

[13] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

Anisotropic Resolution Enhancement for Computed Tomography of Planar Objects

Steffen Kieß, Sven Simon

***Abstract*—In computed tomography scans of large planar objects, the achievable resolution is limited by the requirement to rotate the object by 180° or 360°. Alternative methods like computed laminography allow a higher resolution by not rotating the object, but have difficulty resolving detail in the direction perpendicular to the plane because they lack the information from the projections where the rays go through the plane.**

The method presented in this paper improves the resolution of a CT scan by integrating the images taken for a normal CT, with a relatively low resolution, with high-resolution 2D images taken while the object is very close to the focal spot, similar to images taken for laminography. By combining these images using an iterative reconstruction, a volume can be produced which has a very high resolution in the plane but can also resolve structures perpendicular to the plane.

***Index Terms*—computed tomography, reconstruction, high-resolution, data fusion, laminography**

I. Introduction

In many applications of computed tomography (CT), the objects of interest are planar objects, i.e. objects where the thickness is significantly smaller than the width and the height. Examples for such objects are printed circuit boards (PCBs), chip packages and carbon fiber reinforced plastics. When performing a CT scan, the resolution depends on the distance between the X-ray focal spot and the object. The resolution can be improved by moving the object closer to the source, however CT reconstruction requires that the entire object is visible on the detector for all images. For planar objects this means that the minimum distance between the focal spot and the object – and therefore the resolution – is determined by the width of the object [1]. (E.g. in Figure 1a the object cannot be moved closer to the focal spot because then the object wouldn't fit on the detector anymore.)

For high-resolution 2D images, the object can be rotated so that it is parallel to the detector and moved significantly closer to the source (see Figure 2b, Figure 1c). In this case, only a part of the object will be visible, but by moving the object the entire object or those parts where a high resolution is desired can be scanned (see Figure 1d).

The resulting data set, consisting of both the low-resolution data and of the high-resolution data, can be used as input for an iterative reconstruction algorithm [2]. Iterative reconstruction, unlike the more commonly used filtered backprojection, can be used for data with irregular geometry, in this case the low-resolution images from the CT scan and the high-resolution 2D images.

S. Kieß and S. Simon are with the Institute for Parallel and Distributed Systems, University of Stuttgart, Stuttgart, Germany

An alternative to CT which allows moving the object closer to the focal spot is computed laminography [3]. However, computed laminography has the limitation that it lacks the information of the rays going through the plane and therefore does not fulfill the Tuy condition [4]. The method presented here allows combining the additional data obtained by placing the object close to the focal spot with the data for a full CT.

II. Acquisition Process

A. CT Scan

The acquisition process for the proposed method consists of two steps: First, a normal CT is performed (see Figure 2a). The result of this step is a set of images where the entire object is visible on the detector in all images (see e.g. Figure 1a and Figure 1b).

B. High-resolution images

In the second step, high-resolution images are taken. For this step the object is moved very close to the focal spot (see Figure 2b). Because only a small part of the object will be visible on the detector (see e.g. Figure 1c), the object will be moved until images of either the entire object or of the region of interest where a high resolution is desired have been taken (Figure 1d). In order to reduce the effects the borders of the high-resolution images have during the reconstruction and to provide an increased resolution in the direction perpendicular to the plane, the images are taken overlapping, so that every part of the object is visible in roughly 10×10 high-resolution images.

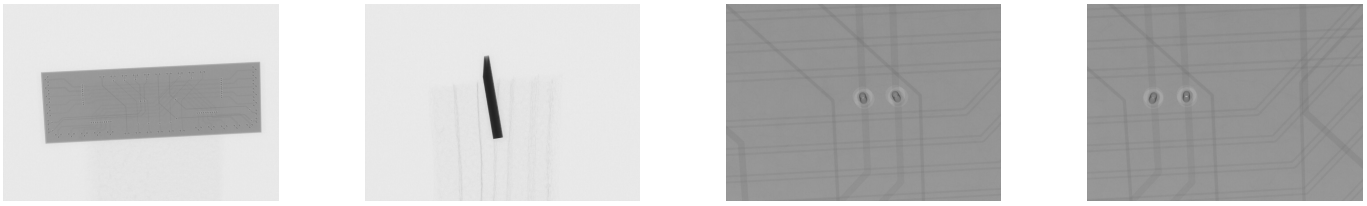
III. Reconstruction Algorithm

The reconstruction is done using an reconstruction iterative algorithm. Unlike the filtered backprojection algorithm, which assumes a regular geometry like cone-beam geometry [5], iterative reconstruction algorithms attempt to provide a general solution of the equation

$$Af = p \tag{1}$$

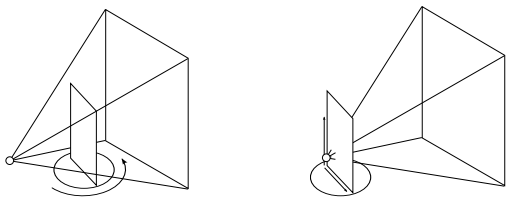
where p is the observed data (i.e. the set of images taken during the scan), f is the volume data and A is the projection matrix which is mapping the volume data f to the projections p. The information about the projection geometry is included in the projection matrix A, so that for reconstructing a data set with an irregular geometry using an iterative algorithm only the projection matrix has to be changed.

The reconstruction method used for this paper is the maximum-likelihood expectation-maximization (MLEM) [6].



(a) Object at 0°. (b) Object at 90°. (c) High-resolution image. (d) High-resolution image.

Fig. 1: (a) and (b): Images for normal CT, (c) and (d): high-resolution 2D images.



(a) (b)

Fig. 2: (a) Standard CT, (b) high-resolution 2D images.

An advantage of MLEM over algebraic reconstruction techniques is that it takes the stochastic nature of the X-ray photons into account. The MLEM starts with an initial assumption f_0 for the volume, often simply an volume where all voxels are set to 1. Then in every iteration the volume is updated according to the following equation:

$$f_{n+1} := f_n \cdot \frac{A^T \frac{p}{Af_n}}{A^T 1} \tag{2}$$

where A^T is the backprojection matrix (the transposed of the projection matrix), 1 is a projection data set where all the pixels of all images are 1 and the multiplications and divisions are done per voxel or pixel. The calculation of Ax and $A^T x$ are normally done by calculating the projection and backprojection directly, as representing the matrices for A and A^T, even in sparse form, would require a huge amount of memory.

For the reconstruction of the high-resolution data set, both the CT data set (acquired in Section II-A), called p_{CT}, and the high-resolution 2D data set (acquired in Section II-B), called p_{high}, are used. The update equation for the MLEM is then

$$f_{n+1} := f_n \cdot \frac{A_{CT}^T \frac{p_{CT}}{A_{CT} f_n} + A_{high}^T \frac{p_{high}}{A_{high} f_n}}{A_{CT}^T 1 + A_{high}^T 1} \tag{3}$$

where A_{CT} and A_{CT}^T are the forward and backward projection matrices for the normal CT data set and A_{high} and A_{high}^T are the forward and backward projection matrices for the high-resolution 2D data set.

IV. Resolution Analysis

In order to get information about the achieved improvement in quality, a simulated scan of a test target was performed. This input data for the simulation was a PCB with a size of 1 cm by 1 cm. The PCB contains a layer with test structures for determining the MTF. A slice of the input data for the simulation can be seen in Figure 3. Using this input data, a projection was performed to produce a normal, low-resolution CT data set and a set of high-resolution 2D images. The setup for the high-resolution images had a magnification factor which was 5 times as high as the magnification factor of the low-resolution images. Then both a normal CT reconstruction (using only the low-resolution data set, see Figure 4) and a high-resolution reconstruction (using both data sets, see Figure 5) was performed.

Fig. 3: Slice of input data for simulation.

In order to evaluate the improvement of resolution achievable with this method, the resolution of the reconstructed slices was determined using two methods shown in [7]. First, the modulation transfer function (MTF) was determined by looking at the sharpness of the transition between white and black in the middle box at the bottom of the images. Second, the contrast (i.e. the ratio between the dark areas minus the undisturbed background to the bright areas minus the undisturbed background) of all the line pairs was evaluated. The normalized MTF values and the line pair contrast values can be seen in Figure 6.

When a feature at a certain spacial frequency is considered visible if the MTF or the line pair contrast is at least 10 %, it is possible to determine the maximum spatial frequency where this condition is fulfilled. Figure 7a shows the spatial frequency where the MTF / the line pairs contrast is at

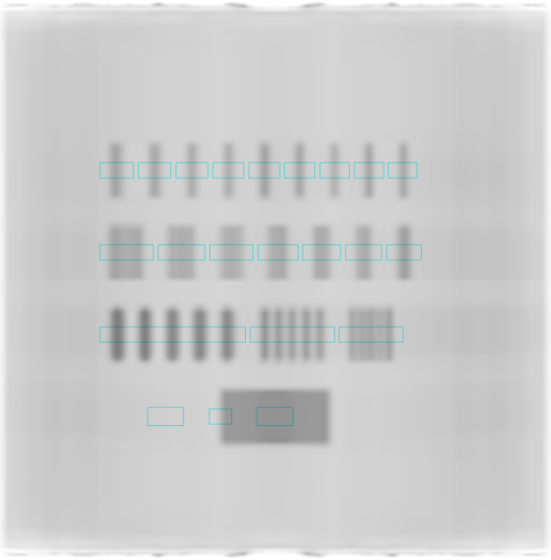


Fig. 4: Slice of reconstructed data without resolution enhancement.

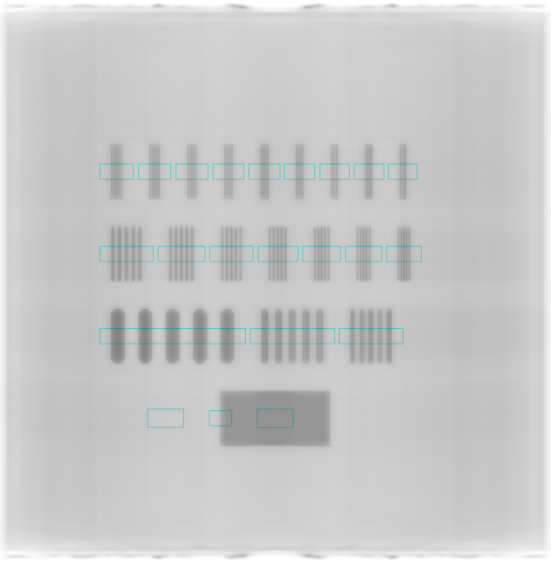


Fig. 5: Slice of reconstructed data with resolution enhancement, magnification for high resolution images is 5x the magnification of normal CT.

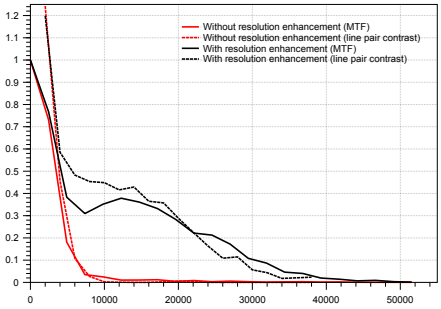
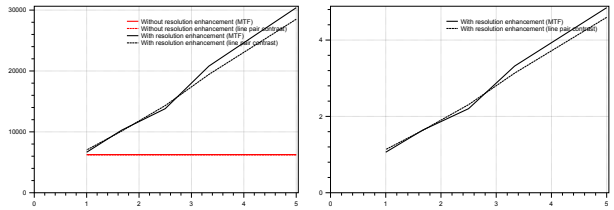


Fig. 6: MTF / line pair contrast depending on spatial frequency.



(a) (b)

Fig. 7: Spatial resolution (where MTF / contrast of line pairs is at least 10%) depending on increase in magnification for high-resolution images, (a) absolute resolution, (b) relative improvement over reconstruction without resolution enhancement.

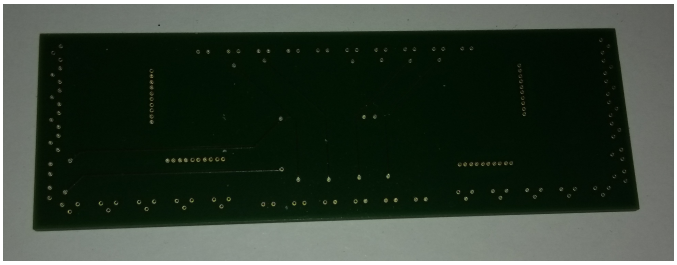


Fig. 8: Photo of object used for experimental verification.

least 10 % for different ratios between the magnification factor of the high-resolution images and the low-resolution images. Figure 7b shows the same information relative to the values for the data set without resolution enhancement, i.e. the improvement achieved by using the high-resolution images. It can be seen that the high-resolution images offer almost a linear improvement of the achieved resolution when the resolution of the high-resolution images is growing, however this does not take into account experimental problems like inaccuracies in the geometry of real-life CT scanners.

V. Experimental Results

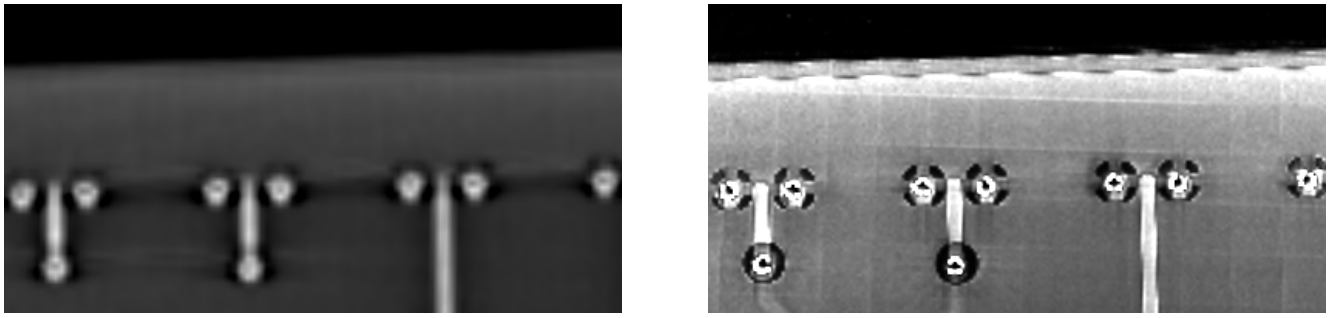
To verify the proposed method experimentally, a scan of the test object seen in Figure 8 was performed according to the method in Section II. The resulting data set, consisting of a CT data set p_{CT} and a high-resolution 2D data set p_{high}, contains the raw images seen in Figure 1.

After the acquisition of the data, a reconstruction of the CT data (p_{CT}) using filtered backprojection was performed:

$$v_{CT} := A_{CT}^T F p_{CT} \tag{4}$$

where F is the filter used for the filtered backprojection. A slice through a part of the data can be seen in Figure 9a. Due to the low resolution of the CT data the details of the vias and transmission lines are blurred.

In a second step, a high-resolution reconstruction using the method presented in Section III was performed. Due to mechanical inaccuracies of the CT scanner, there is a difference between the target position of the manipulator and the actual position of the manipulator, which means that the precise geometry of the high-resolution projections is unknown. Because incorrect geometry information causes strongly visible



(a) (b)

Fig. 9: (a) Slice of filtered backprojection of CT data, (b) slice of reconstruction using CT data and high-resolution 2D images.

artifacts, the geometry data had to be corrected. For this, a projection of the filtered backprojection p_{CT} was performed:

$$r_{high} := A_{high} p_{CT} \tag{5}$$

r_{high} contains a set of images similar to the original high-resolution images p_{high}, but with a lower resolution and with exactly the expected geometry. By determining the rotation and translation between the images in r_{high} and p_{high} with an algorithm similar to [8], the actual geometry of the each of the images in p_{high} can be determined. This geometry can then be used for reconstruction using MLEM.

A slice through a part of the data reconstructed this way can be seen in Figure 9b. It can be seen that integrating the high-resolution images provides a significantly higher level of detail for features like the vias and transmission lines in the slice.

VI. Conclusion

A method for improving the resolution of a 3D volume data set by integrating high-resolution 2D images taken with a reduced focal-object distance has been presented. Quantitative analysis of the MTF using a simulated test structures shows that the resolution of the data set grows almost linearly with the magnification of the high-resolution images. When the magnification for the high-resolution images is increased by a factor of 5, the resolution of the reconstructed volume is improved by a factor of 4.6 measured using the MTF method and by a factor of 4.2 measured using the line-pair method.

The method has also been applied to PCBs. For this PCB, both a normal reconstruction using only the CT raw images and a high-resolution reconstruction using both the CT raw images and additional high-resolution 2D images have been performed. The high-resolution shows a significantly better level of detail than the CT method.

References

[1] T. M. Buzug, *Computed tomography: From photon statistics to modern cone-beam CT*, en. Springer, 2008, ISBN: 9783540394075.

[2] S. Kieß, J. Guhathakurta, J. Hillebrand, J. Denecke, I. Effenberger, and S. Simon, "Computed tomography resolution enhancement by integrating high-resolution 2d x-ray images into the ct reconstruction," in *Int. Symposium on Digital Industrial Radiology and Computed Tomography 2015*, 9 Seiten. [Online]. Available: http://www.ndt.net/search/docs.php3?showForm=OFF&id=18046.

[3] F. Xu, L. Helfen, T. Baumbach, and H. Suhonen, "Comparison of image quality in computed laminography and tomography," *Opt. Express*, vol. 20, no. 2, pp. 794–806, Jan. 2012. DOI: 10.1364/OE.20.000794.

[4] H. K. Tuy, "An inversion formula for cone-beam reconstruction," *SIAM Journal on Applied Mathematics*, vol. 43, no. 3, pp. 546–552, 1983. DOI: 10.1137/0143035. eprint: https://doi.org/10.1137/0143035.

[5] L. A. Feldkamp, L. C. Davis, and J. W. Kress, "Practical cone-beam algorithm," *J. Opt. Soc. Am. A*, vol. 1, no. 6, pp. 612–619, Jun. 1984. DOI: 10.1364/JOSAA.1.000612.

[6] L. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *Medical Imaging, IEEE Transactions on*, vol. 1, no. 2, pp. 113–122, Oct. 1982, ISSN: 0278-0062. DOI: 10.1109/TMI.1982.4307558.

[7] A. Staude and J. Goebbels, "Determining the spatial resolution in computed tomography – comparison of mtf and line-pair structures," in *Int. Symposium on Digital Industrial Radiology and Computed Tomography*, vol. Qualification, 2011. [Online]. Available: http://www.ndt.net/search/docs.php3?showForm=OFF&id=11138.

[8] C. Wilson and J. Theriot, "A correlation-based approach to calculate rotation and translation of moving cells," *Image Processing, IEEE Transactions on*, vol. 15, no. 7, pp. 1939–1951, Jul. 2006, ISSN: 1057-7149. DOI: 10.1109/TIP.2006.873434.

A second derivative based regularization model for limited-angle computed tomography

Jinqiu Xu, Shiwo Deng, Huitao Zhang, Yining Zhu*, and Peng Zhang

***Abstract*—In limited-angle Computed Tomography (CT), the reconstructed image often suffers from severe distortions. In order to deal with this problem, we analyze the features of images reconstructed from limited-angle data and propose a prior knowledge on the second derivative of image. Based on the prior knowledge, we establish an optimization model and the corresponding algorithm for limited-angle CT. Numerical experiments are performed, and the results validate the effectiveness of the proposed method to remove image distortions.**

***Index Terms*—Computed Tomography, limited-angle, prior knowledge, second derivative**

I. Introduction

Computed Tomography (CT) is a nondestructive technique for obtaining interior structure information of an object. It has a wide range of applications in industry, medicine, etc. In the applications, the incomplete data problems occur frequently. We consider the limited-angle CT problem in this paper. The typical examples are the 3D imaging for the large flat objects with high aspect ratio (area-to-thickness) and the reconstruction of clinical image with C-arm equipments, for which limited-angle projection data can be acquired. In this situation, the reconstructed images often suffer from obvious CT value distortions.

In order to reduce the CT value distortions for the limited-angle CT, some kinds of prior information are added to image reconstruction process as additional constraints or regularization terms. The prior information refers to knowledge about the system or the inspected object [1], which includes the object's boundary, shape, gray level, etc. In certain medical and industrial applications, the image is approximately piecewise constant. Many reconstruction methods based on theory of compressed sensing (CS) [2] utilized the sparsity of gradient image as prior information. Among them, a well-known regularization term is the total variation (TV) norm [3]. Many methods utilized the TV norm as regularization term [4]-[9]. Recent years, the minimization of l_0 norm of the image gradient was applied to limited angle reconstruction [10]. It was demonstrated that the l_0 norm works better at preserving edges than TV. These methods mentioned above can suppress artifacts in reconstructed images at a certain degree. However, in practical applications, the reconstructed image are always affected by hardening artifacts or metal artifacts, so, the image does not usually satisfy gradient sparsity.

In order to deal with this problem, we consider use the the l_0 norm of image second derivative as the regularization term. Additionally, we found that when applying conventional algorithms like the Algebraic reconstruction techniques (ART) [11] or Simultaneous Algebraic Reconstruction Technique (SART) [12] on limited-angle data, edges of the reconstructed image should be rather accurately recovered along certain directions, while blurred along the other perpendicular directions. Therefore, we intend to utilize the value of the edge along certain direction to extend the value of internal region. The proposed method utilizes the l_0 norm of image second derivative along certain direction as a regularization term. Numerical experiments verify our method efficiently avoids distortions and provides high-quality reconstructed results, while existing methods suffer from obvious distortions.

II. Method

A. The Limited-angle Problem

In this paper, as preliminary study, we focus on two-dimensional (2D) imaging. The scanning configuration for the limited-angle fan-beam CT is illustrated in Fig. 1. In the Oxy coordinates system, S is the x-ray focal spot, D is the center of the detector and u is the location of the detector units. Let R denotes the distance between S and O, $\vec{x} = (x, y)$, $\vec{\tau}(\beta) = (cos(\beta), sin(\beta))$, $\beta \in (\theta, \pi - \theta)$, $0 < \theta < \frac{\pi}{2}$,

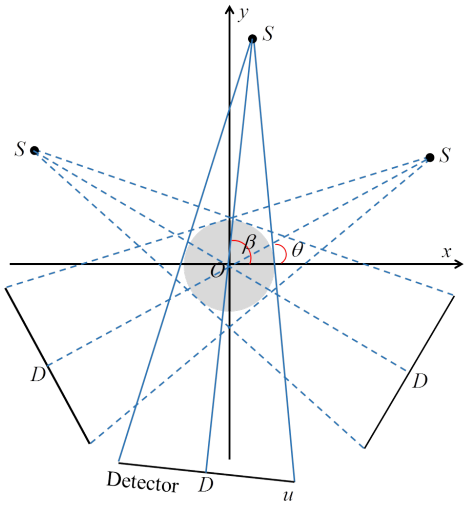


Fig. 1. Scanning angular range configuration.

Jinqiu Xu, Shiwo Deng, Huitao Zhang, Yining Zhu, and Peng Zhang are with the school of Mathematical Sciences, Capital Normal University, Beijing, 100048, China, and are also with Beijing Advanced Innovation Center for Imaging Technology, Beijing, China. (e-mail: jinqiuxu@cnu.edu.cn; 2150501018@cnu.edu.cn; htzhang@cnu.edu.cn; *corresponding author, ynzhu@cnu.edu.cn; pzhang@cnu.edu.cn)

$\vec{v}(u)$ denotes the vector from S to detector unit u and $f(\vec{x})$ represents the the linear attenuation distribution of the object to be scanned. Then fan-beam projection data for $f(\vec{x})$ can be expressed as

$$p(u,\beta) = \int_0^1 f\left(R\vec{\tau}(\beta) + t\vec{v}(u)\right) dt \tag{1}$$

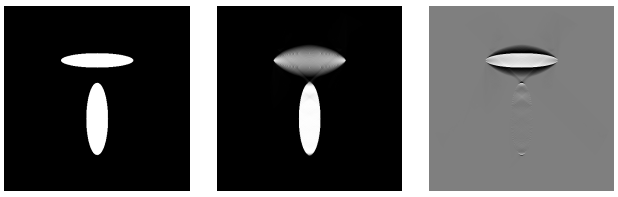
Then solving $f(\vec{x})$ from a set of p in (1) $\left(0 < \theta < \frac{\pi}{2}\right)$ is the limited-angle reconstruction problem. In practical applications, the limited-angle problem can be expressed as following linear system:

$$A\vec{f} = \vec{p} \tag{2}$$

A is the system matrix which represents the forward projection, $\vec{p}$ denotes the projection data vector and $\vec{f} = \{f_n | n = 1, 2, ..., K \times K\}$ is the vector of the reconstructed image, where K is the number of rows or columns. Since the projection data is incomplete which means (2) is undetermined, we could not obtain $\vec{f}$ accurately or with high-quality by using conventional algorithm.

B. Motivation

In this subsection, we will use a simple phantom to analyze the feature of reconstructed image. The simple phantom and the reconstructed result by using SART are shown in Fig. 2(a) and (b), and Fig. 2(c) shows the differences relative to the phantom image. One can see that the edges tangent to the X-Rays should be easy to reconstruct (which are called visable edges), while edges not tangent to any X-Ray would be difficult to reconstruct, which has been proved in [13]. Therefore the gradient or second derivative of the visible edges are reliable. In this paper, we assume the second derivative image is approximately sparse, one can fix the visible edges and diffuse their values to other areas. Whereas the diffusion might introduce striping artifacts, we introduce edge-preserving smoothing operator to suppress such possible artifacts and improve the reconstructed image further.



(a) (b) (c)

Fig. 2. Reconstructed results of the phantom with $\theta = \pi/4$. (a) is the phantom image and (b) is corresponding reconstructed image. The gray scale window is set to [0, 1]. (c) is the difference relative to the phantom image. The gray scale window is set to [-0.5, 0.5].

C. Model

Based on the principles that be discussed above, the proposed model for the limited-angle CT is described in (3).

$$\vec{f}^{*} = \arg\min_{\vec{f}} \quad \left(\|A\vec{f} - \vec{p}\|_2^2 + \lambda\|D_{xx}\vec{f}\|_0 + \beta S\left(\vec{f}\right)\right) \tag{3}$$

where $\|\cdot\|_2$ and $\|\cdot\|_0$ are the l_2 norm and l_0 norm, respectively. $\|A\vec{f} - \vec{p}\|_2^2$ is the data fidelity term, the regularization term $\|D_{xx}\vec{f}\|_0$ is used for propagating the edge values to non-edge points, where $D_{xx}\vec{f}$ is the second derivative operator. The regularization term $S\left(\vec{f}\right)$ aims at edge-preserve smoothing. $S\left(\vec{f}\right)$ can be l_1 norm of gradient, l_0 norm of gradient, l_0 norm of second derivative and median filter along the vertical y-direction. It can also be the l_0 norm of the image gradient. The parameters λ and β are used for balancing the effect of the two regularization terms. The second term satisfies (4),

$$\|D_{xx}\vec{f}\|_0 \triangleq \sum_{i=2}^{K-1} sgn\left(|f_{i-1,j} - 2f_{i,j} + f_{i+1,j}|\right) \tag{4}$$

where sgn denotes the sign function, i, j denote the row and column indices separately and $f_{i,j} = f_n\,(n = (i-1) \times K + j)$.

D. Algorithm

The optimization model (3) contains one data fidelity term and two regularization terms. Inspired by alternating direction method of multipliers (ADMM) [14], we split the model into three sub-problems according to the three terms of (3) separately. Let the initial image reconstructed by SART with 10 iterations denotes the initial image. Suppose that we have known $\vec{f}^{(k)}$ after k iterations for (3), $\vec{f}^{(k+1)}$ is obtained by solving the following three sub-problems:

Subproblem 1:

$$\vec{f}^{(k+1/3)} = \arg\min_{\vec{f}} \quad \left(\|\vec{f}^{(k)} - \vec{f}\|_2^2 + \|A\vec{f} - \vec{p}\|_2^2\right) \tag{5}$$

Subproblem 2:

$$\vec{f}^{(k+2/3)} = \arg\min_{\vec{f}} \quad \left(\|\vec{f}^{(k+1/3)} - \vec{f}\|_2^2 + \lambda\|D_{xx}\vec{f}\|_0\right) \tag{6}$$

Subproblem 3:

$$\vec{f}^{(k+1)} = \arg\min_{\vec{f}} \quad \left(\|\vec{f}^{(k+2/3)} - \vec{f}\|_2^2 + \beta S\left(\vec{f}\right)\right) \tag{7}$$

For Subproblem 1, in this paper, the classical iterative algorithm: SART is used to solve it and the iteration number is one. For subproblem 2, it involves a discrete counting metric, thus we adopt the alternation optimization strategy [15]. One auxiliary variable $\vec{g}$ is introduced corresponding to $D_{xx}\vec{f}$, then subproblem-2 is transformed into optimization problem (8).

$$\arg\min_{\vec{f},\vec{g}} \quad \left(\|\vec{f}^{(k+1/3)} - \vec{f}\|_2^2 + \lambda\|\vec{g}\|_0 + \lambda'\|D_{xx}\vec{f} - \vec{g}\|_2^2\right) \tag{8}$$

where λ' is a weighting parameter. For (8), $\vec{g}$ and $\vec{f}$ are solved in the manner of alternate iteration, so the following two steps need to be solved:

Step 1: computing $\vec{g}$

$$\arg\min_{\vec{g}} \quad \left(\lambda\|\vec{g}\|_0 + \lambda'\|D_{xx}\vec{f} - \vec{g}\|_2^2\right) \tag{9}$$

The solution can be achieved by using hard thresholding shrinkage operator.

Step 2: computing $\vec{f}$

$$\arg\min_{\vec{f}} \quad \left(\|\vec{f}^{(k+1/3)} - \vec{f}\|_2^2 + \lambda^{'}\|D_{xx}\vec{f} - \vec{g}\|_2^2\right) \qquad (10)$$

The solution will be achieved by using the gradient descent method.

For subproblem 3, if $S\left(\vec{f}\right)$ is the l_1 norm of gradient along vertical direction, then the algorithm in [3] can be used. If $S\left(\vec{f}\right)$ is the l_0 norm of second derivative along vertical direction, then its solution algorithm is similar to subproblem 2. If $S\left(\vec{f}\right)$ is $\|D_y\vec{f}\|_0$, it can actually be decoupled into a series of one dimensional problems, then be solved according to the algorithm proposed in [15]. If $S\left(\vec{f}\right)$ is the l_0 norm of image gradient, the equation of solution has been provided in [15].

III. Experiments

A.

In order to verify the effectiveness to remove distortions, robustness with respect to noise of the proposed method, this section performs experiments on two phantoms which are utilized for testing medical and industrial applications, respectively. Phantom 1 is Forbild head phantom and phantom 2 is synthesized by combining ellipses, disks, squares and polygon regions as shown in Fig. 3. TV based method, l_0 of gradient based method were adopted for comparison. The iterations are stopped after 1000 iterations. The scanning parameters are listed in Table I and the reconstruction parameters are listed in Table II. In this subsection, $S\left(\vec{f}\right)$ is chosen as l_0 norm of gradient along the vertical y-direction.

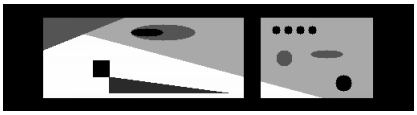


Fig. 3. Phantom 2 that is used for simulation experiments.

TABLE I
Scanning parameters for phantom 1 and 2.

Parameter	Value
Scanning type	Fan-beam Scan
SO	300 mm
SD	600 mm
θ	$\pi/4$
Number of Projections	179
Detector unit width	0.127 mm
Number of detector units	1024
Image Pixel Size	0.1061×0.1061 mm^2

TABLE II
Reconstruction parameters for phantom 1 and 2.

	Phantom	TV based method	l_0 of gradient based method	The proposed method
Noise-free	Phantom 1	$N = 40$ $\alpha = 0.004$	$\lambda_{l_0} = 0.001$	$\lambda = 0.0001$ $\beta = 0.00001$
	Phantom 2	$N = 20$ $\alpha = 0.01$	$\lambda_{l_0} = 0.002$	$\lambda = 0.001$ $\beta = 0.0002$
Noisy	Phantom 1	$N = 40$ $\alpha = 0.008$	$\lambda_{l_0} = 0.004$	$\lambda = 0.002$ $\beta = 0.0003$
	Phantom 2	$N = 20$ $\alpha = 0.02$	$\lambda_{l_0} = 0.005$	$\lambda = 0.004$ $\beta = 0.0003$

Figure 4 and Fig. 5 show the reconstructed results for phantom 1 and phantom 2 using noise-free projections . The blocks in the middle of the reconstructed image are the zoom-in views of the corresponding regions. We can see that TV based method and l_0 of gradient based method introduce various artifacts and the reconstructed images are severely distorted. On the other hand, the proposed method produces near perfect reconstruction.

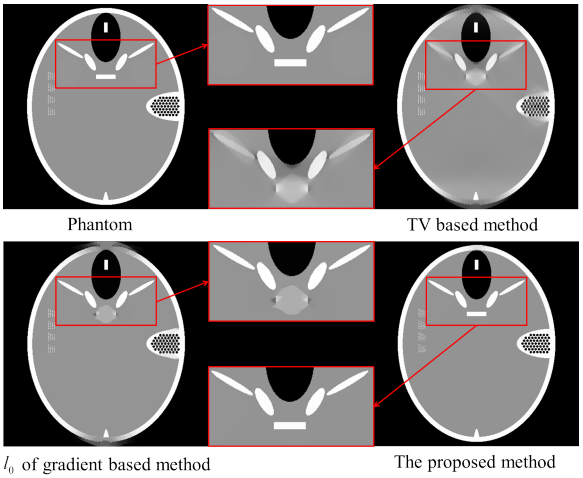


Fig. 4. Reconstructed results of phantom 1 using noise-free projections with $\theta = \pi/4$. The gray scale window is set to [0,1].

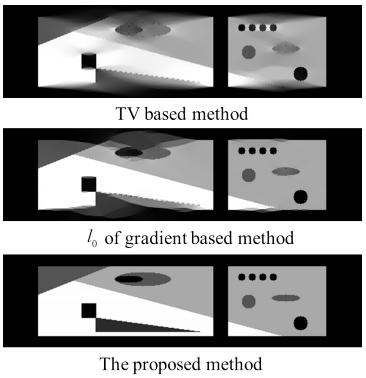


Fig. 5. Reconstructed results of phantom 2 using noise-free projections with $\theta = \pi/4$. The gray scale window is set to [0,1].

The next experiment is to test how the methods perform against noise. The noisy projections are simulated by adding

Poisson noise to the ideal raw data with 10^6 initial photons for each x-ray path. The reconstructed results are shown in Fig.6 and Fig.7. Compared to the noise-free case, the noise could reduce the quality of images reconstructed by TV based method and l_0 of gradient based method, and more artifacts might be introduced. The proposed method seems completely not aware of the noise, and the reconstructed result is almost the same as the noise-free case.

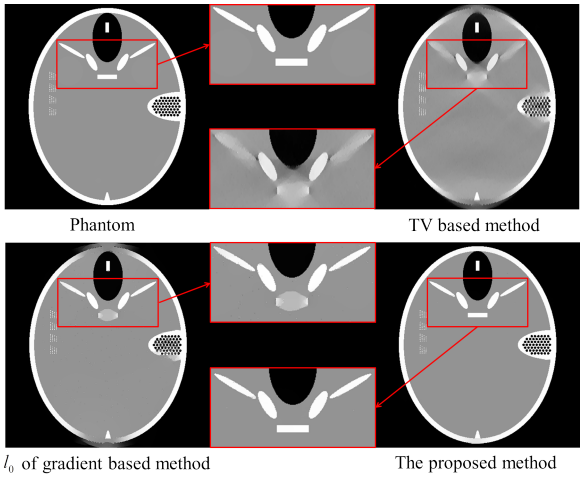


Fig. 6. Reconstructed results of phantom 1 using noisy projections with $\theta = \pi/4$. The gray scale window is set to [0,1].

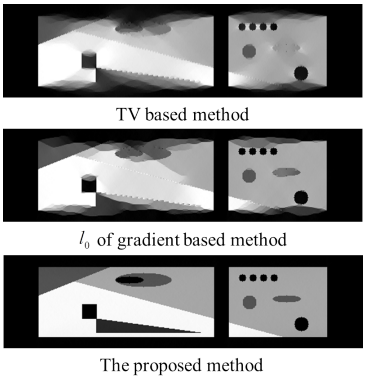


Fig. 7. Reconstructed results of phantom 2 using noisy projections with $\theta = \pi/4$. The gray scale window is set to [0,1].

To assess the reconstruction performances in quantitative way, Table III and IV list the value of Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index (SSIM) of all reconstructed results in Figs.4-7. We can find out that the PSNR and SSIM of the proposed method are consistently higher than those of TV based method and l_0 of gradient based method. Above quantitative and qualitative investigation of results demonstrate the proposed method outperforms other two methods.

TABLE III
PSNR AND SSIM OF THE RECONSTRUCTED RESULTS FOR PHANTOM 1.

	Measure	TV based method	l_0 of gradient based method	The proposed method
Noise-free	PSNR	22.6909	24.4284	**35.8581**
	SSIM	0.9761	0.9841	**0.9994**
Noisy	PSNR	21.9560	23.5912	**32.1339**
	SSIM	0.9715	0.9807	**0.9991**

TABLE IV
PSNR AND SSIM OF THE RECONSTRUCTED RESULTS FOR PHANTOM 2.

	Measure	TV based method	l_0 of gradient based method	The proposed method
Noise-free	PSNR	17.0032	17.8559	**34.7696**
	SSIM	0.9265	0.9326	**0.9994**
Noisy	PSNR	16.0327	16.1885	**31.5604**
	SSIM	0.9176	0.9241	**0.9992**

B.

As mentioned in the section I, in practical application, there always exists hardening artifacts or metal artifacts in the reconstructed image. This subsection studies the robustness with respect to the metal artifacts. $S\left(\vec{f}\right)$ is chosen as l_0 norm of image gradient.

We used the phantom shown in Fig. 8, which is designed according to a real reconstructed image of a monkey, in addition, we add a small circular disk filled with titanium. The phantom image is consist of water ($1.0g/cm^3$), bone ($1.7g/cm^3$) and titanium ($4.54g/cm^3$). We simulated the X-ray spectra of GE Maxiray 125 X-ray tube with tube voltages 120 kV. The mass attenuation coefficients used for simulating polychromatic projections are retrieved from the National Institute of Standard Technology (NIST) tables of X-ray mass attenuation coefficient [16]. In the simulation of polychromatic projections, the sampling subintervals are 1 keV for both X-ray spectra and the mass attenuation coefficient of the three materials. The number of projections is 279, $\theta = \pi/9$, the reconstruction size is 512×512 and the pixel size is $0.0804 \times 0.0804mm^2$. The other scanning parameters are the same as those in Table I.

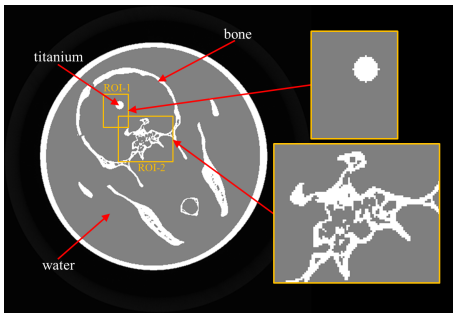


Fig. 8. Phantom 3 that is used for numerical experiments. The gray scale window is set to [0,1].

Figure 9 shows a comparison of the results reconstructed with the TV based method, l_0 of gradient based method and the proposed method. The iterations for these methods are stopped

after 1000 iterations. The reconstruction parameters are listed in Table V. The middle blocks in Fig.9 are the zoom-in views of the corresponding regions. We can see that the TV based method suffers from the strongest artifacts. The results of the l_0 of gradient based method shows slight metal artifacts in ROI-1, and there exists blurring in ROI-2. While in the images reconstructed with the proposed method, there are no evident artifacts. Therefore, the proposed method is superior to the classic reconstruction method for removing metal artifacts.

TABLE V
RECONSTRUCTION PARAMETERS FOR PHANTOM 3.

TV based method	l_0 of gradient based method	The proposed method
$N = 40$ $\alpha = 0.03$	$\lambda_{l_0} = 0.02$	$\lambda = 0.08$ $\beta = 0.005$

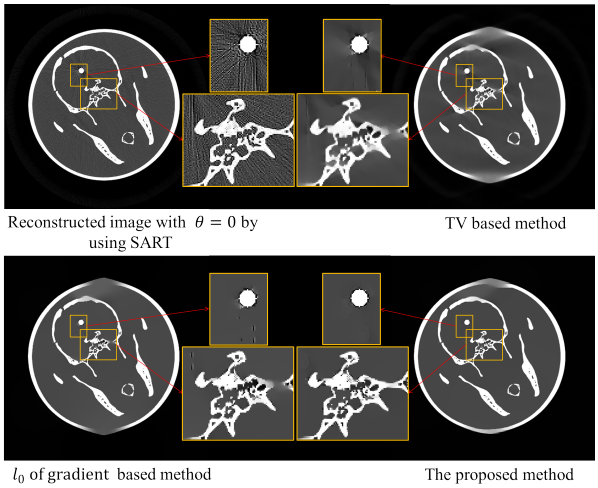


Fig. 9. Reconstructed image of phantom 3. For the reconstructed image by using the TV based method, l_0 of gradient based method and the proposed method, $\theta = \pi/9$. The gray scale window is set to [0,1].

IV. CONCLUSION

In this paper, we establish an optimization model based on second derivative regularization to deal with the distortions introduced by the limited-angle tomography. Numerical experiments are carried out to test the proposed method. The results illustrate the proposed method evidently suppresses distortions, removes metal artifacts and recovers image structures well.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China under the grants 61501310, 61671311 and by the Beijing Municipal Commission of Education grant KM201710028002.

REFERENCES

[1] N.S. OBrien, R.P. Boardman, I. Sinclair, and T. Blumensath, "Recent advances in X-ray cone-beam computed laminograph," J. X-Ray Sci Technol., Vol. 24, no. 5, pp. 691-707, Oct. 2016.

[2] E.J. Candes, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," Commun. Pure Appl. Math., Vol. 59, no. 8, pp. 1207-1223, Mar. 2006.

[3] E.Y. Sidky, C. Kao and X. Pan "Accurate image reconstruction from few-views and limited-angle data in divergentbeam CT," J. X-Ray Sci. Technol., Vol. 14, no. 2, pp. 119-139, Mar. 2006.

[4] E.Y. Sidky, and X. Pan, "Image reconstruction in circular cone-beam computed tomography by total variation minimization," Phys. Med. Biol., Vol. 53, no. 17, pp. 4777-4807, Sep. 2008.

[5] G.H. Chen, J. Tang, and S. Leng, "Prior image constrained compressed sensing (PICCS): a method to accurately reconstruct dynamic CT images from highly undersampled projection data sets," Med. Phys., Vol. 35, no. 8, pp. 660-663, Feb. 2008.

[6] H. Yu, and G. Wang, "A soft-threshold filtering approach for reconstruction from a limited number of projections," Phys. Med. Biol., Vol. 55, no. 13, pp. 3905-3516, Jul. 2010.

[7] L. Ritschl, F. Bergner, C. Fleischmann, and M. Kachelriess, "Improved total variation-based CT image reconstruction applied to clinical data," Phys. Med. Biol., Vol. 56, no. 6, pp. 1545-1561, Mar. 2011.

[8] X. Jin, L. Li, Z. Chen, L. Zhang, and Y. Xing, "Anisotropic total variation for limited-angle CT reconstruction," IEEE Nucl. Sci. Symp. Conf. Rec. (NSS/MIC), Vol. 8506, pp. 2232-2238, Nov. 2010.

[9] Z.Q. Chen, X. Jin, L. Li, and G. Wang,"A limited-angle CT reconstruction method based on anisotropic TV minimization," Phys. Med. Biol., Vol. 58, no. 7, pp. 2119-2141, Apr. 2013.

[10] W. Yu, and L. Zeng, "l_0 gradient minimization based image reconstruction for limited-angle computed tomography," Plos one, Vol. 10, no. 7, pp. e0130793, Jul. 2015.

[11] R. Gordon, R. Bender, and G.T. Herman, "Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and x-ray photography," J. Theor. Biol., Vol. 29, no. 3, pp. 471-478, Dec. 1970.

[12] A.H. Andersen and A.C. Kak, "Simultaneous algebraic reconstruction technique (SART): a superior implementation of the ART algorithm," Ultrason. imaging, Vol. 6, no. 1, pp. 81-94, Jan. 1984.

[13] E.T. Quinto, "Artifacts and visible singularities in limited data X-Ray tomography," Sens. Imaging, Vol. 18, no. 1, pp. 9-23, Jan. 2017.

[14] S. Boyd, N. Parikh, B. P. E Chu, and J. Eckstein, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," Found. and Tre. in Mach. Lea., Vol. 3, no. 1, pp. 1-122, Jan. 2011.

[15] L. Xu, C. Lu, Y. Xu, and J. Jia, "Image smoothing via L0 gradient minimization," ACM Trans. Graph. Vol. 30, no. 6, pp. 1-12, Dec. 2011.

[16] J. H. Hubbell and S. M. Seltzer, "Tables of x-ray mass attenuation coefficients and mass energy-absorption coefficients from 1 kev to 20 mev for elements z = 1 to 92 and 48 additional substances of dosimetric interest," 1989, 1990, 1996.

Image reconstruction method for the exterior problem with 1D edge-preserved diffusion and smoothing

Jinqiu Xu, Zhiqiang Wang, Yunsong Zhao*, and Peng Zhang

***Abstract*—The exterior problem is a special computed tomography (CT) problem for industrial nondestructive inspections. Because scanning data are incomplete, direct application of traditional reconstruction algorithms will blur the images along some directions. Based on the finding that the image blurring is along the polar angle direction, in this paper, we propose an optimization model for the exterior problem, which regularizes the image in polar coordinates instead of the image in Cartesian coordinates. In addition to the data fidelity term, two 1D regularization terms along the polar radius direction and polar angle direction are added respectively to the objective function to perform edge-preserving diffusion and edge-preserving smoothing. The blurred edges are restored gradually by edge-preserving diffusion and edge-preserving smoothing. Numerical experiments verify the validity of the proposed model and the corresponding reconstruction algorithm.**

***Index Terms*—computed tomography, exterior problem, polar coordinates**

I. Introduction

The exterior problem of computed tomography (CT) is a special imaging problem, in which only the exterior region of the object is scanned by x-ray beam. The exterior problem obtains important applications in industrial area [1]-[3], such as the inspection of pipelines.

Because the obtained projections are incomplete [3], the exterior problem is a highly ill-posed inverse problem. The reconstructed image suffers from severe distortions by utilizing conventional reconstruction algorithm such as Algebraic Reconstruction Techniques (ART) or simultaneous algebraic reconstruction technique (SART) [4]-[5]. In order to avoid the distortion, one kind of method is to use the properties of the image or projection data to attempt to recover the missing projection data. Lewitt, Louis and Natterer proposed algorithms to complete the missing projections for the exterior problem, and reconstructed images with standard inversion from the completed projections [6]-[8]. Generally, these algorithms provide better images than the conventional reconstruction algorithms, but still fail to meet the requirements of the industrial applications.

The other kind of method incorporating features of the object or reconstructed image into image reconstruction becomes popular for the ill-posed inverse problem. Quinto proposed an exterior reconstruction algorithm (ERA) which employs a singular value decomposition and the prior information of the scanned object into image reconstruction in [3]. In certain industrial applications, the image gradients are generally sparse. A typical reconstruction algorithm utilized sparsity of image gradients by minimizing the total variation (TV), i.e., the l_1 norm of the image gradient [9]-[10]. Then many algorithms applied the regularization method into CT exterior problem [11]-[13]. Theoretical studies have provided us much deeper understandings on the ill-posed inverse problem, in [14], Quinto et al showed that, for limited-angle problem, features tangent to the x-rays are easy to reconstruct, while features not tangent to any x-ray should be difficult to reconstruct. The edge-recovery ability of conventional iterative methods, such as ART and SART, is angle-dependent. Thus Jin et al proposed the anisotropic total variation (ATV) reconstruction approach in [15] and then Chen et al refined the approach in [16]. Please note that the exterior problem is a special limited-angle problem, with varying scanning angular ranges for different points in the image. Then Guo et al proposed a weighted directional total variation (WDTV) based regularization model to deal with exterior problem and achieve significant improvement on the image quality [17].

Besides the TV regularization approach, the l_0 norm of image gradient has also been applied to limited-angle reconstruction [18].It was demonstrated that the l_0 norm has a better performance on edge preserving than the l_1 norm [18]. In this paper, we propose to use the l_0 norm of the image gradient as the regularization term. One difficulty of applying the l_0 norm of the image gradient is that the l_0 norm counts the number of pixels with non-zeros gradients, and has no anisotropic concept. To overcome this obstacle, we propose to regularize polar images (i.e., images with pixels on polar grids) instead of Cartesian images (i.e., images with pixels on Cartesian grids), and regularize the images with two 1D l_0 norm along the two directions, the polar radius direction and the polar angle direction separately, instead of a 2D l_0 norm. With these two essential points, an imaging model with two regularization terms is proposed. The regularization along the polar radius direction performs an edge-preserving diffusion, while the regularization along the polar angle direction performs an edge-preserving smoothing. Numerical experiments show that our model and algorithm have a better performance on the reconstruction of the edges along the polar radius direction over the existing state-of-the-art algorithms.

Jinqiu Xu, Zhiqiang Wang, Yunsong Zhao, and Peng Zhang are with the school of Mathematical Sciences, Capital Normal University, Beijing, 100048, China, and are also with Beijing Advanced Innovation Center for Imaging Technology, Beijing, China. E-mail: zhao_yunsong@cnu.edu.cn

II. METHODS

A. The Exterior Problem

For simplicity, we consider the exterior problem with a 2D fan-beam scanning configuration. There should be no problem to extend our model and algorithm to 3D cone-beam case.

The scanning configuration for the exterior problem is illustrated in Fig. 1. The object is placed at the origin of the coordinate system Oxy and S is the x-ray focal spot. We can get only the exterior projections, whose corresponding x-ray paths intersect with the scanned object only at the exterior region. Image reconstruction from such exterior projections is the so-called exterior problem. It is obvious that the exterior problem is an incomplete data reconstruction problem. Different from limited-angle problem, the projection angular ranges for different points in the exterior problem are generally different.

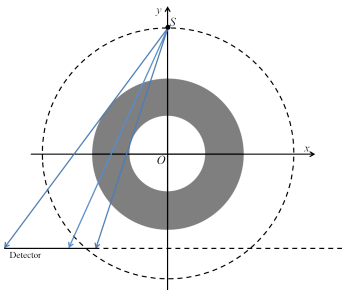


Fig. 1. A 2D fan-beam scanning configuration of the exterior problem. The gray circular ring is the area covered by the fan beam.

B. Optimization Model

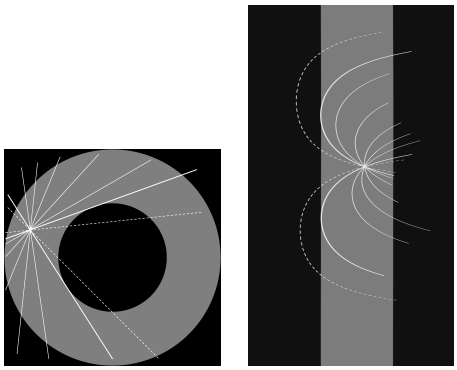
It is proved by Quinto that, when applying conventional reconstruction algorithms on incomplete projection data, edges tangent to the x-ray paths could be recovered easily, while edges tangent to the missing x-ray paths would be blurred [14]. So, for the exterior problem, the edges that are tangent to the polar angle direction could be recovered easily, while the edges that are tangent to the polar radius direction would be recovered hard. Based on this finding, Guo et al propose a weighted directional total variation (WDTV) minimization model to deal with exterior problem [15]. The proposed model has a better performance on artifact reduction and edge preserving than existing state-of-the-art models. But high spatial resolution details are distorted.

As the l_0 norm of image gradient has a better performance than TV on edge preserving [18], So, to reconstruct spatial resolution details, we consider replacing TV with the l_0 norm of image gradient. Additionally, for the exterior problem, scanning angular ranges vary by the points in the image. Thus we reconstruct polar images (i.e., images with pixels on polar grids) instead of Cartesian images (i.e., images with pixels on Cartesian grids) based on the finding that for polar images, the blurs are mainly along the polar angle direction. Fig. 2 gives an example to illustrate this fact. Fig. 2 (a) is the pantom image with 512×512 pixels. Fig. 2 (b) and Fig. 2 (c) are Cartesian image and polar image reconstructed from exterior projections with traditional method, respectively. We can see clearly that the blurs are mainly along the polar angle direction, i.e., the vertical direction in the polar image.

(a) (b) (c)

Fig. 2. (a) is the phantom image. (b) and (c) are Cartesian image and polar image reconstructed from exterior projections with SART, respectively.

Now we introduce our model in detail. Note that, as shown in Fig. 3, the x-ray path along the straight line in subfigure (a) corresponds to the x-ray path along the corresponding curve in subfigure (b). So, a projection of the image is the line integral of the first kind along a curve which is a straight line in the Cartesian coordinate system.



(a) (b)

Fig. 3. The x-ray paths in Cartesian coordinate system and polar coordinate system. The x-ray path along the straight line in Cartesian image (a) corresponds to the path along the corresponding curve in polar image (b).

Suppose that $\vec{f} = \{f_n | n = 1, 2, ..., N\}$ is the discrete image with pixels on polar grids, and let the curved intersection length of the x-ray path indexed with i and the j-th pixel be a_{ij}, we have the corresponding projection

$$p_i = \vec{A}_i \vec{f} \tag{1}$$

where $\vec{A}_i = (a_{i1}, a_{i2}, \ldots, a_{iN})^T$ is a row vector. Then reconstructing image from projections is to solve the following linear system

$$A\vec{f} = \vec{p} \tag{2}$$

where $A = [\vec{A_1}, \vec{A_2}, ..., \vec{A_M}]^T$ and $\vec{p} = [p_1, p_2, ..., p_M]^T$. The proposed optimization model is shown in (3).

$$\vec{f}^* = \arg\min_{\vec{f}} \quad \left(\|A\vec{f} - \vec{p}\|_2^2 + \lambda \|D_h \vec{f}\|_0 + \mu \|D_v \vec{f}\|_0 \right) \quad (3)$$

where $\|\cdot\|_2$ and $\|\cdot\|_0$ are the l_2 norm and l_0 norm of a vector or an image, the first term of the objective function in (3) is the data fidelity term, the second and the third are regularizaton terms, λ and μ balance the contribution of the three terms to the objective function. $D_h\vec{f}$ and $D_v\vec{f}$ are differences along the horizontal and vertical direction of an image, which are defined as

$$\|D_h \vec{f}\|_0 = \sum_{s=2}^{\sqrt{N}} \text{sign}\left(|f_{s,t} - f_{s-1,t}|\right) \quad (4)$$

$$\|D_v \vec{f}\|_0 = \sum_{t=2}^{\sqrt{N}} \text{sign}\left(|f_{s,t} - f_{s,t-1}|\right) \quad (5)$$

Note that we have used double subscript here, where s, t are the row and column indices in the polar coordinate system, respectively.The role of the second term is to restore the edges tagent to polar radius direction, while the role of the third term is to smooth out the possible artifacts introducted by the second term, which will be verified in the section of numerical experiments. So, in numerical experiments, λ should be bigger than μ.

C. The Solving Algorithm

The optimization problem (3) is non-convex. In our practice, the alternating direction method of multipliers (ADMM) framework [19] is employed to develop its solution algorithm, which appears to be convergent and obtains good results.

Let $\vec{f}^{(0)}$ be the initial image and suppose that we have got the reconstruction result $\vec{f}^{(k)}$ after k iterations, the updating scheme for $\vec{f}^{(k+1)}$ can be split into three subproblems:

Subproblem 1:

$$\vec{f}^{(k+1/3)} = \arg\min_{\vec{f}} \quad \left(\|\vec{f}^{(k)} - \vec{f}\|_2^2 + \|A\vec{f} - \vec{p}\|_2^2 \right) \quad (6)$$

Subproblem 2:

$$\vec{f}^{(k+2/3)} = \arg\min_{\vec{f}} \quad \left(\|\vec{f}^{(k+1/3)} - \vec{f}\|_2^2 + \lambda \|D_h \vec{f}\|_0 \right) \quad (7)$$

Subproblem 3:

$$\vec{f}^{(k+1)} = \arg\min_{\vec{f}} \quad \left(\|\vec{f}^{(k+2/3)} - \vec{f}\|_2^2 + \mu \|D_v \vec{f}\|_0 \right) \quad (8)$$

The subproblem 1 can be solved by applying the SART method with $\vec{f}^{(k)}$ as the initial guess. In our tests, one iteration of SART method works well. The subproblem 2 and subproblem 3 can actually be decoupled into a series of one dimensional problems, which are then solved by utilizing the algorithm proposed in [20], with a little extra work to tailor its two-dimensional code for solving one-dimensional problems.

Generally speaking, the images obtained with the above algorithm can be used directly for image analysis, such as automatic defect inspection. But as the pixels of the images are distributed on polar grids, it is not intuitive for human observation. So, in our experiment, we transform the pixels back on Cartesian grids with linear interpolation.

III. Numerical experiments

This section validates the proposed imaging model and the corresponding reconstruction algorithm for the exterior problem. The evaluation includes the spatial, contrast resolution and sensitivity to noise.

We will first use the simple phantom shown in Fig. 2 to illustrate the roles of the second and third regularization terms in the optimaization model (3). The scanning parameters are listed in Table I. When we remove the third term, the reconstructed polar image and the corresponding Cartesian iamge are shown in Fig. 4 (a) and (b), respectively. The reconstruction parameter λ is set to be 0.005. The results indicate that the second term would help to recover most of the features of the true image while it causes some stripping artifacts. Then we add the third term, the reconstructed image is shown in Fig. 4 (c) and (d). The reconstruction parameters λ and μ are set to be 0.005 and 0.0001, respectively. One can see that the third term can effectively remove the artifacts introduced by the second term.

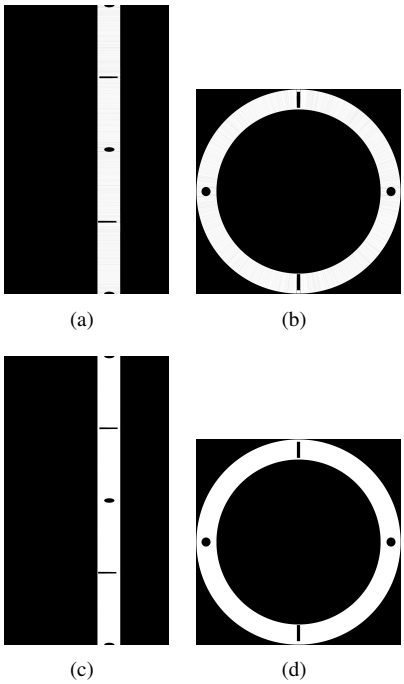


Fig. 4. (a) and (b) are Cartesian image and polar image reconstructed by using optimization model (3) without the third regularization term, respectively. (c) and (d) are the reconstructed Cartesian image and polar image with the third regularization term, respectively. The gray scale window is [0,1].

Then we will use the phantom shown in Fig. 5 to validate the proposed algorithm further. This phantom is designed according to the one used in [17]. It is a cross section of a pipeline with 512×512 pixels. Two kinds of structures are placed on different positions of the pipelines. The first kind of structure containing seven low contrast circles (marked with

ROI-B1 to ROI-B7), with gray value 0.8, 0.85, 0.9, 0.95, 0.97, 0.98 and 0.99 respectively, is used to verify the performance on contrast resolution. The second kind of structure containing five sets of line pairs (marked with ROI-A1 to ROI-A5) with different resolutions and a cross (marked with ROI-A6) is used to verify the performance on spatial resolution.

For comparison, two state-of-the-art algorithms for the exterior problem, i.e., SART + l_0 [18] and WDTV [15] are also implemented. The Peak Signal to Noise Ratio (PSNR) and Root Mean Square Error (RMSE) are calculated to evaluate the algorithms quantitatively.

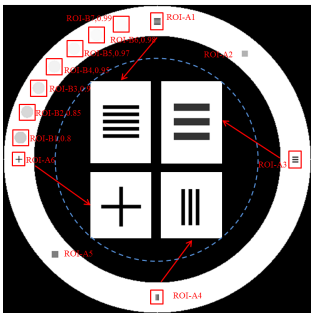


Fig. 5. The phantom used for simulation experiments. The gray scale window is [0,1].

Table I shows the scanning parameters for this phantom. With these parameters, the projection angular range for the inner shell point and the outer shell point are 69.41 and 96.87 degree, respectively. The noisy projections are simulated by adding Poisson noise, with initial photon number 10^6 for each x-ray path, to the noise-free projections. The used projection data are along projection rays that do not intersect the blue circle in Fig. 5.

All the initial images for SART+l_0, WDTV, and the proposed algorithm are the reconstruction result by using SART with 10 iterations. Table II shows the reconstruction parameters for the three algorithms. The parameters λ and μ for the proposed algorithm are the ones defined in (3). The iterations are stopped when the PSNR and the RMSE of the reconstructed images become stable.

TABLE I
SCANNING PARAMETERS.

Parameters	Value
Scanning type	Fan-beam Scan
Detector unit width	0.127 mm
Number of effective detector units	350
X-ray source to the rotation center distance	1950 mm
X-ray source to detector distance	2098 mm
Scanning Angular Range	360 degree
Scanning Angular Interval	0.5 degree
Reconstruction size	512×512 pixels
Image Pixel Size	0.1669×0.1669 mm^2

Figure 6 shows the reconstruction results of the three algorithms. The four blocks in the middle of the reconstructed images are the zoom-in views of the corresponding ROIs. We can see that the SART+l_0 failed to reconstruct ROI-A3 and ROI-A4, as the edges along the radical direction is too long, too much information is lost in the projections. WDTV reconstructs the low resolution line pairs (ROI-A3) successfully, but fail to reconstruct high resolution line pairs (ROI-A4). While the proposed algorithm reconstructs all the line pairs correctly, as the l_0 norm has a stronger edge-preserving ability.

TABLE II
RECONSTRUCTION PARAMETERS.

	SART+l_0	WDTV	The proposed algorithm
Parameters	$\lambda_{l_0} = 0.00001$	$\omega_1 = 1$ $\omega_2 = 0.07$ $\beta = 0.2$ $N_{TV} = 30$	$\lambda = 0.00008$ $\mu = 0.00002$

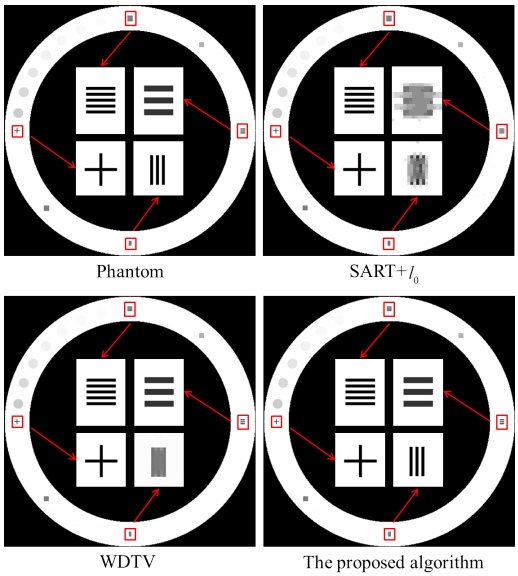


Fig. 6. Reconstruction results of simulated phantom. The gray scale window is [0,1].

Figure 7 shows the ROIs with low contrast. The SART+l_0 algorithm and the proposed algorithm fail to reconstruct the ROI with the lowest contrast. We analyze the reason is that the l_0 norm regularization is too strong.

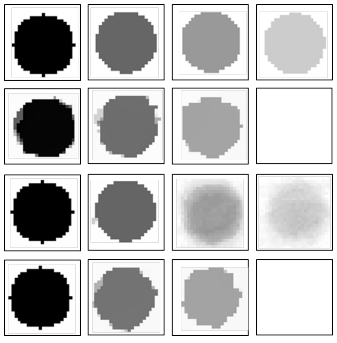


Fig. 7. Reconstruction results from noisy data. The first row is the true images, and the second to the forth row are the results of SART+l_0, WDTV and the proposed algorithm, respectively. The results from left to right correspond to the circle with gray value of 0.95, 0.97, 0.98 and 0.99, respectively. The gray scale window is set to [0.95,1].

To assess the performance of the algorithms in a quantitative way, Table III lists the value of PSNR and RMSE of all the reconstruction results. The index values coincide well with the results shown in Fig. 6.

TABLE III
PSNR AND RMSE OF THE RECONSTRUCTION RESULTS.

	SART+l_0	WDTV	The proposed algorithm
PSNR	36.2716	39.0841	39.1328
RMSE	0.0207	0.0176	0.0108

IV. CONCLUSION

In this paper, an optimization model is proposed to remove the blur and artifacts introduced by the incomplete data of the exterior problem. Based on the observation that the blur of the image reconstructed from the exterior CT data is along the polar angle direction, we reconstruct images with pixels on polar grids, and regularize the images with 1D gradient l_0 norm along the polar angle direction and the polar radius direction, separately. Then the ADMM framework is employed to compute the solution of the optimization model. The experiments on simulated data suggest that the proposed model and algorithm outperform the state-of-the-art algorithms, e.g. SART+l_0 and WDTV, on the reconstruction of details with high spatial resolution and high contrast. So, the proposed model and algorithm are fit for the detection of pipelines. But we have also find that the proposed model and algorithm fail to reconstruct details with very low contrast, which might be caused by the strong regularization of the l_0 norm. Further improvement of this algorithm is ongoing.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China under the grants 61771324 and 61401289.

REFERENCES

[1] E.T. Quinto, "Exterior and limited-angle tomography in non-destructive evaluation," Inverse Probl, Vol. 14, no. 2, pp. 339-353, Dec. 1998.
[2] L.A. Shepp, and S. Srivastava, "Computed tomography of PKM and AKM exit cones," A. T. Technical Journal, Vol. 14, no. 2, pp. 78-88, Jan. 1986.
[3] E.T. Quinto, "Exterior and limited-angle tomography in non-destructive evaluation," Inverse Probl, Vol. 14, no. 2, pp. 339-353, Sep. 1988.
[4] R. Gordon, R. Bender, and G. T. Herman, "Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and x-ray photography," Journal of theoretical Biology, vol. 29, no. 3, pp. 471-481, Aug. 1970.
[5] A.H. Andersen and A.C. Kak, "Simultaneous algebraic reconstruction technique (SART): a superior implementation of the art algorithm," Ultrasonic imaging, vol. 6, no. 1, pp. 81-94, Jan. 1984.
[6] R. Lewitt, R. Bates, "Image-reconstruction from projections. 3. Projection completion methods (theory)," Optik vol, 50, no. 4, pp. 269-278, Apr. 1978.
[7] F. Natterer, K. Hadeler, "Efficient implementation of optimal algorithms in computerized tomography," Mathematical Methods in the Applied Sciences, vol. 2, no. 4, pp. 545-555, Sep. 1980.
[8] A.K. Louis, "Incomplete data problems in x-ray computerized tomography[J]," Numerische Mathematik, vol. 48, no. 3, pp. 251-262, May. 1986.
[9] E.Y. Sidky, C. Kao, and X. Pan, "Accurate image reconstruction from few-views and limited-angle data in divergent-beam CT," J. Xray. Sci. Technol., Vol. 14, no. 2, pp. 119-139, Mar. 2016.
[10] E.Y. Sidky, and X. Pan, "Image reconstruction in circular cone-beam computed tomography by total variation minimization," Phys. Med. Biol., Vol. 53, no. 17, pp. 4777-4807, Sep. 2008.
[11] B. Chen, M. Yang, Z. Zhang, X. Han, J. Bian, E.Y. Sidky, X. Pan, "Constrained TV-minimization reconstruction from exterior CT data," in Proc. NSS-MIS, IEEE, Washington, DC, pp. 1-3, Jan. 2013.
[12] L. Zeng, B. Liu, L. Liu, and C. Xiang, "A new iterative reconstruction algorithm for 2D exterior fan-beam CT," J. Xray. Sci. Technol., Vol. 18, no. 3, pp. 267-277, Jan. 2010.
[13] J. Guo, L. Zeng, and B. Liu, "High-quality image reconstruction from exterior helical cone-beam CT data for NDE of industrial pipelines," Insight, Vol. 53, no. 10, pp. 534-541, Oct. 2011.
[14] E.T. Quinto, "Artifacts and visible singularities in limited data X-Ray tomography," Sens. Imaging, Vol. 18, no. 1, pp. 9-23, Jan. 2017.
[15] X. Jin, L. Li, Z. Chen, L. Zhang, and Y. Xing, "Anisotropic total variation for limited-angle CT reconstruction," IEEE Nucl. Sci. Symp. Conf. Rec. (NSS/MIC), Vol. 8506, pp. 2232-2238, Nov. 2010.
[16] Z.Q. Chen, X. Jin, L. Li, and G. Wang, "A limited-angle CT reconstruction method based on anisotropic TV minimization," Phys. Med. Biol., Vol. 58, no.7, pp. 2119-2141, Apr. 2013.
[17] Y. Guo, L. Zeng, C. Wang, and L. Zhang, "Image reconstruction model for the exterior problem of computed tomography based on weighted directional total variation," Appl. Math. Model., Vol. 52, pp. 358-377, Dec. 2017.
[18] W. Yu, and L. Zeng, "l_0 gradient minimization based image reconstruction for limited-angle computed tomography," Plos one, Vol. 10, no. 7, pp. e0130793, Jul. 2015.
[19] S. Boyd, N. Parikh, B.P.E Chu, and J. Eckstein, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," Found. and Tre. in Mach. Lea., Vol. 3, no. 1, pp. 1-122, Jan. 2011.
[20] L. Xu, C. Lu, Y. Xu, and J. Jia, "Image smoothing via L0 gradient minimization," ACM Trans. Graph., Vol. 30, no. 6, pp. 1-12, Dec. 2011.

Iodine Quantification in Limited Angle Tomography

Koen Michielsen, Alejandro Rodríguez-Ruiz, Ingrid Reiser, James Nagy, and Ioannis Sechopoulos

***Abstract*—Contrast-enhanced mammography has shown promise as a cost-effective alternative to MRI for imaging breast cancer in dense breasts. However, one limitation is the poor quantitation of iodine contrast since the true 3D lesion shape cannot be inferred from the 2D projection. Use of limited angle tomography can potentially overcome this limitation. In this work, we present a two pass reconstruction algorithm with material decomposition designed to obtain quantitative iodine measurements in breast tomosynthesis by segmenting the iodine map generated by the first pass reconstruction using a convolutional neural network, and using this segmentation to restrict the allowed iodine distribution in the second pass of the reconstruction.**

To evaluate the performance of the algorithms, a set of digital breast phantoms with lesions with varying iodine concentration was used. The acquisition of breast tomosynthesis images of these phantoms was simulated and the proposed reconstruction and segmentation algorithms were applied. The resulting iodine concentration estimates increased from 12% to 76% of the true iodine concentration, and the coefficient of variation of the estimates was reduced from 45% to 21%, demonstrating that it is feasible to obtain more quantitative results from contrast enhanced tomosynthesis.

I. Introduction

Digital breast tomosynthesis (DBT) is a pseudo 3D imaging technique that keeps the high in-plane spatial resolution of digital mammography while adding limited vertical resolution. However, DBT images are still only morphological, relying on differences in attenuation and distortions in the shape of tissues for tumors to be detected, and having limited utility post-diagnosis. Obtaining functional information of the breast would allow for not only an increase in the detection and diagnostic performance of DBT, but more importantly could have a significant impact for post-diagnosis clinical tasks.

Functional imaging is currently dominated by magnetic resonance imaging (MRI), an accurate but expensive modality, both in terms of equipment and installation costs, and especially in terms of running cost. In comparison, functional imaging with mammography or DBT would be considerably faster, easier, more accessible, and overall cheaper. However, contrast-enhanced (CE) mammography and DBT lack quantitative capability, making it challenging to accurately and repeatedly measure the functional response of the investigated breast tissue. This is especially important for longitudinal evaluation, such as during therapy response monitoring, where it is important to compare changes in tumor perfusion metrics across months.

K. Michielsen and A. Rodríguez-Ruiz are with the Department of Radiology and Nuclear Medicine, Radboud University Medical Center, Nijmegen, The Netherlands; I. Reiser is with the Department of Radiology, The University of Chicago, Chicago, IL, United States; J. Nagy is with the Department of Mathematics and Computer Science, Emory University, Atlanta, GA, United States; and I. Sechopoulos is with the Department of Radiology and Nuclear Medicine, Radboud University Medical Center, Nijmegen, The Netherlands and the Dutch Expert Center for Screening (LRCB), Nijmegen, The Netherlands.

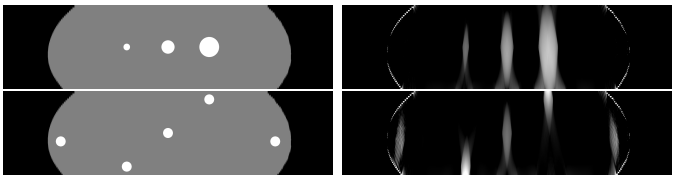


Figure 1. To demonstrate the absence of quantitative reconstructed attenuation values in limited angle tomography, we simulated a phantom with lesions of different sizes and a phantom with lesions at different locations withing the imaged area, shown left. In the reconstruction on the right, it is clear that the reconstructed attenuation depends on both the size of the lesion and the location of the lesion within the reconstructed volume.

The limitations that need to be overcome for CE-DBT to become a quantitative modality are demonstrated in figure 1, where it is clear that targets with equal iodine content but different sizes and locations within the reconstruction field of view are reconstructed with varying apparent amounts of iodine content.

In this work we propose an approach combining a maximum likelihood based polychromatic reconstruction algorithm including material decomposition with a deep learning based segmentation method to allow accurate recovery of iodine concentration from limited angle tomography data and demonstrate its effectiveness in simulated data.

II. Materials & Methods

The approach for our method is to perform a two-pass reconstruction with material decomposition. The iodine component of the first pass reconstruction is automatically segmented with a deep learning convolutional neural network and the resulting segmentation is then used as a mask in the second pass. We assume that the segmentation problem can be solved in practice due to the expected sparsity of the iodine component. The reconstruction and segmentation methods are described in sections II-B and II-C respectively. Sections II-A and II-D describe the simulation and analysis of a set of phantoms used to evaluate the algorithm.

A. Phantom Simulation

To evaluate the reconstruction algorithm, we simulated a total of 1100 breast phantom images. To limit the computational cost, the phantoms consisted of 2D coronal slices extracted from 3D phantoms generated with an algorithm described by Lau et al. [1]. An example is shown in figure 2. These phantoms were indexed with labels for four different materials: skin, adipose tissue, glandular tissue, and Cooper's ligaments. The elemental compositions of these materials, obtained from the work of Hammerstein et al.[2] are listed

Table I
ELEMENTAL MATERIAL COMPOSITIONS.

	Skin	Adipose tissue	Glandular tissue	ε% contrast enhanced lesion	
$\rho(\frac{g}{cm^3})$	1.09	0.93	1.04	1.04	
H	0.098	0.112	0.102	0.102	$\times(1-\varepsilon)$
C	0.178	0.619	0.184	0.184	$\times(1-\varepsilon)$
N	0.05	0.017	0.032	0.032	$\times(1-\varepsilon)$
O	0.667	0.251	0.677	0.677	$\times(1-\varepsilon)$
P	0.001 75	0.000 25	0.001 25	0.001 25	$\times(1-\varepsilon)$
S	0.001 75	0.000 25	0.001 25	0.001 25	$\times(1-\varepsilon)$
K	0.001 75	0.000 25	0.001 25	0.001 25	$\times(1-\varepsilon)$
Ca	0.001 75	0.000 25	0.001 25	0.001 25	$\times(1-\varepsilon)$
I				ε	

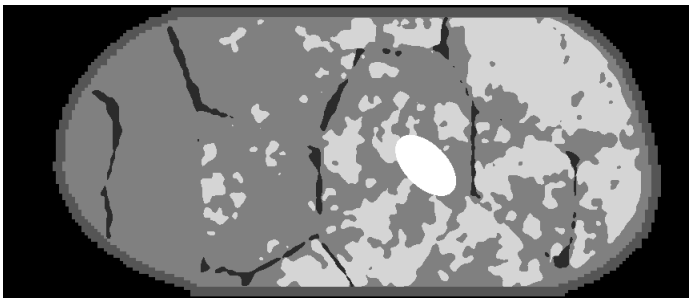


Figure 2. 2D coronal breast phantom containing skin (darkest gray), adipose tissue (dark gray), glandular tissue (light gray), Cooper's ligaments (black), and an iodine enhanced lesion (white).

in table I, which were used to calculate the energy dependent attenuation coefficients, $\mu^{(e)}$, using the software from Boone and Chavez [3]. The composition of Cooper's ligaments was assumed to be identical to that of glandular tissue. The resulting phantoms represented compressed breast thicknesses from $3.0\,\mathrm{cm}$ to $5.6\,\mathrm{cm}$ and widths from $5.8\,\mathrm{cm}$ to $18.0\,\mathrm{cm}$ with an isotropic voxel size of $0.1\,\mathrm{mm}\times0.1\,\mathrm{mm}$.

A randomly generated ellipsoid target consisting of iodinated contrast enhanced glandular tissue was randomly placed within each phantom. The major axis of the lesion varied between $4\,\mathrm{mm}$ and $8\,\mathrm{mm}$, the minor axis varied between $2\,\mathrm{mm}$ and the size of the major axis, and rotation was allowed in any direction. The iodine content of the simulated lesion varied between 0.1% and 0.5% fraction by weight, which was added to the glandular material by rescaling the other elemental fractions, as shown in the last column of table I.

Projection data was generated for all phantoms by simulating a limited angle fan-beam geometry with the center of rotation placed at the bottom center of the phantom. The x-ray source was placed $650\,\mathrm{mm}$ above the center of rotation, and the source-detector distance was $700\,\mathrm{mm}$. A total of 25 equally spaced projections between $-24°$ and $24°$ were calculated, with the detector moving together with the source. The forward model in equation (2) was used with a $39\,\mathrm{kV}$ tungsten spectrum filtered by $0.5\,\mathrm{mm}$ of aluminum.A perfect photon counting detector was simulated with two energy bins, one from $19.5\,\mathrm{keV}$ to $33\,\mathrm{keV}$ and one from $33.5\,\mathrm{keV}$ to $39\,\mathrm{keV}$, and consisting of 2048 elements of $0.1\,\mathrm{mm}$ width. No quantum noise was added to the simulated projection data.

Table II
LIST OF SYMBOLS.

y_i	Measurement for projection line i
$\hat{y}_i$	Forward model for projection line i
b_{ie}	Spectrum for projection line i
$\mu_a^{(e)}$	Linear attenuation of material a at energy e
l_{ij}	Intersection between projection line i and voxel j
w_{aj}	Fraction of material a in voxel j
$\mathcal{L}(\vec{w})$	Log-likelihood cost function

B. Reconstruction Algorithm

The reconstruction algorithm we apply is a modification of a previously published method by Bustamante et al. [4]. It is a maximum likelihood method with material decomposition (named ML-MADE), which features a polychromatic forward model and allows decomposition in selected base materials. The log-likelihood cost function in equation (1) is used with the forward model in equation (2), resulting in the update step in equation (3) after some approximations in the denominator which assume that the reconstruction is started with a good initialization. All symbols are listed in table II. In practice the update steps for each material a and each energy bin e are calculated and applied sequentially, with the loop over the materials within the loop over the energy bins.

$$\mathcal{L}(\vec{w}) = \sum_i y_i \ln \hat{y}_i - \hat{y}_i \tag{1}$$

$$\hat{y}_i(\vec{w}) = \sum_e b_{ie} \exp\left(-\sum_a \mu_a^{(e)} \sum_j l_{ij} w_{aj}\right) \tag{2}$$

$$\Delta w_{aj}^{(n)} = \frac{\sum_i l_{ij} \left(\hat{y}_i - y_i\right) \frac{\psi_i^{(a)}}{\hat{y}_i}}{\sum_i l_{ij} \frac{\psi_i^{(a)}}{\hat{y}_i} \sum_b \psi_i^{(b)} \sum_k l_{ik}} \tag{3}$$

$$\psi_i^{(a)} = \sum_e b_{ie} \mu_a^{(e)} \exp\left(-\sum_b \mu_b^{(e)} \sum_j l_{ij} w_{bj}\right) \tag{4}$$

In cases where we know that the materials a selected for the decomposition match the materials present in the object to be reconstructed, we can add the additional constraint that $\sum_a w_{aj} = 1$ in all voxels by applying the prior cost function with weight β in equation (5).

$$P(\vec{w}) = -\frac{\beta}{2} \sum_j \left(1 - \sum_a w_{aj}\right)^2 \tag{5}$$

The object contour is also used as an additional constraint, as per our previous work [5]. In this instance, the mask α_{aj} is material specific, combining the overall object contour with the material specific segmentation for the iodine component. The new update equation including these constraints is shown in equation (6).

$$\Delta w_{aj}^{(n)} = \frac{\alpha_{aj} \sum_i l_{ij} \left(\hat{y}_i - y_i\right) \frac{\psi_i^{(a)}}{\hat{y}_i} + \alpha_{aj}\beta(1 - \sum_b w_{bj})}{\sum_i l_{ij} \frac{\psi_i^{(a)}}{\hat{y}_i} \sum_b \psi_i^{(b)} \sum_k \alpha_{ak} l_{ik} + \alpha_{aj}\beta(\sum_b w_{bj})} \tag{6}$$

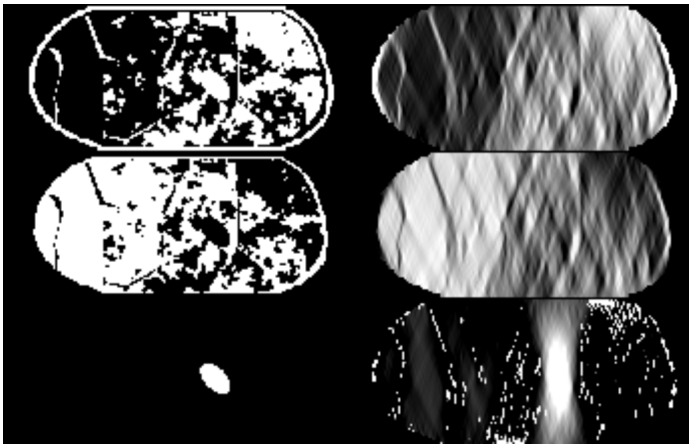


Figure 3. Material decomposition of the phantom shown in figure 2. Glandular tissue (top row), adipose tissue (middle row), and iodine (bottom row). The ground truth is shown in the left column while the results of the first pass reconstruction are shown in the right column.

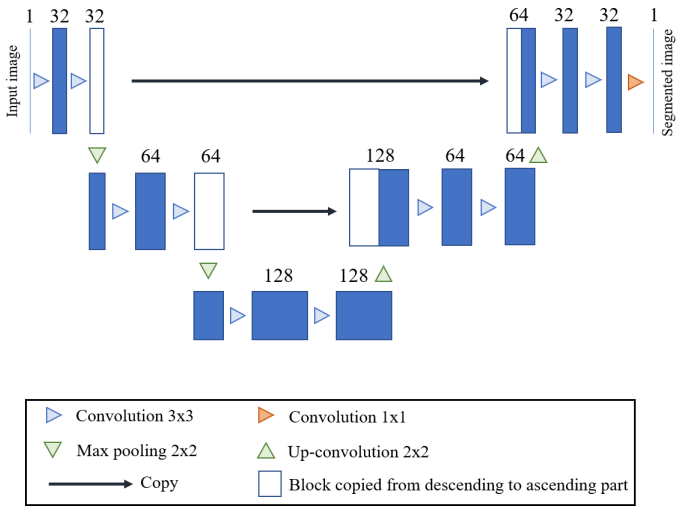


Figure 4. The model based on u-net architecture [6] used to segment the iodine targets on the reconstructed images. Each convolution 3×3 was followed by a batch normalization layer and a leaky rectified linear unit activation layer. The last convolution 1×1 was followed by a sigmoid activation layer.

Each of the two passes of the reconstruction used an initialization of 100 iterations with two base materials (adipose and glandular tissue), followed by 10 iterations using three base materials, adding iodine to the two materials used in the initialization. Each iteration used 5 subsets for each of the two energy bins, resulting in a total of 1100 image updates, and the prior weight β was set to 10^5.

The exact phantom contour was used as binary reconstruction mask $\vec{\alpha}$ for all materials in the first pass reconstruction, and for the adipose and glandular components in the second pass reconstruction. The segmentation resulting from the method described in section II-C was used as mask for the iodine component image in the second pass reconstruction. An example of the three components of the first pass reconstruction of the phantom in figure 2 is shown in figure 3.

C. Neural Network Based Segmentation

The segmentation model was applied locally to each simulated iodine target independently within the complete phantom image. Regions of interest (ROIs) of 256×256 pixels were automatically extracted around each iodine target (both in the first pass reconstruction of the iodine map and in the original ground truth simulated image). In total, 1100 iodine targets were available, and they were split into training (880), validation (110), and evaluation (110) datasets of the model. The model was a deep learning convolutional neural network (CNN) with modified u-net architecture [6], detailed in figure 4. It was trained using the simulated iodine-only phantom ROIs as ground truth binary masks, and using pixel-wise binary cross entropy as loss function. The trained network was applied to the 110 images in the evaluation dataset, and the resulting segmentation probability map was used as a non-binary mask for the iodine component in the second pass reconstruction.

D. Data Analysis

The accuracy of the segmentation model was evaluated using the Dice similarity coefficient (DSC) in equation 7.

$$DSC = \frac{2\left|\mathrm{Mask}^{\mathrm{Model}} \cap \mathrm{Mask}^{\mathrm{Truth}}\right|}{\left|\mathrm{Mask}^{\mathrm{Model}}\right| + \left|\mathrm{Mask}^{\mathrm{Truth}}\right|} \quad (7)$$

To be able to distinguish between the effects of the segmentation and the reconstruction algorithms, the second pass reconstructions were performed without masking the iodine component, with the CNN-based segmentation, and with the true target segmentation.

The reconstructed iodine fraction for each of the 110 evaluation cases was determined by measuring the median pixel value in the target area in the iodine map. The linear regression line was calculated for the results of the different reconstruction masks.

The proportion between the reconstructed iodine fraction and the true iodine content was calculated for all cases to determine the coefficient of variation of this proportion.

III. Results

The median DSC of the segmented masks was 0.95 with interquartile range 0.90 – 0.97. The average absolute difference in predicted iodine fraction between using true and predicted masks was 10%, which was reduced to 4% when considering only cases with DSC larger than 0.90 (75% of cases).

The estimated iodine fractions are plotted as a function of the true iodine fraction in figure 5. The proportion between the measured and true iodine content is shown in figure 6. The slope and confidence interval of the linear regression lines is listed in table III together with the mean and standard deviation of the proportion.

IV. Discussion & Conclusion

The results show that our method results in more quantitative reconstructions, increasing the average recovered proportion from 12% to 70% of the true iodine content, while

Table III
ESTIMATED IODINE FRACTIONS.

	Linear regression		Reconstructed proportion		
	Slope	95% CI	Mean	St.Dev.	(%)
No mask	0.113	0.086 – 0.140	0.115	0.0513	(45%)
CNN mask	0.765	0.714 – 0.816	0.701	0.144	(21%)
True mask	0.772	0.743 – 0.801	0.764	0.0784	(10%)

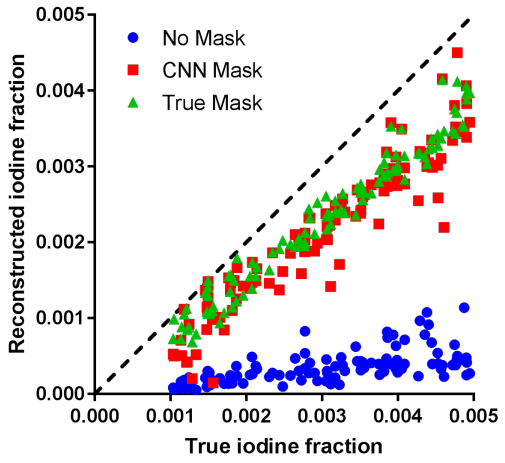


Figure 5. Reconstructed iodine fraction as a function of the true iodine fraction for reconstructions without mask, with a CNN-based segmentation, and with the true mask.

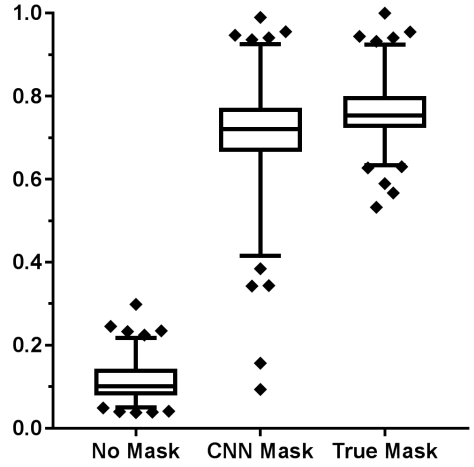


Figure 6. Proportion between the reconstructed iodine fraction and the true iodine fraction for reconstructions without mask, with a CNN-based segmentation, and with the true mask. The boxplot lines are at the 25th, 50th, and 75th percentiles, and the whiskers are at the 5th and 95th percentiles.

reducing the variation from 45% to 21%. Comparing the results using the CNN-based segmentation and with those using the true shape shows that the segmentation quality is sufficient in most cases, resulting in iodine fractions very close to true segmentation. The cases where the segmentation was not very good (roughly 10% of the cases) caused the increase in the coefficient of variation when using the segmented iodine map instead of the true map. The remaining systematic underestimation can thus be attributed to the reconstruction algorithm, which seems to assign attenuation belonging to the iodine component to the other two components, possibly due to convergence to a local maximum.

These results give an indication that our approach to obtain quantitative information from breast tomosynthesis is feasible. They also indicate the gains to be made by improving the reconstruction or segmentation steps of our algorithm. Improving the reconstruction should focus on making sure that the algorithm assigns attenuation to the correct base material when working with a limited angle dataset.

Improving the segmentation method, on the other hand, will ensure that our method works consistently in all cases and thus results in a reduced variation, although given the feasibility of the method demonstrated here, it is now more important to improve the segmentation using more realistic targets in 3D phantoms and real lesions in patient images, rather than trying to obtain a perfectly segmented ellipse in 2D. Of course, it remains to be seen how the segmentation will perform with patient cases, including cases with multiple iodine enhanced lesions and contrast agent circulating in the arteries.

To conclude, we demonstrated that it is possible to get more quantitative data from contrast enhanced breast tomosynthesis by segmenting lesions in a first pass reconstruction, and then limiting the iodine component to the segmented region in the second pass reconstruction.

ACKNOWLEDGEMENTS

This work was supported in part by Grant No. R01CA163746 from the National Cancer Institute, National Institutes of Health.

REFERENCES

[1] B. A. Lau, I. Reiser, R. M. Nishikawa, and P. R. Bakic, "A statistically defined anthropomorphic software breast phantom," *Med. Phys.*, vol. 39, no. 6, pp. 3375–3385, Jun. 2012.

[2] G. R. Hammerstein, D. W. Miller, D. R. White, M. E. Masterson, H. Q. Woodard, and J. S. Laughlin, "Absorbed radiation dose in mammography," *Radiology*, vol. 130, no. 2, pp. 485–491, Feb. 1979.

[3] J. M. Boone and A. E. Chavez, "Comparison of x-ray cross sections for diagnostic and therapeutic medical physics," *Med. Phys.*, vol. 23, no. 12, pp. 1997–2005, Dec. 1996.

[4] V. M. Bustamante, J. G. Nagy, S. S. J. Feng, and I. Sechopoulos, "Iterative Breast Tomosynthesis Image Reconstruction," *SIAM J. Sci. Comput.*, vol. 35, no. 5, pp. S192–S208, 2013.

[5] K. Michielsen, A. Rodríguez-Ruiz, G. Agasthya, and I. Sechopoulos, "Artifact reduction in breast tomosynthesis by including prior knowledge of the compressed breast shape," in *Proc. Intl. Mtg. on Fully 3D Image Recon. in Rad. and Nuc. Med*, 2017, pp. 186–190.

[6] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, ser. Lecture Notes in Computer Science. Springer, Cham, Oct. 2015, pp. 234–241.

FreeCT_ICD: Free, Open-Source MBIR Reconstruction Software for Diagnostic CT

John M. Hoffman, Scott S. Hsieh, Frédéric Noo, and Michael F. McNitt-Gray

***Abstract*—As interest in CT has increased, most recently with the spike in interest around quantitative imaging, more research groups are requiring offline reconstruction software. Recently, FreeCT_wFBP, weighted filtered backprojection reconstruction software, was released targeting clinical third-generation CT scanners, which has historically not been well addressed in the free, open-source software for CT. While FreeCT_wFBP represents an important initial step in the development of software intended for the reconstruction of clinical data sets, most modern scanners are leveraging model-based iterative reconstruction (MBIR) algorithms, which show immense promise for dose reduction. In this work we introduce FreeCT_ICD, the free, open-source, MBIR complement to FreeCT_wFBP. FreeCT_ICD is intended for use with third-generation, helical CT geometries and is capable of reconstructing clinical data sets. It handles clinical scanner features such as flying focal spots, and quarter detector offsets, and can be initialized with a wFBP reconstruction to accelerate convergence. More information and documentation can be found at http://cvib.ucla.edu/freect/ and https://github.com/FreeCT/FreeCT_ICD.**

***Index Terms*—Reconstruction, model-based, iterative, software, open-source**

I. Introduction

CT has established itself as one of the foremost modern imaging modalities in use today. Research interest in CT now spans a wide variety of topics, from hardware innovation (e.g., dual energy CT, photon counting detector, scatter grids, etc.) to quantitative imaging to accelerate and improve diagnosis and detection of disease in patients. While the range of topics is highly diverse and each topic presents its own set of requirements and challenges, reconstruction of raw projection data plays a critical role in every area.

The reconstruction needs of every research group are varied, requiring approaches to cater to factors such as specific system geometry, reconstruction algorithm, and artifact correction. However it is often the case that many of these aspects can be utilized across systems, modifying only the specific system parameters such as distances and detector spacings, or tuning the reconstruction parameters such as kernel or iterative parameters. As a result, open-source reconstruction packages have emerged to support this need. Examples of such packages include such as CONRAD [1], RTK [2], and the Michigan Image Reconstruction Toolbox [3]. While these packages are thorough, well-maintained, and would meet a vast majority of reconstruction needs for most groups, they have, to date, focused primarily on flat-panel cone-beam CT scanners, or more general reconstruction approaches rather than clinical diagnostic systems.

J. Hoffman, S. Hsieh, and M. McNitt-Gray are with the Department of Radiology, UCLA, Los Angeles, CA. F. Noo is with UCAIR, Department of Radiology, University of Utah, Salt Lake City, UT.

With the rise of quantitative imaging and machine learning, specific interest is increasing in reconstructing clinical diagnostic CT data, which, to date, has not been well-addressed by free and open-source CT reconstruction software. Without freely-available, open-source solutions, each group is required to either utilize a clinical scanner for their reconstructions, or is required to implement their own reconstruction software from scratch. For researchers specializing in CT reconstruction, the latter option is not so daunting, however groups looking to leverage CT reconstruction as a means to an end likely do not have the time or expertise. Further complicating work in the world of diagnostic CT are minimally-described, manufacturer-specific techniques designed to improve image quality, such as flying focal spots, and quarter detector offsets. While conceptually many of these are straightforward and have been documented in the literature (e.g., [4]), building an actual implementation can require substantial trial and error before clinical levels of image quality are achieved.

In 2016, we released FreeCT_wFBP providing one option for researchers looking to reconstruct diagnostic third generation CT data [5]. FreeCT_wFBP implemenents weighted filtered backprojection as described in [6] and supports the reconstruction of third-generation clinical diagnostic projection data, including scans acquired with arbitrary pitch values, and flying focal spots. With modern iterative reconstruction algorithms gaining popularity however, groups may not wish to utilize older filtered backprojection methods to conduct quantitative imaging tests. Additionally, as more clinical sites begin to use iterative methods, analyses conducted on filtered backprojection reconstructions are often critiqued for not using the most current approaches.

In this work we introduce FreeCT_ICD: fully 3D, model-based iterative reconstruction software for helical, diagnostic third-generation CT. This software builds on [5] in that it provides a model-based iterative reconstruction software package that addresses the same clinical system geometries and challenges (e.g., flying focal spots) as [5] and is released under the GNU General Public License (GPL) v2.0 in the hopes that it will be useful and educational to the broader CT research community. The purpose of this abstract is to introduce the software and provide sample results; a full discussion of the algorithm details will be provided in future work.

Listing 1. Excerpts of a configuration file used to reconstruct a phantom held in a head holder using FreeCT_ICD. YAML provides a simple, easily readable yet easily parsable format for configuration files.

```
# Paths
sinogram_path: ./n_ffs_1.ptr
output_dir:    ./
output_file:   n_ffs_headholder.img
...

# Iterative recon parameters
wfbp_initialize:        1
penalty:                quadratic
lambda:                 0.1
delta:                  0.005
...

# Scanner Geometry
# Parameters here are FIXED and should
# not change
acquisition_fov:                50.0
n_channels:                     736
num_views_per_turn_without_ffs: 1152
focal_spot_radius:              59.5
source_detector_distance:       108.56
...
```

II. Overview

A. Algorithm

Key approach details for Free_ICD are the following: a stored system matrix is employed with rotating reconstruction grid, which allows reasonable system matrix sizes for use with most computing systems and fast column retrieval from the system matrix. Iterative coordinate descent (ICD) is used to optimize a penalized least-squares objective function with either a quadratic or edge-preserving penalty function. While several works have previously explored the methods for storing the system matrix for clinical CT ([7], [8]), and the approach implemented in FreeCT_ICD uses the rotating slices approach described by Xu et al. in [7], FreeCT_ICD replaces other forward projection approaches with a modified Joseph's method. This approach allows us to leverage the rotating slices for matrix size reduction, minimize discretization errors well known to result from Siddon's method ([9]–[11]), and achieve smaller matrix sizes than the blob-based approach employed by [7].

While the full details of this scheme will be described in future work, briefly, the approach is a 3D extension of the 2D bilinear interpolation scheme described by [12], however with an added further reduction of the final interpolation step to a scheme based on the principle of Joseph's method. This accelerates the computation and reduces the matrix size relative to [12].

B. Software

FreeCT_ICD software is implemented using C++ and is intended primarily for use on Linux systems, however all code and dependencies should be easily portable to Mac and Windows-based systems. Reconstructions are configured using text configuration files written in YAML (https://yaml.org), which is a simple, human-readable "data-serialization" language, similar to markdown languages. A subset of one configuration file is given in listing 1.

FreeCT_ICD only requires two non-standard external libraries: yaml-cpp (https://github.com/jbeder/yaml-cpp) for the parsing of the configuration files, and the Boost libraries (http://www.boost.org/) for some vector functionality and matrix storage. Both libraries are cross platform and available for all major computer operating systems. Finally, FreeCT_ICD utilizes the standard OpenMP libraries (when available) to accelerate iterations. This is a standard parallelization library. This low-dependency design of FreeCT_ICD makes it easy to compile and run on nearly all modern systems.

Some key features of the software are: (1) the inclusion of a penalty term in the objective function as a regularizer, (2) the ability to initialize the MBIR algorithm with a weighted filtered backprojection (wFBP) reconstruction, and (3) support for flying focal spots. Two choices of penalty term are available in the released implementation, a quadratic penalty and an edge-preserving penalty. Because modifying the penalty term is often of interest to researchers, it has been implemented in a highly modular manner to make it easy to configure new penalties for experimentation. Initialization of the reconstruction volume from a wFBP reconstruction is provided by FreeCT_wFBP ([5]) and reduces the number of require iterations for convergence substantially. Because ICD is challenging to parallelize, this is an important addition to achieve reconstruction times usable for research.

While the fundamental ICD optimization is not parallelized in FreeCT_ICD, some level of parallelization has been achieved for the propagation of pixel updates back into the sinogram domain, which helps accelerate overall program execution.

Finally, FreeCT_ICD supports flying focal spots (FFS), a technique used to improve sampling near isocenter and increase axial and longitudinal resolution in the final reconstructed image [4]. While simulated data rarely employs FFS, major manufacturers often employ them in clinical data, and thus any reconstruction package targeting clinical raw data must offer support for FFS. While other works have discussed pathways to adding FFS as a feature (e.g., [8]), FreeCT_ICD at present has support for in-plane FFS (sometimes referred to as "phi" flying focal spot), with implementations of longitudinal (aka "Z") FFS and other configurations (Z+phi and diagonal FFS) planned in the near future.

FreeCT_ICD is being released under the GNU General Public License v2.0 in an effort to encourage further research and education using diagnostic CT. Licensing the software under the GNU GPL v2.0 means that users are free to copy, distribute, and modify the software provided that changes are identified and dated in the source code and any modifications are made freely available under the same license. The full text of the license can be found at https://www.gnu.org/licenses/old-licenses/gpl-2.0.en.html.

TABLE I
MATRIX SIZES, AND CORRESPONDING RECONSTRUCTION AND ACQUISITION PARAMETERS FOR PRESENTED SAMPLE RECONSTRUCTIONS.

Scan	**ACR Phantom**	**Pediatric Chest**
Acquisition Parameters		
Collimation	16 x 1.2mm	16 x 1.2mm
Pitch	1.0	1.0
Flying focal spot	Off	Off
Rotation time [s]	0.33	0.33
Reconstruction Parameters		
wFBP initialization	yes	yes
Voxel grid Dimensions	512 x 512 x 132	512 x 512 x 163
Voxel size [mm]	0.58 x 0.58 x 1.5	0.98 x 0.98 x1.5
FOV radius [mm]	300	500
Edge-preserving parameter	0.005	0.005
Penalty term parameter	0.1	0.1
Iterations	50	50
Matrix size [GB]	8.5	14.6

III. SOFTWARE DEMONSTRATION AND RESULTS

To demonstrate reconstruction quality and other properties of the software package, several sample reconstructions of data acquired on a clinical CT scanner were performed including a reconstruction of the ACR CT Accreditation phantom [13], and a pediatric chest study. Both studies were acquired using a Siemens Definition AS 64 (Siemens Healthineers, Forchheim, Germany).

A. Matrix Sizes

Matrix sizes are impacted by the reconstruction volume and projection data dimensions. Matrix sizes for the reconstructed volume are summarized in Tab. I along with reconstruction and acquisition parameters affecting matrix sizes.

B. Reconstructions

Reconstruction of the ACR phantom using the edge preserving parameter are shown in Fig. 1. The phantom was scanned in a head holder to allow for a smaller reconstructed field of view without incurring truncation artifacts. Three 1.5mm slices were averaged together to obtain a simulated 4.5 mm slice, which is closer to what would be submitted for ACR accreditation.

The slices shown in Fig. 1 meet or exceed the ACR CT accreditation requirements specified in [13], which test HU accuracy, contrast-to-noise ratio (CNR), and uniformity. In particular, CNR was approximately 3.8 which dramatically exceeds the required $>$1.0 needed for ACR accreditation. Additionally, although not shown here, spatial resolution was dramatically improved relative to other reconstruction methods (e.g., wFBP with sharp kernel) however noise levels were similar to those of wFBP with a smooth reconstruction kernel (similar to Siemens B10 or B20). This comparable noise level is highlighted in the pediatric reconstructions shown below (Fig. 2).

The pediatric reconstructions were performed using three different approaches for comparison: a wFBP reconstruction with a smooth reconstruction kernel, a FreeCT_ICD reconstruction with a quadratic penalty, and a FreeCT_ICD reconstruction with an edge-preserving penalty. A comparison of these three reconstructions can be found in Fig. 2.

IV. DISCUSSION AND CONCLUSIONS

Model-based iterative reconstruction (MBIR) has the potential to dramatically reduce radiation doses when applied to diagnostic CT, however the forward projection matrix for clinical systems is often substantially larger than can be stored in typical system memories. Using a previously published rotating slice approach [7], however, we have implemented a MBIR with a stored system matrix that allows for the reconstruction of third-generation, clinical, diagnostic CT scans.

The software is being freely distributed under the GNU GPL v2.0 and is available at https://github.com/FreeCT/FreeCT_ICD. FreeCT_ICD provides a more modern, complementary reconstruction approach to our previously released FreeCT_wFBP making the FreeCT software project more valuable for research that requires the reconstruction of clinical raw projection data.

REFERENCES

[1] A. Maier, H. G. Hofmann, M. Berger, P. Fischer, C. Schwemmer, H. Wu, K. Müller, J. Hornegger, J.-H. Choi, C. Riess, A. Keil, and R. Fahrig *Medical Physics*, no. 11.

[2] S. Rit, M. Vila Oliva, S. Brousmiche, R. Labarbe, D. Sarrut, and G. C. Sharp *Journal of Physics: Conference Series*, no. 1.

[3] J. Fessler, "Michigan Image Reconstruction Toolbox."

[4] T. G. Flohr, K. Stierstorfer, S. Ulzheimer, H. Bruder, A. N. Primak, and C. H. McCollough, "Image reconstruction and image quality evaluation for a 64-slice CT scanner with z-flying focal spot," *Medical Physics*, vol. 32, no. 8, pp. 2536–2547, 2005.

[5] J. Hoffman, S. Young, F. Noo, and M. McNitt-Gray, "Technical Note: FreeCT_wFBP: A robust , efficient , open-source implementation of weighted filtered backprojection for helical , fan-beam CT," *Medical Physics*, vol. 43, no. 3, p. 10 pp., 2016.

[6] K. Stierstorfer, A. Rauscher, J. Boese, H. Bruder, S. Schaller, and T. Flohr, "Weighted FBP - a simple approximate 3D FBP algorithm for multislice spiral CT with good dose usage for arbitrary pitch," *Physics in Medicine and Biology*, vol. 49, no. 11, pp. 2209–2218, 2004.

[7] J. Xu and B. M. W. Tsui, "Iterative image reconstruction in helical cone-beam x-ray CT using a stored system matrix approach.," *Physics in Medicine and Biology*, vol. 57, no. 11, pp. 3477–97, 2012.

[8] M. Guo and H. Gao, "Memory-efficient algorithm for stored projection and backprojection matrix in helical ct," *Medical Physics*, vol. 44, no. 4, pp. 1287–1300, 2017.

[9] W. Zbijewski and F. Beekman, "Characterization and suppression of edge and aliasing artefacts in iterative x-ray CT reconstruction," *Physics in Medicine and Biology*, vol. 49, no. 1, pp. 145–157, 2004.

[10] W. Zbijewski and F. J. Beekman, "Comparison of methods for suppressing edge and aliasing artefacts in iterative x-ray CT reconstruction," *Physics in Medicine and Biology*, vol. 51, no. 7, pp. 1877–1889, 2006.

[11] P. M. Joseph, "An Improved Algorithm for Reprojecting Rays Through Pixel Images," *IEEE Transactions on Medical Imaging*, vol. 1, no. 3, pp. 192–196, 1982.

[12] K. Hahn, H. Schndube, K. Stierstorfer, J. Hornegger, and F. Noo, "A comparison of linear interpolation models for iterative ct reconstruction," *Medical Physics*, vol. 43, no. 12, pp. 6455–6473, 2016.

[13] ACR, "CT Accreditation Phantom Instructions," pp. 1–14, 2013.

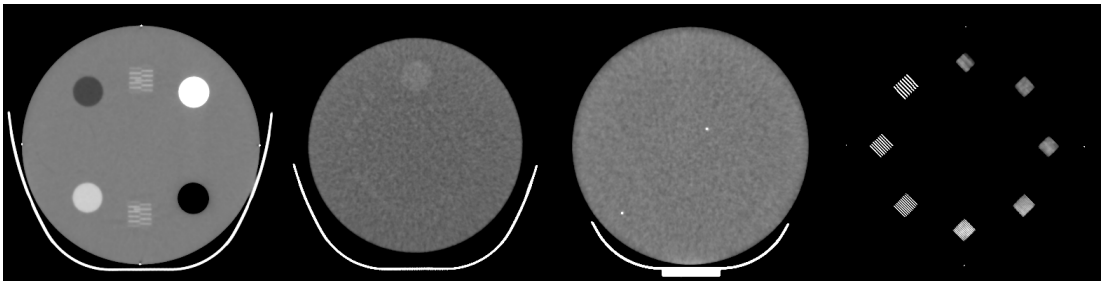


Fig. 1. Reconstruction of ACR CT accreditation phantom (Gammex Model 464, Gammex, Middleton, WI). Central slice of each module is shown. Window level for each slice is adjusted to ACR recommended values. Slightly smaller appearance of phantom in second image (the "low-contrast" module) is due to a ring lower density material that falls outside of the recommended window.

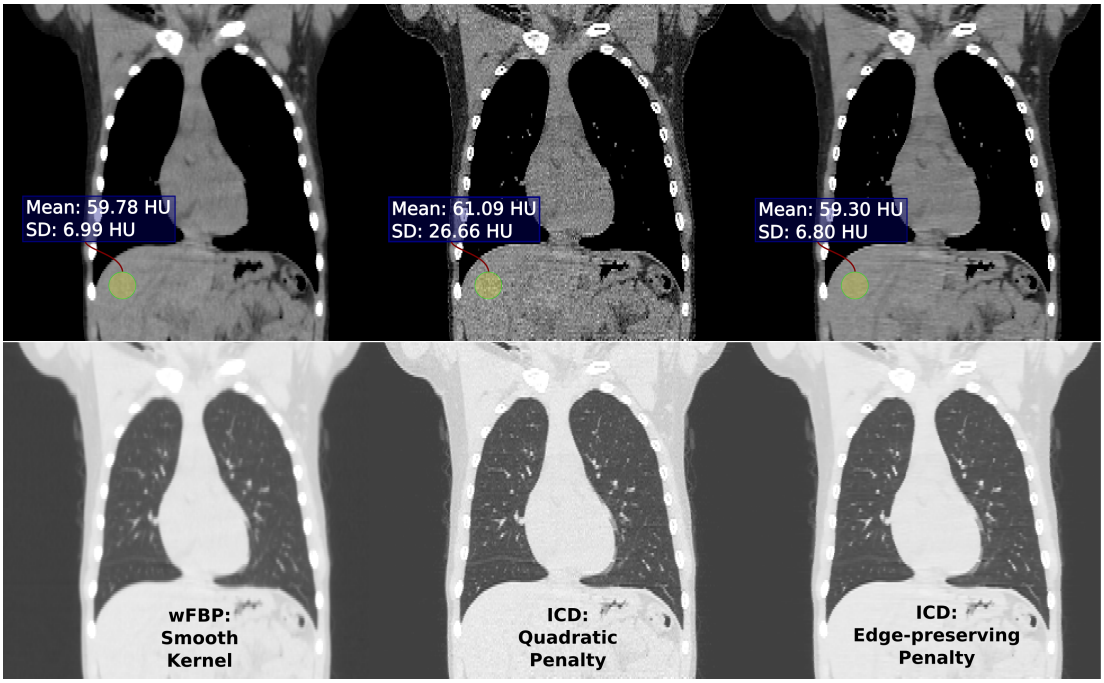


Fig. 2. Coronal reformats of a reconstruction of pediatric chest study. The left column is reconstructed with FreeCT_wFBP using a smooth reconstruction kernel. The middle and right columns are reconstructed using FreeCT_ICD with a quadratic and edge-preserving penalty, respectively. The top row shows the reconstructions with a soft-tissue window, and the bottom row shows the reconstructions with a lung window. The edge-preserving reconstruction clearly demonstrates the improved spatial resolution relative to the wFBP reconstruction while at the same time demonstrating almost the same noise level in the liver, highlighting the potential for noise reduction with MBIR.

Improving GPU Scaling for X-Ray CT

Harini Muthukrishnan, Thomas F. Wenisch, Jeffrey A. Fessler

***Abstract*—Model-based iterative reconstruction (MBIR) for X-Ray CT is computationally expensive, yet highly parallelizable, making it amenable to multi-GPU implementation. However, reconstruction time does not improve linearly with the number of GPUs, mainly due to high inter-GPU copying delays at the end of computation phases. Overlapping copies with computation—by copying incrementally as data are produced—can mitigate copy overhead and improve performance scalability. This paper demonstrates how to perform asynchronous copies using GPU threads initiated via dynamic kernel launch. Our technique enables 90% of copy time to overlap with compute, achieving a speedup of 1.24× (nearing the theoretical bound of 1.31× with instantaneous copies) over conventional cudaMemcpy at the end of compute phases on four Tesla K40m GPUs. Relative to a baseline implementation on a single GPU, our accelerated approach achieves a speedup of 3.46× on four GPUs. We project even higher impact from our technique with more GPUs.**

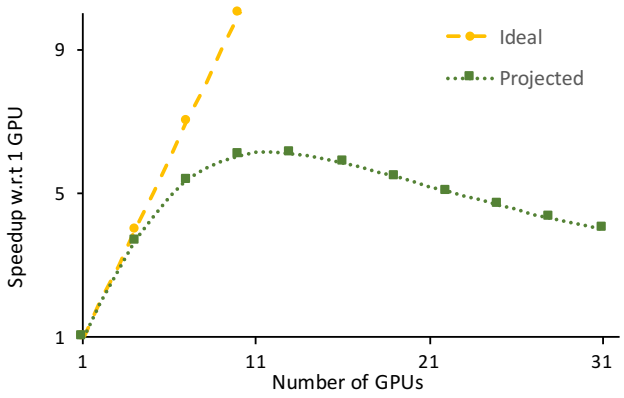


Fig. 1: Ideal and Projected speedups achievable with GPU scaling with conventional multi-GPU implementation.

I. Introduction

Model-based iterative reconstruction (MBIR) for X-Ray CT offers improved image quality at lower radiation doses than Filtered Back Projection (FBP) [1], but at higher computational costs. Although researchers have explored various acceleration techniques, such as using SIMD instructions on CPUs [2, 3] and cloud computing [4, 5], the compute times remain undesirably high for MBIR to be ubiquitous clinically.

Further MBIR acceleration requires increasing both computational resources and memory bandwidth, making a case for employing multiple GPUs [6, 7]. But GPU scaling does not always result in linear speedup. Fig. 1 shows projected speedups as we parallelize a state-of-the-art MBIR algorithm [8] over more GPUs. Using more GPUs initially provides near-linear speedup up to about eight GPUs, as the computational phases of MBIR can be readily partitioned across the GPUs. However, beyond eight GPUS, speedup saturates and then begins to decrease. This reversal arises due to the time spent copying data among GPUs between computational phases. Sinogram and image data must be exchanged all-to-all among the GPUs between phases, yet current systems offer no mechanism to broadcast data, requiring pairwise copies. As a result, even though computation time shrinks, copy time grows and ultimately dominates as the number of GPUs increases.

High copy time can be mitigated by overlapping copying with computation—by "streaming" data from producer GPUs as soon they become ready. By initiating some copies while computation is ongoing, the next computational phase must wait only for straggling data produced at the end of a phase.

While overlapping communication with compute between CPU and GPU has been studied for FBP[9], hiding copies underneath compute in multi-GPU systems entails several challenges: (1) State-of-the-art CPU–GPU systems provide an astonishing diversity of mechanisms to move data from one GPU's memory space to another. Data may be moved by CPU loads/stores through a variety of addressing mechanisms, by GPU loads/stores, by DMA engines integrated on the GPU, and potentially even by other devices on the PCIe bus. These alternatives trade off bandwidth, initiation latency, and disruption to other GPU threads in non-trivial ways; the best approach for our purpose is not obvious. (2) Every CT phase must indicate when enough data has been generated for a copy to start. It is neither clear how to trigger copies, nor at what granularity they should be performed. (3) The CUDA programming model allows enormous freedom in ordering the execution of individual threads. No existing programming interface allows GPU programs to efficiently track production of output data, initiate copies, and await copy completion.

This work focuses on accelerating a penalized weighted least-squares with ordered subsets (PWLS-OS) reconstruction algorithm [8] on a multi-GPU system by hiding copies under computation phases that generate data. We discuss performing copies using GPU threads initiated though dynamic kernel launch as the mechanism best suited for our purpose. We identify the best granularity at which to perform such copies and describe techniques to track data production and copy completion. We also consider how the subdivision of MBIR phases into individual GPU kernels affects the order data is generated and how to orchestrate these to maximize gains from our technique.

Using four Tesla K40m GPUs, we show that a 20-iteration helical CT reconstruction of a 512×512×512 image from 7256 views of size 888×64 takes 11.25 minutes using our copy mechanism. We also demonstrate, using a simple mathematical model, how our technique makes MBIR CT more amenable to further GPU scaling.

Supported in part by NIH Grant U01 EB018753. The authors are with the EECS Department of the University of Michigan, Ann Arbor, MI 48019 USA (email: {harinim, twenisch, fessler}@umich.edu).

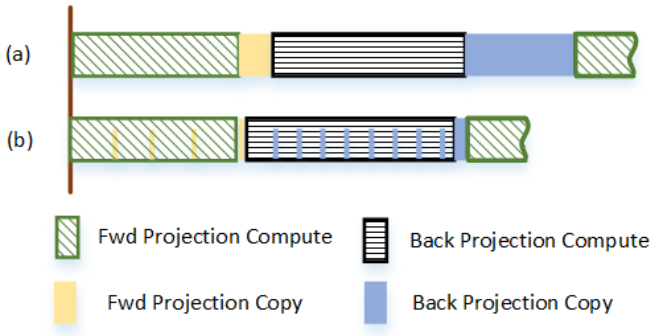


Fig. 2: (a) Conventional: computation and copies in series. (b) Our goal: incremental copying overlapping computation

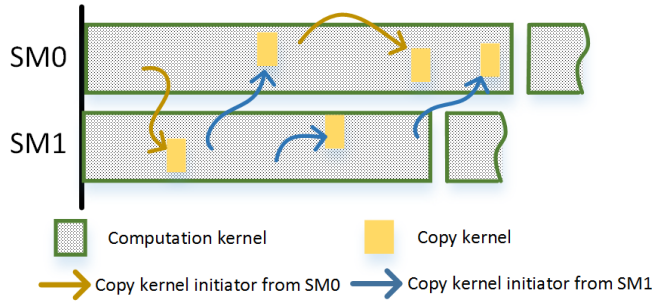


Fig. 3: GPU SM resource utilization and copy initiation. The launch of the initiated copy kernels depends entirely on the hardware scheduler.

II. Methods

A. Background

We reconstruct the image $\widehat{\mathbf{x}}$ by iteratively minimizing the PWLS cost function [5]:

$$\widehat{\mathbf{x}} = \arg\min_{\mathbf{x}\geq 0} \Psi(\mathbf{x}), \qquad \Psi(\mathbf{x}) = \frac{1}{2}\|\mathbf{A}\mathbf{x}-\mathbf{y}\|^2_{\mathbf{W}} + \mathbf{R}(\mathbf{x}), \qquad (1)$$

where $\mathbf{A}$ is the system matrix, $\mathbf{y}$ is the sinogram measurements, $\mathbf{W}$ is the statistical weighting and $\mathbf{R}$ is the regularizer.

Each iteration updates the current image estimate $(\mathbf{x}^{(n)})$ using the following gradient of Ψ:

$$\nabla\Psi(\mathbf{x}^{(n)}) = \mathbf{A}^T\mathbf{W}(\mathbf{A}\mathbf{x}^{(n)}-\mathbf{y}) + \nabla\mathbf{R}(\mathbf{x}^{(n)}). \qquad (2)$$

Each iteration comprises four phases as shown in Figure 4. Each phase can be formulated to admit considerable parallelism over image voxels or detector values in the sinogram, making it well-suited to using the enormous compute capabilities of modern GPUs [10].

GPUs have multiple Streaming Multiprocessors that concurrently execute many threads [11]. These hardware elements execute a GPU kernel (programmed in CUDA) organized as blocks, warps and threads [12]. Kernel ordering is programmer-controlled, but ordering of blocks and warps is left to the hardware scheduler.

In our GPU implementation of PWLS-OS, we parallelize forward projection by partitioning views across multiple GPUs, employing one GPU thread to compute the value of one detector residual. Every GPU broadcasts its generated residual values to every other GPU, since back projection uses all of them. We then parallelize computation of the other phases by partitioning the y plane across multiple GPUs. Each thread performs the back projection, regularization and update of one image voxel. The GPUs then copy the corresponding partial image to other GPUs before the next iteration commences.

B. Overlapping copy with compute

As shown in Figure 1, although PWLS-OS is highly amenable to parallelization, GPU scaling does not yield linear speedup due to time taken to perform all-to-all copies of the detector residual at the end of forward projection and of the image voxels at the end of the update phase. The total amount of data copied increases linearly with number of GPUs; although each GPU produces fewer values, they must copy to more destinations. Since using more GPUs also reduces compute time per GPU, scaling beyond 12 GPUs results in copy time that exceeds compute time.

The available copy bandwidth is bound by the interconnect (PCIe3.0 and PCIe4.0 pose theoretical limits of 16GB/s and 32GB/s, respectively, though substantially lower sustained throughput is achievable in practice) and cannot be increased without hardware enhancements. However, a careful study of the hardware utilization pattern during the reconstruction process indicates that the interconnect remains idle during the computation phases and is utilized to its practical limit only during the ensuing copy phase, when no compute takes place. Hence, one way to decrease the reconstruction time is to initiate copies of data generated early in the computation phase while the rest is still being computed, overlapping computation and copying, as depicted in Figure 2.

To overlap compute and copy, the copy mechanism must impose minimal overhead and interference on the computation. Our concurrent work [13] identifies GPU thread-based copy using dynamic kernel launch as the mechanism best suited for our purpose. In this technique, the GPU threads that perform computation also trigger copies after generating a data chunk (e.g., a set of sinogram bins) of the desired granularity. The copies are issued via *dynamic kernel launch*, wherein a copy kernel is launched from within the main CT computation kernels—a capability introduced in CUDA 5.0 [14]. The copy kernel is scheduled by the GPU hardware scheduler and copies the data chunk to the other GPUs as shown in Figure 3. As computation proceeds, copy kernels are triggered, scheduled by hardware, and executed.

Although our technique incurs minimal copy initiation overhead, the copy kernels nevertheless use GPU resources that might otherwise have been used by CT computation kernels, and hence indirectly delay computation. Hence, invoking the copy kernel at an appropriate frequency while ensuring that the maximum amount of copy time is hidden behind compute becomes crucial. We address this challenge by carefully tuning the granularity at which copies are initiated. Since this granularity depends only on the CT geometry and not on patient features, it can be determined in advance via a parameter-sweep over a sample image.

Choosing the right granularity to track data production is also important, as tracking at too low granularity (such as thread granularity) would increase the overhead of tracking,

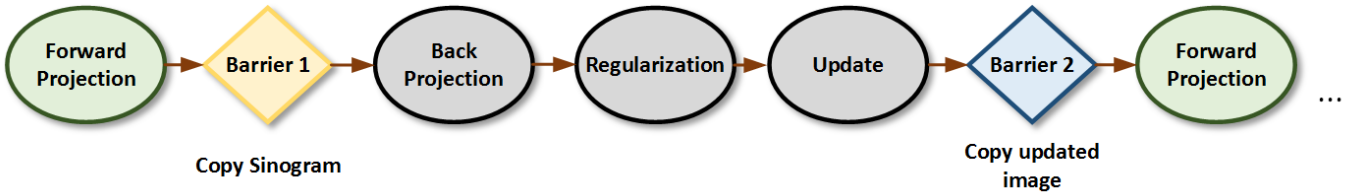


Fig. 4: Phases of X-Ray CT MBIR. After forward projection, the GPUs exchange their generated sinogram data, followed by three computational phases and an image update. The GPUs then exchange image data and the algorithm repeats until convergence.

slowing computation, while tracking at too high granularity might not ensure enough compute-copy overlap. We perform the tracking at the granularity of GPU thread blocks. Once consecutive thread blocks produce a chunk of data of the desired granularity, a dynamic copy kernel initiates the copies.

Our final design decision entails selecting the best implementation to track data generation. Proper design of this mechanism is crucial because the CUDA programming model offers enormous freedom to the scheduler with respect to block ordering, so blocks may complete in any order. We employ an atomic counter-based approach, wherein each data chunk is assigned a corresponding atomic counter, initialized to the number of blocks that contribute data to the chunk. The first thread of each block waits until all its sibling threads complete, then decrements the counter using an atomic decrement instruction. When the counter reaches zero, it indicates that the corresponding group of consecutive blocks is complete. The thread then initiates a dynamic copy kernel for the corresponding data chunk. Although the counter-based approach uses atomic accesses that are inherently slower than normal reads and writes, it performed better than alternatives (e.g., dedicated threads that poll for chunk completion).

C. Sequencing data generation

To ensure that the reconstruction method can effectively use the copy strategy discussed above, it is important to structure the kernels to perform all the computation corresponding to a particular data element in quick succession, producing data elements incrementally rather than performing multiple updates to all data elements during kernel execution. The goal is to ensure that data chunks are available as early as possible to maximize copy-compute overlap. While the original forward projection code computed the detector residuals in succession, the back projection code updated the relevant voxels using one view before updating the voxels using the next view, i.e., an outer loop over views. This led to voxels being ready for copy only during the processing of the last view, leaving very little room for compute-copy overlap. We restructured the code to generate the voxels in succession by having each voxel loop over the relevant views that contribute to it, thus ensuring voxel values are produced incrementally.

III. Experimental Results

We report on our multi-GPU PWLS-OS implementation. We validate the GPU implementation against a CPU baseline, which it matches to within 0.0289HU (Hounsfield units). Our test system comprises four Tesla K40m GPUs, each having

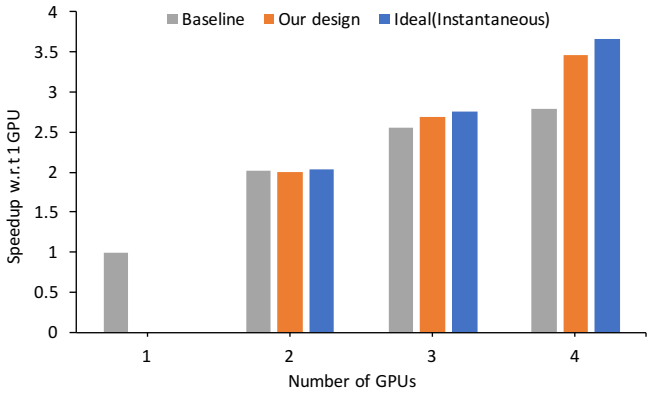


Fig. 5: Measured speedups achieved through GPU scaling

2880 CUDA cores, 11.9GB global memory and 4KB shared memory per block. Each GPU is capable of performing a peer access to the other GPUs. They reside on a PCIe3.0 bus that also interfaces them with the host.

We simulated 9-turn helical CT data with pitch 63/64 and 7256 views of size 64 rows by 888 channels, and reconstructed a 512×512×512 voxel image volume over a 512 mm transaxial field of view (FOV) with 0.625 mm slice thickness using the separable footprint projector [15] and [8] with 24 subsets.

In our design, the GPU threads initiated through dynamic kernel launch perform copies, as explained in Section II-B. The copies are performed for 8kB data chunks for detector residuals and 16kB for image voxels. We compare our approach against the baseline wherein data is copied after compute phases using cudaMemcpy Peer-to-Peer [16].

Figure 5 shows speedup achieved with respect to a single GPU for three cases: (1) Baseline, (2) Our design, and (3) Instantaneous copies (copies are performed in zero time). The blue (Instantaneous) bars indicate the theoretical limit on the performance gains achievable via compute-copy overlap. With only two GPUs, copy overhead is negligible and there is little performance difference between the baseline and the ideal. However, the potential and realized gains grow rapidly with further scaling. Our approach realizes 94% of the theoretical opportunity for four GPUs, achieving a speedup of 1.24× over the baseline and 3.46× over a single GPU.

Our design falls short of the opportunity available with instantaneous copies for two reasons: (1) The dynamic copy kernels require some GPU execution bandwidth, slightly delaying execution of CT kernel threads. (2) A 100% copy-compute overlap is not possible because the data generated by the final blocks is copied after computation is complete.

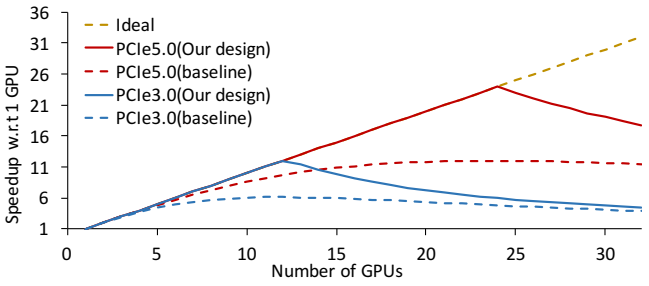


Fig. 6: Effect of GPU scaling on speedup for different interconnect generations

IV. Scalability Analysis

Our approach holds the potential to unlock even higher performance scalability on future, larger multi-GPU systems. To analyze this potential, we develop a simple analytic model that predicts the impact of our approach with more GPUs and faster GPU interconnects. We base our scalability model on the following observations: (1) As previously explained, the four phases of X-ray CT are amenable to GPU scaling. For simplicity, we assume that the compute time of individual algorithmic phases scales linearly with the number of GPUs, although practical implementations typically fall a bit short of ideal linear scaling. (2) An all-to-all broadcast must occur at the end of the forward projection and update phases. Thus, as the number of GPUs increases, the total data to be copied, and hence the time required for copy, increases, since the data must be copied to additional GPUs.

A. Scalability Model

At the end of forward projection, the GPUs must exchange the portions of the sinogram each generated. The total bytes copied is the product of the sinogram size ($\text{Size}_{\text{sino}}$), and the number of destination GPUs ($n_{\text{gpu}} - 1$). The total time for the copy depends upon the interconnect technology, which we model simply as a 'copy time per byte (t_{perbyte})' bandwidth, as expressed in the following equation:

$$\text{Copytime}_1 = (n_{\text{gpu}} - 1) \ * \ \text{Size}_{\text{sino}} \ * \ t_{\text{perbyte}}.$$

After the update phase, the GPUs must exchange the image portions each generated. Hence the total bytes copied is the product of the total image size ($\text{Size}_{\text{image}}$) and the number of destination GPUs ($n_{\text{gpu}} - 1$):

$$\text{Copytime}_2 = (n_{\text{gpu}} - 1) \ * \ \text{Size}_{\text{image}} \ * \ t_{\text{perbyte}}.$$

B. Discussion

Using this simple model, we estimate the impact of our design on CT reconstruction performance. Figure 6 shows the projected speedup of our design and that of the baseline for different GPU counts against a single GPU implementation for two different assumptions on interconnect bandwidth (unidirectional transfer bandwidths of 16GB/sec for PCIe 3.0 and an estimated 64GB/s for PCIe 5.0).

For both assumptions on interconnect bandwidth, our approach enables performance scalability to a much larger number of GPUs than the baseline. For PCIe 3.0, baseline performance saturates at about 6× speedup (over a single GPU) with ten GPUs. Above six GPUs, growth in copy time exceeds reductions in computation time. In contrast, our approach enables near-ideal scaling up to twelve GPUs. Beyond this point, copy time grows to the point where it exceeds compute time and can no longer be hidden.

Higher interconnect bandwidth under PCIe 5.0 reduces copy time, enabling greater performance scalability for both the baseline and our method. However, our technique still drastically increases the scalability potential. Our technique enables near-ideal scaling up to 24 GPUs, while the baseline saturates at about 12× speedup with 20 GPUs. Newer GPUs will further decrease compute time by the time PCIe 5.0 becomes available, making our solution even more relevant.

V. Summary and Conclusion

In this paper, we proposed compute-copy overlap using warp based copies to overcome the communication bottleneck of GPU scaling of CT MBIR. We demonstrated that our technique offers a 3.46× speedup over a single GPU implementation on a 4 GPU system, and makes CT MBIR more amenable to further GPU scaling.

References

[1] J.-B. Thibault, K. Sauer, C. Bouman, and J. Hsieh, "A three-dimensional statistical approach to improved image quality for multi-slice helical CT," *Med. Phys.*, vol. 34, no. 11, pp. 4526–44, Nov. 2007.

[2] K. Zeng, E. Bai, and G. Wang, "A fast CT reconstruction scheme for a general multi-core PC," *Intl. J. Biomedical Im.*, vol. 2007, p. 29160, 2007.

[3] R. Sampson, M. G. McGaffin, T. F. Wenisch, and J. A. Fessler, "Investigating multi-threaded SIMD for helical CT reconstruction on a CPU," in *Proc. 4th Intl. Mtg. on image formation in X-ray CT*, 2016, pp. 275–8.

[4] J. Ni, X. Li, T. He, and G. Wang, "Review of parallel computing techniques for computed tomography image reconstruction," *Current Medical Imaging Reviews*, vol. 2, no. 4, pp. 405–14, Nov. 2006.

[5] J. Rosen, J. Wu, T. Wenisch, and J. Fessler, "Iterative helical CT reconstruction in the cloud for ten dollars in five minutes," in *Proc. Intl. Mtg. on Fully 3D Image Recon. in Rad. and Nuc. Med*, 2013, pp. 241–4.

[6] M. McGaffin and J. A. Fessler, "Alternating dual updates algorithm for X-ray CT reconstruction on the GPU," *IEEE Trans. Computational Imaging*, vol. 1, no. 3, pp. 186–99, Sep. 2015.

[7] B. Jang, D. Kaeli, S. Do, and H. Pien, "Multi GPU implementation of iterative tomographic reconstruction algorithms," in *Proc. IEEE Intl. Symp. Biomed. Imag.*, 2009, pp. 185–8.

[8] D. Kim, D. Pal, J.-B. Thibault, and J. A. Fessler, "Improved ordered subsets algorithm for 3D X-ray CT image reconstruction," *Proc. 2nd Intl. Mtg. on image formation in X-ray CT*, pp. 378–81, 2012.

[9] T. Zinßer and B. Keck, "Systematic performance optimization of cone-beam back-projection on the Kepler architecture," in *Proc. Intl. Mtg. on Fully 3D Image Recon. in Rad. and Nuc. Med*, 2013, pp. 225–8.

[10] K. Mueller, F. Xu, and N. Neophytou, "Why do commodity graphics hardware boards (GPUs) work so well for acceleration of computed tomography?" in *Proc. SPIE 6498 Comp. Imag.*, 2007, p. 64980N.

[11] E. Lindholm, J. Nickolls, S. Oberman, and J. Montrym, "NVIDIA Tesla: A unified graphics and computing architecture," *IEEE micro*, vol. 28, no. 2, 2008.

[12] D. B. Kirk and W. H. Wen-Mei, *Programming massively parallel processors: a hands-on approach*. Morgan Kaufmann, 2016.

[13] H. Muthukrishnan, J. A. Fessler, and T. F. Wenisch, "SUBLIME: Hiding copies in multi-GPU systems," 2018, submitted.

[14] S. Jones, "Introduction to dynamic parallelism," in *GPU Technology Conf. Presentation S*, vol. 338, 2012, p. 2012.

[15] Y. Long, J. A. Fessler, and J. M. Balter, "3D forward and back-projection for X-ray CT using separable footprints," *IEEE Tr. Med. Im.*, vol. 29, no. 11, pp. 1839–1850, 2010.

[16] T. Schroeder, "Peer-to-peer and unified virtual addressing," in *GPU Technology Conference, NVIDIA*, 2011.

Emission EM Look-Alike Algorithms for X-Ray CT and Other Applications

Larry Zeng

***Abstract* — In emission tomography, the EM (expectation maximization) is easy to use with only one parameter to adjust — the number of iterations. On the other hand, the EM algorithms for transmission tomography are not so user-friendly and have many problems. This paper develops a family of emission-EM-look-alike algorithms. One of them can be applied to transmission tomography such as the x-ray CT (computed tomography).**

I. Introduction

THE EM (expectation maximization) methodology is a general approach to compute maximum likelihood estimates by using iterative techniques [1]. There are many EM algorithms. The most famous EM algorithm in medical imaging community is the one for emission tomography [2-5]. The emission EM algorithm uses a multiplicative form to update the image; it has a built-in property to enforce the image non-negativity and Poisson noise nature in the data. It is efficient to implement and stable. It has no adjustable parameters other than the number of iteration. It is safe to state that it is the most favorite iterative algorithm in nuclear medicine.

An EM algorithm for transmission tomography was developed in [5] by Lange and Carson. Unlike its emission tomography counterpart, this EM algorithm for transmission tomography has many drawbacks. It is complicated to compute and slow in convergence [6]. There are other versions of the transmission EM algorithms, but there are no fundamental improvements [7-10].

This paper develops a family of emission-EM-look-alike algorithms. They are iterative algorithms in the form of multiplicative image update, which intrinsically enforces the image non-negativity. Each member of the family has its unique noise model as explained in detail in the next section.

II. Methods

2.1 The emission EM algorithm

The starting point of our development is the emission EM algorithm

$$x_i^{(k+1)} = \frac{x_i^{(k)}}{\sum_j a_{ji}} \sum_j a_{ji} \left(\frac{p_j}{\sum_n a_{jn} x_n^{(k)}} \right), \tag{1}$$

where $x_i^{(k)}$ is the i^{th} image pixel at the k^{th} iteration, P_j is the j^{th} line-integral (ray-sum) measurement value, and A_{ji} is the contribution of the i^{th} image pixel to the j^{th} measurement. The summation over the index n is the projector and the summation over the index j is the backprojector. Expression (1) is in the form of multiplicative image update, and it can be re-written in the form of additive image update

$$\begin{aligned}
x_i^{(k+1)} &= \frac{x_i^{(k)}}{\sum_j a_{ji}} \sum_j a_{ji} \left(\frac{p_j}{\sum_n a_{jn} x_n^{(k)}} \right) \\
&= \frac{x_i^{(k)}}{\sum_j a_{ji}} \sum_j a_{ji} \left(\frac{p_j - \sum_n a_{jn} x_n^{(k)} + \sum_n a_{jn} x_n^{(k)}}{\sum_n a_{jn} x_n^{(k)}} \right) \\
&= \frac{x_i^{(k)}}{\sum_j a_{ji}} \sum_j a_{ji} \left(\frac{p_j - \sum_n a_{jn} x_n^{(k)}}{\sum_n a_{jn} x_n^{(k)}} \right) + \frac{x_i^{(k)}}{\sum_j a_{ji}} \sum_j a_{ji} \left(\frac{\sum_n a_{jn} x_n^{(k)}}{\sum_n a_{jn} x_n^{(k)}} \right) \\
&= \frac{x_i^{(k)}}{\sum_j a_{ji}} \sum_j a_{ji} \frac{1}{\sum_n a_{jn} x_n^{(k)}} \left(p_j - \sum_n a_{jn} x_n^{(k)} \right) + x_i^{(k)} \\
&= x_i^{(k)} + \lambda_i^{(k)} \sum_j a_{ji} w_j^{(k)} \left(p_j - \sum_n a_{jn} x_n^{(k)} \right)
\end{aligned} \tag{2}$$

The last line of (2) is in the form of the iterative Landweber algorithm, which is a gradient decent algorithm. In (2),

$$\lambda_i^{(k)} = \frac{x_i^{(k)}}{\sum_j a_{ji}} \tag{3}$$

is the relaxation parameter, which is also known as the step size, and

$$w_j^{(k)} = \frac{1}{\sum_n a_{jn} x_n^{(k)}} \tag{4}$$

is the weighting factor for the *j*th projection ray. For the Poisson noise model, the noise variance is the mean value of the ray sum.

We can make two observations from (3) and (4). First, the EM algorithm's step size $\lambda_i^{(k)}$ is scaled by the image pixel value $x_i^{(k)}$ at the *k*th iteration. The step is larger for objects with larger image values. Therefore, the bright lesions converge faster than the dark lesions.

Second, the weighting factor $w_j^{(k)}$ is the reciprocal of the estimated mean value of the *j*th ray by the reconstruction at

Gengsheng L. Zeng is with Department of Engineering, Weber State University, Ogden, UT 84408 USA (telephone: 801-626-6864, e-mail: larryzeng@weber.edu) and also with Department of Radiology and Imaging Sciences, University of Utah, Salt Lake City, UT 84108 USA (telephone: 801-581-3918, e-mail: larry.zeng@hsc.utah.edu).

the *k*th iteration. The mean value is the same as the variance for Poisson noise.

2.2 Modification of the emission EM algorithm

We propose to introduce a new scaling factor $s_j^{(k)}$ in the backprojector for each projection ray *j*. Thus the original emission EM algorithm (1) becomes

$$x_i^{(k+1)} = \frac{x_i^{(k)}}{\sum_j a_{ji} s_j^{(k)}} \sum_j a_{ji} s_j^{(k)} \left(\frac{p_j}{\sum_n a_{jn} x_n^{(k)}}\right). \quad (5)$$

As a consequence, the additive version (2) becomes

$$\begin{aligned} x_i^{(k+1)} &= \frac{x_i^{(k)}}{\sum_j a_{ji} s_j^{(k)}} \sum_j a_{ji} s_j^{(k)} \left(\frac{p_j}{\sum_n a_{jn} x_n^{(k)}}\right) \\ &= \frac{x_i^{(k)}}{\sum_j a_{ji} s_j^{(k)}} \sum_j a_{ji} s_j^{(k)} \left(\frac{p_j - \sum_n a_{jn} x_n^{(k)} + \sum_n a_{jn} x_n^{(k)}}{\sum_n a_{jn} x_n^{(k)}}\right) \\ &= \frac{x_i^{(k)}}{\sum_j a_{ji} s_j^{(k)}} \sum_j a_{ji} s_j^{(k)} \left(\frac{p_j - \sum_n a_{jn} x_n^{(k)}}{\sum_n a_{jn} x_n^{(k)}}\right) \\ &\quad + \frac{x_i^{(k)}}{\sum_j a_{ji} s_j^{(k)}} \sum_j a_{ji} s_j^{(k)} \left(\frac{\sum_n a_{jn} x_n^{(k)}}{\sum_n a_{jn} x_n^{(k)}}\right) \\ &= x_i^{(k)} + \frac{x_i^{(k)}}{\sum_j a_{ji} s_j^{(k)}} \sum_j a_{ji} s_j^{(k)} \frac{1}{\sum_n a_{jn} x_n^{(k)}} \left(p_j - \sum_n a_{jn} x_n^{(k)}\right) \\ &= x_i^{(k)} + \hat{\lambda}_i^{(k)} \sum_j a_{ji} \hat{w}_j^{(k)} \left(p_j - \sum_n a_{jn} x_n^{(k)}\right) \end{aligned} \quad (6)$$

with a new step size and a new weighting factor:

$$\hat{\lambda}_i^{(k)} = \frac{x_i^{(k)}}{\sum_j a_{ji} s_j^{(k)}} \quad (7)$$

and

$$\hat{w}_j^{(k)} = \frac{s_j^{(k)}}{\sum_n a_{jn} x_n^{(k)}}. \quad (8)$$

Eq. (5) represents a family of emission-EM-look-alike algorithms depending on the definition of the new scaling factor $s_j^{(k)}$, as explained in the following special cases (or examples).

2.3 Special case 1: No weighting

Let's consider a hypothetical imaging system, in which the noise in the sinagram is identically distributed with the same variance. No noise weighting should be used for the image reconstruction algorithm. Thus we require the weighting factor (8) to be

$$\hat{w}_j^{(k)} = \frac{s_j^{(k)}}{\sum_n a_{jn} x_n^{(k)}} = 1 \quad (9)$$

which results in

$$s_j^{(k)} = \sum_n a_{jn} x_n^{(k)}. \quad (10)$$

Substituting (10) into (5) yields

$$\begin{aligned} x_i^{(k+1)} &= \frac{x_i^{(k)}}{\sum_j a_{ji} (\sum_n a_{jn} x_n^{(k)})} \sum_j a_{ji} (\sum_n a_{jn} x_n^{(k)}) \left(\frac{p_j}{\sum_n a_{jn} x_n^{(k)}}\right) \\ &= x_i^{(k)} \frac{\sum_j a_{ji} p_j}{\sum_j a_{ji} \sum_n a_{jn} x_n^{(k)}}. \end{aligned} \quad (11)$$

In (11), $\sum_j a_{ji} p_j$ is the backprojection of the measured sonogram and $\sum_j a_{ji} \sum_n a_{jn} x_n^{(k)}$ is the backprojection of the forward projection of the current estimate of the image.

2.4 Special case 2: Modified Emission

Let's consider another hypothetical imaging system, in which the noise in the sinogram is not quite Poisson distributed, but the variance is given as

$$var = \bar{p}_j^{\gamma}, \quad (12)$$

where $\bar{p}_j$ is the mean value of the sinogram p_j and γ is a constant. When $\gamma = 1$, this case is the classic situation of emission tomography using a Poisson noise model.

Using the EM strategy of estimating $\bar{p}_j$ by $\sum_n a_{jn} x_n^{(k)}$, the weighting factor (8) for our modified emission case is given as

$$\hat{w}_j^{(k)} = \frac{s_j^{(k)}}{\sum_n a_{jn} x_n^{(k)}} = \frac{1}{(\sum_n a_{jn} x_n^{(k)})^{\gamma}}, \quad (13)$$

which results in

$$s_j^{(k)} = (\sum_n a_{jn} x_n^{(k)})^{1-\gamma}. \quad (14)$$

Substituting (14) into (5) yields

$$\begin{aligned} x_i^{(k+1)} &= \frac{x_i^{(k)}}{\sum_j a_{ji} (\sum_n a_{jn} x_n^{(k)})^{1-\gamma}} \sum_j a_{ji} (\sum_n a_{jn} x_n^{(k)})^{1-\gamma} \left(\frac{p_j}{\sum_n a_{jn} x_n^{(k)}}\right) \\ &= \frac{x_i^{(k)}}{\sum_j a_{ji} (\sum_n a_{jn} x_n^{(k)})^{1-\gamma}} \sum_j a_{ji} \frac{p_j}{(\sum_n a_{jn} x_n^{(k)})^{\gamma}}. \end{aligned} \quad (15)$$

As expected, when $\gamma = 1$, (15) is the famous emission EM algorithm.

2.5 Special case 3: Transmission

In transmission tomography, the sinogram variance is proportional to the exponential function of the sinogram's mean value, that is,

$$var \propto \exp(\bar{p}_j). \quad (16)$$

The weighting factor can be chosen as the reciprocal of the variance. Using the EM strategy of estimating $\bar{p}_j$ by $\sum_n a_{jn} x_n^{(k)}$, the weighting factor (8) for the transmission tomography case is given as

$$\hat{w}_j^{(k)} = \frac{s_j^{(k)}}{\sum_n a_{jn} x_n^{(k)}} = \exp(-\sum_n a_{jn} x_n^{(k)}), \quad (17)$$

which results in

$$s_j^{(k)} = (\sum_n a_{jn} x_n^{(k)}) \exp(-\sum_n a_{jn} x_n^{(k)}). \quad (18)$$

Substituting (18) into (5) yields

$$x_i^{(k+1)} = x_i^{(k)} \frac{\sum_j a_{ji} p_j \exp(-\sum_n a_{jn} x_n^{(k)})}{\sum_j a_{ji} (\sum_n a_{jn} x_n^{(k)}) \exp(-\sum_n a_{jn} x_n^{(k)})}. \quad (19)$$

Eq. (19) is the main result of this paper. It is an emission EM look-alike algorithm for the transmission tomography. It is in the form of multiplicative image update, has an intrinsic non-negativity constraint, and weights the sinogram with (17). Most importantly, (19) is user friendly and easy to implement.

2.6 Special case 4: Modified Transmission

Again, this is another hypothetical imaging system, in which the noise in the sinogram has a variance given by

$$var \propto \exp(\gamma \bar{p}_j), \quad (20)$$

where $\bar{p}_j$ is the mean value of the sinogram p_j and γ is a constant. When $\gamma = 1$, this case is the classic situation of transmission tomography using a Poisson noise model for its pre-log data.

According to the noise model (20), the weighting factor can be chosen as

2.8

$$\hat{w}_j^{(k)} = \frac{s_j^{(k)}}{\sum_n a_{jn} x_n^{(k)}} = \exp(-\gamma \sum_n a_{jn} x_n^{(k)}), \quad (21)$$

which results in

$$s_j^{(k)} = (\sum_n a_{jn} x_n^{(k)}) \exp(-\gamma \sum_n a_{jn} x_n^{(k)}). \quad (21)$$

Substituting (21) into (5) yields

$$x_i^{(k+1)} = x_i^{(k)} \frac{\sum_j a_{ji} p_j \exp(-\gamma \sum_n a_{jn} x_n^{(k)})}{\sum_j a_{ji} (\sum_n a_{jn} x_n^{(k)}) \exp(-\gamma \sum_n a_{jn} x_n^{(k)})}. \quad (22)$$

When $\gamma = 1$, this degenerates to the usual transmission tomography case (19). When $\gamma = 0$, this is the no-weighting case as (11).

From the examples above, a large family of emission EM look-alike image reconstruction algorithms can be developed as long a noise variance function is provided.

2.7 Convergence

It is natural to ask whether the proposed family of algorithms converge or not. The additive form (6) of this family of algorithm is a Landweber algorithm, which is a gradient descent algorithm. The convergence of the Landweber algorithm is guaranteed if the step size is small enough. The upper bound of the step size is $2/\sigma_{max}$, where σ_{max} is the maximum eigenvalue of the combined projection/backprojection matrix. However, it is not an easy task to evaluate σ_{max}. An empirically stable Landweber algorithm uses a relaxation parameter as $1/(\sum_j a_{ji} w_j \sum_n a_{jn})$ and the algorithm is expressed as [11]

$$x_i^{(k+1)} = x_i^{(k)} + \alpha \frac{\sum_j a_{ji} w_j (p_j - \sum_n a_{jn} x_n^{(k)})}{\sum_j a_{ji} w_j \sum_n a_{jn}}. \quad (23)$$

If we replace the weighting factor w_j by $s_j^{(k)} / \sum_n a_{jn} x_n^{(k)}$ as in (8), then the relaxation parameter in (23) becomes

$$\frac{1}{\sum_j a_{ji} w_j \sum_n a_{jn}} = \frac{1}{\sum_j a_{ji} \frac{s_j^{(k)}}{\sum_n a_{jn} x_n^{(k)}} \sum_n a_{jn}} \approx \frac{1}{\sum_j a_{ji} \frac{s_j^{(k)}}{x_i^{(k)} \sum_n a_{jn}} \sum_n a_{jn}} = \frac{x_i^{(k)}}{\sum_j a_{ji} s_j^{(k)}} = \hat{\lambda}_i^{(k)} \quad (24)$$

where $\hat{\lambda}_i^{(k)}$ is given in (7). The approximation in (24) may become exact only for certain values $x_i^{(k)}$. The approximate relationship (24) is not a mathematical proof that the emission EM look-alike family has the similar convergence property as the Landweber algorithm. We do not yet have a rigorous convergence analysis for this family. This paper uses some computer simulations to demonstrate some cases indicating the similar convergence properties between the emission EM look-alike family and the iterative Landweber algorithm.

2.7 Setup of Computer Simulations

The computer simulations in this paper use a two-dimensional (2D) uniform elliptical phantom containing 2 small hot lesions and 2 cold lesions. A parallel-beam imaging geometry is assumed. The image array was 180 x 180, the number of views was 180 over 180°, and the number of detection channels was 180. To avoid inverse-problem crime, the projector used to generate projection sinograms is different from the projector used for image reconstruction. In sinogram date generation, the image was first up-sampled 10 times and the detector was also up-sampled 10 times. The noiseless sonogram data was then down-sampled 10 times and noise was introduced. Negative sinogram values were set to zero.

Results from 3 noise types are presented in the next section: the emission Poisson noise, the uniformly distributed Gaussian noise, and the transmission Gaussian noise with variance proportional to the exponential function of the sinogram value.

III. RESULTS

3.1 Emission Poisson Noise Case

The purpose of this case is to establish the convergence relationship between the well-known emission EM algorithm and the Landweber algorithm, which will be used as a benchmark for the other two noise models. The comparison results are shown in Fig. 1, where the line-profiles are obtained from noiseless data reconstructions along row #72 to compare resolution recovery. The Landweber algorithm uses 100 iterations. The EM algorithm uses 30 iterations to match the hot lesion convergence and 80 iterations to match the cold lesion convergence.

3.2 Stationary Noise Case

There is no noise weighting used in the case. The purpose of this case is to show that the noise weighting does not affect the convergence property much. The results are shown in Fig. 2.

3.3 Transmission Noise Case

The transmission noise case is the main result of this paper, as displayed in Fig. 3. The proposed transmission algorithm converges almost at the same rate as the emission EM algorithm.

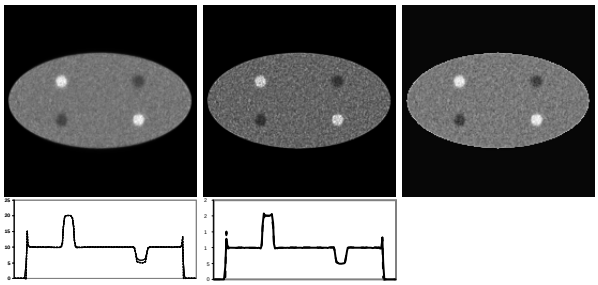


Figure 1. Images reconstructed from sinogram with emission noise model. Left image: 30 iterations of EM. Middle image: 80 iterations of EM. Right image: 100 iterations of Landweber.

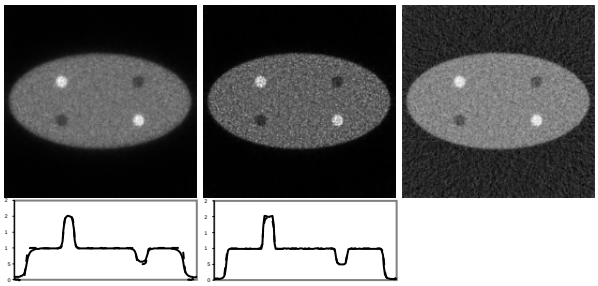


Figure 2. Images reconstructed from sinogram with stationary Gaussian noise model. Left image: 30 iterations of EM look-alike. Middle image: 80 iterations of EM look-alike. Right image: 100 iterations of Landweber.

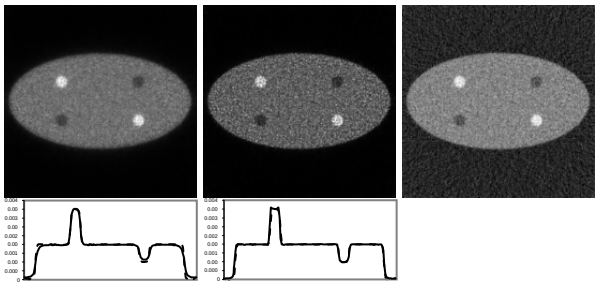


Figure 3. Images reconstructed from sinogram with transmission noise model. Left image: 30 iterations of EM look-alike. Middle image: 80 iterations of EM look-alike. Right image: 100 iterations of Landweber.

IV. Conclusion

The emission EM algorithm has wide applications in nuclear medicine due to its good noise control property, intrinsic non-negativity constraint, and ease of implementation. The emission EM algorithm is a nonlinear algorithm and its convergence rate is non-stationary. If written in the additive update form, it is easy to observe that the step size of the algorithm depends on the image intensity. It converges faster for bright objects and slower for dark objects. Its noise weighting factor is estimated by the current iteration of the image solution, while the gradient descent algorithms (such as the Landweber algorithm) use the sinogram data to calculate the weighting factors.

All these properties mentioned above are preserved by our newly suggested emission EM look-alike algorithms, which are developed in an *ad hoc* manner. For any noise variance function, an emission EM look-alike algorithm can be easily devised by following the steps in Section II.

One important application is the algorithm for transmission tomography. According to our computer simulations, it convergence rate is comparable to that of the iterative Landweber algorithm or the emission EM algorithm.

References

[1] A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum likelihood from complete data via the EM algorithm,” *Journal of the Royal Statistical Society: Series B*, vol. 39(1): pp. 1-38, 1977.

[2] W. H. Richardson, “Bayesian-based iterative method of image restoration,” *JOSA*, vol. 62(1), pp. 55-59, 1972, doi: 10.1364/JOSA.62.000055.

[3] L. B. Lucy, “An iterative technique for the rectification of observed distributions,” *Astronomical Journal*, vol. 79(6), pp. 745-754, 1974, doi: 10.1086/111605.

[4] L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Trans. Med. Imag.*, vol. 1, pp. 113–122, 1982, doi: 10.1109/TMI.1982.4307558, PMID 18238264

[5] K. Langer and R. Carson, "EM reconstruction algorithms for emission and transmission tomography," *J. Comp. Ass. Tomogr.*, vol. 8, pp. 302-316, 1984.

[6] K. Lange and J. A. Fessler, “Globally convergent algorithms for maximum a posteriori transmission tomography,” *IEEE Trans. Imag. Proc.*, vol. 4, pp. 1430-1438, 1995.

[7] J. T. Kent and C. Wright, “Some suggestions for transmission tomography based on the EM algorithm,” in *Stochastic Models, Statistical Methods, and Algorithms in Imag. Analysis* (M. P. Barone, A. Frigessi, ed.), vol. 74 of *Lecture Notes in Statistics*, pp. 219-232, New York: Springer, 1992.

[8] J. A. Browne and T. J. Holmes, “Developments with maximum likelihood x-ray computed tomography,” *IEEE Trans. Med. Imag.*, vol. 12, pp. 40-52, 1992.

[9] J. M. Ollinger and G. Johns, “The use of maximum a-posteriori and maximum likelihood transmission images for attenuation correction PET,” in *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf.*, vol. 2, pp. 1185-1187, 1992.

[10] J. M. Ollinger, “Maximum likelihood reconstruction of transmission images in emission computed tomography via the EM algorithm,” *IEEE Trans. Med. Imag.*, vol. 13, pp. 89-101, 1994.

[11] G. L. Zeng and W. Wang, “Noise weighting with an exponent for transmission CT,” *Biomedical Physics & Engineering Express*, vol. 2, no. 045004, 2016.

Material Decomposition using Spectral Prior Image Constrained Compressed Sensing

Shengzhen Tao, Kishore Rajendran, Cynthia H. McCollough, and Shuai Leng

***Abstract*—Multi-energy CT obtains X-ray attenuation measurements at two or more energy spectra, allowing quantification of materials with different compositions. The material decomposition process is intrinsically susceptible to noise amplification due to its ill-conditioned nature. Hence, basis material-specific images can be contaminated by the presence of strong noise, which compromises the conspicuity of small objects and may cause loss of information regarding structural details and key pathology. In this work, we describe an image domain Material Decomposition framework using Spectral Prior Image Constrained Compressed Sensing (MD-SPICCS). The proposed framework exploits the structural redundancy between the individual basis material-specific images and the spectral energy-bin or energy-threshold-specific CT images to retain structural details in denoised basis material-specific images. Our initial phantom results demonstrate that MD-SPICCS can reduce iodine quantification root-mean-square error (RMSE) from 3.9 mg/mL to 1.0 mg/mL, while preserving structural details and spatial resolution in the basis material-specific images.**

I. Introduction

Spectral CT complements conventional CT by obtaining X-ray attenuation measurements at two or more energy spectra, which allows identification and quantification of materials with different compositions [1]. This process, typically known as material decomposition, is the basis for many spectral CT applications, such as calcium removal, virtual non-contrast imaging, kidney stone characterization, and differentiation between gout and pseudo-gout [2-4].

Various approaches to material decomposition have been developed since this concept was first described [5]. Decomposition can be performed by exploiting the knowledge of the fundamental physical interactions for X-ray attenuation, such as the Compton and photoelectric effects [6]. Alternatively, the spectral CT measurements can be decomposed into a series of basis material-specific images [7], such as calcium, iodine and water images. Based on the mechanism used to obtain the spectral CT information, the material decomposition can be performed in either the projection domain or image domain [7, 8]. In this work, we focus on basis material decomposition in the image domain.

The conventional strategy for generating basis material-specific images is based on matrix inversion or least-square fitting between the underlying basis material-specific images and the measured spectral CT images (energy-bin or energy-threshold CT images). However, due to its ill-conditioned nature, the material decomposition problem is intrinsically susceptible to noise amplification. The noise in basis material-specific images can compromise the conspicuity of small objects and hinder the delineation of anatomical regions of interest and associated pathology.

In this work, we describe an image-domain Material Decomposition framework using Spectral Prior Image Constrained Compressed Sensing, abbreviated as MD-SPICCS. Inspired by Prior Image Constrained Compressed Sensing (PICCS) technique used in a variety of CT imaging settings, such as in dynamic CT [9, 10] and for dose reduction [11], the proposed framework exploits the structural redundancy between the individual basis material-specific images and the measured spectral (energy-bin or energy-threshold) CT images to reduce the noise present in the basis material-specific images while preserving structural details and spatial resolution. The presence of this structural redundancy can be appreciated by the fact that the individual basis material-specific images can be expressed as a linear combination of the measured spectral CT images, and that the spectral CT images at different energy spectra also exhibit strong structural similarity [12]. The proposed framework directly addresses material decomposition by formulating this process as a regularized optimization problem. It is applied in the image domain and does not require access to projection data, and therefore can be applied when well-registered multi-energy projection data are not readily available, such as in a conventional dual-source, dual-energy CT.

II. Method

A. Forward Model

For image-based material decomposition, the spectral (energy-bin or energy-threshold) CT images reconstructed from data acquired using different energy spectra are decomposed into a series of basis material-specific images. Denoting $w_i(n)$ as the mass density of the i-th basis material $(1 \le i \le M)$ for the n-th pixel $(1 \le n \le N)$, the effective linear attenuation measurements from a spectral CT acquisition can be expressed as follows:

$$\mu_{eff}(n,e) = \sum_{i=1}^{M} \frac{\mu}{\rho}(e,i)\, w_i(n) \tag{1}$$

where $\mu_{eff}(n,e)$ denotes the effective linear attenuation

S. Tao, K. Rajendran, C. H. McCollough, and S. Leng are with the Department of Radiology, Mayo Clinic, Rochester, MN 55902 USA. Please direct all correspondence to S. Leng (Leng.Shuai@mayo.edu).

coefficient for the e-th energy spectra ($1 \le e \le E$) and the n-th pixel; $\mu/\rho(e,i)$ is the mass attenuation coefficients of the i-th material, and N, M, and E denote the number of pixels, basis materials, and energy spectra for a spectral CT acquisition. Let $\mathbf{M}_0$ be an E by M matrix with $\mathbf{M}_0[e,i] = \mu/\rho(e,i)$. Equation 1 can be reformatted into a linear algebraic form as:

$$\mathbf{f} = \left(\mathbf{M}_0 \otimes \mathbf{I}_N\right)\mathbf{w} = \mathbf{M}\mathbf{w} \tag{2}$$

where $\mathbf{w} = [\mathbf{w}_1^T, \mathbf{w}_2^T, \ldots, \mathbf{w}_M^T]^T$ denotes the basis material-specific image concatenated into a single vector, $\mathbf{f} = [\mathbf{f}_1^T, \mathbf{f}_2^T, \ldots, \mathbf{f}_E^T]^T$ is the vectorized spectral CT images of different energy spectra. $\mathbf{M} = \mathbf{M}_0 \otimes \mathbf{I}_N$ denotes the system matrix that maps basis material-specific images to measured spectral CT images; $\mathbf{I}_N$ is an $N \times N$ identity matrix, and $\otimes$ denotes the Kronecker product.

The standard image-domain basis-material decomposition calculates the basis material-specific image through a least-square fitting process, i.e.:

$$\mathbf{w}_0 = \arg\min_{\mathbf{w}} \left\|\mathbf{f} - \mathbf{M}\mathbf{w}\right\|_2^2 \tag{3}$$

which has a non-iterative solution that can be formulated into matrix inversion and is given by:

$$\mathbf{w}_0 = \left(\mathbf{M}^T\mathbf{M}\right)^{-1}\mathbf{M}^T\mathbf{f} \tag{4}$$

B. MD-SPICCS

The original PICCS was developed as an image reconstruction framework from projection data. It formulates the image reconstruction process as an optimization procedure and incorporates a prior image into the total variation (TV)-based compressed sensing objective function [10]. The regularization functional of the PICCS framework is listed as follows:

$$f_{\text{PICCS}}(\mathbf{x}) = c\left\|\mathbf{x}\right\|_{\text{TV}} + (1-c)\left\|\mathbf{x} - \mathbf{x}_{\text{prior}}\right\|_{\text{TV}} \tag{5}$$

where $\mathbf{x}$ is an image to be reconstructed, $\mathbf{x}_{\text{prior}}$ is its prior image, and $\|\cdot\|_{TV}$ is a sparsifying transform based on TV. Among many applications, PICCS has been shown to be able to reconstruct dynamic CT data series from undersampled projection data [9, 10]. The concept of PICCS has also been extended to the case of spectral CT imaging using the photon-counting detectors (PCD) used in this work. This previous work has successfully demonstrated the use of spectral prior information in the reconstruction of narrow energy bin PCD-CT images from projection data [12]. In this case, the full spectrum, high SNR CT image reconstructed with all the available photons was used as prior image ($\mathbf{x}_{\text{prior}}$). Despite the contrast difference between the narrow energy-bin CT image ($\mathbf{x}$) and the prior image ($\mathbf{x}_{\text{prior}}$), their structural similarity allows successful reconstruction of narrow-bin energy-specific images with high spectral fidelity at reduced noise levels.

Inspired by the original PICCS, we propose an image-domain material decomposition framework that exploits the structural redundancy between the individual basis material-specific images and a spectral CT prior image. Note that, different from the original PICCS algorithm, which involves the reconstruction of anatomical CT image from raw projection data, the proposed image-domain material decomposition is focused on the generation of basis material-specific images using spectral CT images measured with different energy spectra. In the proposed framework, the basis material-specific images are obtained by minimizing the following objective function:

$$\mathbf{w}_0 = \arg\min_{\mathbf{w}}\left(\sum_{i=1}^{M} f_{\text{MD-SPICCS}}^{i}(\mathbf{w}_i)\right) + \lambda\left\|\mathbf{f} - \mathbf{M}\mathbf{w}\right\|_2^2 \quad s.t.\ \ \mathbf{w}_i = \mathbf{P}_i\mathbf{w} \tag{6}$$

with

$$f_{\text{MD-SPICCS}}^{i}(\mathbf{w}_i) = c\left\|\mathbf{w}_i\right\|_{\text{TV}} + (1-c)\left\|\mathbf{w}_i - \Gamma_i\left(\mathbf{f}_{\text{prior}}\right)\right\|_{\text{TV}} \tag{7}$$

where $\mathbf{P}_i$ denotes a binary diagonal matrix selecting the i-th basis material-specific image ($\mathbf{w}_i$) from $\mathbf{w}$, i.e., $\mathbf{w} = \mathbf{P}_i\,\mathbf{w}$; $f_{\text{MDSPICCS}}^{i}(\mathbf{w}_i)$ denotes the regularization term directly applied to the i-th basis material-specific image ($\mathbf{w}_i$); $\lambda \ge 0$ is a parameter that controls the relative preference for the regularization term and the data fidelity term, $\|\mathbf{f} - \mathbf{M}\mathbf{w}\|_2^2$.

The regularization term $f_{MD-SPICCS}^{i}(\mathbf{w}_i)$ consists of two components. The $\|\mathbf{w}_i\|_{TV}$ term directly minimizes the TV of each material-specific image, while the $\left\|\mathbf{w}_i - \Gamma_i(\mathbf{f}_{prior})\right\|_{TV}$ term applies a sparsifying transform to the difference between the basis material-specific image and its corresponding basis material prior image ($\Gamma_i(\mathbf{f}_{prior})$) that is obtained from applying a transformation $\Gamma_i(\cdot)$ to a spectral CT prior image ($\mathbf{f}_{prior}$). The algorithmic parameter, c ($0 \le c \le 1$) controls the relative weight between the two components in the regularization. By including $\left\|\mathbf{w}_i - \Gamma_i(\mathbf{f}_{prior})\right\|_{TV}$ as a regularization term, it is assumed that the individual basis material-specific image $\mathbf{w}_i$ and the basis material prior image $\Gamma_i(\mathbf{f}_{prior})$ share structural redundancy, so that after removing the contribution of each basis material ($\mathbf{w}_i$), the residual image is still sparse after TV transform.

Ideally, the measured spectral CT prior image $\mathbf{f}_{prior}$ should be a high SNR image. On a PCD system, the CT image reconstructed using all the available photons of different energy spectra typically has the highest SNR, which can be used for $\mathbf{f}_{\text{prior}}$. On a conventional energy-integrating-detector (EID) based system, such as the dual-source, dual-energy CT system, the individual image associated with each source, or the "mixed" image which is a weighted combination of the two source images, can be used as the spectral CT prior image ($\mathbf{f}_{\text{prior}}$). The $\Gamma_i(\mathbf{f}_{\text{prior}})$ is chosen as a linear transformation and is a function of material type:

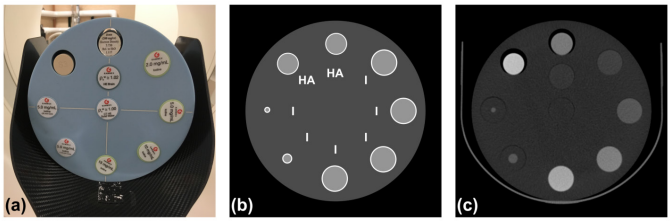


Fig. 1. (a) Multi-energy CT phantom on the head holder; (b) phantom inserts configuration. There are six iodine (I) inserts (four large inserts with 2, 5, 10, 15 mg/mL iodine, and two smaller 5 mg/mL inserts with a diameter of 5/10 mm) and two hydroxyapatite (HA) inserts (200, 400 mg/mL) contained within the phantom base made from solid water material; (c) an example of the spectral CT image acquired using a research whole-body PCD-CT system [14].

$$\Gamma_i\left(\mathbf{f}_{\text{prior}}\right)=\frac{\mathbf{f}_{\text{prior}}}{\frac{\mu}{\rho}(\text{prior},i)} \tag{8}$$

where $\mu/\rho(\text{prior},i)$ is a scaling factor that converts the spectral CT prior image (i.e., $\mathbf{f}_{prior}$, in CT number or linear attenuation coefficients) to a basis material prior image (i.e., $\Gamma_i(\mathbf{f}_{prior})$, in mass fraction) that is on the same scale of the basis material-specific image for the i-th basis material type ($\mathbf{w}_i$). The physical meaning of $\mu/\rho(\text{prior},i)$ is similar to the empirical mass attenuation coefficient ($\mu/\rho(e,i)$) in Eq. 1, and can be obtained from the same calibration scan using data from the corresponding energy spectrum.

C. Experimental Methods

To test the proposed MD-SPICCS framework, the head section of a multi-energy calibration phantom (Gammex Inc., Middleton, WI) was scanned on a research PCD CT system [13] using two energy thresholds. The phantom setup is shown in Fig. 1. The head section of the phantom measures 20.0 cm in diameter and 16.5 cm in depth, and is made from solid water material. The iodine inserts from the Gammex phantom with concentrations of 2, 5, 10, and 15 mg/mL were scanned, together with hydroxyapatite (HA) inserts with concentrations of 200 and 400 mg/mL. Gammex phantom inserts with an iodine core (5 mg/mL) of 5 and 10 mm diameters were also included to evaluate the sensitivity of algorithm within a small region.

The PCD system was designed based on the same platform of the second generation dual source, dual energy CT system (Siemens Healthcare GmbH, Forchheim, Germany). This system consists of two X-ray tube-detector pairs, with one tube coupled to a conventional EID-based subsystem and the other to a PCD subsystem. The EID and PCD subsystems have a scan field-of-view (FOV) of 500 and 275 mm, respectively, and a longitudinal coverage of 38.4 mm (64×0.6 mm detector rows) and 16 mm (32×0.5 mm detector rows). Currently, two scan protocols are available for the PCD subsystem, which are designed for head and body scans, respectively. With the body protocol, an additional beam filter is used.

The phantom scans were performed using the head protocol and sequential acquisition at 120 kV and 110 mAs. The two energy thresholds were set at 30 and 52 keV. The energy bin

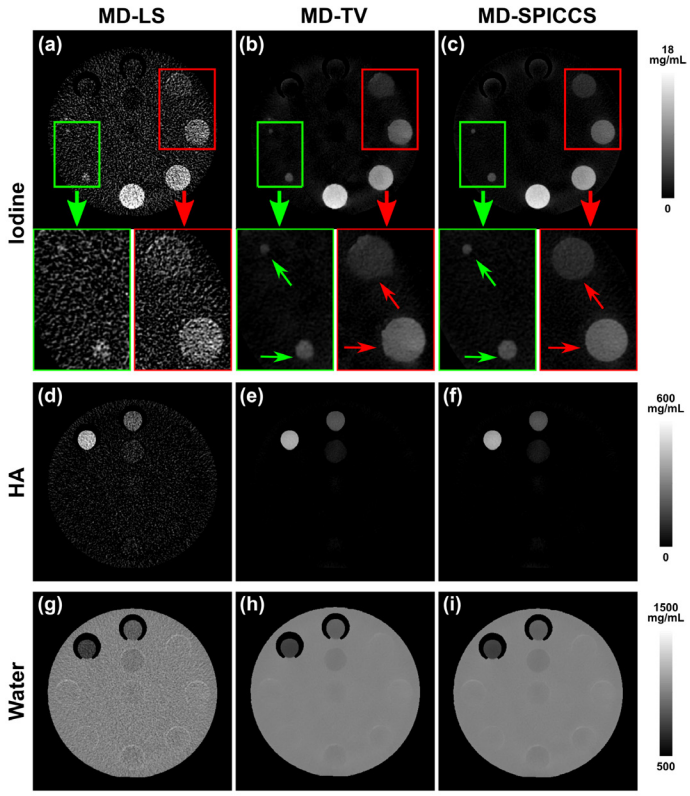


Fig. 2. Iodine (a-c), hydroxyapatite (HA, d-f), and water-specific images (g-i) calculated from the phantom data set using the standard, least-squares based material decomposition (MD-LS, a,d,g), the total variation (TV) regularized material decomposition (MD-TV, b,e,h), and the proposed MD-SPICCS framework with c = 0.6 (c,f,i). The magnified insets in a-c show the low-contrast iodine inserts with a concentration of 2 mg/mL (red inset, upper insert) and 5 (red inset, lower insert) mg/mL from the iodine-specific image, as well as that of the small iodine inserts (5 mg/mL) with a diameter of 5 mm (green inset, upper insert) and 10 mm (green inset, lower insert).

images (30-52 and 52-120 keV) and low-energy threshold images (30-120 keV) were reconstructed using a medium-smooth quantitative kernel (D30) onto a 275 mm FOV and 512×512 matrix size with image thickness of 3 mm. The mass attenuation coefficients for each basis material (i.e., $\mu/\rho(e,i)$ in Eq. 1) were measured from a separate calibration scan using materials of known concentrations.

D. Data Processing

Material decomposition was performed assuming three basis materials (hydroxyapatite, iodine, water) with a volume conservation constraint using the energy bin images. The optimization problem shown in Equations 6 and 7 was solved iteratively using the alternating direction method of multipliers (ADMM) routine [14] with 500 iterations, which was found to be sufficient for the algorithm to converge. The high SNR low-energy threshold image (30-120 keV) reconstructed using all available photons was used as the prior image ($\mathbf{f}_{\text{prior}}$). The parameter c was set as 0.6, which yields reasonable denoising performance. As a comparison, the standard least-squares based material decomposition, and the TV-regularzied material decomposition (c = 1 in Eq. 7) were also performed. The same $\lambda = 1 \times 10^3$ (empirically determined) was used for both cases.

Noise reduction algorithms, especially TV-based algorithm can change image texture and create patchy looking images.

To alleviate this effect, a linear blending technique to combine the basis material-specific images from MD-SPICCS and the standard decomposition was adopted to maintain the original image texture while reducing noise. The final ouput of MD-SPICCS ($\mathbf{w}_{\text{output}}$) is then experessed as:

$$\mathbf{w}_{\text{output}} = \alpha \mathbf{w}_{\text{MD-SPICCS}} + (1-\alpha)\mathbf{w}_{\text{LS}} \tag{9}$$

where α is the blending factor and was chosen as 0.9, $\mathbf{w}_{\text{MD-SPICCS}}$ and $\mathbf{w}_{\text{LS}}$ denote the basis material-specific images obtained by solving Eqs. 6 and 7, and by using the standard least-squares (LS) based approach in Eq. 4, respectively.

III. Results

The basis material-specific (iodine, HA, water) images obtained using the standard least-squares based material decomposition (MD-LS), the TV regularized material decomposition (MD-TV), as well as the proposed MD-SPICCS algorithm with c = 0.6 are shown in Fig. 2. Due to the intrinsic noise amplification effect of the material decomposition process, the results of MD-LS contain strong noises in the material specific images, which compromise the conspicuity of phantom inserts, especially for the low-concentration iodine inserts, as well as the inserts with smaller diameters. The use of TV regularization can substantially reduce the noise in the individual material-specific images (a,d,g vs. b,e,h in Fig. 2). However, it also introduces undesirable image blurring while denoising. Figure 2 also shows the magnified insets of the low-contrast iodine inserts with a concentration of 2 mg/mL (red panel, upper insert) and 5 mg/mL (red panel, lower insert) generated using different methods, as well as that of the small iodine inserts (5 mg/mL) with a diameter of 5 mm (green panel, upper insert) and 10 mm (green panel, lower insert). Note the distorted and blurred boundaries around the low-concentration inserts, as well as the reduced iodine contrast for small iodine inserts using MD-TV (see arrows in magnified insets). The basis material-specific images generated using the proposed MD-SPICCS are also shown in Fig. 2. Compared to MD-TV, MD-SPICCS achieved similar denoising performance for all basis materials. Table 1 summarizes the root-mean-square-error (RMSE) measured based on all iodine and HA inserts, which demonstrate that MD-SPICCS can achieve a similar level of overall quantitative accuracy and denoising performance as MD-TV while better preserving the structural integrity of the low-concentration or small-diameter inserts, as exemplified by the sharper boundaries of iodine inserts (see arrows in magnified insets).

TABLE I
RMSE MEASUREMENTS

	MD-LS	MD-TV	MD-SPICCS
Iodine (mg/mL)	3.9	1.0	1.0
HA (mg/mL)	80.8	20.0	20.5
Note	-	Blurred boundary around the low-concentration iodine inserts; reduced contrast for small iodine inserts	Improved structural integrity of the low-concentration or small-diameter iodine inserts

IV. Conclusion

In this work, a material decomposition framework using spectral prior image constraint compressed sensing was developed. The proposed framework exploits the structural redundancy between the individual basis material-specific images and the measured source spectral CT images, and can reduce iodine quantification RMSE (from 3.9 to 1.0 mg/mL) while preserving structural integrity and low-contrast details in the basis material-specific images.

References

[1] C. H. McCollough, S. A. Leng, L. F. Yu, and J. G. Fletcher, "Dual- and Multi-Energy CT: Principles, Technical Approaches, and Clinical Applications," *Radiology,* vol. 276, pp. 637-653, Sep 2015.

[2] D. Morhard, C. Fink, A. Graser, M. F. Reiser, C. Becker, and T. R. C. Johnson, "Cervical and Cranial Computed Tomographic Angiography With Automated Bone Removal Dual Energy Computed Tomography Versus Standard Computed Tomography," *Investigative Radiology,* vol. 44, pp. 293-297, May 2009.

[3] A. Korn, B. Bender, C. Thomas, S. Danz, M. Fenchel, T. Nagele*, et al.*, "Dual energy CTA of the carotid bifurcation: Advantage of plaque subtraction for assessment of grade of the stenosis and morphology," *European Journal of Radiology,* vol. 80, pp. E120-E125, Nov 2011.

[4] J. Ferda, M. Novak, H. Mirka, J. Baxa, E. Ferdova, A. Bednarova*, et al.*, "The assessment of intracranial bleeding with virtual unenhanced imaging by means of dual-energy CT angiography," *European Radiology,* vol. 19, pp. 2518-2522, Oct 2009.

[5] G. N. Hounsfield, "Computerized transverse axial scanning (tomography). I. Description of system," *Brisith Journal Radiology,* vol. 46, pp. 1016-1022, 1973.

[6] R. E. Alvarez and A. Macovski, "Energy-Selective Reconstructions in X-Ray Computerized Tomography," *Physics in Medicine and Biology,* vol. 21, pp. 733-744, 1976.

[7] X. Liu, L. F. Yu, A. N. Primak, and C. H. McCollough, "Quantitative imaging of element composition and mass fraction using dual-energy CT: Three-material decomposition," *Medical Physics,* vol. 36, pp. 1602-1609, May 2009.

[8] W. A. Kalender, W. H. Perman, J. R. Vetter, and E. Klotz, "Evaluation of a Prototype Dual-Energy Computed Tomographic Apparatus .1. Phantom Studies," *Medical Physics,* vol. 13, pp. 334-339, May-Jun 1986.

[9] S. Leng, J. Tang, J. Zambelli, B. Nett, R. Tolakanahalli, and G. H. Chen, "High temporal resolution and streak-free four-dimensional cone-beam computed tomography," *Physics in Medicine and Biology,* vol. 53, pp. 5653-5673, Oct 21 2008.

[10] G. H. Chen, J. Tang, and S. Leng, "Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets," *Medical Physics,* vol. 35, pp. 660-663, Feb 2008.

[11] Z. Qi, S. Huang, B. Nett, J. Tang, K. Yang, J. Boone*, et al.*, "Dramatic Noise Reduction and Potential Radiation Dose Reduction in Breast Cone-Beam CT Imaging Using Prior Image Constrained Compressed Sensing (PICCS)," *Medical Physics,* vol. 37, Jun 2010.

[12] Z. C. Yu, S. Leng, Z. B. Li, and C. H. McCollough, "Spectral prior image constrained compressed sensing (spectral PICCS) for photon-counting computed tomography," *Physics in Medicine and Biology,* vol. 61, pp. 6707-6732, Sep 21 2016.

[13] Z. C. Yu, S. Leng, S. M. Jorgensen, Z. B. Li, R. Gutjahr, B. Y. Chen*, et al.*, "Evaluation of conventional imaging performance in a research whole-body CT system with a photon-counting detector array," *Physics in Medicine and Biology,* vol. 61, pp. 1572-1595, Feb 21 2016.

[14] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed Optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning,* vol. 3, pp. 1-122, 2010.

Three material decomposition for spectral computed tomography enabled by block-diagonal step-preconditioning

Emil Y. Sidky[1], Rina Foygel Barber[2], Taly Gilat-Schmidt[3], and Xiaochuan Pan[1]

***Abstract*—A potential application for spectral computed tomography (CT) with multi-energy-window photon-counting detectors is quantitative medical imaging with K-edge contrast agents [1]. Image reconstruction for spectral CT with such contrast agents necessitates expression of the X-ray linear attenuation map in at least three expansion functions, for example, bone/water/K-edge-material or photo-electric-process/Compton-process/K-edge-material. The use of three expansion functions can result in slow convergence for iterative image reconstruction (IIR) algorithms applied to spectral CT. We propose a block-diagonal step-preconditioner for use with a primal-dual iterative image reconstruction framework that we have been developing for spectral CT. We demonstrate the advantage of the new step-preconditioner on a sensitive spectral CT simulation where the test object has low concentration of Gadolinium (Gd) contrast agent and the X-ray attenuation map is represented by three materials - PMMA, a soft-tissue equivalent, Aluminum, a bone equivalent, and Gd.**

I. Introduction

We have been developing a general algorithm framework for one-step spectral CT image reconstruction (OSSCIR) that we have applied to experimental data acquired employing a spectral CT system with photon-counting detectors [2]. The OSSCIR algorithm framework involves direct one-step image reconstruction of basis material maps from energy-windowed X-ray transmission data. The one-step approach contrasts with standard two-step processing where the photon transmission data is converted to material sinograms followed by image reconstruction to material maps [1]. The one-step approach enables unconventional scan configurations where the transmission rays need not be co-registered for all energy-windows [3], and the image reconstruction process can be regularized by applying constraints directly to the material maps. Implementing OSSCIR consists of: (1) specifying the material maps with an optimization problem that includes a nonconvex data discrepancy term with convex constraints, and (2) solution of the nonconvex optimization problem by the mirrored convex/concave (MOCCA) algorithm [4,5].

MOCCA is the heart of the OSSCIR framework. It is an extension of the Chambolle-Pock primal-dual (CPPD) algorithm for large-scale convex optimization [6,7]. The MOCCA extension applies to certain forms of large-scale nonconvex optimization composed of a smooth nonconvex objective function and convex nonsmooth functions, such as convex constraints. The design of MOCCA is based on the idea that for some classes of nonconvex smooth objective functions the difficulty for algorithm design results from local saddle points and not local minima. Local saddle points have directions of negative curvature that can result in spurious update steps. Accordingly, a MOCCA iteration consists of constructing a local convex quadratic approximation to the objective function, removing directions of negative curvature, and performing a CPPD step on this approximation.

An important aspect of MOCCA is the diagonal step-preconditioner (SPC) for CPPD proposed by Pock and Chambolle [8]. Because the convex approximation to the objective function is changing at every iteration, the CPPD step length parameters need to be recomputed at every iteration. The step lengths of diagonal-SPC CPPD Ref. [8] can be computed at the cost of two additional matrix-vector product operations, which is equivalent to an additional forward- and back-projection per iteration for CT IIR.

In this contribution, we extend diagonal SPC to block-diagonal SPC that effectively counteracts slow convergence due to the near linear dependence from the basis material attenuation curves. In our original work on spectral CT IIR, we had already encountered slow convergence rates with two-material expansion of the attenuation map, and in that work we proposed μ-preconditioning (μ-PC), where the materials expansion set is transformed to an orthogonal set of functions in X-ray energy. The μ-PC transformation was effective at improving convergence rates.

In attacking three-materials expansion sets, μ-PC also improves convergence, but in this case the convergence issue is more acute than the two-materials case. In our original application of MOCCA to spectral CT in Ref. [5], we successfully demonstrated one-step reconstruction for three

[1]The University of Chicago, Department of Radiology MC-2026, 5841 S. Maryland Avenue, Chicago IL, 60637.

[2]The University of Chicago, Department of Statistics, 5734 S. University Avenue, Chicago IL, 60637.

[3]Marquette University, Department of Biomedical Engineering, PO Box 1881, Milwaukee WI, 53201.

materials, but the simulation modeled five ideal photon-counting spectral response windows with sharp boundaries and no window overlap. The three-material simulation we consider here involves only four windows with realistic spectral responses that have significant overlap with each other. Accordingly, the worse conditioning of the realistic setup can impact convergence. We propose a block-diagonal SPC that has slightly more computational overhead per iteration but dramatically improves convergence of MOCCA in the spectral CT setting with three basis materials and realistic spectral responses.

We briefly summarize OSSCIR and MOCCA with μ-preconditioning; and introduce the new block-diagonal preconditioner in Sec. II. The improvement in convergence gained by the new preconditioner is demonstrated in Sec. III on a challenging, idealized spectral CT simulation.

II. Methods

As in Ref. [5], the spectral CT data model is written

$$I_{w,\ell} = \int S_{w,\ell}(E) \exp\left[-\int_\ell \mu(E,\vec{r}(t))dt\right] dE, \quad (1)$$

where $I_{w,\ell}$ is the transmitted X-ray photon fluence along ray ℓ in energy window w; t is a parameter indicating location along ℓ); $S_{w,\ell}(E)$ is the spectral response; and $\mu(E,\vec{r}(t))$ is the energy and spatially dependent linear X-ray attenuation coefficient.

We employ a standard material-expansion decomposition to model the attenuation map

$$\mu(E,\vec{r}(t)) = \sum_m \left(\frac{\mu_m(E)}{\rho_m}\right) \rho_m f_m(\vec{r}[t]), \quad (2)$$

where ρ_m is the density of material m; $\mu_m(E)/\rho_m$ is the mass attenuation coefficient of material m; and $f_m(\vec{r})$ is the spatial map for material m.

To obtain the final discrete data model, we combine Eq. (1) with Eq. (2); normalize the spectral response; and discretize all integrations. The detected counts model becomes

$$\hat{c}_{w,\ell}(f) = N_{w,\ell} \sum_i s_{w,\ell,i} \exp\left(-\sum_{m,k} \mu_{m,i} X_{\ell,k} f_{k,m}\right),$$

where $N_{w,\ell}$ is the total number of incident photons along ray ℓ in energy window w; $s_{w,\ell,i}$ is the normalized spectral response, i.e. $\sum_i s_{w,\ell,i} = 1$; i indexes the energy E_i; $X_{\ell,k}$ represents X-ray projection along the ray ℓ); and $f_{k,m}$ is the pixelized material map with k and m indexing pixel and expansion-material, respectively. The spectral responses are assumed known, and the goal is to reconstruct the material maps f from measured counts data c.

Transmission Poisson likelihood maximization: Maximizing the transmission Poisson likelihood is equivalent to minimizing the Kullback-Leibler distance between the counts data, c, and counts model, $\hat{c}(f)$,

$$D_{\text{TPL}}(c,\hat{c}(f)) = \sum_{w,\ell} \left[\hat{c}_{w,\ell}(f) - c_{w,\ell} - c_{w,\ell} \log \frac{\hat{c}_{w,\ell}(f)}{c_{w,\ell}}\right], \quad (3)$$

where $c_{w,\ell}$ are the measured counts in energy window w along ray ℓ. This objective function is nonconvex as can be verified by computing the Hessian (the multivariable second derivative) of $D_{\text{TPL}}(c,\hat{c}(f))$ with respect to f. The nonlinearity of $\hat{c}(f)$ as a function of f gives rise to directions of negative curvature in $D_{\text{TPL}}(c,\hat{c}(f))$.

The MOCCA algorithm is designed to minimize the nonconvex $D_{\text{TPL}}(c,\hat{c}(f))$ objective function and the pseudo-code for doing so is given in Eqs. (47)-(52) in Ref. [5]. The algorithm results from making a local convex quadratic approximation to Eq. (3). In the neighborhood of an expansion point f_0, we approximate $D_{\text{TPL}}(c,\hat{c}(f))$ with

$$D_{\text{TPL}}(c,\hat{c}(f)) \approx Q(K(f_0)f),$$

where the precise form of the quadratic function Q is specified in Ref. [5]. The matrix $K(f)$ comes from the gradient of $D_{\text{TPL}}(c,\hat{c}(f))$, expressed here in component form:

$$\frac{\partial D_{\text{TPL}}(c,\hat{c}(f))}{\partial f_{k,m}} = \sum_{w,\ell} (c_{w,\ell} - \hat{c}_{w,\ell}(f)) K_{w\ell,km}(f).$$

The rows of $K(f)$ index the data space consisting of energy windows, w, and rays, ℓ, and the columns index the image space consisting of materials, m, and pixels, k.

Step lengths of MOCCA and μ-PC: The MOCCA algorithm is primal-dual as it is based on the diagonal-SPC CPPD. Following Refs. [5,8], the step lengths for the dual and primal updates are

$$\Sigma_{w\ell} = \frac{1/\lambda}{\sum_{m,k} |K_{w\ell,mk}(f_0)|}, \quad T_{mk} = \frac{\lambda}{\sum_{w,\ell} |K_{w\ell,mk}(f_0)|},$$

respectively, and λ is a step size ratio parameter that must be tuned. In our previous work (Ref. [5]), we found that faster convergence can be obtained by applying mu-PC to the materials basis, which transforms it to an orthogonal basis; in this new formulation of the optimization problem, the step lengths are computed the same way as before by substituting the new matrix $K(f_0)$ calculated in this transformed basis.

A m-block diagonal SPC for MOCCA applied to spectral CT: The condition on Σ and T that leads to convergence for SPC CPPD is that the matrix

$$B = \begin{pmatrix} T^{-1} & -K^\top \\ -K & \Sigma^{-1} \end{pmatrix}$$

is positive semi-definite, i.e. $v^\top B v \geq 0$ for any vector v. In designing step-matrices Σ and T for MOCCA, we respect

the constraint imposed by positive definiteness of B with $K(f_0)$ changing at each iteration.

We propose a m-block diagonal SPC for Σ and T that is motivated by preserving invariance to rotations of the materials expansion set; in other words, the output of the algorithm would be identical regardless of any rotation applied to the selected basis of materials, which is a natural property that is not satisfied by the μ-PC method. In the process of developing μ-PC we had noticed sensitive convergence behavior simply by performing such rotations. This sensitivity was traced to the diagonal PC strategy for σ and τ. The proposed step matrices are

$$\left(\Sigma^{-1}\right)_{w\ell,w'\ell'} = \frac{1}{\lambda}\sum_k \sqrt{\sum_m K^2_{w\ell,mk}(f_0)}\ \mathbf{I}_{w\ell,w'\ell'}$$

for the dual step and

$$\left(T^{-1}\right)_{mk,m'k'} = \lambda \sum_{w,\ell} \frac{K_{w\ell,mk}(f_0)K_{w\ell,m'k}(f_0)}{\sqrt{\sum_{m''} K^2_{w\ell,m''k}(f_0)}}\mathbf{I}_{k,k'},$$

for the primal step. As before, the Σ^{-1} matrix is diagonal, and inverting to find Σ only involves computing the reciprocal of the diagonal elements. The new definition of T^{-1}, however, is diagonal only in k, k' and each diagonal element indexed by k consists of an $m \times m$ block. Inversion to find Σ thus involves inversion of an $m \times m$ matrix where each entry is a N_k-length vector, where N_k is the total number of pixels in a single material map. The inversion of such an $m \times m$ matrix is feasible, because the number of expansion materials is low. In this work in fact we use $N_m = 3$. The matrix inversion must be computed at every iteration because $K(f_0)$ is a function of the expansion center, which changes at every iteration for our application of MOCCA. The overhead in inverting the 3x3 blocks is negligible in comparison with the computationally intensive X-ray forward- and back-projections.

III. Results

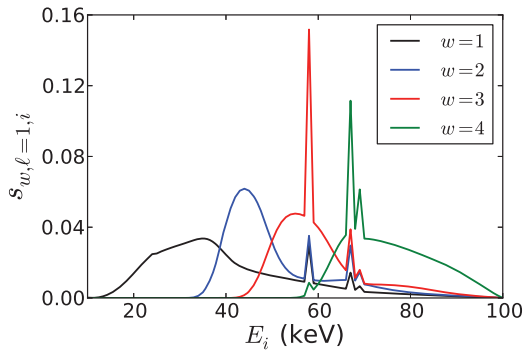


Fig. 1. Realistic X-ray normalized spectral response curves for 4-window spectral CT with a photon-counting detector. Shown is the response curves for the first detector pixel; other pixels have slight variations from these curves.

Spectral CT counts data are generated based on a simulation of our bench-top X-ray system including a photon-counting detector with 192 pixels. Mean transmitted photon counts acquired in four energy windows are computed based on spectra generated from calibration of our system. The precise spectra vary as a function of detector pixel, and example spectra are shown in Fig. 1. For the spectral CT data, 200 projections are generated from a phantom simulation of one of our physical test objects: a 6.35cm-diameter Poly(methyl methacrylate) (PMMA) cylinder with four inserted rods including, PMMA Air (empty), Teflon, and low-density polyethylene (LDPE) inserts. In the empty insert, Gd contrast agent is included at a density fraction of 0.003 (Note this is only possible in simulation). An Aluminum/PMMA/Gd materials expansion set is used form image reconstruction, and the corresponding material maps of the phantom are shown in Fig. 2.

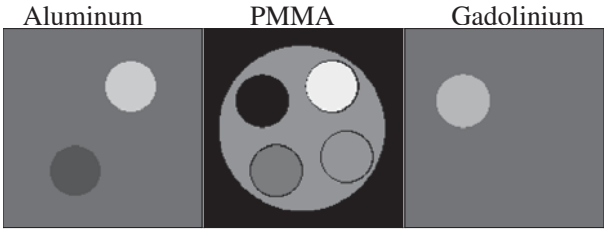


Fig. 2. Rods phantom decomposed into Aluminum, PMMA, and Gd maps. The structure of the phantom is most easily visible in the PMMA map, where the PMMA background cylinder is clearly visible. The rods, clockwise from the upper left are: Gd at a density fraction of 0.003, Teflon, PMMA, and LDPE. The Gd "rod" is only visible in the Gd map. The display windows are [-0.1,0.2], [0.5,1.5], and [-0.003,0.006] for Aluminum, PMMA, and Gd maps, respectively.

The test data are the noiseless mean counts, and the goal of this "inverse crime" set up is to characterize MOCCA convergence for μ-PC and m-block diagonal SPC by observing the accurate recovery of the test object. The difficulty of the problem lies in the fact that we employ realistic spectra that include non-flux-dependent physical factors that blur the sharp energy-window borders. The blurred spectra have realistic overlap with each other as opposed to ideal spectral responses with no overlap.

In Fig. 3, we display the D_{TPL} data discrepancy as a function of iteration number for both PC strategies. [1] In each case the λ parameter is tuned for most rapid convergence in this quantity. Both versions of MOCCA are run for 2,000 iterations and in this example it is clear that m-block diagonal SPC outperforms μ-PC. Not shown is the result for MOCCA with diagonal SPC, which exhibits divergent behavior for all tested λ values. Divergent behavior can

[1] In performing these calculations, we found it useful to employ a soft-exponential (sexp) function parameterized with a cut-off c: $\text{sexp}(c;x) = \exp(x)$ for $x \leq c$, and $\text{sexp}(c;x) = \exp(c) + (x-c)\exp(c)$ for $x > c$. Positive values of x in the spectral CT model are unphysical because they correspond to an increase in X-ray intensity along the transmission ray. Yet large positive values of x may occur at early iterations, potentially causing an overflow. We use $c = 10$ in the present results.

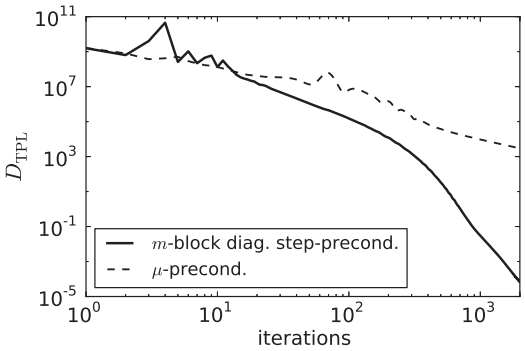


Fig. 3. The log-log plot shows convergence of $D_{\mathrm{TPL}}(c, \hat{c}(f^{(n)}))$, where $f^{(n)}$ is the material map estimates at iteration n. The curves show results for MOCCA with μ-PC and with m-block diagonal SPC.

occur with MOCCA, when only a single "inner loop" is performed [4,5]. Due to efficiency constraints, we aim to operate MOCCA with parameter and preconditioning choices that allow its operation without nested inner and outer loops.

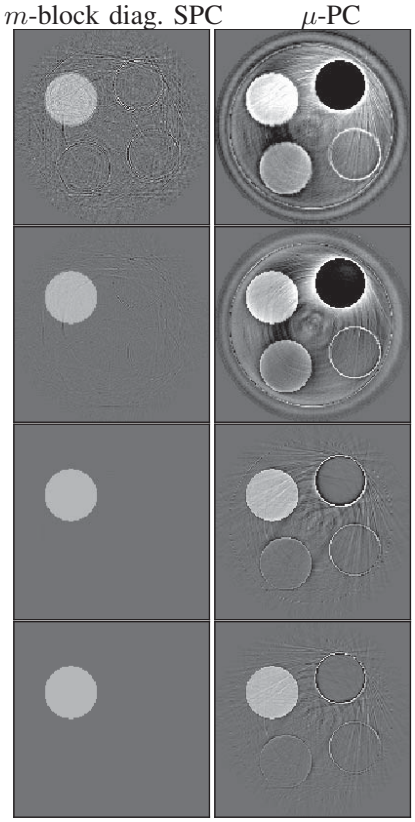


Fig. 4. Gd material maps at various iteration numbers for MOCCA with the new m-block diagonal SPC and with μ-PC. From top to bottom the iteration numbers are: 100, 200, 1000, and 2000. The display window is [-0.003,0.006] for all panels.

Of particular interest for convergence studies, in this case, is the Gd material map. It has such low density that lack of convergence is obvious in visualizing the corresponding images. In Fig. 4, we display a series of intermediate estimates of the Gd map for both preconditioning methods. Of particular interest is the fact that at 100 iterations the proposed m-block method has little contamination from the PMMA and aluminum maps, while μ-PC shows significant bleed-through from the other expansion materials at 100 and 200 iterations. From the images series it is also clear that the m-block method achieves accurate Gd recovery much earlier than μ-PC. We also note that the artifact patterns are rather complex at intermediate iterations; this results from the variations of spectral response across detector pixels.

IV. Summary

We propose a new m-block diagonal step-preconditioner for use with MOCCA applied to spectral CT. In these preliminary convergence studies we have primarily been concerned with K-edge imaging with the use of a three-material expansion set: a soft-tissue equivalent, a bone equivalent, and Gd contrast agent. In this setting, the new preconditioner enables MOCCA to be applied effectively for one-step reconstruction of three three material maps from four-window photon-counting data with realistic spectral responses. At the conference, we will also present experimental results on our K-edge imaging phantom using MOCCA with m-block diagonal step-preconditioning.

V. Acknowledgment

RFB is supported by an Alfred P. Sloan Fellowship and by NSF award DMS-1654076. This work is also supported in part by NIH Grant Nos. R01-EB018102, and R01-CA182264. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

References

[1] J. P. Schlomka, E. Roessl, R. Dorscheid, S. Dill, G. Martens, T. Istel, C. Bäumer, C. Herrmann, R. Steadman, G G. Zeitler, A. Livne, and R. Proksa, "Experimental feasibility of multi-energy photon-counting K-edge imaging in pre-clinical computed tomography," *Phys. Med. Biol.*, vol. 53, no. 15, pp. 4031–4048, 2008.

[2] T. G. Schmidt, R. F. Barber, and E. Y. Sidky, "A spectral ct method to directly estimate basis material maps from experimental photon-counting data," *IEEE Trans. Med. Imag.*, pp. 1808–1819, 2017.

[3] B. Chen, Z. Zhang, E. Y. Sidky, D. Xia, and X. Pan, "Image reconstruction and scan configurations enabled by optimization-based algorithms in multispectral CT," *Phys. Med. Biol.*, vol. 62, pp. 8763–8793, 2017.

[4] R. F. Barber and E. Y. Sidky, "MOCCA: mirrored convex/concave optimization for nonconvex composite functions," 2015, http://arxiv.org/abs/1510.08842.

[5] R. F. Barber, E. Y. Sidky, T. Gilat-Schmidt, and X. Pan, "An algorithm for constrained one-step inversion of spectral CT data," *Phys. Med. Biol.*, vol. 61, pp. 3784–3818, 2015.

[6] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Math. Imag. Vis.*, vol. 40, pp. 120–145, 2011.

[7] E. Y. Sidky, J. H. Jørgensen, and X. Pan, "Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm," *Phys. Med. Biol.*, vol. 57, pp. 3065–3091, 2012.

[8] T. Pock and A. Chambolle, "Diagonal preconditioning for first order primal-dual algorithms in convex optimization," in *2011 IEEE International Conference on Computer Vision (ICCV)*, 2011, pp. 1762–1769.

Digital charge summing for photon counting detectors

Scott S. Hsieh

***Abstract*—Photon counting detectors are an emerging technology for next-generation CT architectures but suffer from a variety of limitations that limit their performance. One of the most significant obstacles today is the phenomenon of charge sharing. Charge sharing can be adequately corrected using charge summing circuitry, but existing approaches to charge summing are complex to implement, and several prototype detectors being investigated for diagnostic CT have not implemented charge summing. The purpose of this work is to evaluate a variation of charge summing, digital charge summing, that operates only after digitization. Because it uses only digital logic, it may be simpler and more practical to implement. We evaluate digital charge summing in simulations to show that a substantial performance improvement may be possible in basis material imaging.**

***Index Terms*—photon counting detectors, charge sharing**

I. Introduction

PHOTON counting technology has seen increasing interest within the CT community (1). Unlike traditional, energy-integrating detectors, single x-ray photon counting detectors are able to discriminate individual photons as they arrive on the detector and, in many cases, to determine their approximate energy. Photon counting detectors have been on the rise for many years, but only recently has the count rate of these detectors been sufficient for possible clinical use (2). The flux in diagnostic CT can reach approximately 1000 million photons (counts) per second per square millimeter (Mcps/mm^2). However, the flux that actually reaches the detector in most applications is typically much smaller due to the attenuation of the patient. Also, it is known that the characteristic (or maximum) count rate of the detector needs to be several times greater than the actual flux for the detector to achieve good performance (3), although different forms of detection logic could theoretically alleviate this performance penalty (4). When the flux approaches or exceeds the characteristic count rate, loss of counts and pileup ensues, distorting the spectrum and causing, for example, multiple low energy photons arriving in close spatiotemporal proximity to be perceived instead as a single high energy photon.

To alleviate this non-ideality, recent photon counting detector prototypes have taken the approach of building detectors with a small pixel pitch. In one prototype, the detector pixel size for the photon counting detector is approximately 0.05 mm^2, compared to 1 mm^2 for a traditional energy-integrating detector pixel (5). One effect of this small pixel size is the potential for dramatically increased spatial resolution (6). Other approaches have been proposed for reducing the count rate requirements (2, 7, 8), but a strength of a small pixel size is that it is a hardware solution that is simple and provides resilience to high count rates. The negative impact of the small pixel size, however, is charge sharing. In charge sharing, a single high-energy photon will produce charge that spills over and is detected in multiple pixels. Charge sharing has a dramatic effect on spectral performance.

One potent solution for charge sharing is charge summing technology, found in some detector prototypes such as Medipix3 (9-12). In the long term, this form of charge summing could greatly improve the spectroscopic performance of photon counting detectors. However, the technical challenges of incorporating this type of inter-pixel communication are significant because it requires addition of limited charge in the analog domain. The purpose of this work is to examine an alternative, with inter-pixel communication in the digital domain, or "digital charge summing."

II. Methods

Traditional charge summing circuitry sums up the charge in adjacent pixels (e.g., a 2x2 block of pixels) and uses logic to determine which block of pixels is an appropriate maximum that should be sent to the comparator for digitization. In digital charge summing, logic is applied only after digitization via the comparator.

Charge sharing creates several types of pathological detection pathways. Ideally, for a two-bin detector, every low energy photon is recorded in the low bin, and every high energy photon is recorded in the high bin. Because of charge sharing, low energy photons may spread their energy between multiple pixels and hence be undetected. High energy photons may spread their energy and hence be detected as a single low energy photon. Alternatively, a high energy photon may spread its energy between multiple pixels and be detected as two low-energy photons. Digital charge summing is only able to remedy this final pathway, whereas analog charge summing can improve characteristics of all three of these pathways. Figure 1 depicts digital charge summing, showing a case where it provides an improvement and a case where it does not provide

Scott S. Hsieh is with the University of California, Los Angeles.

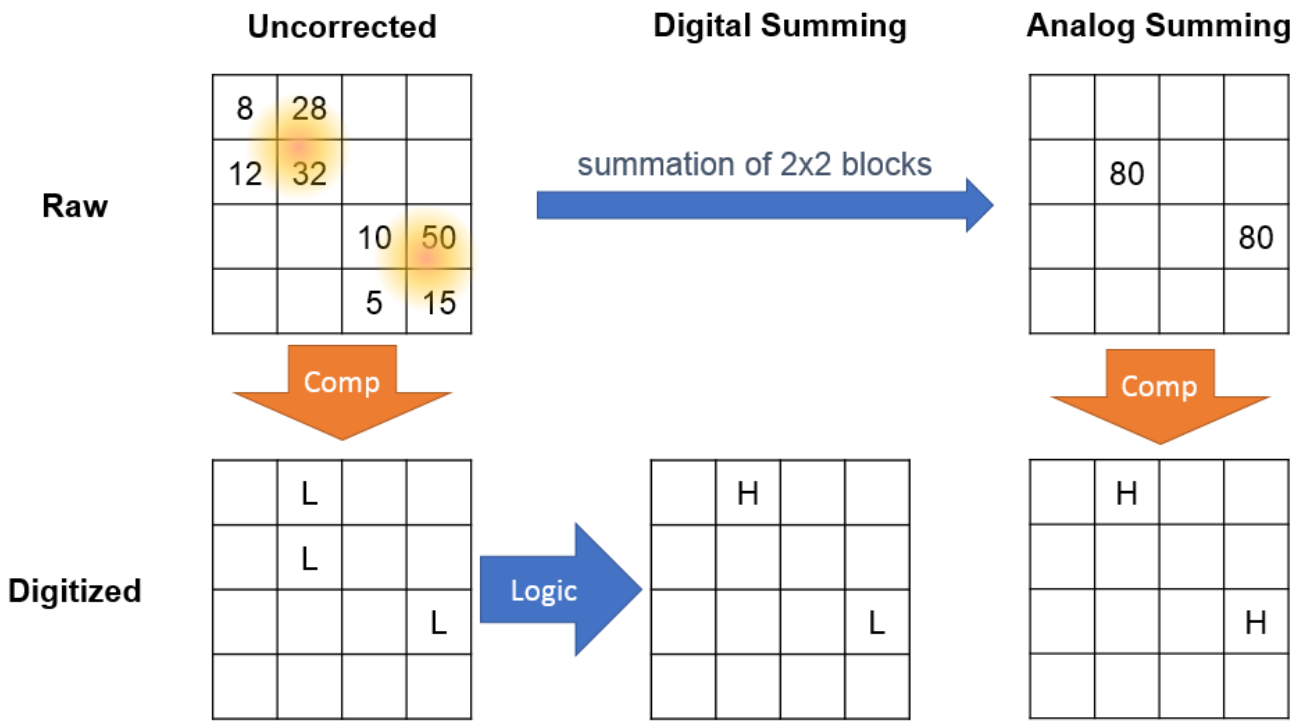


Figure 1. Two 80 keV photons are detected on a photon counting detector, and their charges are shared. (Top left) Charge deposited in each pixel of a 4x4 detector subsection is shown. (Bottom left) In an uncorrected case, each charge is compared to a bank of comparators (low bin at 25 keV, high bin at 65 keV) and converted into a deposition in the (L)ow bin or the (H)igh bin. (Bottom center) In digital charge summing, simultaneous events in adjacent pixels is interpreted via digital logic into a single event in one of the pixels at a higher bin. (Top right) However, in analog charge summing, 2x2 blocks of pixels are dynamically summed together prior to digitization via the comparator. (Bottom right) After digitization, analog charge summing correctly recovers both high energy events.

an improvement compared to the uncorrected case.

A. Simulations

All experiments presented in this work are done in simulation and were performed in the low flux (no pileup) regime. The simulations are performed using Monte Carlo transport of photons with a CdTe substrate. The detector substrate is modeled as being a 1.6 mm slab. The pixel size is 0.25 mm by 0.25 mm. Photons taken from a 120 kVp spectrum (13) are first filtered by the object thickness and then impinge on the substrate. The transport of the photons is handled using GEANT4. Photons interacting with CdTe undergo a variety of possible interaction mechanisms, including Compton scattering or photoelectric effect. Excess energy deposited into the CdTe is converted instantaneously into a cloud of charges that is distributed as a Gaussian with a characteristic standard deviation of 40 microns. These charges are assumed to migrate vertically and any charge within the 0.25 mm by 0.25 mm boundary of a pixel is assumed to be collected within the pixel. The charges within each pixel were then converted into events at low or high energies. The threshold for low energy was at 25 keV, and the threshold for high energy was at 65 keV, following settings used in one prototype detector (14).

We sampled the spectrum from 30 to 120 keV in 5 keV increments and simulated 20,000 impinging photons at each energy level. This allowed us to precalculate a table of possible outcomes for an arriving photon. The table output was the presence of a low or high energy event in a 5x5 window of pixels centered on the pixel being irradiated. This implicitly captures the spatial correlations present with charge sharing.

The performance of digital charge summing depends on the input parameters, so we make a few comments on our choices. Modeling perfect collection within pixel boundaries is not completely accurate. In practice, this depends on the pulse shaping time and the magnitude of ballistic deficit, with incomplete collection likely to occur at the boundaries between pixels. The size of the charge cloud depends on the electrical field and other such factors and could equally well be modeled as a sphere rather than a Gaussian. The detector pixel pitch is a particularly important parameter, and some other designs call for larger pixels so that charge sharing effects are reduced. However, use of charge summing will reduce the effective characteristic count rate, so in these designs we expect that a smaller pixel size that is equal to or less than 0.25 mm would

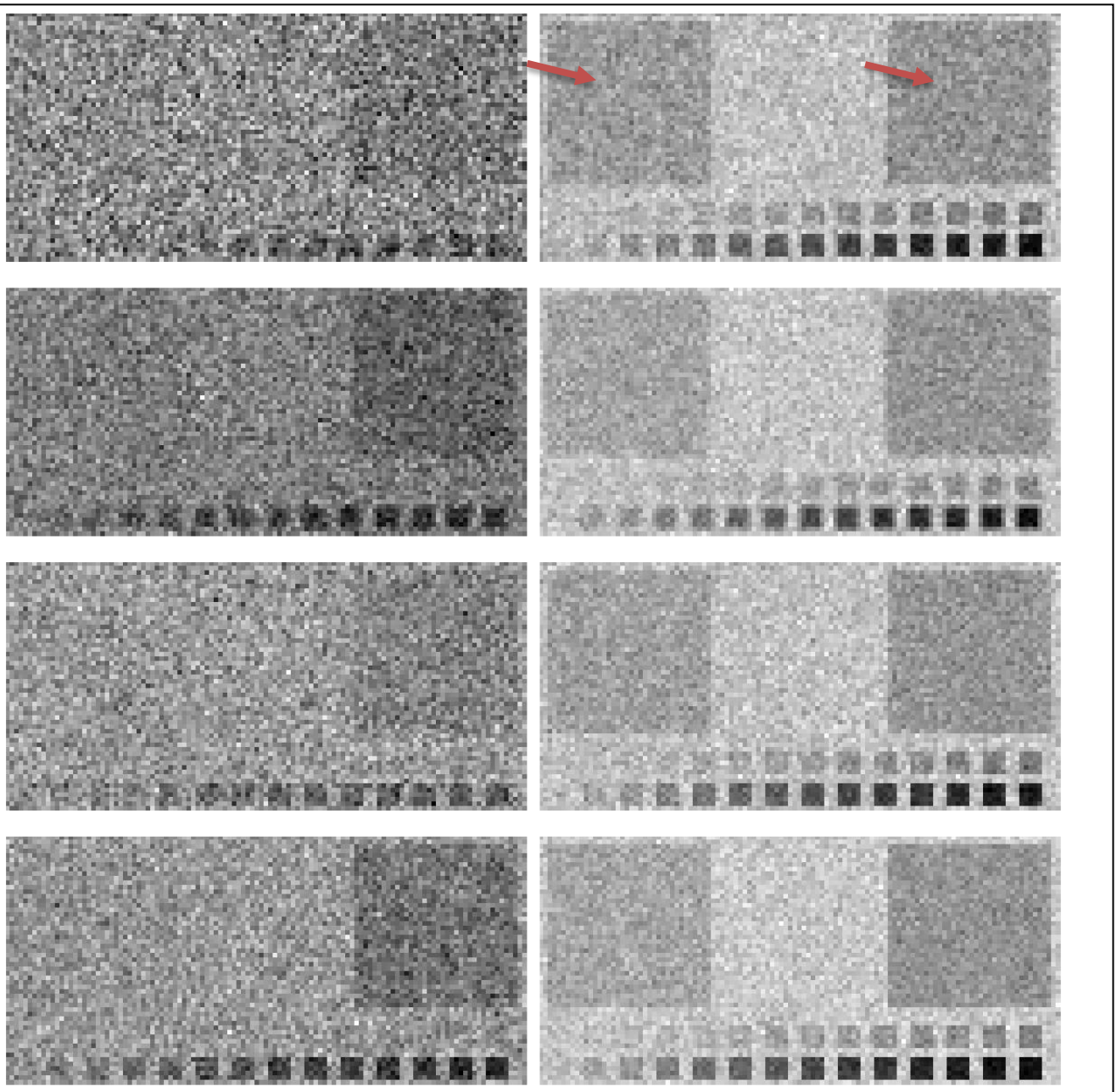


Figure 2. Simulated images with the two bin photon counting detector, using different forms of charge summing correction. The object includes a low contrast targets of water and of iodine, shown with red arrows in the top right picture. Each row represents a different detector design, and the column represents a different way to reconstruct the same data (i.e., a different combination of the low and high energy bins). From top to bottom, the rows are: uncorrected (no charge summing), analog charge summing, digital charge summing, and completely ideal detector. The left column corresponds to the linear combination which cancels water (i.e., iodine quantification task or basis material imaging) so that the water target disappears. The right column corresponds to the combination that maximizes iodine CNR.

be used.

B. Experiments

A numerical phantom was created consisting of 20 cm of water. The phantom was imaged only in projection mode. Two large squares of contrast were included of both water (0.5 cm) and iodine material (0.02 g/cm^2) to serve as low contrast targets. In addition to this, a row of small targets of variable water and iodine contrast were included to help assess detectability. Photons were filtered by the object and then arrived on the detector, at which point they were converted into events using the precomputed model.

III. Results

Figure 2 shows simulated images after passage through the phantom. An ROI was drawn in the iodine target and its CNR was compared to a background ROI. Table 1 shows the CNR estimated from this experiment.

The strongest benefits occur for the water canceled or iodine quantification task. An improvement of CNR of 42% is seen compared to no correction, which corresponds to a 2x dose efficiency improvement. Compared to digital charge summing, analog charge summing provides another 41% improvement, which again doubles dose efficiency.

Table 1. Contrast-noise ratio of the large iodine target in Figure 2 compared to the background for different detector designs, for the two different tasks.

	Water Cancel	Iodine CNR
Uncorrected	0.60	2.6
Analog charge summing	1.2	2.9
Digital charge summing	0.85	2.5
Ideal detector	1.45	3.1

IV. Discussion

Charge sharing is a significant problem for current photon counting detectors. Although existing "analog" charge summing techniques provide the best improvement, a digital charge summing technique may offer implementation advantages. A preliminary investigation showed that the improvement for iodine basis material imaging was significant, with a doubling in dose efficiency compared to no correction. This is approximately half the benefit of analog charge summing. Benefits in other forms of CNR are small.

Existing work has shown little to no benefit of photon counting detectors for dual energy imaging, at least when compared to conventional approaches that maximize spectral separation such dual source approach with filtration (15). Digital charge summing provides a possible and potentially practical pathway for photon counting to overtake conventional dual energy. Further studies are necessary to validate the initial improvements seen here, to confirm that they generalize to different object sizes, different number of detector bins and different models of charge spread.

The author would like to acknowledge Martin Sjolin for his invaluable contribution to this work.

REFERENCES

1. Taguchi K, Iwanczyk JS. Vision 20/20: Single photon counting x-ray detectors in medical imaging. *Med Phys.* 2013; 40: 100901.

2. Taguchi K, Srivastava S, Kudo H, Barber WC. In: Enabling photon counting clinical x-ray CT. Nuclear science symposium conference record (NSS/MIC), 2009 IEEE; IEEE; 2009. p. 3581-3585.

3. Wang AS, Harrison D, Lobastov V, Tkaczyk JE. Pulse pileup statistics for energy discriminating photon counting x-ray detectors. *Med Phys.* 2011; 38: 4265.

4. Hsieh SS, Pelc NJ. In: Pulse detection logic for multibin photon counting detectors: Beyond the simple comparator. SPIE medical imaging; International Society for Optics and Photonics; 2015. p. 941210-941210-6.

5. Kappler S, Hannemann T, Kraft E, Kreisler B, Niederloehner D, Stierstorfer K, Flohr T. In: First results from a hybrid prototype CT scanner for exploring benefits of quantum-counting in clinical CT. SPIE medical imaging; International Society for Optics and Photonics; 2012. p. 83130X-83130X-11.

6. Leng S, Yu Z, Halaweish A, Kappler S, Hahn K, Henning A, Li Z, Lane J, Levin DL, Jorgensen S. Dose-efficient ultrahigh-resolution scan mode using a photon counting detector computed tomography system. *Journal of Medical Imaging.* 2016; 3(4): 043504-043504.

7. Hsieh SS, Pelc NJ. A dynamic attenuator improves spectral imaging with energy-discriminating, photon counting detectors. *Medical Imaging, IEEE Transactions on.* 2015; 34(3): 729-739.

8. Hsieh SS, Pelc NJ. Improving pulse detection in multibin photon-counting detectors. *Journal of Medical Imaging.* 2016; 3(2): 023505-023505.

9. Gimenez E, Ballabriga R, Campbell M, Horswell I, Llopart X, Marchal J, Sawhney K, Tartoni N, Turecek D. Study of charge-sharing in MEDIPIX3 using a micro-focused synchrotron beam. *Journal of Instrumentation.* 2011; 6(01): C01031.

10. Ballabriga R, Campbell M, Heijne E, Llopart X, Tlustos L. The Medipix3 prototype, a pixel readout chip working in single photon counting mode with improved spectrometric performance. *IEEE Trans Nucl Sci.* 2007; 54(5): 1824-1829.

11. Koenig T, Hamann E, Procz S, Ballabriga R, Cecilia A, Zuber M, Llopart X, Campbell M, Fauler A, Baumbach T. Charge summing in spectroscopic x-ray detectors with high-Z sensors. *IEEE Trans Nucl Sci.* 2013; 60(6): 4713-4718.

12. Koenig T, Zuber M, Hamann E, Cecilia A, Ballabriga R, Campbell M, Ruat M, Tlustos L, Fauler A, Fiederle M. How spectroscopic x-ray imaging benefits from inter-pixel communication. *Phys Med Biol.* 2014; 59(20): 6195.

13. Punnoose J, Xu J, Sisniega A, Zbijewski W, Siewerdsen J. Technical note: Spektr 3.0—A computational tool for x- ray spectrum modeling and analysis. *Med Phys.* 2016; 43(8): 4711-4717.

14. Gutjahr R, Halaweish AF, Yu Z, Leng S, Yu L, Li Z, Jorgensen SM, Ritman EL, Kappler S, McCollough CH. Human imaging with photon counting-based computed tomography at clinical dose levels: Contrast-to-noise ratio and cadaver studies. *Invest Radiol.* 2016 Jul; 51(7): 421-429. PMCID: PMC4899181.

15. Faby S, Kuchenbecker S, Sawall S, Simons D, Schlemmer H, Lell M, Kachelrieß M. Performance of today's dual energy CT and future multi energy CT in virtual non-contrast imaging and in iodine quantification: A simulation study. *Med Phys.* 2015; 42(7): 4349-4366.

Application of the X-ray Transmittance Modeling-based Three-step Algorithm to Experimental Data from a Prototype PCD-CT System

Okkyun Lee*, Christoph Polster, Steffen Kappler, Kishore Rajendran,
Cynthia H. McCollough, Shuai Leng, and Katsuyuki Taguchi

***Abstract*—Recently, we have developed an x-ray transmittance modeling-based three-step algorithm for photon counting detector (PCD)-based computed tomography (CT). Starting with count data and integrating a PCD spectral distortion model, the three-step algorithm compensates for the spectral distortion of the PCD, achieves almost unbiased and minimum variance estimator, and is computationally efficient due to the linearized estimation steps using the x-ray transmittance model. In this paper, we apply the three-step algorithm to an actual PCD-CT system and evaluate the method by experiments using water-equivalent phantom with iodine and bone inserts. We compare the results to those of the conventional image-based material decomposition method, and FBP applied to the threshold data of the system.**

I. Introduction

Photon counting detector (PCD)-based computed tomography (CT) exploits the spectral information from each PCD pixel for more accurate material decomposition including K-edge imaging and has a great potential in many clinical applications [1]. However, one of the challenges is the spectral distortion in the PCD due to physical effects such as charge sharing and K-escape fluorescence. Conventional approaches to compensate for the spectral distortion include 1) model-based method using maximum likelihood (ML) estimator [2] and 2) calibration-based one such as the A-table method [3]. While the ML estimator guarantees an asymptotic optimality in estimating basis line-integrals (or thicknesses of materials), a substantial computational burden is a concern in practice. On the other hand, the calibration-based methods efficiently estimate basis line-integrals; however, it heavily relies on the calibration process and how to handle various calibration materials in practice can be a challenge.

Recently, we proposed a three-step algorithm as a fast alternative to the ML estimator in the model-based approach and demonstrated that the bias and noise are comparable to those of the ML and that it is more stable than the ML when photon counts are low [4, 5]. The main idea is the use of a novel x-ray transmittance model: The x-ray transmittance, $\exp\left(-\int \mu_a(E,\mathbf{r})d\mathbf{r}\right)$, not μ_a but the exponential function itself, can be approximated by a linear combination of energy-dependent basis functions where the number of bases is limited by the number of energy bins in the PCD. Using the x-ray transmittance model, the original nonlinear counts model can be linearized with unknown coefficients of the bases and then the three-step algorithm can be derived as follows: 1) Estimating the unknown bases coefficients from the linearized counts model and then calculating the x-ray transmittance, 2) estimating the basis line-integral using the least squares fitting after applying the $-\ln$ operation to the estimated x-ray transmittance, and 3) correcting for a bias using pre-calculated look-up tables.

In the previous work [4, 5], the three-step algorithm has been assessed by using numerical simulations; in the current work, we apply it to an actual PCD-CT system. The output data from the actual PCD-CT system is calibrated using water-equivalent phantoms to address some of the pre-processing performed on the data such as pixel uniformity correction, and the calibrated data is then fed into the three-step algorithm. We model the spectral distortion in the PCD with a parallel cascaded systems analysis which was validated with Monte Carlo simulation [6] and use it in the three-step algorithm to compensate it. We also use detector pixel-specific incident spectrum to address different thicknesses of bow-tie filter that each incident x-ray experiences. The details of the proposed method are described in Section II. Experimental conditions, conventional image-based material decomposition (IMD), and how to evaluate the methods are described in Section III. Validation of the proposed method with test phantoms are shown in Section IV with comparisons to the IMD and filtered back-projection (FBP) applied to the threshold data. Conclusions are provided in Section V.

II. Proposed Method

A. Measurement Model

The linear attenuation coefficients $\mu_a(\mathbf{r}, E)$ at position $\mathbf{r}$ and energy E can be modeled by a linear combination of two energy-dependent functions (e.g., photoelectric effect and Compton scattering) in the absence of K-edge materials. If we select linear attenuation coefficients of soft and hard tissues

O. Lee and K. Taguchi are with the Department of Radiology and Radiological Science, Johns Hopkins University School of Medicine, Baltimore, MD (okkyun2@gmail.com; ktaguchi@jhmi.edu).

C. Polster and S. Kappler are with the Siemens Healthineer, Forchheim, Germany.

K. Rajendran, C. H. McCollough, and S. Leng are with the Department of Radiology, Mayo Clinic, Rochester, MN.

*Corresponding author: O. Lee

This research was supported by Siemens grants No. JHU-2016-CT-1-01-Taguchi-C0022.

such as water ($\mu_W(E)$) and bone ($\mu_B(E)$) for those functions, then the j[1]-th line-integral of $\mu_a(\mathbf{r}, E)$ can be formulated by

$$\int_{\Omega_j} \mu_a(\mathbf{r}, E) d\mathbf{r} = v_{(W,j)}\mu_W(E) + v_{(B,j)}\mu_B(E) = \boldsymbol{\mu}(E)\mathbf{v}_j, \quad (1)$$

where Ω_j is the path for the j-th line-integral, and $v_{(W,j)}$ and $v_{(B,j)}$ are the thicknesses of water and bone, respectively, along the path; thus, $\boldsymbol{\mu}(E) = [\mu_W(E), \mu_B(E)]$ and $\mathbf{v}_j = [v_{(W,j)}, v_{(B,j)}]^T$. Using (1), the expected photon counts $\bar{y}_{(b,c)}(\mathbf{v}_j)$ measured at the b-th energy bin and the detector pixel c for the j-th line-integral can be formulated by[2]

$$\bar{y}_{(b,c)}(\mathbf{v}_j) = \int_0^\infty S_b(E) S_c(E) \exp\left(-\boldsymbol{\mu}(E)\mathbf{v}_j\right) dE, \quad (2)$$

for $b = 1, 2, \cdots, N_b$ and $c = 1, 2, \cdots, N_c$, where N_b and N_c are the number of bins and pixels, respectively. $S_b(E)$ is the energy-bin dependent function to model the spectral distortion in PCD [6] and $S_c(E)$ is the incident x-ray spectrum after passing a bow-tie filter with a certain thickness of the filter in the direction toward the detector pixel c. The normalized noisy counts for the j-th line-integral can then be described as follows:

$$\mathbf{y}_c^N(\mathbf{v}_j) = \frac{\bar{\mathbf{y}}_c(\mathbf{v}_j)}{\bar{\mathbf{y}}_c(\mathbf{0})} + \mathbf{n}_{(c,j)}, \quad (3)$$

where $\bar{\mathbf{y}}_c(\mathbf{v}_j) = [\bar{y}_{(1,c)}(\mathbf{v}_j), \bar{y}_{(2,c)}(\mathbf{v}_j), \cdots, \bar{y}_{(N_b,c)}(\mathbf{v}_j)]^T \in \mathbb{R}^{N_b \times 1}$, and $\mathbf{n}_{(c,j)} \in \mathbb{R}^{N_b \times 1}$ is a noise vector. Note that we normalized the counts by $\bar{\mathbf{y}}_c(\mathbf{0})$ (i.e., counts without object) so that a possible mismatch of the total number of incident photons in the model can be canceled out.

B. Three-step Algorithm

The three-step algorithm can be derived by modeling the x-ray transmittance: $\exp\left(-\boldsymbol{\mu}(E)\mathbf{v}_j\right)$ can be approximated by a linear combination of a few of energy-dependent basis functions $\{D_k(E)\}_{k=1}^{N_k}$ and it can be formulated by

$$X_j(E) = \exp\left(-\boldsymbol{\mu}(E)\mathbf{v}_j\right) \approx \sum_{k=1}^{N_k} \theta_{(j,k)} D_k(E), \text{ for } N_k \leq N_b, \quad (4)$$

where $\{D_k(E)\}_{k=1}^{N_k}$ can be either low-order polynomials or pre-calculated from a low-rank approximation [4, 5]. If we substitute (4) into (2), then the normalized counts model in (3) can be linearized with the unknown coefficients θ as follows:

$$\mathbf{y}_c^N(\mathbf{v}_j) \approx B_c \boldsymbol{\theta}_j + \mathbf{n}_{(c,j)}, \quad (5)$$

where $\boldsymbol{\theta}_j = [\theta_{(j,1)}, \theta_{(j,2)}, \cdots, \theta_{(j,N_k)}]^T \in \mathbb{R}^{N_k \times 1}$, and $B_c \in \mathbb{R}^{N_b \times N_k}$ is a sensing matrix whose b-th row and k-th column component is given by

$$[B_c]_{(b,k)} = \int_0^\infty S_c(E) S_b(E) D_k(E) dE \;/\; \bar{y}_{(b,c)}(\mathbf{0}). \quad (6)$$

Using the linearized model (5), we can describe the three-step algorithm as follows:

· **1st step**:

$$\hat{\boldsymbol{\theta}}_j = \arg\min_{\boldsymbol{\theta}} \left\| \mathbf{y}_c^N(\mathbf{v}_j) - B_c\boldsymbol{\theta} \right\|^2 + \lambda \left\| K\boldsymbol{\theta} \right\|^2, \quad (7)$$

where $K \in \mathbb{R}^{N_k \times N_k}$ is a regularization matrix.

· **2nd step**:

$$\hat{\mathbf{v}}_j = \arg\min_{\mathbf{v}} \left\| \ln\left(\sum_{k=1}^{N_k} \hat{\theta}_{(j,k)} D_k(E) \right) + \boldsymbol{\mu}(E)\mathbf{v} \right\|^2. \quad (8)$$

· **3rd step**:

$$\hat{\mathbf{v}}_j^i = \hat{\mathbf{v}}_j^{i-1} - BCT^i\left(\hat{\mathbf{v}}_j^{i-1}\right), \text{ for } i = 1, \cdots, N_{iter}, \quad (9)$$

where $\hat{\mathbf{v}}_j^0 = \hat{\mathbf{v}}_j$ and $\{BCT^i(\cdot)\}_{i=1}^{N_{iter}}$ is the pre-calculated bias look-up tables. For more details of theoretical analysis and practical implementation, see [4, 5].

III. Experimental Setup

In this section, we describe scan conditions, calibration settings for the three-step method and the image-based material decomposition (IMD) method to which the three-step method is compared, and evaluations with test phantoms.

• Scan conditions: We used a prototype, whole-body PCD-CT system (SOMATOM CounT; Siemens Healthineer; Forchheim, Germany) for the experiments. The number of detector pixels was 480 for $\sim$28.78° and the number of projection views was 2,304 over 2π with 1 sec of rotation time. We set the energy thresholds at $20, 38, 52, 80\,[\mathrm{keV}]$ and performed scans with $140\,\mathrm{kVp}$ and $100\,\mathrm{mA}$.

• Calibration phantoms (three-step method): Cylindrical water-equivalent phantoms (QRM, Moehrendorf, Germany) with diameters of $10\,\mathrm{cm}$ and $20\,\mathrm{cm}$ with two air holes.

Each phantom was scanned twice with different offsets from the iso-center, $2.5\,\mathrm{cm}$ and $5\,\mathrm{cm}$, to obtain a sufficient amount of attenuations for each detector pixel. These phantoms were used to calibrate the output data from the actual PCD-CT system. It was also used to estimate the noise in the forward model (3), under the assumption that the noise follows a zero-mean Gaussian distribution, and synthesize noisy data sets for generating bias correction tables in the third-step of the algorithm.

• Calibration phantoms (IMD): Cylindrical water-equivalent phantom with a diameter of $20\,\mathrm{cm}$ ($2.5\,\mathrm{cm}$ offset from the iso-center) with iodine/bone insertions.

We used water as one of the calibration material and a linear combination of water and bone as the second material. The linear combination was calculated from the result of the three-step algorithm applied to the data for this calibration phantom, and the area for the calibration of the second material was defined as the center of the phantom where the bone was filled with unknown density. During the calibration process (i.e., estimating weighting coefficients; see [7] for details), two energy bin data were used to fit the two calibration materials, and the other two were used to minimize the noise [7].

• Test phantoms: Cylindrical water-equivalent phantoms with a diameter of $20\,\mathrm{cm}$ (without center offset) with and

[1]We used a single index j for simplicity. In practice, it can be represented by a combination of indices of such as detector pixel and projection view.

[2]We neglected the pulse pileup effect in this model since it may be severe only near the boundary of the object [1].

without iodine/bone insertions.

• CounT data processing: Seven polychromatic CT images in total were reconstructed from energy bin and threshold data using FBP, which we call "poly CT images." The basis images of water and bone were estimated using IMD and the four poly CT images from four energy bins as the input. Monoenergetic CT images at desirable energies were synthesized by a linear combination of the two basis images, which we call "mono CT images."

• Proposed method: Basis sinograms of water and bone were estimated using the three-step method and the four (calibrated) energy bin data as the input. The corresponding basis images were reconstructed using FBP, and mono CT images were generated by a linear combination of the two basis images.

• Data analysis: Bias and standard deviation (STD) of the following data were measured: basis images, poly CT images, and mono CT images obtained by CounT, IMD, and the proposed method.

IV. Results

Figures 1 and 2 present water and bone basis images of the test phantoms each obtained by both the IMD and the three-step method. The holes of the phantom had the air for Fig. 1 and iodine/bone inserts for Fig. 2. The three-step method had a smaller bias than the IMD did with both of the phantoms; and the three-step method had smaller noise than the IMD did when the phantom had the water-equivalent material only (Fig. 1), while the IMD had less noise than the three-step method with iodine and bone inserts (Fig. 2). Note that the test phantom with inserts is strikingly similar to the calibration phantom for the IMD (the only difference was the 2.5 cm of offset from the iso-center); thus, it might overestimate the performance for the IMD.

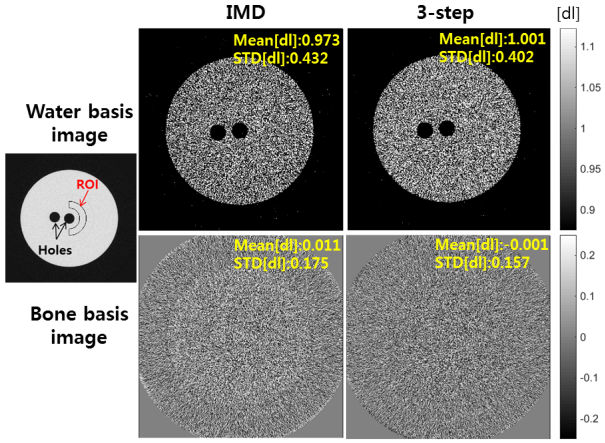


Fig. 1: Water and bone basis images of the test phantom with the air holes, reconstructed using the IMD and the three-step method. The mean and standard deviation (STD) values are calculated at the region of interest (ROI) indicated in the far left image.

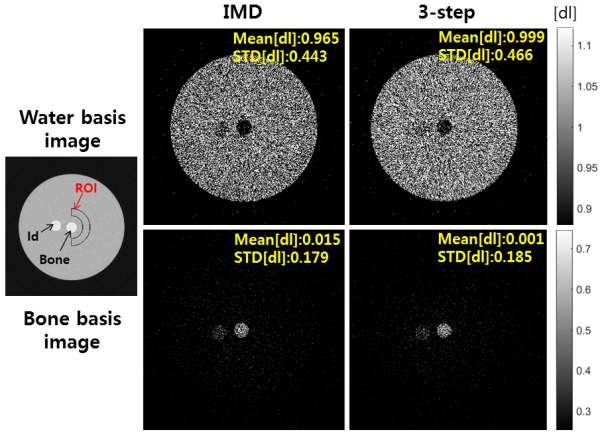


Fig. 2: Water and bone basis images of the test phantom with iodine/bone inserts, reconstructed using the IMD and the three-step method.

Figures 3 and 4 illustrate reconstructed poly CT image (>20 keV) and mono CT images (at 71 keV) from the IMD and the three-step method for the test phantom without and with inserts, respectively. Bias and STD measured in the ROI (see Fig. 2) of the CT images show that all the methods were almost unbiased and that the STDs of the mono CT images from the three-step method were close to those of the poly CT images (>20 keV) which have the largest SNR. STD of the mono CT images from the IMD for both test phantoms were smaller than other methods due to 1) the noise minimization process during the calibration of the IMD and 2) possibly the similar conditions between the test phantoms and calibration phantoms used in the IMD. Both the poly CT image and the mono CT image (IMD) showed decreased pixel values near the boundary of the phantoms (without/with inserts; thin arrows) and a slight beam hardening artifact between the two inserts (thick arrows). In contrast, the mono CT image from the three-step algorithm showed nearly no bias throughout the image.

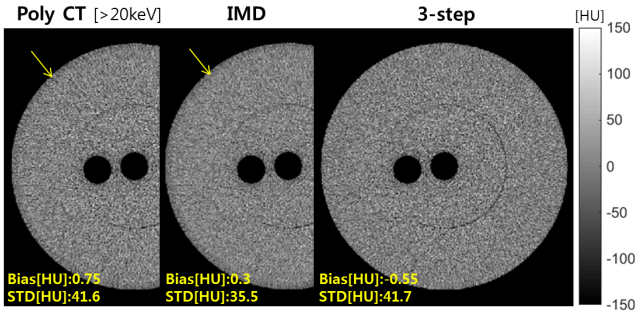


Fig. 3: Reconstructed poly CT image (>20 keV) and mono CT images (at 71 keV) from the IMD and the three-step method for the test phantom without inserts. Decreased pixel values near the boundary of the phantom is indicated by thin arrows. Bias and STD of the poly and mono CT images were measured in the ROI (see Fig. 2).

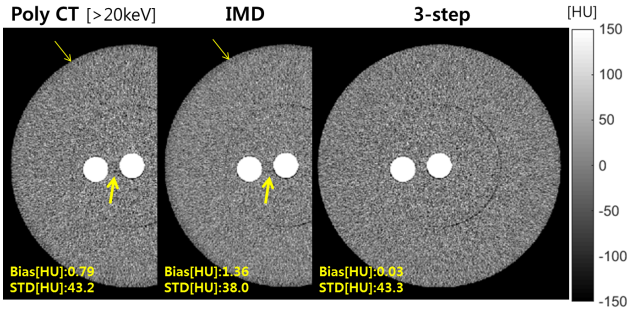


Fig. 4: Reconstructed poly CT image ($>20\,\mathrm{keV}$) and mono CT images (at $71\,\mathrm{keV}$) from the IMD and the three-step method for the test phantom with inserts. Decreased pixel values near the boundary of the phantom and possible beam hardening artifact between two inserts are indicated by thin and thick arrows, respectively.

V. Conclusions

We applied the previously developed three-step algorithm to the actual PCD-CT system. Water calibration phantoms were used to calibrate the output of the system to feed it into the three-step algorithm and to estimate the noise variance in the measurements for generating bias correction tables. The reconstructed images of the test phantoms using the proposed method were almost unbiased, had similar noise to the poly CT images, and had fewer artifacts than those of the image-based material decomposition and poly CT images. The proposed method was also computationally efficient: It took only $\sim 12\,\mathrm{sec}$ using Intel Xeon CPU (dual) 2.00GHz for estimating sinograms ($2{,}304 \times 480$) of water and bone.

Acknowledgment

We appreciate the support of Drs. Thomas G. Flohr, Matthew K. Fuld, and George S. K. Fung in Siemens Healthcare during the course of the study.

References

[1] K. Taguchi and J. S. Iwanczyk, "Vision 20/20: single photon counting x-ray detectors in medical imaging," *Med. Phys.*, vol. 40, no. 10, pp. 100 901(1–19), October 2013.

[2] J. P. Schlomka, E. Roessl, R. Dorscheid, S. Dill, G. Martens, T. Istel, C. Bäumer, C. Herrmann, R. Steadman, G. Zeitler, A. Livne, and R. Proksa, "Experimental feasibility of multi-energy photon-counting K-edge imaging in pre-clinical computed tomography," *Phys. Med. Biol.*, vol. 53, pp. 4031–4047, 2008.

[3] R. E. Alvarez, "Estimator for photon counting energy selective x-ray imaging with multibin pulse height analysis," *Med. Phys.*, vol. 38, no. 5, pp. 2324–2334, May 2011.

[4] O. Lee, S. Kappler, C. Polster, and K. Taguchi, "Estimation of basis line-integrals in a spectral distortion-modeled photon counting detector using low-order polynomial approximation of x-ray transmittance," *IEEE Trans. on Medical Imaging*, vol. 36, no. 2, pp. 560–573, Feb. 2017.

[5] ——, "Estimation of basis line-integrals in a spectral distortion-modeled photon counting detector using low-rank approximation-based x-ray transmittance modeling: K-edge imaging application," *IEEE Trans. on Medical Imaging*, vol. 36, no. 11, pp. 2389–2403, Nov. 2017.

[6] K. Taguchi, K. Stierstorfer, C. Polster, O. Lee, and S. Kappler, "Spatio-energetic cross-talk in photon counting detectors: Numerical detector model (PcTK) and workflow for CT image quality assessment," *Med. Phys.*, (*Under Revision*).

[7] S. Faby, S. Kuchenbecker, S. Sawall, D. Simons, H. P. Schlemmer, M. Lell, and M. Kachelriess, "Performance of todays dual energy CT and future multi energy CT in virtual non-contrast imaging and in iodine quantification: A simulation study," *Med. Phys.*, vol. 42, no. 7, pp. 4349–4366, July 2015.

Photon-counting CT Reconstruction using Total Image Constrained Diffusion Tensor

Shanzhou Niu, Xiaokun Huang, Jianhua Ma, and Jing Wang

***Abstract*—Photon-counting CT (PCCT) is a promising imaging technique because of its ability to differentiate and identify different materials with high energy resolution using narrow energy bins. However, the narrow energy bin images will be affected by serious noise because of limited amount of photons in the corresponding energy bins compared to the full spectrum. To reconstruct high quality photon-counting CT images, we present a total image constrained diffusion tensor (TICDT) for statistical iterative reconstruction (SIR) based on a penalized weighted least-squares (PWLS) principle, which is referred to as 'PWLS-TICDT'. Specifically, the TICDT exploits the correlation in energy domain to improve the PCCT reconstruction performance. An alternating optimization algorithm is adopted to solve the associated objective function. Qualitative and quantitative studies were conducted to evaluate PWLS-TICDT method using the XCAT phantom. Results from numerical simulation study show that the presented PWLS-TICDT method achieves noticeable gains over the competing methods in terms of noise suppression and structural information preservation. Moreover, the narrow energy bin images reconstructed by PWLS-TICDT method can produce more accurate basis material decomposition.**

***Index Terms*—Photon-counting CT, image reconstruction, PWLS, diffusion tensor**

I. Introduction

PHOTON-COUNTING CT is a promising imaging technique because of its ability to differentiate and identify different materials. The photon-counting detector (PCD) is used in this technique to increase the energy resolution and material specificity. The PCD can count individual x-ray photons and sort out them into different energy bins based on their energies and preset energy thresholds, which separates a whole polychromatic x-ray spectrum into several energy bins [1, 2]. Multiple projection datasets with different spectral information can be acquired from a single scan, which results in improved energy resolution and efficient material decomposition. However, given the amount of total photons, the amount of photons in narrow energy bin will be smaller, and the noise in each energy bin image will be dramatically increased [3-5]. The image quality of material decomposition will be degraded, which would compromise the identification of small material component differences.

In this work, the noise properties of projection data is incorporated into statistical iterative reconstruction (SIR) for spectral CT with narrow energy bins. The SIR algorithms usually consist of two terms. The first term describes the statistics of projection data and the second term characterizes a prior information to regularize the result. Sparsity regularizations have been widely applied to CT reconstruction and demonstrated superior performance compared with classical analytic reconstruction algorithms [6, 7]. These sparsity-regularized methods can be directly applied to spectral CT in each energy bin individually. Xu et al. presented an interior tomography method with TV minimization to reconstruct each energy bin images individually [8]. Zeng et al. proposed a structure tensor TV based SIR method for spectral CT reconstruction that using the high-order derivatives information to eliminate the staircase effect [9]. Zhao et al. developed a tight frame based reconstruction method for spectral breast CT using fewer projection views, which can produce accurate material composition quantification of breast tissue [10]. The limitation of these methods is that the abundant redundancy between different energy bin images is not incorporated into reconstruction.

To exploit the structure correlations in energy domain, we present a total image constrained diffusion tensor (TICDT) for spectral CT reconstruction. Specifically, we first reconstruct a high-quality total image from integrated projection data, and construct a new diffusion tensor with this total image and energy bin image that characterizes structure correlation in energy domain. Then, the TICDT is incorporated into SIR, and the energy bin image can be enhanced using the redundant information in the total image while preserving the primary features in narrow energy bin image. Furthermore, an alternating optimization algorithm is adopted to solve the associated objective function. The TICDT method has two major advantages. First, it maintains structure features of the total image even when the noise level of energy bin image is extremely high.

This Work is supported in party by the Cancer Prevention and Research Institute of Texas (RP160661), US National Institutes of Health (R01 EB020366), National Natural Science Foundation of China (11701097, 61571214, U1708267), Natural Science Foundation of Jiangxi Province (2016BAB212055), Science and Technology Program of Jiangxi Education Committee (GJJ150994), and Science and Technology Program of Guangzhou (201510010039).

S. Niu, X. Huang, and J. Wang are with Department of Radiation Oncology, University of Texas Southwestern Medical Center, Dallas, TX 75235, USA. S. Niu is also with School of Mathematics and Computer Science, Gannan Normal University, Ganzhou 341000, China.

J. Ma is with the School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China.

Second, it can detect matching directions of diffusion tensor between different images, achieving substantial noise reduction with preserved edge and structure information.

II. Methods and Materials

A. PWLS image reconstruction

Consider a sequence of log-transformed projection data y_{mi} corresponding to $m=1,2,\cdots,M$ energy bins and $i=1,2,\cdots,N$ projection views, where M is the number of energy bins, N_1 is the total number of rays at each projection view, N_2 is the total number of projection views, and $N=N_1\times N_2$. When the energy bin width is narrow enough, the measurement in any energy bin can be approximated as a mono-energetic projection. For each energy bin m, the variance of the noise in projection data y_{mi} can be determined by the following exponential formula [11, 12]

$$\lambda_{mi}^2=\exp(\bar{y}_{mi})/I_{mi}, \tag{1}$$

where I_{mi} is the incident photon number at energy bin m and projection ray i, $\bar{y}_{mi}$ and λ_{mi}^2 is the mean and variance of projection at energy bin m and detector channel i.

Based on the properties of the noise at energy bin m, the PWLS image reconstruction with penalty term $R(\mu_m)$ can be formulated as

$$\min_{\mu_m\geq 0}(y_m-A\mu_m)^T\Sigma^{-1}(y_m-A\mu_m)+\beta R(\mu_m), \tag{2}$$

where y_m is the vector of sinogram data (projection after logarithm transform) at energy bin m, μ_m is the corresponding vector of attenuation map to be reconstructed, and A denotes the projection or system matrix which is determined by a fast ray-tracing method. The symbol T represents matrix transpose operator. Σ is a diagonal matrix with the ith element of λ_{mi}^2, which can be obtained from the measurement y_{mi} according to Eq. (1). The penalty or smoothing parameter $\beta>0$ controls the tradeoff between the data fidelity and penalty terms.

B. Tensor based anisotropic diffusion

Using diffusion tensor, Weickert [13] rewrote the anisotropic diffusion process as follows

$$\partial_t u=\mathrm{div}\left(\mathbf{D}(\mathbf{J}_\rho(\nabla u_\sigma))\nabla u\right), \tag{3}$$

where $\mathbf{D}(\mathbf{J}_\rho(\nabla u_\sigma))$ denotes a diffusion tensor on the basis of the structure tensor $\mathbf{J}_\rho(\nabla u_\sigma)$. The structure tensor is defined as a positive semidefinite symmetric matrix

$$\mathbf{J}_\rho(\nabla u_\sigma)=K_\rho*\left(\nabla u_\sigma\nabla u_\sigma^T\right)=\begin{pmatrix}K_\rho*\left(\partial_x u_\sigma\right)^2 & K_\rho*\left(\partial_x u_\sigma\partial_y u_\sigma\right)\\ K_\rho*\left(\partial_x u_\sigma\partial_y u_\sigma\right) & K_\rho*\left(\partial_y u_\sigma\right)^2\end{pmatrix}, \tag{4}$$

where for stable evolution, a Gaussian kernel K_σ is convolved with the image u, i.e., $u_\sigma=K_\sigma*u$.

With the eigenvectors of structure tensor in Eq. (4), the diffusion tensor can be given as

$$\mathbf{D}(\mathbf{J}_\rho(\nabla u_\sigma))=\begin{pmatrix}v_1 & v_2\end{pmatrix}\begin{pmatrix}\gamma_1 & 0\\ 0 & \gamma_2\end{pmatrix}\begin{pmatrix}v_1 & v_2\end{pmatrix}^T=\begin{pmatrix}D_{11} & D_{12}\\ D_{12} & D_{22}\end{pmatrix}, \tag{5}$$

$$D_{ij}=\sum_{n=1,2}\gamma_n v_{ni}v_{nj}. \tag{6}$$

The diffusion tensor $\mathbf{D}$ possesses two new eigenvalues $\gamma_{1,2}$, which determined the strength of smoothing with preferred direction $v_{1,2}$. In this study, the eigenvalues $\gamma_{1,2}$ relating to normal and tangential directions are given by

$$\gamma_1=\exp\left(\left|\nabla u\right|^2/\delta^2\right),\gamma_2=1, \tag{7}$$

where $\delta>0$ is a threshold that determines the strength of smoothing in normal direction, and the smoothing in tangential direction is performed by setting $\gamma_2=1$.

C. Total image constrained diffusion tensor

The proposed TICDT consists of two steps: 1) to reconstruct a total image from the integrated projection data; and 2) to incorporate the high-quality total image into image reconstruction using diffusion tensor.

The structure information from total image μ_* (reconstructed from the integrated projection using TV minimization) is embedded into energy bin image μ_m by means of diffusion tensor. The target image μ_m can be enhanced with the supplementary information from total image μ_*, while preserving the important features in target image. Let $\mathbf{D}_m$ and $\mathbf{D}_*$ be the diffusion tensor for images μ_m and μ_* respectively using Eq. (5). Using the intensity and geometrical information of both images, we can construct a total image constrained diffusion tensor (TICDT)

$$\mathbf{D}_{m*}=s\mathbf{D}_m+(1-s)\mathbf{D}_*, \tag{8}$$

where $s\in[0,1]$ is a spatially variant that determines the relative contribution between these two diffusion tensors. In this study, we follow the strategy as suggested by [14]:

$$s=\sin\left(\mathrm{abs}\left(\arccos\left(v_1(\mu_m)\cdot v_1(\mu_*)\right)-\frac{\pi}{2}\right)\right), \tag{9}$$

where $v_1(\mu_m)\cdot v_1(\mu_*)$ is the component-wise inner production of the normalized eigenvectors from images μ_m and μ_*.

D. PWLS-TICDT reconstruction

For the PWLS reconstruction in Eq. (2), it can be iteratively minimized by the following two steps:

$$\mu_m^{k+1/2}=\mu_m^k-\frac{1}{L}A^T\Sigma^{-1}\left(A\mu_m^k-y_m\right), \tag{10}$$

$$\mu_m^{k+1} = \arg\min_{\mu_m} \varphi(\mu_m) = \frac{L}{2}\left\|\mu_m - \mu_m^{k+1/2}\right\|_2^2 + \beta R(\mu_m), \quad (11)$$

where L is a Lipschitz constant, which is set to be the maximum eigenvalue of the matrix $A^T\Sigma^{-1}A$, k is the index of iteration steps.

The intention of these two steps is summarized as follows. First, a gradient descent algorithm is applied to the data fidelity term to obtain an intermediate solution $\mu_m^{k+1/2}$ as described in Eq. (10). This intermediate solution may contain artifact and noise, since no penalty is performed. Second, an iterative gradient descent algorithm is used to solve Eq. (11):

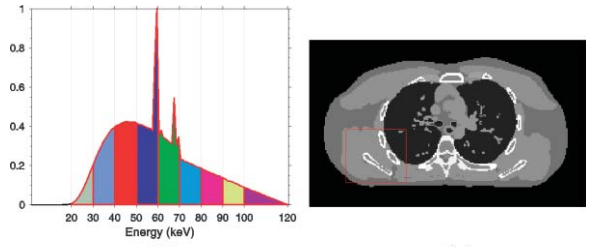
$$\begin{cases} \mu_m^{l+1} = \mu_m^{l} + \tau\varphi'(\mu_m) \\ \mu_m^{l=0} = \mu_m^{k+1/2} \end{cases}. \quad (12)$$

Using the TICDT based anisotropic diffusion, the first-order derivative of the function $\varphi(\mu_m)$ can be formulated as follows:

$$\varphi'(\mu_m) = L\left(\mu_m - \mu_m^{k+1/2}\right) + \beta \operatorname{div}\left(\mathbf{D}_{m*}\nabla\mu_m\right). \quad (13)$$

E. Numerical simulation

In this study, a 120 kV x-ray spectrum was generated using the TASMICS method [15]. This spectrum was divided into nine energy bins: [0, 30] keV, [30, 40] keV, [40, 50] keV, [50, 60] keV, [70, 80] keV, [90, 100] keV, and [100, 120] keV. The incident x-ray intensity was distributed to each energy bin according to the weights calculated from the simulated spectrum. The normalized x-ray spectrum and nine energy bins are shown in Fig. 1. The anthropomorphic XCAT phantom [16] was used to simulate a PCD-based spectral CT imaging. This phantom (Fig. 2(a)) is of 256×256 array size, and the pixel size is 2 mm. Using our previously developed simulation method [17], multi-energy projection data can be obtained by assuming 1.0×10^5 photons emitted from the x-ray source of a fan-beam CT scanner. The distance from the source to detector is 949 mm and the distance from the rotation center to detector is 408 mmm. The number of channels per projection view is 888 with 984 projection views evenly spanned on a circular orbit for a 2π rotation. The size of each detector channel is 1.0 mm.



(a) (b)

Figure 1. Numerical simulation: (a) 120 kV x-ray spectrum with nine energy bins; (b) XCAT phantom.

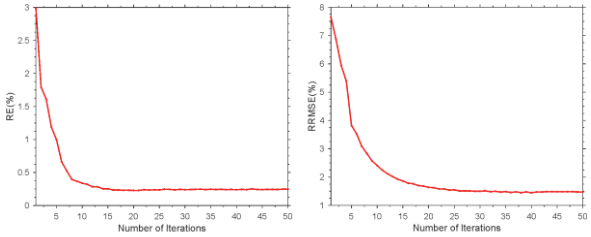
F. Comparison method

To validate and evaluation the performance of PWLS-TICDT method, the diffusion tensor based PWLS (PWLS-DT) method and the PWLS-PINL method describe in [18] were adopted for comparison. The high-quality total image μ_* is also used as the prior image for PWLS-PINL method.

III. Results

A. Convergence analysis

Figure 2 illustrates the convergence of the proposed PWLS-TICDT method with a representative energy bin (65 keV). Fig. 9(a) and 9(b) show the relative error (RE) and root mean square error (RMSE) curves versus number of iterations, respectively. We can observe that PWLS-TICDT method can converge to a steady solution after enough iterations in terms of RE and RMSE measures.



(a) (b)

Figure 2. Convergence curves for the PWLS-TICDT reconstruction: (a) RE curve versus the number of iterations; (b) RMSE curve versus the number of iterations.

B. Visual Evaluation

Figure 3 shows images of the XCAT phantom, with columns 1–3 showing 25, 65, and 110 keV energy bin images, respectively. Not only the severe noise are suppressed in PWLS-TICDT results, but the edges are well preserved for bone and structures inside of the lung. The PWLS-TICDT method achieves the best performance in terms of noise reduction and structure detail preservation.

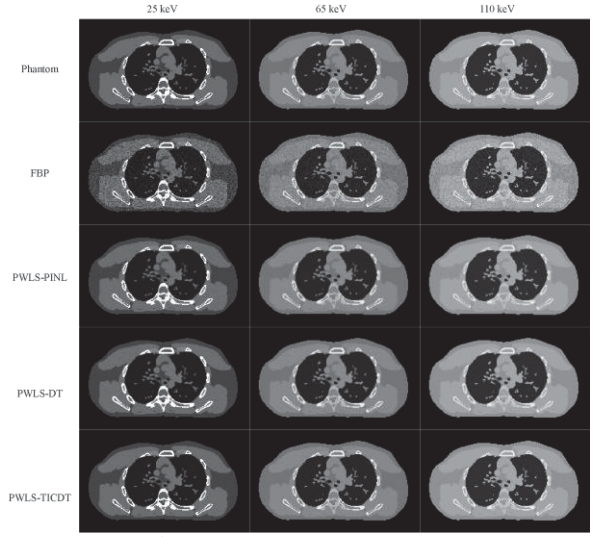


Figure 3. Results of XCAT phantom correspond to three representative energy bins. From the left to right columns the display winds are [0, 0.26], [0, 0.08], and [0, 0.055] mm^{-1}.

C. Structural Similarity Study

To further show the difference of the above reconstructed image, zoomed ROI (indicated by the red and blue squares in Fig. 1(b)) are displayed in Figs. 4. To quantitatively evaluate the performance of these reconstruction method, the corresponding SSIM values on this ROO at each energy bin are depicted in Fig. 5. The average SSIM value of the images reconstructed by the PWLS-PINL method, PWLS-DT method, and PWLS-TICDT method are 0.9468, 0.9409, and 0.9839, respectively. The results demonstrate that PWLS-TICDT method achieves noticeable gains over the PWLS-PNNL and PWLS-DT methods in terms of noise reduction and structure detail preservation.

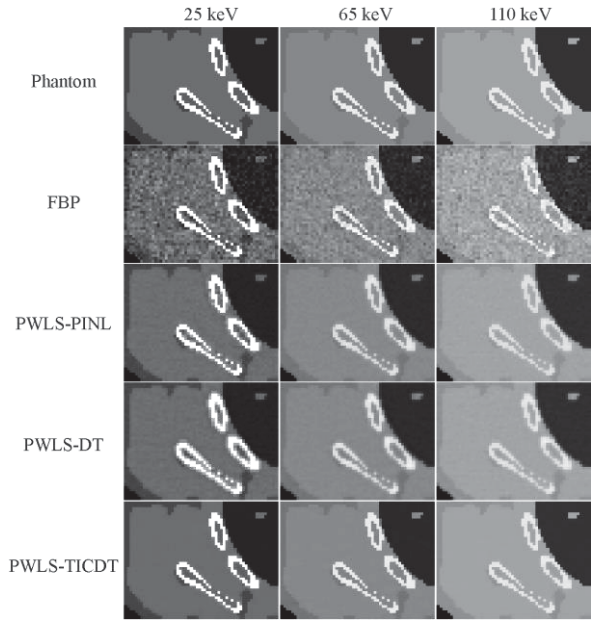


Figure 4. Zoomed-in views of the ROI (indicated by the red square in Fig. 1(b)) in results of XCAT phantom.

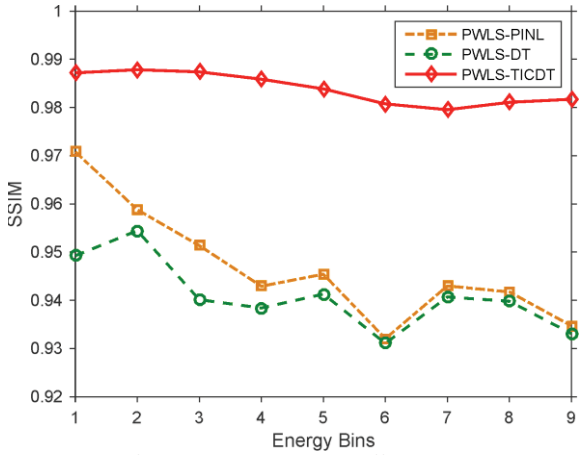


Fig. 5. Curves of the SSIM values versus different energy bins on the ROI shown in Fig. 4.

D. Material Decomposition

To evaluate the performance of the propose PWLS-TICDT method for material decomposition, the reconstructed images were decomposed into three basis materials, i.e., adipose tissue, striated muscle, and cortical bone. The decomposed images from phantom, FBP, PWLS-PINL, PWLS-DT, and PWLS-TICDT images are presented in Fig. 6. To quantitatively evaluation the decomposed images, the SSIM values of each material image are shown in Fig. 7. The PWLS-TICDT method has the greatest SSIM value in all the basis materials, and the SSIM values decrease gradually form PWLS-PINL/PWLS-DT method to FBP algorithm. This indicates that the high quality of PWLS-TICDT reconstruction produces more accurate material decomposition results.

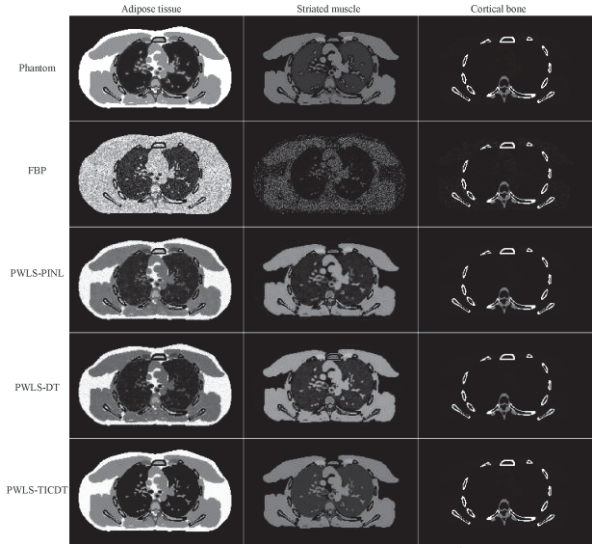


Figure 6. The basis material decomposition images from the XCAT phantom. All images are displayed in the same window: [0, 1].

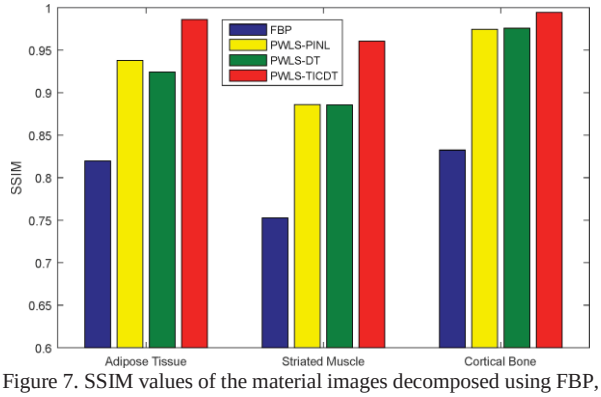


Figure 7. SSIM values of the material images decomposed using FBP, PWLS-PINL, PWLS-DT, and PWLS-TICDT results in Fig. 6.

IV. Conclusion

In this work, we presented an iterative PWLS-TICDT method to enhance image quality of PCCT using narrow energy bins. Based on PWLS principle, the presented PWLS-TICDT can explicitly exploit the redundant information in the energy domain. This redundant information is embedded into reconstruction by rotating and scaling the combined diffusion tensor of high-quality total image and noisy target image. The PWLS-

TICDT method can detect the matching normal and tangential directions of diffusion tensor fields in total image and target image that results in substantial noise reduction in narrow energy bin images without sacrificing edges detail and spatial resolution, as demonstrated in results section.

REFERENCES

1 Schmidt, T.G., Zimmerman, K.C., and Sidky, E.Y.: 'The effects of extending the spectral information acquired by a photon-counting detector for spectral CT', Physics in medicine and biology, 2015, 60, (4), pp. 1583-1600

2 Shikhaliev, P.M.: 'Energy-resolved computed tomography: first experimental results', Physics in medicine and biology, 2008, 53, (20), pp. 5595-5613

3 Leng, S., Yu, L., Wang, J., Fletcher, J.G., Mistretta, C.A., and McCollough, C.H.: 'Noise reduction in spectral CT: reducing dose and breaking the trade-off between image noise and energy bin selection', Medical physics, 2011, 38, (9), pp. 4946-4957

4 Niu, S., Yu, G., Ma, J., and Wang, J.: 'Nonlocal low-rank and sparse matrix decomposition for spectral CT reconstruction', Inverse problems, 2018, 34, (2), pp. 024003

5 Yu, Z., Leng, S., Li, Z., and McCollough, C.H.: 'Spectral prior image constrained compressed sensing (spectral PICCS) for photon-counting computed tomography', Physics in medicine and biology, 2016, 61, (18), pp. 6707-6732

6 Niu, S., Huang, J., Bian, Z., Zeng, D., Chen, W., Yu, G., Liang, Z., and Ma, J.: 'Iterative reconstruction for sparse-view x-ray CT using alpha-divergence constrained total generalized variation minimization', Journal of X-ray science and technology, 2017, 25, (4), pp. 673-688

7 Niu, S., Gao, Y., Bian, Z., Huang, J., Chen, W., Yu, G., Liang, Z., and Ma, J.: 'Sparse-view x-ray CT reconstruction via total generalized variation regularization', Physics in medicine and biology, 2014, 59, (12), pp. 2997-3017

8 Xu, Q., Yu, H., Bennett, J., He, P., Zainon, R., Doesburg, R., Opie, A., Walsh, M., Shen, H., Butler, A., Butler, P., Mou, X., and Wang, G.: 'Image reconstruction for hybrid true-color micro-CT', IEEE transactions on bio-medical engineering, 2012, 59, (6), pp. 1711-1719

9 Zeng, D., Gao, Y., Huang, J., Bian, Z., Zhang, H., Lu, L., and Ma, J.: 'Penalized weighted least-squares approach for multienergy computed tomography image reconstruction via structure tensor total variation regularization', Computerized medical imaging and graphics : the official journal of the Computerized Medical Imaging Society, 2016, 53, pp. 19-29

10 Zhao, B., Gao, H., Ding, H., and Molloi, S.: 'Tight-frame based iterative image reconstruction for spectral breast CT', Medical physics, 2013, 40, (3), pp. 031905

11 Wang, J., Lu, H., Liang, Z., Eremina, D., Zhang, G., Wang, S., Chen, J., and Manzione, J.: 'An experimental study on the noise properties of x-ray CT sinogram data in Radon space', Physics in medicine and biology, 2008, 53, (12), pp. 3327-3341

12 Wang, J., Li, T., and Xing, L.: 'Iterative image reconstruction for CBCT using edge-preserving prior', Medical physics, 2009, 36, (1), pp. 252-260

13 Weickert, J.: 'Anisotropic diffusion in image processing' (Teubner Stuttgart, 1998. 1998)

14 Kazantsev, D., Ourselin, S., Hutton, B.F., Dobson, K.J., Kaestner, A.P., Lionheart, W.R.B., Withers, P.J., Lee, P.D., and Arridge, S.R.: 'A novel technique to incorporate structural prior information into multi-modal tomographic reconstruction', Inverse problems, 2014, 30, (6), pp. 065004

15 Hernandez, A.M., and Boone, J.M.: 'Tungsten anode spectral model using interpolating cubic splines: unfiltered x-ray spectra from 20 kV to 640 kV', Medical physics, 2014, 41, (4), pp. 042101

16 Segars, W.P., Sturgeon, G., Mendonca, S., Grimes, J., and Tsui, B.M.: '4D XCAT phantom for multimodality imaging research', Medical physics, 2010, 37, (9), pp. 4902-4915

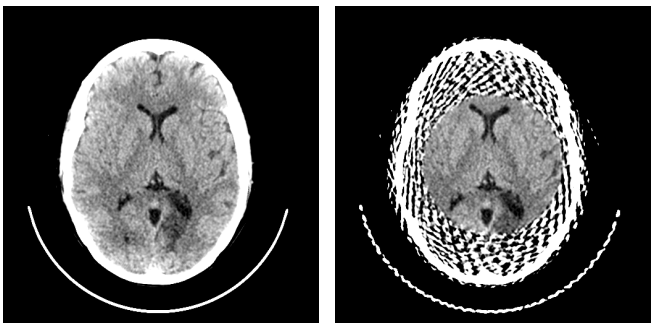
17 Zeng, D., Huang, J., Zhang, H., Bian, Z., Niu, S., Zhang, Z., Feng, Q., Chen, W., and Ma, J.: 'Spectral CT Image Restoration via an Average Image-Induced Nonlocal Means Filter', IEEE transactions on bio-medical engineering, 2016, 63, (5), pp. 1044-1057

18 Zhang, H., Huang, J., Ma, J., Bian, Z., Feng, Q., Lu, H., Liang, Z., and Chen, W.: 'Iterative reconstruction for x-ray computed tomography using prior-image induced nonlocal regularization', IEEE transactions on bio-medical engineering, 2014, 61, (9), pp. 2367-2378

Parallel-beam ROI reconstruction with differentiated backprojection and angularly subsampled complementary sinograms

Aymeric Reshef, Tina Nikoukhah, Cyril Riddell, Yves Trousset, Saïd Ladjal, and Isabelle Bloch

***Abstract*—Recently, we introduced a parallel-beam two-pass analytical reconstruction that allows truncation to be accounted for in the image domain rather than the projection domain. In particular, we showed that backprojection of a vastly angularly undersampled sinogram of un-truncated data could be used to extrapolate the backprojection of a finely sampled, fully truncated sinogram of the same object to perform more accurate region-of-interest (ROI) imaging. The same extrapolation idea can be performed using differentiated backprojection (DBP). The goal of this study is to give a general DBP-based formula when reconstructing a finite set of projections in parallel geometry. We discuss the discretization of this formula, in particular when the image grid size is large with respect to the number of projections, and we show how it can be applied to our extrapolation problem.**



(a) Reference image (b) With 45 un-truncated projections

Fig. 1: Dual-rotation reconstruction with a two-pass Hilbert-transformed DBP (DBP-HT-2) [3]. Window width: 50 HU.

I. Introduction

We address the classic case of interior tomography, when truncation due to the limited size of the detector defines a centered region-of-interest (ROI) of an object for which no complete projection is available. Without additional a priori knowledge, exact reconstruction of this ROI is not possible from the truncated data. Even though the amount of required a priori data may be considered tiny [1], it is still usually unavailable. A second approach is to complement the truncated data with a second acquisition. In the context of very high-resolution imaging, only the ROI is scanned at very high resolution, because the cost of scanning the entire object is prohibitive either in terms of scanning time, dose, or both. A second acquisition encompassing the full object at standard resolution removes the truncation artifacts through extrapolation of the truncated high-resolution data with the lower-resolution complete data. In the context of C-arm CBCT, we proposed a dual-rotation scheme, where a complementary acquisition of un-truncated data at same resolution but angularly vastly undersampled removes truncation artifacts using a standard least-square iterative reconstruction [2]. The iterative approach makes full use of the two sets of data without any explicit data extrapolation. For an analytical alternative, the extrapolation must be explicit and we showed that it is most conveniently achieved in the image space after backprojection of both data sets [3] (Fig. 1). Here, we look at the specific case of extrapolating the differentiated backprojection (DBP) within the context of a general DBP-based reconstruction formula for parallel-beam geometry.

The general DBP formula is presented in Section II, along with the DBP-based extrapolation for ROI imaging. The specific case of using standard DBP images is studied in Section III. Simulations on forward projections of a clinical head CT slice are described in Section IV. The results are shown in Section V.

A. Reshef (corresponding author: aymeric.reshef@ge.com) is with GE Healthcare, Buc, France, and LTCI, Télécom ParisTech, Université Paris-Saclay, Paris, France. T. Nikoukhah, C. Riddell and Y. Trousset are with GE Healthcare, Buc, France. S. Ladjal and I. Bloch are with LTCI, Télécom ParisTech, Université Paris-Saclay, France. This work was supported by the CIFRE grant No. 873/2014 from the French Association Nationale de la Recherche et de la Technologie (ANRT).

II. General DBP reconstruction formula

A. Notations

We parametrize the orientation of a parallel beam by its direction $\boldsymbol{\theta} = (\cos\theta, \sin\theta)^T$, where $\theta \in [0, \pi]$. The detector is orthogonal to the beam and oriented along $\boldsymbol{\theta}^\perp = (-\sin\theta, \cos\theta)^T$. Any point $\boldsymbol{x}$ is thus related to its detector coordinate $u_\theta(\boldsymbol{x})$ by equation $u_\theta(\boldsymbol{x}) = \boldsymbol{x} \cdot \boldsymbol{\theta}^\perp$. The Hilbert transform of a 1D signal is denoted $\mathcal{H}$, and we denote $\mathcal{H}_\alpha$, the 2D filter that applies $\mathcal{H}$ over all the lines colinear to $\boldsymbol{\alpha}^\perp$. We denote p_θ, the projection available at angle θ, and p'_θ the differentiation of p_θ along $\boldsymbol{\theta}^\perp$. We denote $\mathcal{B}_\theta$ the backprojection operator from angle θ, defined as: $\mathcal{B}_\theta[q](\boldsymbol{x}) = q(u_\theta(\boldsymbol{x}))$ for any projection q. For a set of angular positions Θ, we define p'_Θ as the collection $\{p'_\theta\}_{\theta\in\Theta}$. We further write $\sigma^\alpha_\Theta \otimes p'_\Theta = \{\mathrm{sgn}(\boldsymbol{\alpha}\cdot\boldsymbol{\theta})\cdot p'_\theta\}_{\theta\in\Theta}$, and $\mathcal{B}_\Theta[p'_\Theta] = \int_\Theta \mathcal{B}_\theta[p'_\theta]\,\mathrm{d}\theta$.

We consider a finite set of projections over interval $[0, \pi]$. The interval is split into N angular sectors of aperture $\Delta\theta = \frac{\pi}{N}$. We denote $\Theta = \left\{\theta_n = \left(n - \frac{1}{2}\right)\Delta\theta, n = 1, \cdots, N\right\}$ the set of acquired angular positions.

B. Splitted DBP formula

Given a partition $\{\Theta_k\}_{k=1,\cdots,K}$ of Θ, one can reconstruct image f through:

$$\begin{aligned} f &= \frac{1}{2\pi}\sum_{k=1}^{K}\mathcal{H}_{\alpha_k}\mathcal{B}_{\Theta_k}\left[\sigma_{\Theta_k}^{\alpha_k}\otimes p'_{\Theta_k}\right] \\ &= \frac{1}{2\pi}\sum_{n=1}^{N}\mathcal{B}_{\theta_n}\mathcal{H}\left[p_{\theta_n}\right], \end{aligned} \tag{1}$$

where $\alpha_k \in [0,\pi]$ is such that $\boldsymbol{\alpha}_k \cdot \boldsymbol{\theta} = \cos(\theta-\alpha_k) \neq 0$ for all $\theta \in \Theta_k$. The most right handside is the standard filtered backprojection (FBP) and the result holds because one can replace Hilbert transform $\mathcal{H}$ prior to backprojection by $\mathcal{H}_{\alpha_k}$ after backprojection (see Appendix). If all partition sets Θ_k are singletons ($K=N$), we have one filtering direction per projection, typically $\boldsymbol{\theta}_k^{\perp}$, as in standard FBP. On the other hand, if $K=1$, the only admissible filtering direction is given by $\alpha_1 = \frac{\pi}{2}$, resulting in a filtering step along the x-axis, as in the single-pass Hilbert-transformed DBP (DBP-HT-1) algorithm [4].

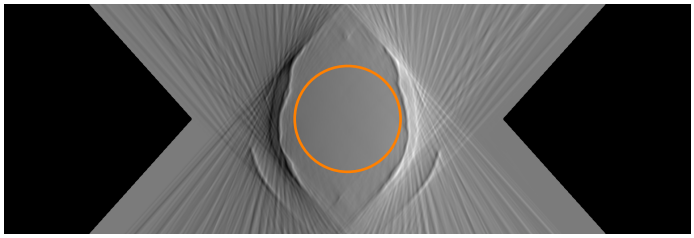
For $K=2$, we divide Θ into frontal views and lateral views. Frontal views correspond to $\Theta_{\mathrm{FRT}} = \Theta \cap \left[\frac{\pi}{4},\frac{3\pi}{4}\right]$. We select $\alpha_{\mathrm{FRT}} = \frac{\pi}{2}$, yielding a horizontal filtering along the y-axis. Lateral views correspond to $\Theta_{\mathrm{LAT}} = \Theta \setminus \Theta_{\mathrm{FRT}}$. We select $\alpha_{\mathrm{LAT}} = 0$, yielding a vertical filtering along the y-axis. We call this reconstruction method the two-pass Hilbert-transformed DBP (DBP-HT-2). This common formula proves that DBT-HT-2, DBP-HT-1, and FBP, applied to a finite set of projections, yield the same reconstruction.

C. DBP extrapolation for ROI imaging

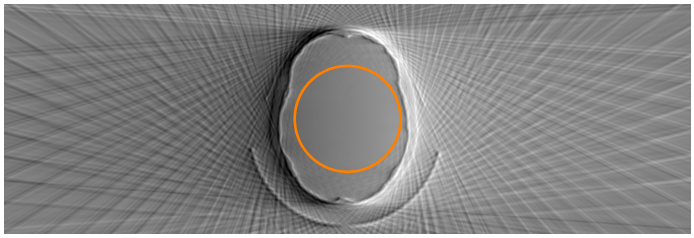
When considering two acquisitions from the same detector with different angular sampling but equal detector resolution, the complete data cannot be used to directly extrapolate the truncated ones because such data are not available for all angles. Alternatively, the DBP of each sinogram correspond to two images: one that is uniformly sampled but at a low rate, and one that is finely sampled but not uniformly over the image space due to truncation of the projections. In Fig. 2, we illustrate the merge of both DBP images. The central field of view (FOV) corresponds to the ROI (area inside the orange circle); it is made of the finely sampled DBP where the sampling is also uniform. Outside the ROI, the coarsely sampled DBP is used instead because it is uniformly sampled. At the edges of the ROI, a smooth transition is ensured using the radial weighting function:

$$\eta(r) = \frac{1}{2}\left(1-\cos\left(\pi\cdot\frac{r-r_\Omega}{\Delta r}\right)\right), \tag{2}$$

where r_Ω denotes the radius of the ROI, and Δr is the transition zone radial width. The DBP extrapolation actually consists of K extrapolations, one per partition subset. Fig. 2a illustrates such extrapolation for the frontal views of DBP-HT-2. Hilbert-filtering for all rows crossing the ROI will not be truncated in this case, since these rows have a finite support. In contrast, when using $K=1$ and DBP-HT-1 (Fig. 2b), all Hilbert lines are truncated.

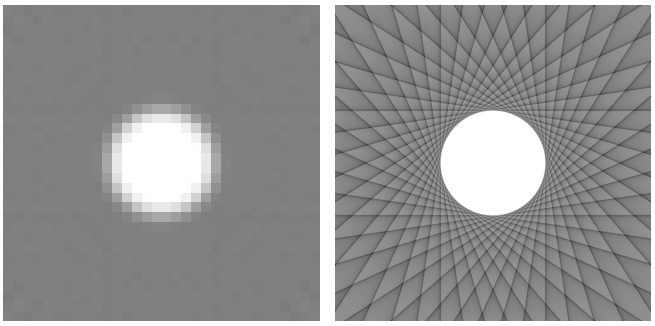


(a) DBP-HT-2, $\Theta_2 = \left[\frac{\pi}{4},\frac{3\pi}{4}\right]$



(b) DBP-HT-1, $\Theta_1 = [0,\pi]$

Fig. 2: Dual-rotation DBP extrapolation for ROI imaging. (a) The Hilbert transform can be computed for all horizontal lines crossing the circular ROI (orange circle). (b) The Hilbert transform is truncated for all horizontal lines crossing the ROI.



(a) $N=22$, $M=32$ (b) $N=22$, $M=1024$

Fig. 3: Angular sampling, resolution and image support.

III. ROI IMAGING WITH DBP-HT-1

A. Discretization

Reconstruction from a finite set of projections means estimating the image from a finite set of angular samples in the Fourier space according to the central-slice theorem. Because the images have a finite support, the finite Fourier sampling leads to sampling artifacts, which depends on the defined sampling of the images themselves, typically two-dimensional grids of size $M \times M$. This point is illustrated in Fig. 3 where the image is reconstructed from a set of $N=22$ projections using Eq. (1) with $K=N$ over two different grids ($M=32$ and $M=1024$).

When comparing both images, we note the classical streak artifacts on the right high-resolution image that are not visible on the left. A second consideration needs to be mentioned: the sampling artifacts cover the whole grid beyond the object and are actually truncated by the finite size of the image grid. In reality, the issue is the same on both images, but with a different intensity.

B. Undersampling

The general formula in Eq. (1) does not lead to equivalent results for any image f because the computation of the Hilbert transform is sensitive to truncation. When using $K > 1$, it is always possible to define filtering directions $\boldsymbol{\alpha}_k^{\perp}$ such that the Hilbert transform $\mathcal{H}_{\alpha_k}$ is applied to non-truncated signals if the original projections are not truncated. For the case $K = 1$, this is not true anymore, but if one assumes that the result of the Hilbert transform has a finite support, then the truncated Hilbert transform can be applied [4]. Since, as already mentioned, the reconstruction of an image from a finite set of projections has an infinite support, DBP-HT-1 cannot be computed. In practice, there exists a large enough number of projections such that the sampling artifacts can be neglected and the support considered finite. In the case of significant under-sampling ($N \ll M$), the sampling artifacts are significant over the whole reconstruction grid and beyond. However, one solution lies in Fig. 3 itself: there exists a coarser reconstruction grid of sampling $M' \approx N$, or equivalently, there exists a low-pass version of the DBP that can be used to reconstruct a low-pass version of f.

IV. Simulations

A diagnostic CT slice of a brain (Fig. 1a) was forward-projected to simulate 720 noise-free parallel-beam projections of 576 bins, sampling uniformly $[0, \pi]$. Truncated data p_{T} consist of the 720 projections with a centered digital truncation corresponding to a centered, circular 2D region-of-interest Ω of diameter 256 pixels. Full-FOV projections $p_{\mathrm{F},s}$ consist of N_{F} samples of the 720 original projections, again uniformly distributed over $[0, \pi]$. We write $N_{\mathrm{F}} = 720/s$, where s is the down-sampling factor; we used $s = 2^q$ with q integer varying from 0 ($s = 1$) to 7 ($s = 128$). The images are sampled on a 512×512 square grid. In order to use DBP-HT-1, the DBP of each sinogram is computed, that of $p_{\mathrm{F},s}$ being further filtered by 2D Gaussian filtering parameterized by its standard deviation σ, that is varied from 0 (no smoothing) to 40 pixels by steps of 5 pixels. The DBP are merged using the weighting function of Eq. (2) with $\Delta r = 15$ pixels. All Hilbert transforms are computed as proposed in [5]. The image quality was assessed by the mean relative error (MRE) between each reconstructed image $f_{\sigma,s}$ with respect to the reference image $f_{\mathrm{ref}} = f_{0,1}$, computed over Ω, as:

$$\mathrm{MRE}_{\sigma,s} = \frac{1}{\mathrm{Card}(\Omega)} \sum_{\boldsymbol{x} \in \Omega} \frac{|f_{\sigma,s}(\boldsymbol{x}) - f_{\mathrm{ref}}(\boldsymbol{x})|}{|f_{\mathrm{ref}}(\boldsymbol{x})|}. \qquad (3)$$

V. Results

The mean relative errors are shown in Fig. 4. For each value of s, the optimal smoothing parameter σ^* achieving the minimum MRE is indicated with a green dot. These values are also recalled in Table I. Although for $N_{\mathrm{F}} \geq 90$ ($s \leq 16$) the minimum MRE values seem to follow a smooth curve on the plane (σ, s), they are significantly higher and towards much higher σ values when N_{F} drops to as few as 22 ($s \geq 32$). This shows that filtering is not sufficient when s is too high.

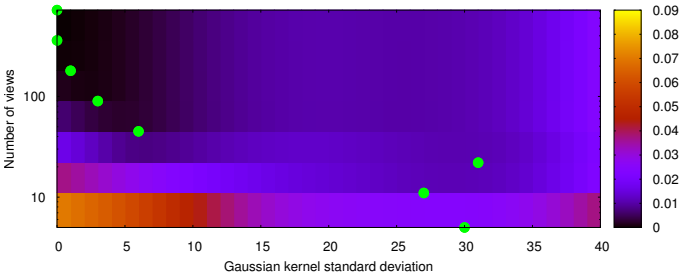


Fig. 4: MRE inside the ROI as a function of angular sampling and of the Gaussian smoothing filter.

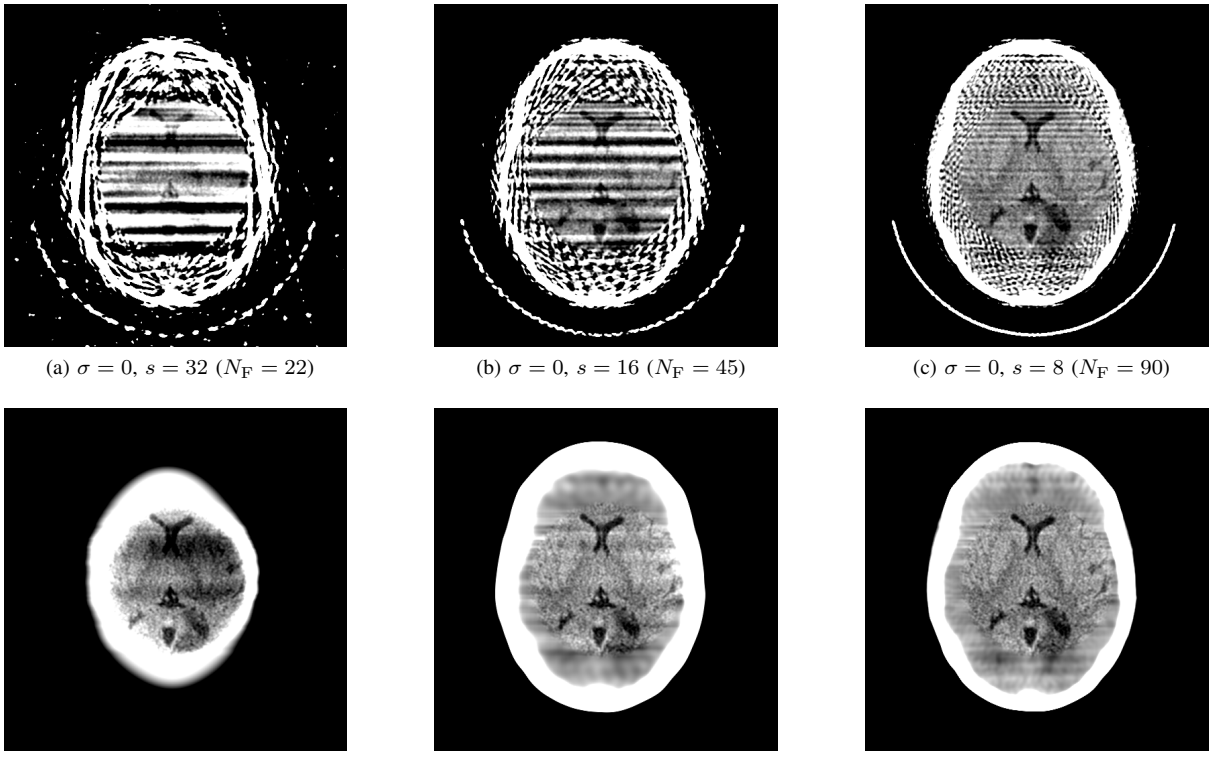
s	N_{F}	σ^*	$\mathrm{MRE}_{\sigma^*,s}$ (%)
1	720	0	0.00
2	360	0	$8 \cdot 10^{-3}$
4	180	1	0.04
8	90	3	0.10
16	45	6	0.30
32	22	31	1.10
64	11	27	1.11
128	5	30	3.72

Table I: Optimal smoothing parameter σ and corresponding MRE for each angular subsampling ratio s.

Reconstructed images with the optimal values σ^* according to the MRE criterion are shown in Fig. 5, that compares reconstruction $f_{0,s}$, that is, without Gaussian smoothing, to $f_{\sigma^*,s}$. For all values of $s > 2$, the absence of Gaussian smoothing resulted in reconstructed ROI suffering from horizontal streaks due to the invalid inversion of the Hilbert transform, even with 90 full-FOV projections. For $s < 32$, the optimal smoothing strongly reduced the impact of horizontal streaks in the ROI, resulting in values of the MRE of 0.30% for $s = 16$ (Fig. 5e) and 0.10% for $s = 8$ (Fig. 5f). For $s = 32$ (Fig. 5d), the strong Gaussian smoothing is shown to reduce the effect of the horizontal streaks, but the reconstructed image suffers from residual low-frequency non-uniformities. For higher values of s this issue worsened (images not shown).

VI. Discussion

In this study, we introduced a general formula for DBP-based reconstruction in parallel-beam geometry and applied it to solve the interior problem with a few extra full-FOV projections. This led us to investigate the case of reconstructing vastly angularly subsampled acquisitions with DBP-HT-1. Our study shows that because the key requirement of a finite support is not met, this algorithm cannot be used "as is". An alternative two-pass method (DBP-HT-2) does not suffer from this issue. However, we showed that subsampling could be mitigated by Gaussian smoothing of the areas outside the ROI. This is important because DBP-HT-1 is the only algorithm covered by our formula that allows for DBP extrapolation when reconstructing a single line, whereas DBP-HT-2 is applicable to reconstructing the full ROI only.



(a) $\sigma = 0$, $s = 32$ ($N_{\mathrm{F}} = 22$) (b) $\sigma = 0$, $s = 16$ ($N_{\mathrm{F}} = 45$) (c) $\sigma = 0$, $s = 8$ ($N_{\mathrm{F}} = 90$)

(d) $\sigma = 31$, $s = 32$ ($N_{\mathrm{F}} = 22$) (e) $\sigma = 6$, $s = 16$ ($N_{\mathrm{F}} = 45$) (f) $\sigma = 3$, $s = 8$ ($N_{\mathrm{F}} = 90$)

Fig. 5: Dual-rotation reconstruction with DBP extrapolation, $K = 1$, no smoothing (top) and optimal smoothing (bottom) of the full-FOV projections. Window width: 50 HU.

APPENDIX
PROOF OF EQ. (1)

Let $\theta \in [0, \pi]$ and $b_\theta = \mathcal{B}_\theta \left[p'_\theta\right]$. Let $\boldsymbol{x} \in \mathbb{R}^2$, and $\alpha \in [0, \pi]$ such that $\boldsymbol{\theta} \cdot \boldsymbol{\alpha} \neq 0$. We write:

$$b_\theta(\boldsymbol{x}) = p'_\theta(\boldsymbol{x} \cdot \boldsymbol{\theta}^\perp) = p'_\theta\left(A(\boldsymbol{x} \cdot \boldsymbol{\alpha}^\perp) + B\right), \tag{4}$$

where $A = \boldsymbol{\theta} \cdot \boldsymbol{\alpha} \neq 0$ and $B = (-\boldsymbol{x} \cdot \boldsymbol{\alpha})\boldsymbol{\theta} \cdot \boldsymbol{\alpha}^\perp$. The Fourier transform of $\mathcal{A}_{A,B}[p'_\theta] : u \mapsto p'_\theta(Au + B)$ is related to the Fourier transform of p'_θ through:

$$\mathcal{F}\left[\mathcal{A}_{A,B}[p'_\theta]\right](\rho) = \frac{1}{|A|}\mathcal{F}\left[p'_\theta\right]\left(\frac{\rho}{A}\right) e^{2i\pi\frac{\rho}{A}B}. \tag{5}$$

Hence, applying $\mathcal{H}_\alpha$ to b_θ is equivalent to multiplying the right-hand side of Eq. (5) by $-i\,\mathrm{sgn}(\rho) = -i\,\mathrm{sgn}(A)\,\mathrm{sgn}\left(\frac{\rho}{A}\right)$, prior to taking the inverse Fourier transform:

$$\mathcal{H}_\alpha\left[b_\theta\right](\boldsymbol{x}) = \frac{1}{A}\int_{-\infty}^{+\infty} -i\,\mathrm{sgn}\left(\frac{\rho}{A}\right)\mathcal{F}\left[p'_\theta\right]\left(\frac{\rho}{A}\right) e^{2i\pi\frac{\rho}{A}(A\boldsymbol{x}\cdot\boldsymbol{\alpha}^\perp + B)}\mathrm{d}\rho. \tag{6}$$

Taking the change of variables $\rho' = \frac{\rho}{A}$ yields:

$$\mathcal{H}_\alpha\left[b_\theta\right](\boldsymbol{x}) = \mathrm{sgn}(A)\int_{-\infty}^{+\infty} -i\,\mathrm{sgn}\left(\rho'\right)\mathcal{F}\left[p'_\theta\right]\left(\rho'\right) e^{2i\pi\rho'(A\boldsymbol{x}\cdot\boldsymbol{\alpha}^\perp + B)}\mathrm{d}\rho'. \tag{7}$$

The right-hand side of Eq. (7) is equal to $\mathrm{sgn}(A)\mathcal{B}_\theta\mathcal{H}\left[p'_\theta\right]$. In other words:

$$\mathcal{H}_\alpha\mathcal{B}_\theta\left[\mathrm{sgn}(\boldsymbol{\theta} \cdot \boldsymbol{\alpha}) \cdot p'_\theta\right] = \mathcal{B}_\theta\mathcal{H}\left[p'_\theta\right]. \tag{8}$$

If $\Theta_k = \{\theta_{k,1}, \cdots, \theta_{k,\mathrm{Card}(\Theta_k)}\}$ is a partition subset of Θ, where $\theta_{k,1} < \cdots < \theta_{k,\mathrm{Card}(\Theta_k)}$, we can choose a common admissible angle α such that $\boldsymbol{\alpha} \cdot \boldsymbol{\theta} \neq 0$ and Eq. (8) holds for all $\theta \in \Theta_k$. We denote α_k this common value, so that:

$$\mathcal{H}_{\alpha_k}\mathcal{B}_{\Theta_k}\left[\sigma_{\Theta_k}^{\alpha_k} \otimes p'_{\Theta_k}\right] = \sum_{\theta\in\Theta_k}\mathcal{B}_\theta\mathcal{H}\left[p'_\theta\right]. \tag{9}$$

Summing the contributions of all the partition subsets yields:

$$\sum_{k=1}^{K}\mathcal{H}_{\alpha_k}\mathcal{B}_{\Theta_k}\left[\sigma_{\Theta_k}^{\alpha_k} \otimes p'_{\Theta_k}\right] = \sum_{\theta\in\Theta}\mathcal{B}_\theta\mathcal{H}\left[p'_\theta\right], \tag{10}$$

which concludes the proof.

REFERENCES

[1] H. Kudo, M. Courdurier, F. Noo, and M. Defrise, "Tiny a priori knowledge solves the interior problem in computed tomography," *Physics in Medicine and Biology*, vol. 53, no. 9, p. 2207, 2008.

[2] A. Reshef, C. Riddell, Y. Trousset, S. Ladjal, and I. Bloch, "Dual-rotation C-arm cone-beam computed tomography to increase low-contrast detection," *Medical Physics*, vol. 44, no. 9, pp. e164–e173, 2017.

[3] ——, "Divergent-beam backprojection-filtration formula with applications to region-of-interest imaging," in *Proceedings of the 5th CT Meeting*, Salt Lake City, USA, May 2018.

[4] F. Noo, R. Clackdoyle, and J. D. Pack, "A two-step Hilbert transform method for 2D image reconstruction," *Physics in Medicine and Biology*, vol. 49, no. 17, p. 3903, 2004.

[5] J. You and G. L. Zeng, "Explicit finite inverse Hilbert transforms," *Inverse Problems*, vol. 22, no. 3, p. L7, 2006.

Image Reconstruction with Two Native Focal Spots for z-Flying Focal Spot Tomography

Hongbin Guo

***Abstract*—z-flying focal spot technique is well-accepted for modern CT industry, which can increase the image slice sensitive profile (SSP) and reduce helical windmill artifacts. However it works only for limited FOV. This paper propose a new algorithm to get over this limitation by using the native two focal spots to perform the image reconstruction instead of combining the two sets of data. Thus the used geometry is valid everywhere in contrast to limit FOV. The algorithm has been validated with offset zFFS scans, for which the object is far from the iso center.**

***Index Terms*—z-Flying Focal Spot, Computer Tomography, Image reconstruction.**

I. Introduction

In X-ray computed tomography (CT) imaging systems, the x-ray tube generates high speed electrons from the filament negative cathode. These electrons fly toward the positive target anode, in which the energy of the electrons are converted to X-rays. In conventional CT scanners, the X-ray emits from one focal spot on the anode plate. For multi-row scanners, to increase the resolution and reduce or remove under-sampling related image artifacts the so-called "Flying focal spot(FFS)", i.e. the focal spot is periodically moved among certain given positions, can be employed. The in-plane focal spot motion, βFFS [1], can increase resolution of transverse planes while the motion in z-direction, zFFS, can increase axial resolution, [2], [3], [4]. This study focuses on the zFFS.

In regular single focal spot cone beam (CB) system, the sampling interval in z-direction is the same as detector height of each row. The practical detector height is restrained by production technology and cost. This physical limit may cause windmill of the helical scans when we scan high contrast region, [5]. zFFS strategy can increase the sampling rate in z-direction, thus it not only can boost the z-resolution but also reduce helical windmill artifacts, [2], [3].

The zFFS strategy has been proposed for about one decade, the CT scanner venders have produced products to implement the focal spot wobbling idea. To the best of our knowledge, the current image reconstruction methods for zFFS scanning treat the detector readings from the alternating two focal spots as interleave sampling, i.e. group the two sets of data to one set by interleaving the rows of each consecutive (odd and even) reading pair to build one sinogram with double number of rows. Then the combined data is used for image reconstruction by regular single focal spot geometry, either use native fan geometry or rebin the data to parallel geometry, [2], [3], [6]. This type of reconstruction method has two limitations:

The author is affiliated with FMI Medical Systems, Inc. Solon, OH 44139, USA. Communication email address: Hongbin.Guo@fmimaging.com

1) Small FOV: rebuilding the data sets by interleaving assumes that the rays from two focal spot are stacked alternatively in z-direction, which is true for limited FOV. Apparently this assumption does not hold for the voxels close to focal spots. In fact, the assumption only hold for FOV at bout 200 mm for most commercial scanners. This limitation can be seen clearly from Figure 2.
2) Inaccuracy: Not only for voxels out of the limited FOV the interleaved data may cause mistakes, for voxels within the limited FOV the interleave strategy also may introduce inaccuracy since the "perfect" equal spaced z-interleave for the combined data only happens at z-axis, the far the voxel from the z-axis the worse the violation of the equal space assumption.

In this study, we propose to use the native two focal spots to perform the image reconstruction without combine the two sets of data but keep the original focal spot geometry. By doing so we use the exact native geometry of each rays, thus the object to be reconstructed can include voxels beyond the limited FOV and the reconstructed images will be more accurate. The rest of the paper is organization as follows. In Section II we detail our algorithm step by step. The proposed algorithm is validated by physical head phantom scanning data with offset in Section III. Conclusions are presented in Section IV.

II. METHODS

A. Geometry description

This subsection will describe the zFFS geometry. X-ray tube is illustrated in Figure 1, where the high energy electrons are emitted from the cathode and "bent" by an electric field toward the focal spot. In normal single focal spot case, the electronics hit the target at f_0 while zFFS has two focal spots f_1 and f_2 on the anode. In zFFS data collection procedure, the electrons bomb f_1 and f_2 alternatively while the tube rotate about the iso of the scanner. The f_1 and f_2 are away from the virtual focal spot f_0 by $\pm\Delta R$ in y-direction and $\pm\Delta Z$ in z-direction. If f_0 has coordinate $(-R, 0)$ in YZ plane then f_1 and f_2 are located at

$$\begin{aligned} f_1: &\quad (-R_1, \Delta Z), \\ f_2: &\quad (-R_2, -\Delta Z), \end{aligned}$$

here

$$\begin{aligned} R_1 &= R + \Delta R, \\ R_2 &= R - \Delta R. \end{aligned}$$

Figure 2 plots all rays from focal spots f_1 and f_2 to one detector pilar, i.e. one channel and all rows. When plotting the figure the geometry of the Minfound Scinticare 64 row scanner is used, which has 64 row detectors each having height $T_{det} = 0.625mm$ along z-axis, thus the z-coverage at z-axis is 40mm, and cone angle about $\pm 2^o$. The focal to detector distance, S_{F2D}, and the iso to detector distance, S_{I2D}, are 950.42mm and 392.42mm respectively.

Let β_i be the tube position angle and γ_j the fan angle of rays from f_0 to detectors. The β_i are usually equal spaced so do the γ_j since the detector arc has f_0 as its circular center. However, for zFFS the fan angle distributions for focal spot f_1, $\gamma_j^{(1)}$, and focal spot f_2, $\gamma_j^{(2)}$, are not equal spaced since neither f_1 nor f_2 is the circular center of the detector arc.

From Figure 2, it is clear that

1) The interleave property of the two set 64 rays is only true for the points beyond the red curve, at which the two set rays cross each other. This is the reason that current zFFS reconstruction limits its FOV.
2) The interleaved rays are not equal spaced at z-direction. They are close to equal spaced around z-axis and crossed at the red curve and the detector plane.
3) When the image points close to the focal spots the interleave property is severely violated.

To create nearly equal spaced rays along z-axis we need to control the focal spot positions by grid potential such that

$$\Delta Z = \frac{T_{det}}{4} \cdot \frac{S_{F2D}}{S_{I2D}}.$$

Since the anode surface has the tilde angle ϕ, usually $7 - 9^o$, then

$$\Delta R = \Delta Z \cdot \arctan(\phi).$$

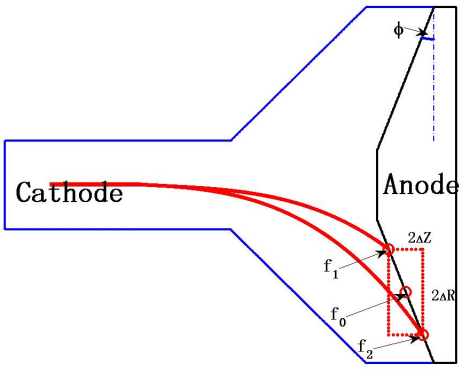


Fig. 1: X-ray tube and z-flying focal spot. For normal scan, electrons fly from the cathode towards the anode focal spot and X-rays are generated at the focal spot f_0. In z-flying focal spot scan electrons are controlled to hit two targets f_1 and f_2 at anode alternatively.

B. Rebinning Algorithm

The native focal spot based CB reconstruction is presented with an analytical type reconstruction algorithm, specifically FDK (Feldkamp - Davis - Kress) type algorithms. To simplify the problem we focus on circular cone beam (CCB) case only. In our algorithm we first perform fan beam to parallel beam rebinning for each focal spot separately. However, the azimuthal rebinning formula,

$$\theta = \beta + \gamma$$

for ideal focal spot f_0 does not hold for zFFS since the focal spots in zFFS scanning are not the center of the detector arc. Thus to perform rebinning we need to map the rays from deflected source, SA in Figure 3 for example, to rays emitted from ideal source circle with associated source position angles and fan angles. Figure 3 illustrates the relationship of this mapping for focal spot f_1. The angles are related with the following formula

$$\begin{aligned} \beta &= \beta_0 + \epsilon, \text{ and } \gamma = \gamma_0 - \epsilon, && \text{for focal spot} f_1, \\ \beta &= \beta_0 - \epsilon, \text{ and } \gamma = \gamma_0 + \epsilon, && \text{for focal spot} f_2. \end{aligned}$$

Here β_0 and γ_0 are angles associated with ideal focal spot S_0. The adjustment angle ϵ is determined by using the "law of sines",

$$\begin{aligned} \epsilon &= \pi - \gamma - \arcsin\left(\frac{R_1 \cdot \sin(\gamma)}{R}\right), \quad \text{for focal spot} f_1, \\ \epsilon &= \gamma - \arcsin\left(\frac{R_2 \cdot \sin(\gamma)}{R}\right), \quad \text{for focal spot} f_2. \end{aligned}$$

The radial rebinning is the same as normal focal spot data rebinning. Just need to pay more attention to the accurate fan angle, $\gamma_j^{(1)}$ and $\gamma_j^{(2)}$, calculation before using $t_1 = R_1 \sin(\gamma_j^{(1)})$ and $t_2 = R_2 \sin(\gamma_j^{(2)})$ since the fan angles are not equal spaced.

Finally, the target parallel angels and channels of rebinning two data sets should be the same (t, θ) so that the interpolation over the complete parallel data can be performed easily and accurately in backprojection step.

C. The reconstruction Algorithm

We present our the complete FDK type zFFS reconstruction algorithm for CCB data in Algorithm 1, the backprojection section, which is the core content of the proposed algorithm, is listed independently in Algorithm 2. The data sorting step 1c in Algorithm 2, i.e. row order associated with the two focal, may vary for voxels from different region. For voxels far from the sources at the given projection angle we will have the interleave order as used in conventional zFFS data merge strategy, however this is not the case for voxels close to sources, see Figure 2.

III. EXPERIMENTS AND RESULTS

The zFFS effects on SSP enhancement and helical windmill artifacts reduction for limited FOV have already be validated in literatures,[3], [4], we will focus on the feasibility of zFFS

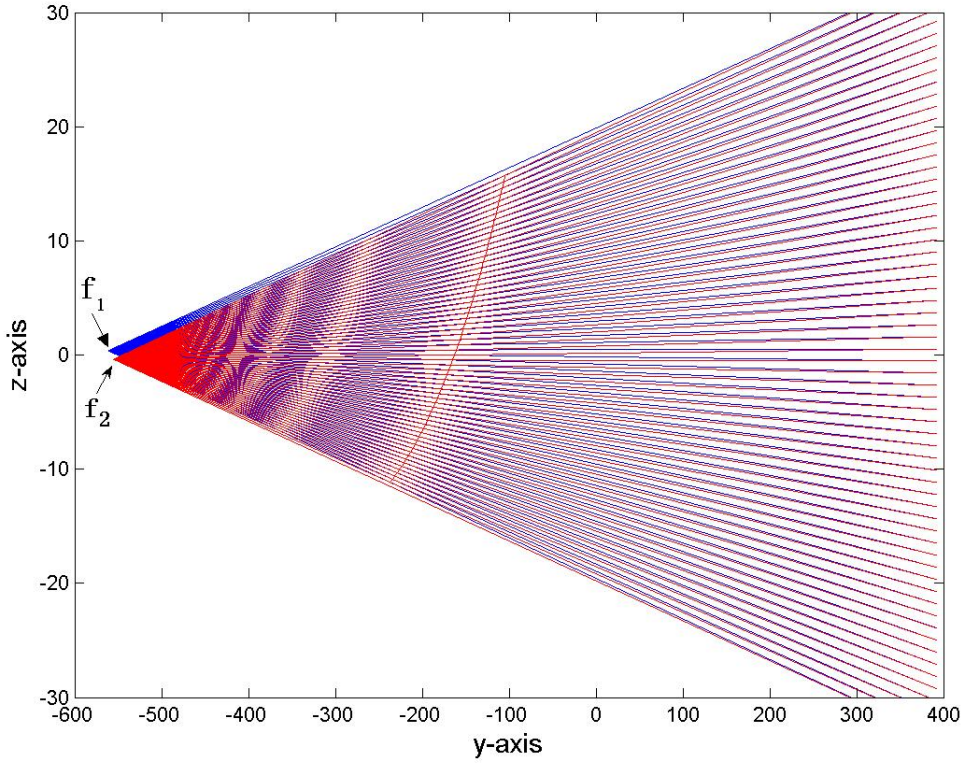


Fig. 2: The Rays emit from two focal spots and arrive to the same detector row pilar. The true geometry of the Minfound Scinticare 64 row scanner is used to plot these rays. The rays from focal spots f_1 and f_2 are interleaved for limited region: the portion beyond the red curve.

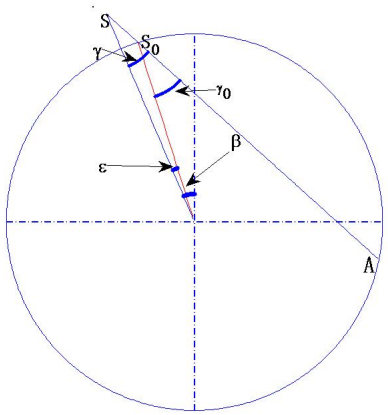


Fig. 3: The angle relationship for z-flying focal spot f_1 and the corresponding "ideal" focal spot f_0. β_0 and γ_0 are angles associated with ideal focal spot S_0, and β and γ are angles associated with focal spot S, i.e. f_1.

Algorithm 1 Reconstruction for zFFS CCB data by weighted rebinning FDK Algorithm.

1: Rebinning the zFFS CCB projection data $p_1(\gamma_1, v_1, \beta_1)$ and $p_2(\gamma_2, v_2, \beta_2)$ from fan geometry to cone parallel geometry with the same target channel/angle (t, θ) for each row v to produce $p_1(t, v_1, \theta)$ and $p_2(t, v_2, \theta)$.
2: Filtering the rebinned projection by kernel $h(t)$.
3: Applying cosine weight to obtain $q_1(t, v_1, \theta)$ and $q_2(t, v_2, \theta)$.
4: Performing weighted backprojection, Algorithm 2, to reconstruct the image volume.

reconstruction for large FOV or objects away from iso center by using the proposed zFFS reconstruction algorithm. Experimental data is obtained with the geometry of the Minfound Scinticare 64 row scanner, which has 64 row detectors each having height 0.625mm along z-axis, thus the z-coverage at z-axis is 40mm, and 840 channels at x-axis direction with arc angle 54.4^o. The scanner has maximum FOV 500mm, cone angle about $\pm 2^o$. Experimental scans are set with protocol: SAS, 120 kVp, 200mA and 1440 projections per rotation at a speed 1 second per rotation. 720 projections for two zFFS focal spots respectively.

Head phantom was scanned by CCB protocol with 100 mm **shift** from iso-center and the zFFS CCB data is reconstructed

Algorithm 2 Weighted backprojection for zFFS CCB data.

$q_1(t, v_1, \theta)$ and $q_2(t, v_2, \theta)$ are the parallel filtered projection data for two focal spots. Pseud-code of the weighted backprojection for every voxel $\mathbf{x}$ is listed below:

- For views $\theta_k \in [0, \pi), k = 1, 2, \cdots, K$
  1) For half turns $m = 0, 1$
     a) For the given voxel $\mathbf{x}$ determine the channel position $t_{\mathbf{x}}$
     b) Apply interpolation for rows of the two parallel data sets, close to the pixel $\mathbf{x}$, to generate $q_1(t_{\mathbf{x}}, v_1, \theta_k)$ and $q_2(t_{\mathbf{x}}, v_2, \theta_k)$.
     c) Build a single z-pilar vector by sorting the two parallel ray sets according to their z-coordinates of the intersection points with vertical line $x =$**x**(1), y= **x**(2).
     d) Perform z-interpolation at $z = \mathbf{x}(3)$ to obtain $q(t_{\mathbf{x}}, z, \theta_k)$
     e) Compute weight accordingly.
  2) Normalize weights over half turns.
  3) $f(\mathbf{x}) = f(\mathbf{x}) + \sum_{m=0}^{1} \mathbf{w}(m, k, \mathbf{x}) * q(t_{\mathbf{x}}, z, \theta_k)$, where $\mathbf{w}(0, k, \mathbf{x}) + \mathbf{w}(1, k, \mathbf{x}) = 1$.
- End For views

(a)

Fig. 4: Head phantom scan with zFFS protocol. The image is reconstructed by the proposed zFFS algorithm. The slice 104 of 128 slices is presented with window WL 40HU and WW 300HU.

with FOV 350mm with 100mm reconstruction center shift correspondingly. The total number of slices of the reconstructed images is 128 slices with separation 0.3125mm. The sinogram and reconstructed image are not processed with complete calibration/correction, e.g. no bone beam hardening correction. The focus of the experiment is the feasibility of the Algorithm 1 for large FOV reconstruction. The slice 104 is presented in Figure 4 with window WL 40HU and WW 300HU. Since we lower the patient table 100mm down before scanning the phantom, the bottom part is close to the sources when the tube spin to the range around 6 o'clock. In other words, the bottom part is out of the FOV "limitation" for the conventional zFFS reconstruction algorithm. The image is reconstructed by Algorithm 1 does not show strong artifacts at its bottom region where the pixels are close to the source.

IV. CONCLUSION

We proposed a new reconstruction algorithm for zFFS CCB scans. The proposed algorithm use the native two focal spots to perform the image reconstruction instead of combining the two sets of data. By doing so we use the exact native geometry of each rays, thus the object to be reconstructed can include voxels beyond the limited FOV. The algorithm has been validated with offset zFFS scan of a physical head phantom. The image reconstructed by Algorithm 1 does not show artifacts due to data misplacement for large FOV. This suggest that the proposed algorithm has overcome the FOV limitation for zFFS scans.

REFERENCES

[1] J. Hsieh, M. A. Gard, and S. Gravelle, "Reconstruction technique for focal spot wobbling," in *Medical Imaging VI: Image Processing*, 1992.

[2] M. Kachelriess, M. Knaup, C. Penssel, and W. A. Kalender, "Flying focal spot (ffs) in cone-beam ct," *IEEE Transactions on Nuclear Science*, vol. 53, no. 3, pp. 1238–1247, 2006.

[3] Y. Kyriakou, M. Kachelrieβ, M. Knaup, J. U. Krause, and W. A. Kalender, "Impact of the z-flying focal spot on resolution and artifact behavior for a 64-slice spiral ct scanner," *European Radiology*, vol. 16, no. 6, pp. 1206–1215, Jun 2006.

[4] T. G. Flohr, K. Stierstorfer, S. Ulzheimer, H. Bruder, A. N. Primak, and C. H. McCollough, "Image reconstruction and image quality evaluation for a 64-slice ct scanner with z-flying focal spot," *Medical Physics*, vol. 32, no. 8, pp. 2536–2547, 2005.

[5] K. Taguchi, H. Aradate, Y. Saito, I. Zmora, K. S. Han, and M. D. Silver, "The cause of the artifact in 4-slice helical computed tomography," *Medical Physics*, vol. 31, no. 7, pp. 2033–2037, 2004.

[6] I. HEIN and A. Zamyatin, "Interpolation interlacing based data upsampling algorithm for cone-beam x-ray ct flying focal spot projection data," May 8 2012, uS Patent 8,175,218. [Online]. Available: https://encrypted.google.com/patents/US8175218

Analytical Statistical Reconstruction Algorithm with the Direct Use of Projections Performed in Spiral Cone-beam Scanners

Robert Cierniak

***Abstract*—This paper is concerned with the originally formulated 3D statistical model-based iterative reconstruction algorithm for spiral cone-beam x-ray tomography. The conception proposed here is based on a formulation of the reconstruction problem as a shift invariant system. This algorithm significantly improves the quality of the subsequently reconstructed images, so allowing a decrease in the x-ray dose absorbed by a patient. The analytical roots of the algorithm proposed here permit a decrease in the complexity of the reconstruction problem in comparison with other model- based iterative approaches. In this paper, we proved that this statistical approach, originally formulated for parallel beam geometry, can be adapted for helical cone-beam geometry of scanner, with the direct use of projections. Computer simulations have shown that the reconstruction algorithm presented here outperforms conventional analytical methods with regard to the image quality obtained.**

I. Introduction

Nowadays, the most significant problem in medical computer tomography is the development of image reconstruction algorithms from projections which would enable the reduction of the impact of measurement noise on the quality of tomography images. This kind of approach is intended to improve image quality and, in consequence, reduce the dose of X-ray radiation while at the same time preserving an appropriate level of quality in the tomography images. The concept has found its application in the form of statistical reconstruction algorithms. One of the most interesting from the scietific and practical point of view, an approach, called MBIR (Model-Based Iterative Reconstruction), is presented in publications like [1], where a probabilistic model of the measurement signals is described analytically. The objective in those solutions was devised according to an algebraic scheme of the signal processing for formulating the reconstruction problem [2]. An algebraic scheme has been selected in this case for one very obvious reason - the measurement noise can be modelled relatively easily, because each measurement is considered separately. Such a scheme adds significant calculative complexity to the problem. The time for image reconstruction becomes difficut from the practical point of view. For instance, if the image resolution is assumed to be I×I pixels, the complexity of the algebraic problem is of the level of $\mathrm{I}\times\mathrm{I}\times number\_of\_measurements\times number\_of\_cross-sections$ (in 3D tomography); a multiple of I to the power of four in total.

Corresponding author: Robert Cierniak, Institute of Computational Intelligence, Czestochowa University of Technology, Armii Krajowej 36, 42-200 Czestochowa, Poland, e-mail: cierniak@kik.pcz.czest.pl

The difficulties mentioned above connected with the use of an algebraic methodology can be decreased by using an analytical strategy of reconstructed image processing. In previous papers we have shown how to formulate the analytical reconstruction problem consistent with the ML methodology for parallel scanner geometry [4]. This strategy has been used for fan-beams [3], and finally for the spiral cone-beam scanner [5]. However, an approach to the reformulation of the reconstruction problem from parallel to real scanner geometries, called rebinning, was applied there. Much more popular 3D reconstruction methods, which are implemented in practice, are FDK (Feldkamp)-type algorithms that use projections obtained from spiral cone-beam scanners directly (see e.g. [6]). In this paper, we present a mathematical derivation of a method for the direct (i.e. without rebinning) adaptation of spiral cone-beam projections to the statistical analytical reconstruction algorithm originally formulated by us.

II. Adaptation of the 2D Analytical Approximate Reconstruction Problem to Spiral Cone-beam Projections

A foundation for our conception of the model-based interative statistical algorithm is the 2D analytical approximate reconstruction problem which was originally formulated for a parallel scanner geometry [3] of scanner, as follows:

$$\mu_{\min}=\arg\min_{\mu}\left(\frac{n_0}{2}\sum_{i=1}^{\mathrm{I}}\sum_{j=1}^{\mathrm{J}}\cdot\right. \tag{1}$$
$$\left.\cdot\left(\sum_{\bar{i}}\sum_{\bar{j}}\mu\left(x_{\bar{i}},y_{\bar{j}}\right)\cdot h_{\Delta i,\Delta j}-\tilde{\mu}\left(x_i,y_j\right)\right)^2\right),$$

where coefficients $h_{\Delta i,\Delta j}$ are precaculated in the numerical way according to the following relation:

$$h_{\Delta i,\Delta j}=\Delta_\alpha\sum_{\psi=0}^{\Psi-1}int\left(\Delta i\cos\psi\Delta_\alpha+\Delta j\sin\psi\Delta_\alpha\right), \tag{2}$$

and $\tilde{\mu}\left(i,j\right)$ is an image obtained by way of a back-projection operation; $int\left(\Delta s\right)$ is an interpolation function used in the back-projection operation; every projection is carried out after a rotation by Δ_α.

Above presented shift-invariant system is much better conditioned than quadratic form used in algebraic aproaches [7],

and can be a starting point for the design of a 3D iterative reconstruction algorithm for spiral cone-beam scanner geometry. One of the principal reconstruction methods devised for the cone-beam spiral scanner is the generalized FDK algorithm. In the traditional FDK approach, the cone-beam projections are filtered and then back-projected in three dimensions. This methodology is adapted to our original iterative model-based reconstruction concept.

Taking into consideration the definition of the two-dimensional inverse Fourier transform, and the frequential form of the relation between the original image of a cross-section of an examined object represented by function $\mu(x,y)$ and the image obtained after the back-projection operation $\tilde{\mu}(x,y)$, we obtain:

$$\tilde{\mu}(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{|f|} M(f_1,f_2)\, e^{j2\pi(f_1 x + f_2 y)} df_1 df_2, \quad (3)$$

which, after converting to polar coordinates and using the projection slice theorem (taking into account a full revolution of the projection system), takes the form:

$$\tilde{\mu}(x,y) = \frac{1}{2}\int_{-\pi}^{\pi}\int_{-\infty}^{\infty} \bar{P}(f,\alpha^p)\, e^{j2\pi f(x\cos\alpha^p + y\sin\alpha^p)} df\, d\alpha^p. \quad (4)$$

Then, after transferring the projections into the spatial domain, and arranging the right hand side of the formula and changing the order of integration, we getwe have the formula:

$$\tilde{\mu}(x,y) = \frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\pi}^{\pi} \bar{p}^p(s,\alpha^p)\, e^{j2\pi f(x\cos\alpha^p + y\sin\alpha^p - s)} d\alpha^p ds df, \quad (5)$$

where $\bar{p}^p(s,\alpha^p)$ are projections obtained in a hypothetical parallel scanner (after interpolation).

Next, after converting the attenuation function into polar coordinates, we obtain:

$$\tilde{\mu}(r\cos\phi, r\sin\phi) = \frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\pi}^{\pi} \bar{p}^p(s,\alpha^p)\, e^{j2\pi f[r\cos(\alpha^p-\phi)-s]} d\alpha^p ds df. \quad (6)$$

In our considerations, we should also take into account the application of the interpolation function used during the back-projection operation, which should be placed appropriately (a frequency representation of this function) in the formula above, as follows:

$$\breve{\mu}(x,y) = \frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\pi}^{\pi} INT(f)\, p^p(s,\alpha^p)\, e^{j2\pi f[r\cos(\alpha^p-\phi)-s]} d\alpha^p ds df. \quad (7)$$

After suitable transformation we obtain a relationship for the fan-beam image reconstruction method:

$$\breve{\mu}(x,y) = \frac{\mathrm{R}_f}{2}\int_{0}^{2\pi}\int_{-\beta_m}^{\beta_m} p^f(\beta,\alpha^f) \frac{\cos\beta}{\dot{u}^2}\int_{-\infty}^{\infty} INT(f)\, e^{j2\pi f \dot{u}\sin(\dot{\beta}-\beta)} df\, d\beta\, d\alpha^f, \quad (8)$$

where $p^f(\beta,\alpha^f)$ are projections obtained in a hypothetical fan-beam scanner, and

$$\dot{u} = \left(x\cos\alpha^f + y\sin\alpha^f\right)^2 + \left(R_f + x\sin\alpha^f - y\cos\alpha^f\right)^2. \quad (9)$$

There is a several serious drawbacks associated with the use of the fan-beam reconstruction method formulated like this. It stems from the dependence of equation (9) on the parameter $\dot{u}$, which poses certain practical problems when carrying out the calculations during the reconstruction process. Instead of a simple formula for the convolution kernel, it now becomes necessary to determine a different form of the kernel for every point of the object's cross-section. This is because $\dot{u}$ represents the distance of the point (r,ϕ) from the radiation source. Therefore, by changing the angle α^f, we also change $\dot{u}$. The appropriate adjustment is based on a term in equation (9), which is reproduced here in a suitably amended form:

$$int(s) = \int_{-\infty}^{\infty} INT(f)\, e^{j2\pi f \dot{u}\sin(\dot{\beta}-\beta)} df. \quad (10)$$

In this equation, the integration is carried out with respect to the frequency f. The next step will be to make a substitution for f, using the following expression:

$$f^f = \frac{f\cdot\dot{u}\cdot\sin\beta}{\mathrm{R}_f\cdot\beta}. \quad (11)$$

If at the same time we change the limits of integration, the convolving function will be modified to:

$$int^f(\beta) = \frac{\mathrm{R}_f\cdot\beta}{\dot{u}\cdot\sin\beta}\int_{-\infty}^{\infty} INT\left(\frac{f^f\cdot f_0}{f_0^f}\right) e^{j2\pi f^f \mathrm{R}_f\beta} df^f, \quad (12)$$

where

$$f_0^f = \frac{f_0\cdot\dot{u}\cdot\sin\beta}{\mathrm{R}_f\cdot\beta}. \quad (13)$$

Unfortunately, even here we encounter problems caused by the dependence of the cut-off frequency f_0^f on the parameter $\dot{u}$. On the other hand, if we were to establish a constant value for f_0^f it would mean that the reconstruction process for the point (r,ϕ) would have a different resolution (determined by the value of the cut-off frequency f_0) for every angle α^f. However, if we put aside the assumption of uniform resolution for the resulting reconstructed image, then, by manipulating the values $\dot{u}$ and f_0, the varying value of f_0^f can be fixed as:

$$f_0^f = \mathrm{f}_0^f = \frac{1}{\mathrm{R}_f \cdot \Delta_\beta}. \tag{14}$$

Let us assume that we apply a linear interpolation function in formula (8). The frequency form of the linear interpolation function is given by this formula:

$$INT_L\left(f\right) = \frac{\sin^2\left(\pi f \Delta_s\right)}{\left(\pi f \Delta_s\right)^2}. \tag{15}$$

Taking into account in the formula (12) the proposed interpolation function given by (15), we obtain the following relation:

$$int_L^f\left(\beta\right) = \frac{\mathrm{R}_f \cdot \beta}{\dot{u} \cdot \sin\beta} \begin{cases} \frac{1}{\Delta_s'}\left(1 - \frac{\mathrm{R}_f|\beta|}{\Delta_s'}\right) & \text{for} \quad |\beta| \le \Delta_s' \\ 0 & \text{for} \quad |\beta| \ge \Delta_s' \end{cases}, \tag{16}$$

where $\Delta_s' = f_0 / f_0^f$, and next, bearing in mind relations (14), it leads immediately to:

$$int_L^f\left(\beta\right) = \frac{\beta}{\dot{u} \cdot \sin\beta} \begin{cases} \frac{\Delta_s}{\Delta_\beta}\left(1 - \frac{\Delta_s|\beta|}{\Delta_\beta}\right) & \text{for} \quad |\beta| \le \frac{\Delta_\beta}{\Delta_s} \\ 0 & \text{for} \quad |\beta| \ge \frac{\Delta_\beta}{\Delta_s} \end{cases}. \tag{17}$$

Finally, if we assume that $\Delta_s = 1$, it gives

$$int_L^f\left(\beta\right) = \frac{\beta}{\dot{u} \cdot \sin\beta} int_L\left(\beta\right), \tag{18}$$

where

$$int_L\left(\beta\right) = \begin{cases} \frac{1}{\Delta_\beta}\left(1 - \frac{|\beta|}{\Delta_\beta}\right) & \text{for} \quad |\beta| \le \Delta_\beta \\ 0 & \text{for} \quad |\beta| \ge \Delta_\beta \end{cases}. \tag{19}$$

In consequence, returning to the formula (9), we obtain

$$\breve{\mu}\left(x,y\right) = \frac{1}{2}\int_0^{2\pi}\int_{-\beta_m}^{\beta_m} p^f\left(\beta,\alpha^f\right) \frac{\mathrm{R}_f \cos\beta}{2\dot{u}} \frac{\Delta\beta}{\sin\Delta\beta} int_L\left(\Delta\beta\right) d\beta d\alpha^f. \tag{20}$$

Fortunately, we can linearize relation (21) by considering expressions inside the integration, namely $\frac{\Delta\beta}{\sin\Delta\beta}$.

In the case of linear interpolation we use only line of integrals from the neighborhood of a given pixel (x,y), then $\Delta\beta \le \Delta_\beta$, and $\sin\Delta\beta \approxeq \Delta\beta$. Additionally, it is possible to omit the term $\frac{\mathrm{R}_f \cos\beta}{2\dot{u}}$ taking into account the fact that each projection value $p^f\left(\beta,\alpha^f\right)$ has its equivalent $p^f\left(-\beta,\alpha^f+\pi+2\beta\right)$, as shown in Figure 1.

Because of this we can notice that the sum of this pair of projections is proportional to $\frac{\dot{u}_1+\dot{u}_2}{4\dot{u}_1} + \frac{\dot{u}_1+\dot{u}_2}{4\dot{u}_2} = \frac{(\dot{u}_1+\dot{u}_2)^2}{4\dot{u}_1\dot{u}_2}$. This means that for $\dot{u}_1 \approxeq \dot{u}_2$ this factor is equal to 1, and finally, we can write

$$\breve{\mu}\left(x,y\right) \approxeq \frac{1}{2}\int_0^{2\pi}\int_{-\beta_m}^{\beta_m} p^f\left(\beta,\alpha^f\right) int_L\left(\Delta\beta\right) d\beta d\alpha^f, \tag{21}$$

which is consistent with a form of the formula of the back-projection operation for parallel beams.

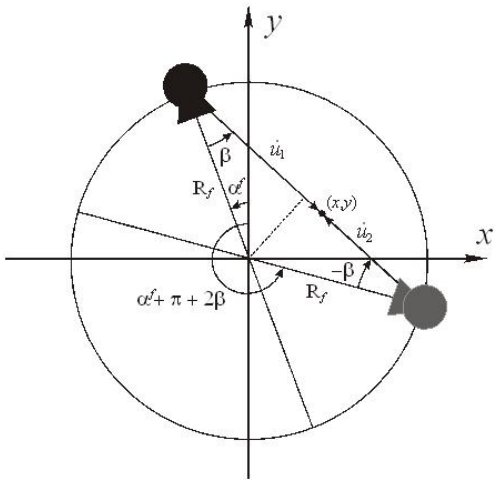


Fig. 1. Selecting complementary projection values

Moreover, if we assume that rays, i.e. integral lines defining $p^f\left(\beta,\alpha^f\right)$, from the hypothetical fan-beam geometry pass through almost the same tissues as rays from cone-beam geometry ($p^h\left(\beta,\alpha^h,z_k\right)$), the projection values associated with these rays will be related to the corresponding path lengths through the tissues. Because of this, we can derive the correction factor by using the following relation:

$$p^f\left(\beta,\alpha^f\right) = p^h\left(\beta,\alpha^h,z_k\right) CORR = p^h\left(\beta,\alpha^h,z_k\right) \frac{R_{fd}}{\sqrt{R_{fd}^2 + z_k^2}}, \tag{22}$$

where R_{fd} is the source-to-detector distance; z_k is the transverse position on the screen where a given ray is detected.

The geometry of this method of determining the correction factor is shown in Figure 2.

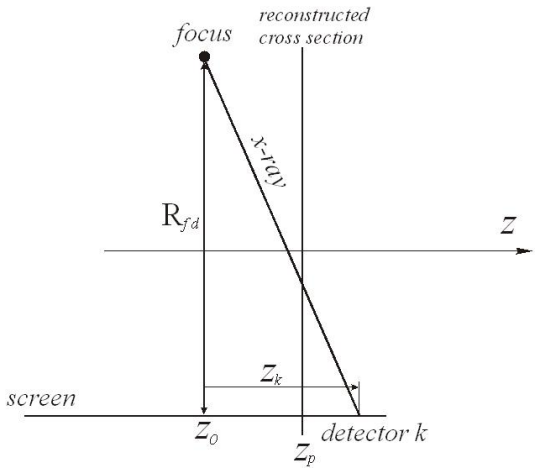


Fig. 2. The geometry of the cosine correction factor

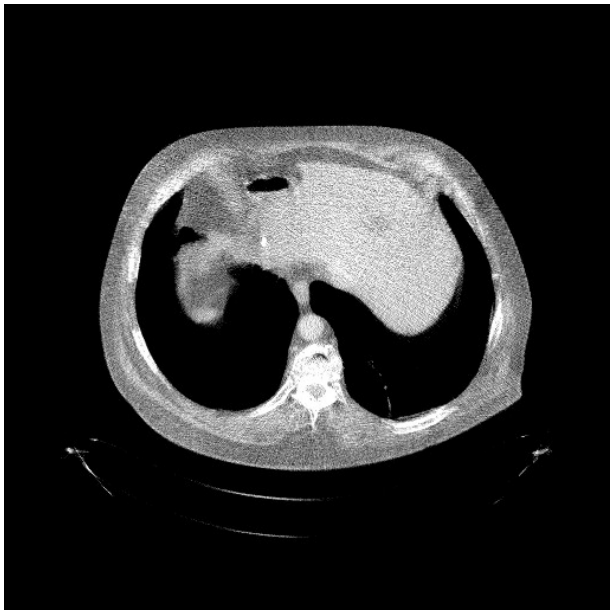
Finally, formula (21) can be used directly to obtain a reference image for the analytical statistical iterative reconstruction algorithm presented by the formula (2), which was originally formulated for parallel beam scanner geometry, as follows

$$\breve{\mu}(x,y) \approxeq \frac{1}{2}\int_{0}^{2\pi}\int_{-\beta_m}^{\beta_m} CORR{\cdot}p^h\left(\beta,\alpha^h,z_k\right) int_L\left(\Delta\beta\right) d\beta d\alpha^f. \tag{23}$$

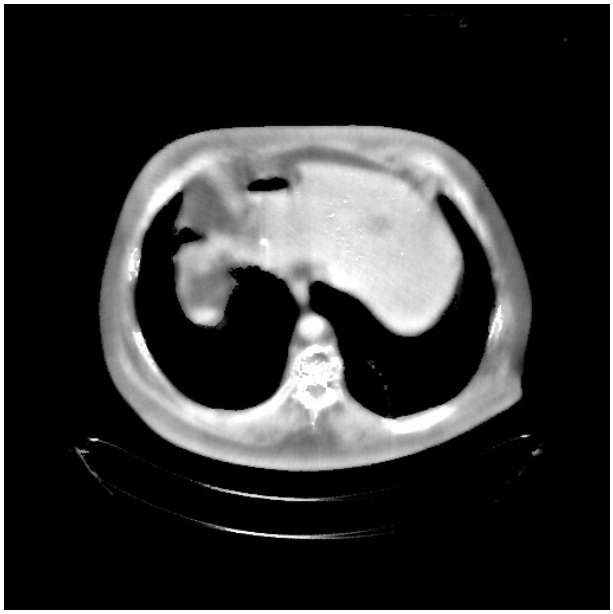
III. Experimental results

In our computer simulations, we have used projections obtained from a helical scanner Somatom Definition AS+ (Siemens Healthcare), with the following parematers: reference tube potentia 120kVp and quality reference effective 200mAs, $R_{fd} = 1085.6mm$ (SDD - Source-to-Detector Distance); $\mathrm{R}_f = 595\mathrm{mm}$ (SOD - Source-to-AOR Distance); number of views per rotation $\Psi = 1152$; number of pixels in detector panel 736; detector dimensions $1.09\mathrm{mm} \times 1.28\mathrm{mm}$. During the experiments, the size of the processed image was fixed at 512×512 pixels. The matrix of the coefficients $h_{\Delta i,\Delta j}$ were precomputed before the reconstruction process was started, and these coefficients were fixed for the subsequent processing. The image obtained after back-projection operation was then subjected to a process of reconstruction (optimization) using an iterative procedure. The starting point of this procedure was choosen as a result of using a reconstruction FBP algorithm. It is worth noting that our reconstruction procedure was performed without any regularization regarding the objective function described by (2).

View of the reconstructed images after 30000 iterations are presented (Table 3(a). For comparison, the image reconstructed by a standard FBP reconstruction method (Table 3(b)) is also presented.



(a)



(b)

Fig. 3. View of the images reconstructed image using the standard FBP (a); reconstructed image using the method described in this paper after 30000 iterations (b) ($C = 45$, $W = 600$)

IV. Conclusion

We have shown in this paper fully feasible statistical reconstruction algorithm for helical cone-beam scanner. It is proved that this statistical approach, originally formulated for parallel beam geometry, can be adapted for heica cone-beam geometry, without any filtration and any rebinning. Simulations have been conducted, which prove that our reconstruction method can be very fast (first of all thanks to the use of FFT algorithms) and gives satisfactory results with suppressed noise, without introducing any additional regularization term, using only an early stopping regularization strategy.

V. Acknowledgments

This work was partly supported by The National Centre for Research and Development in Poland (Research Project POIR.01.01.01-00-0463/17).

References

[1] Bouman, C., Sauer, K., *A unified approach to statistical tomography using coordinate descent optimization*, IEEE Tran. Image Processing, vol. 5, pp. 480–492, 1996.

[2] Thibault, J. -B., Sauer, K., Bouman, C., Hsieh, J., *A three-dimensional statistical approach to improved image quality for multislice helical CT*, Medical Physics, vol. 34, No. 11, pp. 4526–4544, 2007.

[3] Cierniak, R., *An analytical iterative statistical algorithm for image reconstruction from projections*, Applied Mathematics and Computer Science, vol. 24, pp. 7–17, 2014.

[4] Cierniak, R., *New neural network algorithm for image reconstruction from fan-beam projections*, Neurocomputing, vol. 72, pp. 3238–3244, 2009.

[5] Cierniak, R., *A three-dimensional neural network based approach to the image reconstruction from projections problem*, in: Rutkowski, L., Tadeusiewicz, R., Zadeh, L.A., urada, J. (eds.) LNCS, vol. 6113, pp. 505–514. Springer, Heidelberg, 2010.

[6] Cierniak, R., Knas, M., *An analytical statistical approach to the 3D reconstruction problem*, in Proc. of the 12th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine, pp. 521–524, July 2013.

[7] Cierniak, R., Lorent, A., *Comparison of algebraic and analytical approaches to the formulation of the statistical model-based reconstruction problem for x-ray computed tomography*, Computerized Medical Imaging and Graphics, vol. 52, pp. 19–27, 2016.

GPU-Based Tools for Multi-Channel X-ray CT Reconstruction

Darin P. Clark and Cristian T. Badea

***Abstract*—The established superiority of iterative reconstruction techniques in preclinical and clinical x-ray CT is heavily under-utilized in routine imaging applications. The primary reason for this disparity is the substantial increase in data processing time inherent with iterative reconstruction methods. This problem is further exacerbated in multi-channel CT reconstruction problems (e.g. cardiac CT, spectral CT) where the gap between the amount and fidelity of the data acquired and the amount of data to be reconstructed is often exaggerated. To routinely reap benefits such as reduced radiation dose and relaxed sampling constraints in these advanced applications, efficient, scalable, and high-fidelity CT reconstruction tools are required. We report on matched, distance-driven projection and backprojection operators optimized for memory-efficient, parallel execution on multiple Nvidia GPUs. The 3D operators support several customizable features, including projection filtration, affine registration, arbitrary source trajectories for both flat-panel and cylindrical x-ray detectors, volumetric masking, and projection weighting. These operators are part of a new CT reconstruction toolkit which includes a MATLAB mex interface for rapid prototyping, GPU-based implementations of several popular regularization algorithms (bilateral filtration, sparse dictionary coding, patch-based SVT), and dozens of support functions for pipelining CT data processing operations. We are currently finalizing this toolkit for open-source distribution, with plans to add a python interface for interoperability with popular machine learning packages.**

I. Introduction

X-ray CT is a vital tool for numerous imaging applications in the areas of medical diagnosis, non-destructive testing, and basic science research. Generally, these applications rely on analytical reconstruction techniques which provide sufficient reconstruction quality when x-ray exposure and geometric constraints are satisfied. Within the scope of clinical imaging, however, there is strong motivation to reduce the radiation dose associated with imaging due to the potential dangers associated with radiation exposure. These concerns are amplified when considering advanced CT imaging applications such as dual energy CT and cine cardiac CT. Similar concerns exist for *in vivo*, preclinical imaging applications, particularly in longitudinal studies involving radio-sensitive organs and/or radiation therapy, where imaging dose may be considered a confounding variable.

To compensate for the loss of image quality commonly associated with low-dose CT scanning protocols, significant commercial and academic effort has been devoted to regularization strategies, including post-process denoising and iterative reconstruction. Historically, these regularization strategies have seen limited routine use because of their associated computational cost. More recently, however, the rise of low-cost, GPU-based scientific computing and of open-source code distribution has led to the availability of several high-quality CT reconstruction packages [1, 2]. Continued development of these packages and of GPU computing resources will undoubtedly increase the utilization of iterative CT reconstruction techniques as tools for dose reduction, scan time reduction, handling irregular scanning geometries, etc.

In this work, we introduce our own CT reconstruction tools. Unlike existing reconstruction packages, our tools have been developed to solve multi-channel CT reconstruction problems such as multi-energy photon-counting CT [3], spectrally-resolved cardiac CT [4], and 4D myocardial perfusion [5] in small animal models. Unique challenges associated with these problems and with *in vivo* data acquired in preclinical model systems have motivated significant effort to reconstruction and regularization operators robust to high levels of projection undersampling, irregular angular sampling, and noise. Specifically, to handle ill-conditioned CT reconstruction problems we have adopted matched, distance-driven projection and backprojection operators [6] and the biconjugate gradient stabilized convex solver [7], which both exhibit a high degree of stability in solving regularized, iterative reconstruction problems. To combat the computation time associated with these operations, we have ported them and supporting regularizers to run on Nvidia GPUs using the CUDA API. In the remainder of this abstract, we detail how we combine these operators within the split Bregman framework to solve a wide variety of multi-channel CT reconstruction problems.

II. Methods

A. CT Reconstruction Framework

We pose the problem of multi-channel reconstruction as follows:

All work was performed at the Duke Center for In Vivo Microscopy, an NIH/NIBIB national Biomedical Technology Resource Center (P41 EB015897). This work was also supported by the NIH National Cancer Institute (R01 CA196667, U24 CA220245) and an NVIDIA Academic GPU Grant.

D. P. Clark and C. T. Badea are with the Center for In Vivo Microscopy, Duke University Medical Center, Dept. of Radiology, Durham, NC 27710 USA (e-mail: cristian.badea@duke.edu).

$$\mathrm{X} = [\mathrm{X}_{t=1} \quad \ldots \quad \mathrm{X}_{t=\mathrm{N}_t}], \tag{1}$$

$$\mathrm{X} = \underset{\mathrm{X}}{\operatorname{argmin}} \tfrac{1}{2} \sum_t \|\mathrm{R}\mathrm{X}_t - \mathbf{y}\|_{\mathbf{Q}_t}^2 + \lambda \|\mathrm{X}\|_{\mathrm{Reg}}, \tag{2}$$

$$\|\mathrm{R}\mathrm{X}_t - \mathbf{y}\|_{\mathbf{Q}_t}^2 = (\mathrm{R}\mathrm{X}_t - \mathbf{y})^{\mathrm{T}} \mathbf{Q}_t (\mathrm{R}\mathrm{X}_t - \mathbf{y}). \tag{3}$$

X is a matrix of column-vectorized, reconstructed volumes for each of N_t channels, indexed by t (eq. 1). Each channel of X is reconstructed subject to a channel-specific, weighted least-squares data fidelity term and one (or more) multi-channel regularizers (eq. 2, "Reg"). The data fidelity term compares the projection of each channel, $\mathrm{R}\mathrm{X}_t$, (R^{T}, backprojection) with the vectorized projection data, $\mathbf{y}$. The array of least-squares weights, $\mathbf{Q}$, weights the contribution of projections to the reconstruction of each channel (eq. 3). These channel-specific weights can by multiplied by additional line-integral specific weights to incorporate statistical weighting [8] and/or additional subset-specific weights to accelerate the convergence of reconstruction with ordered subsets [9].

Returning to eq. 2, data fidelity is enforced subject to one or more multi-channel regularizers ("Reg"). Enforcing regularity between channels links the otherwise channel-specific reconstruction sub-problems. Within the split Bregman framework (fig. 1, [10, 11]), multi-channel CT reconstruction problems of this form (eq. 2) can be solved efficiently when the following regularization subproblem can be solved efficiently:

$$\mathrm{D} = \underset{\mathrm{D}}{\operatorname{argmin}} \tfrac{1}{2} \|\mathrm{X} + \mathrm{V} - \mathrm{D}\|_{\mathrm{F}}^2 + \gamma \|\mathrm{D}\|_{\mathrm{Reg}}, \tag{4}$$

where the "F" subscript denotes the Frobenius norm (vectorized L2 norm). Additional derivation details using similar notation can be found in our previous work [12]. Convenient analytical solutions to eq. 4 (and closely related equations) exist when enforcing sparsity in a wavelet transform domain ("Reg": L1 norm) or low column rank ("Reg": nuclear norm) [11]. Regularizers such as bilateral filtration can also be shown to reduce the cost associated with eq. 4 ("Reg": bilateral total variation, [13, 14]). Practically, even regularizers such as sparse dictionary coding [3, 15] and rank-sparse kernel regression [3], which solve less well-defined regularization sub-problems, can yield robust reconstruction algorithms.

Fig. 1 summarizes the split Bregman method with the add-residual-back strategy, including our own extensions to facilitate multi-channel CT reconstruction. Prior to iterative reconstruction, initialization is performed (steps 1-2). During step 1, an initial estimate is produced for the reconstruction of each channel by solving independent, weighted least-squares reconstruction problems. Producing these initial estimates with the same weights, $\mathbf{Q}$, and with the same convex solver used during the iterative portion of the algorithm (step 5) speeds the algorithm's convergence. Furthermore, this initialization procedure enables robust estimation of an appropriate regularization parameter for each channel, $\boldsymbol{\mu}_t$ (step 2). Specifically, $\boldsymbol{\mu}$ is computed from the data and scalar multipliers, $\boldsymbol{\alpha}$, which take on predictable values ($\boldsymbol{\alpha} \in [0.001 \quad 0.01]$). For spectral CT reconstruction, $\boldsymbol{\alpha}$ can be further scaled to account for differential noise levels between spectral channels [3].

Following initialization, iterative reconstruction (steps 3-5) proceeds with consecutive regularization (step 3), residual update (step 4), and data fidelity update (step 5) steps. Step 3 applies regularization to $\mathrm{X} + \mathrm{V}$, yielding a regularized version of X, D (consistent with eq. 4; V is initialized with zeros). Regularization (Reg()) is performed subject to one or more hyperparameters (e.g. h) which control regularization strength, neighborhood size, sparsity level, etc. as appropriate for the employed regularization. Typically, the regularization parameters λ and γ are not explicitly defined (eqs. 2, 4), but rather the regularization strength is scaled based on empirical measurements taken from data (e.g. using the median absolute deviation to estimate the noise level, [3]). Combining this with our strategy for scaling the $\boldsymbol{\mu}$ regularization parameters (step 2) greatly facilitates parameter selection for multi-channel CT reconstruction problems. Following regularization, step 4 updates the regularization residuals, V, which were removed from $\mathrm{X} + \mathrm{V}$ by the previous regularization step.

Step 5 updates the reconstruction of each channel, X_t, subject to the data fidelity and a convex proxy for the regularization. Typically, we rely on the biconjugate gradient stabilized method (BiCGSTAB, [7]) or its L variant [16] to update X, as these methods demonstrate very reliable convergence across a range of regularization methods, projection weighting schemes, and parameter values within this framework. Notably, since step 5 updates each channel independently, it is possible to expedite multi-channel reconstruction by updating multiple channels in parallel. Depending on the regularization used, the proposed framework can converge in as little as two to three iterations of steps 3-5.

Solve: $\mathrm{X} = \underset{\mathrm{X}}{\operatorname{argmin}} \tfrac{1}{2} \sum_t \|\mathrm{R}\mathrm{X}_t - \mathbf{y}\|_{\mathbf{Q}_t}^2 + \lambda \|\mathrm{X}\|_{\mathrm{Reg}}$

Inputs: $\mathbf{Q}, \mathbf{y}, \boldsymbol{\alpha}, \mathrm{h}$

Initialization:

1: $\mathrm{X} = \underset{\mathrm{X}}{\operatorname{argmin}} \tfrac{1}{2} \sum_t \|\mathrm{R}\mathrm{X}_t - \mathbf{y}\|_{\mathbf{Q}_t}^2$

2: $\boldsymbol{\mu}_t = \boldsymbol{\alpha}_t \frac{\|\mathrm{R}^{\mathrm{T}} \mathrm{Q}_t \mathbf{y}\|_2}{\|\mathrm{X}_t\|_2}$

While not converged ...

3: $\mathrm{D} = \mathrm{Reg}(\mathrm{X} + \mathrm{V}, \mathrm{h})$

4: $\mathrm{V} = \mathrm{X} + \mathrm{V} - \mathrm{D}$

5: $\mathrm{X} = \underset{\mathrm{X}}{\operatorname{argmin}} \sum_t \left[\tfrac{1}{2} \|\mathrm{R}\mathrm{X}_t - \mathbf{y}\|_{\mathbf{Q}_t}^2 + \tfrac{\boldsymbol{\mu}_t}{2} \|\mathrm{X}_t + \mathrm{V}_t - \mathrm{D}_t\|_2^2 \right]$

End While

Fig. 1. Generalized, multi-channel CT reconstruction framework. Indexing channels by *t*, the method solves for a series of reconstructed volumes, X_t, subject to one (or more) multi-channel regularizers ("Reg"). During initialization, an initial estimate is produced for each channel by solving independent, weighted least-squares reconstruction problems (weights $\mathbf{Q}_t$; step 1). The regularization parameter for each channel, $\boldsymbol{\mu}_t$, is scaled based on the data and a vector of scalar multipliers, $\boldsymbol{\alpha}$, which can account for differential noise levels between input channels (step 2). Iterative reconstruction (steps 3-5) then proceeds with consecutive regularization (step 3), residual update (step 4), and data fidelity update (step 5) steps.

B. *Reconstruction Operators*

At the core of our implementation of the split Bregman method and our multi-channel CT reconstruction toolkit are matched, distance-driven projection and backprojection operators optimized for memory-efficient, parallel execution on one or more Nvidia GPUs. Specifically, our operators combine the flexibility of a linear projection matrix defined in homogenous coordinates [17], with the numerical accuracy and stability provided by Long et al.'s distance-driven, separable footprint operators (flat-panel detector, SF-TT; cylindrical detector, SF-TR) [6], and the computational speed provided by the latest Nvidia GPU hardware and the CUDA API. Both the projection and backprojection operators are voxel-centric, enhancing numerical stability, since the same weights are used for both operations, and facilitating localized reconstruction operations through volumetric masking. Combined with the use of the BiCGSTAB solver for algebraic reconstruction, these operators exhibit a high degree of numerical stability in solving difficult CT reconstruction problems. Given the computational expense associated with these methods, multi-GPU parallelization and ordered subsets are supported to enhance computational speed.

In the current version of the toolkit, analytical and iterative reconstructions are performed from MATLAB scripts which call projection, backprojection, and regularization operators. These GPU-based operators, which are implemented using the C programming language and the CUDA runtime API, interface with MATLAB through MATLAB's MEX interface. The modularity of these operators facilitates understanding of sample reconstruction code included in the toolkit, while also enabling rapid prototyping of new reconstruction algorithms. To minimize computational overhead associated with separate calls for each operation, geometric parameters and memory on the host system and GPU device(s) are allocated prior to reconstruction and are shared between the projection and backprojection operators. As described in detail in fig. 2, the memory buffers allocated on the GPU are smaller than the total problem size, allowing large reconstruction problems to be solved with a limited amount of available memory on each GPU. To combat the loss of computational efficiency associated with the use of memory buffers, each GPU executes code on two (or more) GPU streams. These streams overlap memory transfer and compute operations to hide most of the latency associated with memory transfer operations.

C. *Regularization Operators*

Currently, our reconstruction toolkit includes several GPU-based, multi-channel regularization operators which can be called from MATLAB for post-reconstruction denoising or for direct incorporation into the split Bregman framework (fig. 1, step 3). Most commonly, we perform piece-wise constant regularization with bilateral filtration (BF, [18]). Our implementation of BF intrinsically scales its regularization strength based on the noise level measured in the data and can be called to jointly filter an arbitrary number of spectral and/or temporal channels. Spectral regularization operations compute a single 3D smoothing (range) kernel from multiple spectral channels to robustly preserve edge features while removing noise in low contrast CT data. Temporal regularization operations expand the filtration neighborhood across several consecutive time points to enforce spatio-temporal gradient sparsity [12]. Most recently, we have presented a regularization strategy called rank-sparse kernel regression, which maximizes BF performance while greatly simplifying parameter selection for both temporal and spectral regularization problems [3].

In addition to BF, we have implemented sparse dictionary coding [19]. Our GPU-based implementation uses a learned dictionary to sparsely code vectorized, multi-channel volume patches of variable size, performing image denoising subject to coefficient sparsity and error constraints. We have also implemented 3D patch-based singular value thresholding (pSVT, [20]), which enforces localized low rank constraints to enforce redundancy between reconstruction channels. Our implementation of pSVT supports non-convex thresholding parameters to more explicitly enforce low rank. Notably, while we have employed these regularization operators almost exclusively for CT reconstruction, we have had success in applying them to imaging data acquired with other modalities such as MRI.

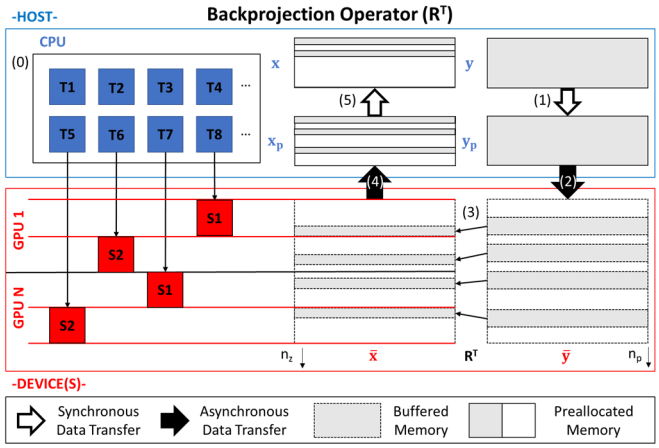


Fig. 2. Multi-GPU implementation of distance-driven backprojection (R^T). (0) Prior to computation, an initialization procedure statically allocates pinned memory on the host ($\mathbf{x}_p$: reconstructed volume; $\mathbf{y}_p$: projection data) and fixed-sized memory buffers on each GPU ($\bar{\mathbf{x}}$, $\bar{\mathbf{y}}$). Multi-GPU computing is coordinated by CPU threads (T#) on the host computer which are paired with GPU streams (S#) spread across each of N GPU devices. GPU streams run concurrently to backproject different sub-volumes of the output reconstruction. (1) Following initialization, backprojection is performed by copying the source projection data into the statically allocated, pinned memory, $\mathbf{y}_p$. (2) The inner loop of each GPU stream loops through the projection data, asynchronously transferring pinned projection data into a private memory buffer on the stream's associated GPU. (3) The outer loop of each GPU stream loops through fixed size chunks of the stream's associated sub-volume, backprojecting voxels at the intersection of the buffered volume data and the buffered projection data. (4) When all of the projection data has been backprojected into the current volume chunk, the volume chunk is asynchronously transferred to the statically allocated, reconstructed volume on the host, $\mathbf{x}_p$. (5) Events recorded by the CUDA API coordinate the transfer of completed volume chunks from the pinned memory volume, $\mathbf{x}_p$, to the final reconstructed volume, $\mathbf{x}$, prior to the overall completion of the backprojection operation.

III. Results

To illustrate the performance of our multi-channel reconstruction toolkit, we applied the split Bregman framework to iteratively reconstruct a dual-energy micro-CT data set. The data was acquired *in vivo*, in a C57BL/6 mouse which was injected with liposomal iodine [21] prior to imaging (following protocols approved by the Duke University Animal Care and Use Committee). The dual energy data was acquired with our dual-source micro-CT scanner [3], which now uses two Dexela 1512 flat-panel, CMOS x-ray detectors (CsI scintillator; 75 μm pixels; PerkinElmer, Inc.). The projection data was acquired with G297 tungsten x-ray sources (Varian Medical Systems) at 40 kVp (50 mA, 25 ms) and 50 kVp (80 mA, 12.5 ms, 0.15 mm Cu). 720 projections (1136x760 pixels; ~1.2x magnification) were acquired per kVp over 360° and used to reconstruct 512^3 volumes with 88-μm, isotropic voxels. Reconstruction was performed on an Ubuntu Linux workstation with two Intel Xeon E5-2650 v4 processors, 256 GB of system RAM, and four Nvidia Titan Xp GPUs. Regularization (fig. 1, step 3) was performed with a dual-energy, convolutional neural network constructed and trained with the Keras python library and similar in structure to the RED-CNN [22]. Data fidelity updates (fig. 1, steps 1, 5) were performed with the BiCGSTAB(L) algorithm (L=2) with three iterations over three subsets per update and energy.

Fig. 3 summarizes the reconstruction results, including comparable analytical, weighted algebraic, and iterative reconstruction results, as well as material decomposition results. Reconstruction required approximately 1GB of GPU RAM on each GPU used for computation and approximately 25GB of system RAM. The initialization and data fidelity update steps each took ~170, 268, and 449 seconds per energy using 4, 2, and 1 GPU(s), respectively, illustrating some degree of overhead associated with the use of additional GPUs. CNN-based regularization took ~200 seconds for both energies combined. In total, performing initialization (4 GPUs), two CNN-based regularization operations (1 GPU), and two data fidelity updates (4 GPUs) took ~25 minutes.

IV. Conclusions and Ongoing Work

Over the last 5+ years, we have developed highly specialized tools for multi-channel x-ray CT reconstruction. However, the impact of publications employing these tools has been somewhat limited by their complexity and the inherent difficulty in adequately describing them in written form. Following the trend in open-source code distribution to improve dissemination and reproducibility in science, we are preparing our first release of our multi-channel CT reconstruction toolkit, tentatively scheduled to coincide with the conference presentation of this work. In preparation for this release, we are preparing documentation and sample scripts to aid outside researchers in making use of our tools. It is our hope that these tools will enable new classes of multi-channel CT imaging such as clinical photon-counting CT, spectrally-resolved cardiac imaging, and 4D myocardial perfusion.

In addition to preparing for the initial release of the toolkit, we are currently working to identify and alleviate bottlenecks which inhibit linear scaling of our reconstruction operator performance when multiple GPUs are used for computation. We are also working to bridge the gap between the MATLAB interface we have developed for our GPU-based operators and the Python interface popular with many machine learning packages. We look forward to reporting on this issue and our preliminary experiments with CNN-based regularization in our conference presentation.

V. Acknowledgements

The GPU-based projection and backprojection operators included in this toolkit were originally built upon ray-based operators developed by Samuel M. Johnston [23]. The toolkit also relies on external code from the Michigan image reconstruction toolbox (MIRT), the NIfTI tools MATLAB toolbox, and the OMP-Box v10 and KSVD-Box v13 MATLAB toolboxes.

VI. References

[1] A. Biguri, M. Dosanjh, S. Hancock, and M. Soleimani, "TIGRE: a MATLAB-GPU toolbox for CBCT image reconstruction," *Biomedical Physics & Engineering Express,* vol. 2, no. 5, p. 055010, 2016.

[2] W. van Aarle *et al.*, "Fast and flexible X-ray tomography using the ASTRA toolbox," *Optics express,* vol. 24, no. 22, pp. 25129-25147, 2016.

[3] D. Clark and C. Badea, "Hybrid spectral CT reconstruction," *PloS one,* vol. 12, no. 7, p. e0180324, 2017.

[4] D. Clark and C. Badea, "Joint regularization for spectro-temporal CT reconstruction," *SPIE Medical Imaging,* pp. 97834E-97834E-11, 2016.

[5] D. Clark and C. Badea, "Low-dose, 4D myocardial perfusion with x-ray micro-CT," in *SPIE Medical Imaging*, 2017, pp. 101322L-101322L-8: International Society for Optics and Photonics.

[6] Y. Long, J. A. Fessler, and J. M. Balter, "3D forward and back-projection for X-ray CT using separable footprints," *IEEE Transactions on Medical Imaging,* vol. 29, no. 11, pp. 1839-1850, 2010.

[7] H. A. Van der Vorst, "Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems," *SIAM Journal on Scientific and Statistical Computing,* vol. 13, no. 2, pp. 631-644, 1992.

[8] J. Wang, T. Li, H. Lu, and Z. Liang, "Penalized weighted least-squares approach to sinogram noise reduction and image reconstruction for low-dose X-ray computed tomography," *IEEE Transactions on Medical Imaging,* vol. 25, no. 10, pp. 1272-1283, 2006.

[9] H. M. Hudson and R. S. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE transactions on medical imaging,* vol. 13, no. 4, pp. 601-609, 1994.

[10] T. Goldstein and S. Osher, "The split Bregman method for L1-regularized problems," *SIAM Journal on Imaging Sciences,* vol. 2, no. 2, pp. 323-343, 2009.

[11] H. Gao, H. Yu, S. Osher, and G. Wang, "Multi-energy CT based on a prior rank, intensity and sparsity model (PRISM)," *Inverse Problems,* vol. 27, no. 11, p. 115012, 2011.

[12] D. P. Clark, C.-L. Lee, D. G. Kirsch, and C. T. Badea, "Spectrotemporal CT data acquisition and reconstruction at low dose," *Medical Physics,* vol. 42, no. 11, pp. 6317-6336, 2015.

[13] M. Elad, "On the origin of the bilateral filter and ways to improve it," *IEEE Transactions on Image Processing,* vol. 11, no. 10, pp. 1141-1151, 2002.

[14] S. Farsiu, M. D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," *IEEE Transactions on Image Processing,* vol. 13, no. 10, pp. 1327-1344, 2004.

[15] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing,* vol. 54, no. 11, pp. 4311-4322, 2006.

[16] G. L. Sleijpen and D. R. Fokkema, "BiCGstab (l) for linear equations involving unsymmetric matrices with complex spectrum," *Electronic Transactions on Numerical Analysis,* vol. 1, no. 11, p. 2000, 1993.

[17] R. R. Galigekere, K. Wiesent, and D. W. Holdsworth, "Cone-beam reprojection using projection-matrices," *IEEE Transactions on Medical Imaging,* vol. 22, no. 10, pp. 1202-1214, 2003.

[18] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Computer Vision, 1998. Sixth International Conference on*, 1998, pp. 839-846: IEEE.

[19] R. Rubinstein, M. Zibulevsky, and M. Elad, "Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit," *Cs Technion,* vol. 40, no. 8, pp. 1-15, 2008.

[20] K. Kim *et al.*, "Sparse-view spectral CT reconstruction using spectral patch-based low-rank penalty," *IEEE Transactions on Medical Imaging,* vol. 34, no. 3, pp. 748-760, 2015.

[21] S. Mukundan Jr *et al.*, "A liposomal nanoscale contrast agent for preclinical CT in mice," *American Journal of Roentgenology,* vol. 186, no. 2, pp. 300-307, 2006.

[22] H. Chen *et al.*, "Low-dose CT with a residual encoder-decoder convolutional neural network," *IEEE transactions on medical imaging,* vol. 36, no. 12, pp. 2524-2535, 2017.

[23] S. Johnston, G. Johnson, and C. Badea, "GPU-based iterative reconstruction with total variation minimization for micro-CT," in *Proc. SPIE*, 2010, vol. 7622, p. 762238.

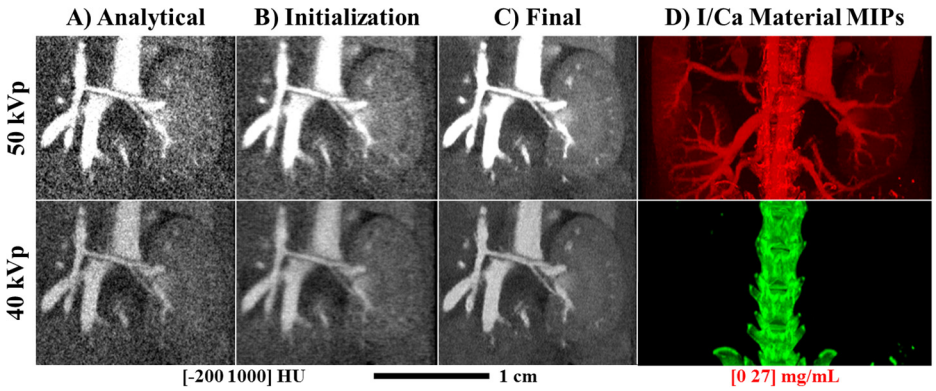


Fig. 3. Multi-channel reconstruction of *in vivo*, dual energy micro-CT data acquired in a mouse. (A) Matching 2D coronal slices through the right kidney following analytical reconstruction (FDK with Ram-Lak filter) of the 50 and 40 kVp data sets (rows). (B) Matching algebraic reconstruction results used for initialization of the split Bregman method (fig. 1, step 1). (C) Matching, final iterative reconstruction results following two regularization operations and two data fidelity updates (fig. 1, steps 3-5). (D) Iodine (I, red) and calcium (Ca, green) material decomposition results displayed as 100-slice maximum intensity projections (MIPs). Note that all results are shown over a limited field of view relative to the full reconstructed volumes.

Low-Dose CT Image Restoration Using Full-Dose Patch Database

Yuanke Zhang*, Dong Zeng, Junyan Rong, Jianhua Ma, Yuxiang Xing, Peng Gao, Tianshuai Liu, Yingmei Wang, Zhengrong Liang, Hongbing Lu

***Abstract*—This study develops a full-dose CT (FdCT) patch database induced nonlocal means (NLM) paradigm (fdpiNLM) to capture the valuable structure information from the FdCT scans. The innovation lies in the construction of an offline FdCT patch database and the online patch-search scheme integrated with the NLM operation. Specifically, the offline database is composed of 3D patches extracted from existing full-dose lung scans, avoiding repeated scans of the same patient. The online patch-search selects FdCT patches with structures similar to the target patch and then the NLM is utilized to extract redundancy information from the selected patches for regularization of the LdCT target patch. In case of no similar reference patches are found in the database, patches from local search window of the LdCT image are instead used for NLM regularization, avoiding introducing false prior structures from the FdCT images. The effectiveness of the proposed algorithm is valuated and validated by clinical lung cancer studies.**

***Index Terms*—low-dose CT, image restoration, full-dose CT database, nonlocal means (NLM).**

I. Introduction

The risk of X-ray radiation induced genetic, cancerous and other diseases has raised growing concerns to patients and operators[1]. Low-dose CT (LdCT) technique can reduce the radiation to the patient efficiently. However, the quality of the resulting image may be degraded with serious noise and streak artifacts [2]. To address this problem, various noise-reduction strategies were proposed, including statistic-based iterative reconstruction (SIR) approaches [3], pre-processing methods [4] and post-processing algorithms [5].

Recently, using the valuable prior structure features from full-dose CT (FdCT) scans for LdCT image restoration/reconstruction has become a noticeable research interests. Most of the full-dose scans assisted methods share the common ideas of first registering the previous FdCT scan of the same patient with the LdCT scan, and then extracting the intensity information of pixels in corresponding regions of FdCT scan as prior knowledge [6-10]. While these methods have been successful in many cases, such a previous FdCT scan of the same patient may not always be available. Moreover, as misalignments usually occur among the image series due to patients' respiration, internal motion or pathological changes of the tissues/organs, even after registration of the image series, false prior structures may be introduced into the current LdCT image. More details of such cases would be demonstrated in the experiment studies.

This study presented a strategy to utilize existing FdCT scans of different patients to restore the current LdCT image. The innovation is the construction of an offline patch database from FdCT scans and the use of online patch-search scheme integrated with the nonlocal means (NLM) [11] operation. Specifically, the offline patch database is composed of 3D patches extracted from full-dose CT scans of different patients. Given a 3D target patch to be restored in the LdCT images, the online patch-search first selects patches whose structure is similar to the target patch from the database as the reference patches, and then NLM is employed on the patches to regularize the target patch through a weighted average scheme. The proposed method does not need repeated scanning of the same patient and can avoid introducing false prior structures from FdCT images.

The rest of this paper is organized as follows. In section II, the proposed FdCT patch database induced NLM strategy (fdpiNLM) is presented in details. Section III evaluates the proposed algorithm with clinical lung cancer studies. Finally, a conclusion is given in section IV.

II. Method

The flowchart of the proposed fdpiNLM strategy contains three major steps: offline patch database construction, online patch-search, and online NLM regularization. In the

This work was supported in part by the China Postdoctoral Science Foundation under Grant No. 2017M613348, in part by the National Key Research and Development Program of China under Grant Nos. 2017YFC0107400, 2017YFC0107403, in part by the National Natural Science Foundation of China under Grant Nos. 61572283, 81230035.

Yuanke Zhang is with the School of Biomedical Engineering, Fourth Military Medical University, Xi'an, Shaanxi 710032, China (e-mail: yuankezhang@163.com).

Dong Zeng and Jianhua Ma are with the Department of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China.

Junyan Rong, Peng Gao, Tianshuai Liu and Hongbing Lu are with the School of Biomedical Engineering, Fourth Military Medical University, Xi'an, Shaanxi 710032, China.

Yuxiang Xing is with the Department of Engineering Physics, Tsinghua University, Beijing 100084, China.

Yingmei Wang is with the School of Mathematics and Statistics, Shandong University of Technology, Zibo, Shandong 255049, China.

Zhengrong Liang is with the Departments of Radiology and Biomedical Engineering, State University of New York at Stony Brook, Stony Brook, NY 11794 USA.

following subsections, we describe each step in detail.

A. Offline Patch Database Construction

In the proposed scheme, a 3D patch, which is modeled as a voxel and its $K\times K\times K$ cubic nearest neighbors, is used to reflect local spatial structure of a voxel. To construct the offline database, several full dose lung scans were first gathered from different patients. For each patient' FdCT scans, 3D patches were extracted automatically with a sliding distance of one voxel on the 3D volume. Then all the 3D patches from different patients were stacked into a reference database, denoted by Ω.

B. Online Patch Search

By representing a 3D patch as a column vector, the ith patch in the database Ω is denoted by $\mathbf{x}_i^{\mathrm{Ref}} = [x_{i,1}^{\mathrm{Ref}}, x_{i,2}^{\mathrm{Ref}}, \cdots, x_{i,M}^{\mathrm{Ref}}]^T$, where $M=K^3$, K is the patch size, and $i=1,2,\cdots, |\Omega|$. The target patch to be restored in the LdCT image is modeled as a vector variable, and denoted by $\mathbf{x}^{\mathrm{Ld}} = [x_1^{\mathrm{Ld}}, x_2^{\mathrm{Ld}}, \cdots, x_M^{\mathrm{Ld}}]^T$. The nearest neighbor grouping method was employed to select patches from the database. Specifically, the Euclidean distance between $\mathbf{x}^{\mathrm{Ld}}$ and a patch $\mathbf{x}_i^{\mathrm{Ref}}$ ($i=1,2,\cdots|\Omega|$) in the database is computed and then compared with a preset threshold T. Patches with distance less than the threshold are chosen as reference patch candidates for $\mathbf{x}^{\mathrm{Ld}}$.

For the case that the number of selected candidates is larger than $10M$ ($M=K^3$), experiment results indicated that selecting at least the $10M$ most similar candidates as reference patches could provide relatively robust regularization of the LdCT target patch by NLM. If the number of selected candidates is less than $10M$, it indicates that there are no enough similar patches in the offline database. In this case, instead of using FdCT reference patches, patches from the local search window of the LdCT image are used for regularization of the target patch by NLM.

Suppose finally N ($N=10M$) reference patches are selected, we denote them by a two-dimensional matrix as

$$\mathbf{X}^{\mathrm{Ref}} = (\mathbf{x}_1^{\mathrm{Ref}}, \mathbf{x}_2^{\mathrm{Ref}}, \cdots, \mathbf{x}_N^{\mathrm{Ref}}),\ i=1, 2, \cdots, N. \tag{1}$$

where each reference patch is a column vector of length M, $\mathbf{x}_i^{\mathrm{Ref}} = [x_{i,1}^{\mathrm{Ref}}, x_{i,2}^{\mathrm{Ref}}, \cdots, x_{i,M}^{\mathrm{Ref}}]^T$.

C. Target Patch Regularization by NLM

The NLM concept was first proposed by Buades et al. [11], which relies on the information redundancy within the patches. In this paper, we employ the NLM operation to extract prior information from the FdCT reference patches for the restoration of the LdCT target patch.

With the NLM operation, the target patch $\mathbf{x}^{\mathrm{Ld}}$ can be regularized as

$$\hat{\mathbf{x}}^{\mathrm{Ld}} = \sum_{i=1}^{N} w_i \mathbf{x}_i^{\mathrm{Ref}} \tag{2}$$

where w_i is the weight assigned to $\mathbf{x}_i^{\mathrm{Ref}}$, which satisfies the conditions of $0 \le w_i \le 1$ and $\sum_{i=1}^{N} w_i = 1$. It can be computed as

$$w_i = \frac{1}{Z}\exp\left(-\frac{\left\|\mathbf{x}^{\mathrm{Ld}} - \mathbf{x}_i^{\mathrm{Ref}}\right\|_2^2}{h^2}\right) \tag{3}$$

where the term $Z = \sum_{i=1}^{N}\exp\left(-\left\|\mathbf{x}^{\mathrm{Ld}} - \mathbf{x}_i^{\mathrm{Ref}}\right\|_2^2 / h^2\right)$ is a normalizing factor, and the term $\left\|\mathbf{x}^{\mathrm{Ld}} - \mathbf{x}_i^{\mathrm{Ref}}\right\|_2^2$ is the Euclidean distance between $\mathbf{x}^{\mathrm{Ld}}$ and $\mathbf{x}_i^{\mathrm{Ref}}$, which has been calculated previously in Subsection II.*B*. In Eq.(3), h is a preset parameter that controls the overall smoothness of the NLM filtering.

D. An Alternative If No Similar Patch Found in the Database

To avoid introducing false prior structures from the FdCT images or losing true structures in the LdCT image, it is essential to detect whether there are similar patches in the offline database or not, for a target patch. In Subsection II.*B*, we have given the criterion for how to detect such a case. For target patch that no similar reference patches be found in the database, we instead use noisy patches from local search window of the LdCT image for the NLM regularization. The rest of the steps are essentially the same as those described in Subsection II.*C*.

III. Experiment Results

Clinical CT-guided lung nodule needle biopsy studies were used to evaluate and validate the proposed algorithm. 11 patients (denoted with patient #1 to patient #11) were recruited under informed consents after the approval by the Institutional Review Board. The full-dose and low-dose CT scans were performed with Siemens Sensation 16 CT scanner at X-ray tube voltage of 120 kVp and tube current of 100 mAs and 20 mAs, respectively. Other scanning parameters were as follows: 0.5 s per gantry rotation, helical pitch of 1 mm, 16×0.75 mm collimation, 5 mm slice thickness, 2 mm reconstruction slice thickness, 1 mm reconstruction slice interval, without using automatic exposure control (AEC). The reconstructed image was of 512×512 size. The averaged CT dose indexes (CTDIvol) recorded for the 120kVp/100mAs and 120kVp/20mAs scans were 10.31 mGy and 2.09 mGy, respectively.11 patients were classified into a training group (10 patients: patient #1 to patient #10) and a test group (patient #11). Based on the patient scans, the clinical studies were performed to evaluate the performance of the proposed fdpiNLM algorithm.

A. Construction of the Offline Database

Total 432 slices of 120kvp/100mAs lung CT images constructed by FBP"B60f" from the patients #1 to #10 were

used to serve as reference scans. We first extracted 3×3×3 patches from each patient's scan series with a sliding distance of one voxel, then stacked them into one training patch set Ω. Considering the structure redundancy of patches in the entire set, we instead use a randomly selected small subset with 0.5% of Ω to form the offline training database (with 318672 3D patches), which can reduce the computational load in online patch search step.

B. Visual Evaluation

In this pilot clinical study, the 120 kVp/20 mAs low-dose scan from the patient in the validate group (patient #11) were chosen for the evaluations. One slice of the LdCT scan was shown in Fig.1(b). The corresponding slices of a 120kVp/100mAs full-dose scan from the same patient were used as the gold standard. Fig.1(a) shows the corresponding slice of the FdCT scan. The FBP with Siemens kernel "B60f" was employed to reconstruct both the full-dose and low-dose scans.

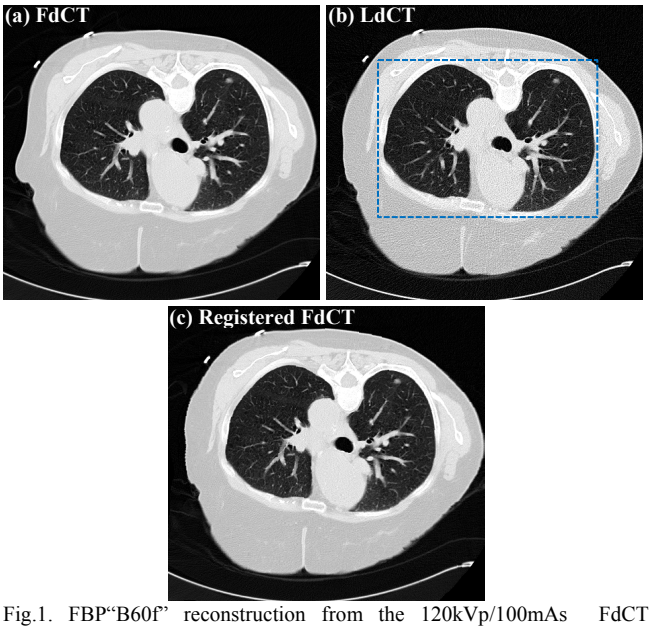


Fig.1. FBP"B60f" reconstruction from the 120kVp/100mAs FdCT sinogram and the 120kVp/20mAs LdCT sinogram. (a) FdCT image. (b) LdCT image. (c) Registered FdCT image with the LdCT image, for the ndiNLM algorithm. The display window is [-1024, 176] HU.

For comparison, the traditional NLM algorithm [11] and the previous normal-dose scan induced nonlocal means (ndiNLM) algorithm [8] were also evaluated. In the NLM algorithm, the parameters were set as: patch size of 5x5, search window size of 41x41, and the filtering parameter of 6000. The ndiNLM algorithm employs the NLM operation to extract the intensity information of corresponding pixels of the registered previous full-dose scan. Here, we adopted the state-of-arts scale-invariant-feature transform flow (SIFT-flow) algorithm [12] for the registration of the FdCT scan with the corresponding slices of LdCT scan. The registration result was shown in Fig.1(c). The other parameters were set as: patch size of 5x5, search window size of 41x41, and the filtering parameter of 1000. In the proposed fdpiNLM algorithm, the parameters were set as: patch size of 3x3x3, the threshold T of 5×10^3, and the filtering parameter of 1000.

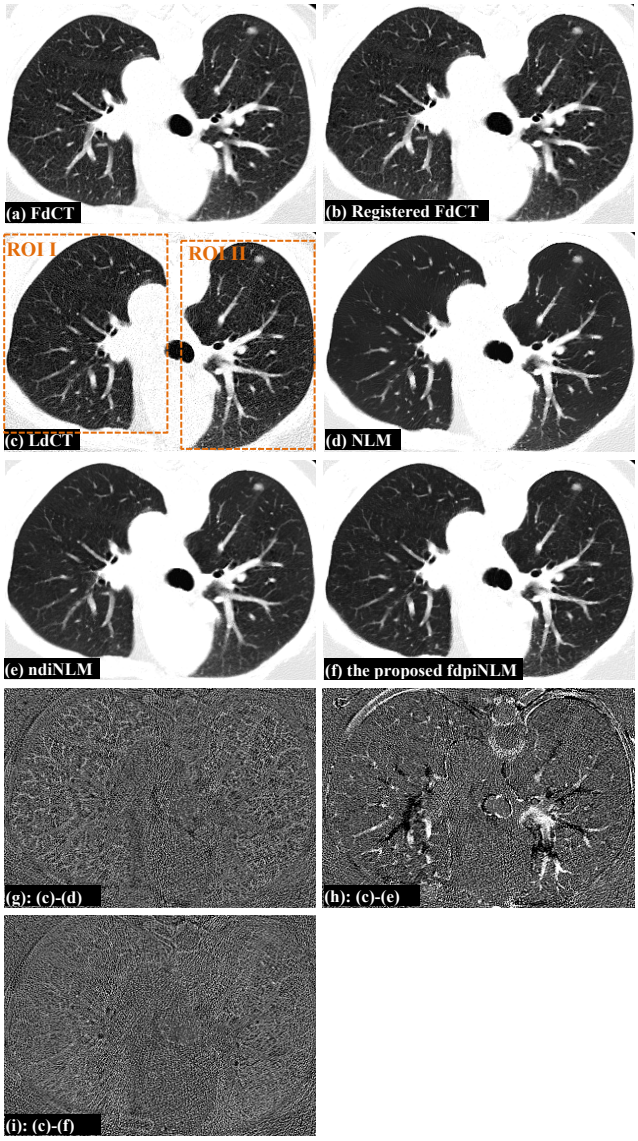


Fig.2. Processing results of the selected lung region (outlined by blue dotted rectangles in Fig.1(b)) of the 120 kVp/20 mAs low-dose scan. (a) FdCT image. (b) Registered FdCT image with the LdCT image, for the ndiNLM algorithm. (c) LdCT image. (d) NLM processed result. (e) ndiNLM processed result with (b) as the previous FdCT image. (f) The proposed fpdiNLM processed result. (g)-(i) are the difference images between the LdCT image and images processed by the NLM, ndiNLM, and the proposed fpdiNLM algorithm, respectively. The display window of (a)-(f) is [-1024, -24] HU, and the display window of (g)-(i) is [-100, 200] HU.

Figures 2 and 3 show the processing results and associated difference images of the selected lung region (outlined by blue dotted rectangles in Fig.1(b)) of one slice of the scan, with a lung display window of [-1024, -24] HU. It can be observed in Figs.2(d), (g) and the fourth column of Fig.3 that by using the NLM algorithm, the noise was suppressed, but streak artifacts and over-smoothing could be simultaneously observed in the lung regions. We can observe in Figs.2(e), (h) and the fifth column of Fig.3 that by using the ndiNLM method, the noise and streak artifacts were better suppressed, but serious inconsistence between the original LdCT image

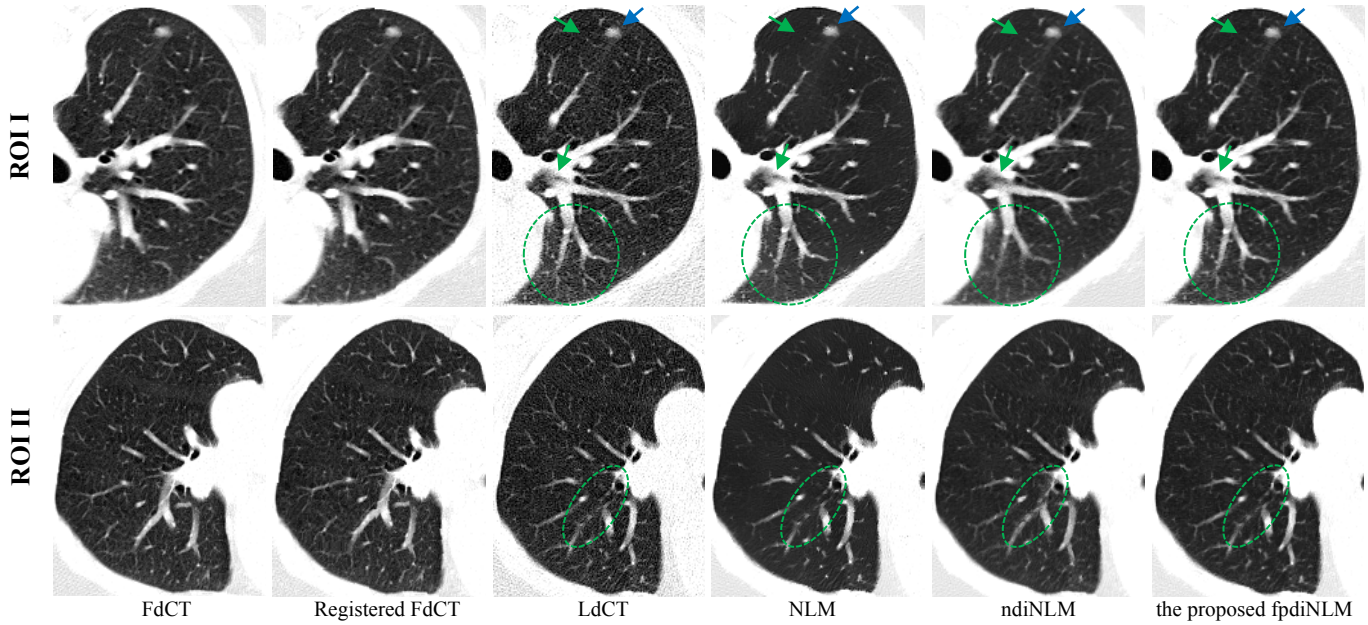


Fig.3. Zoomed-in views. From left to right: FdCT image, registered FdCT image with the LdCT image, LdCT image, NLM processed result, ndiNLM processed result, and the proposed fpdiNLM processed result, respectively. The display window is [-1024, -24] HU.

and the ndiNLM processed result could be observed (please see Fig.2(h) and regions indicated by green arrows and green circles in Fig.3). Figs.2(a)-(c) indicated that the inconsistence was induced by the misalignments between the FdCT image and the LdCT image, due to patients' respiration and/or internal motion. We can observe in Figs.2(f), (i) and the sixth column of Fig.3 that the proposed fpdiNLM algorithm performs much better in both noise/streak artifacts suppression and details/textures preservation compared with NLM and ndiNLM, and produces a visual effect similar to the full-dose reference image, especially for the ROI of the lung nodule (indicated by the blue arrow in Fig.3).

IV. Conclusion

This study developed a full-dose patch database induced NLM paradigm to capture the valuble structure information from the FdCT scans, which can better preserve the detail textrues of the LdCT image. The proposed paradigm contains three major components, which are the offline FdCT patch database construction, online patch-search, and online NLM regularization. Through constructing the offline database with clean patches extracted from the existing FdCT scans, valuable prior knowledge of high quality could be obtained, avoiding repeated scans of the same patient. The online patch-search selected FdCT patches with structures similar to the target patch and then the NLM operation was utilized to extract redundancy information from the selected patches for regularization of the LdCT target patch. For target patch that no similar reference patches were found in the database, patches from local search window of the LdCT image were instead used for NLM regularization, avoiding introducing false prior structures from the FdCT images.

We have implemented the proposed fdpiNLM algorithm with the Matlab parallel pool of 20 local workers in a PC workstation with Intel Xeon CPU (20 core, 2.5G Hz) and 64G RAM memory. With the patch size of 3x3x3 and the training database of 318672 patches, the algorithm took approximately 56 minutes to process a 512×512×30 CT volume.

References

[1] D. J. Brenner and E. J. Hall, "Current concepts - Computed tomography - An increasing source of radiation exposure," N. Engl. J. Med., 2007, 357(22): 2277-2284.

[2] Hsieh J, "Adaptive streak artifact reduction in computed tomography resulting from excessive X-ray photon noise," Med. Phy., 1998, 25(3): 2139–2147.

[3] Idris A. Elbakri, Jeffrey A. Fessler, "Statistical Image Reconstruction for Polyenergetic X-Ray Computed Tomography," IEEE Trans. Med. Imaging., 2002, 21(2): 89-99.

[4] Y. Zhang, J. Zhang, and H. Lu, "Statistical sinogram smoothing for low-dose CT with segmentation-based adaptive filtering," IEEE Trans. Nucl. Sci., 2010, 57(5): 2587-2598.

[5] Y. Chen, L. Shi, Q. Feng, J. Yang, H. Shu, L. Luo, J. Coatrieux, and W. Chen, "Artifact suppressed dictionary learning for low-dose CT image processing," IEEE Trans. Med. Imag., 2014, 33(12): 2271-2292.

[6] P. Theriault-Lauzier, J. Tang, and G. Chen, "Prior image constrained compressed sensing: Implementation and performance evaluation," Med. Phys., 2012, 39(1): 66–80.

[7] H. Dang, A. Wang, M. Sussman, J. Siewerdsen, and J. Stayman, "DPIRPLE: A joint estimation framework for deformable registration and penalized likelihood CT image reconstruction using prior images," Phys. Med. Biol., 2014, vol. 59: 4799–4826.

[8] J. Ma, J. Huang, Q. Feng, H. Zhang, H. Lu, Z. Liang, and W. Chen, "Low-dose CT image restoration using previous normal-dose scan," Med. Phys., 2011, 38(10): 5713–5731.

[9] Xu W, Mueller K. Efficient low-dose CT artifact mitigation using an artifact-matched prior scan. Medical Physics. 2012, 39(8): 4748-4760.

[10] Xu W, Ha S, Mueller K. Database-assisted low-dose CT image restoration. Medical Physics. 2013, 40(3): 031109-1:7.

[11] Buades A, Coll B, Morel J M, "A review of image denoising algorithms, with a new one," Multiscale Modeling and Simulation, 2005, 4(2): 490-530.

[12] C. Liu, J. Yuen, and A. Torralba, "SIFT flow: Dense correspondence across scenes and its applications," IEEE Trans. Pattern Anal. Mach. Intell., 2011, 33(5): 978–994.

Adaptive Edge Preserve Filter for Streak Artifacts Reduction in Computed Tomography

Chuang Miao, Ph.D., Hongbin Guo, Ph.D. and Aziz Ikhlef, Ph.D.

***Abstract*—in computed tomography (CT), streaking artifacts resulting from x-ray photon starvation reduce diagnostic image quality. In order to obtain an acceptable diagnostic image, a higher tube current is often selected results in higher dose to patients and lower scanner throughput. In this paper, we propose an adaptive edge preserve filter in radon domain to reduce streaks resulting from lack of photons while preserve edges containing diagnostic information. We first model the noise characteristics of CT imaging in photon domain and radon domain. An adaptive bilateral edge preserve filter is then designed in radon domain. Its parameters such as the window size, filtering parameters in spatial and range domains are adaptively selected based the noise characteristics in photon domain. Clinical studies have been performed to demonstrate the effectiveness of the proposed filter in suppressing streak artifacts, preserving edges and maintaining spatial resolution.**

Keywords—edge preserve filtering; image processing; computed tomography; streak reduction.

I. Introduction

CT image artifacts are annoying to radiologist because many artifacts can mimic diseases and cause misdiagnosis[1]. The CT image artifacts are mainly caused by the mismatch between projection measurements and reconstruction model. The mismatch and therefore the artifacts could be result from patient motion, imperfect system calibration, data under-sampling, metal object, and x-ray photon starvation. Many endeavors have been made by previous researchers to combat CT artifacts[2]–[6].

In this paper, we are interested in the streak artifacts reduction caused by photon starvation which is very common when scanning shoulder, pelvis or big patients. One way to combat this problem is to increase the x-ray tube voltage that can increase the maximum x-ray photon energy and penetration. However, this can decrease the low contrast detectability and may not a valid option in clinical. Another approach is to increase the x-ray tube current to generate more photons. However, this approach will result in a higher patient dose and a decreased x-ray tube throughput. The image processing technique is a promising approach to suppress the streak artifacts effectively. For clinical application, a majority of measurements are beyond the range of x-ray photon starvation and the reconstruction from those measurements does not introduce streak artifacts. Hence, a filtering operation to the overall measurements will have a significant impact on the spatial resolution. To overcome this problem, an adaptive trimmed mean (ATM) filter was proposed by Hsieh [7]. The ATM parameters such as the filtering length and trimming percentage are adaptively selected based on the measured signal intensity. The lower the measured signal is, the larger the filtering window size. The filtering process is then performed in projection domain. Though the ATM achieves a good balance between streaks reduction and spatial resolution preservation, the ATM tends to blur edges and introduce artifacts to soft tissue around high intensity materials such as bone or contrast agent. To overcome the shortcomings of ATM, we proposed an adaptive edge preserve bilateral filter (ABF). The window size and filtering parameters are adaptively selected to adapt to the signal intensity.

The rest of this paper is organized as follows: in Section II, we will build a model for the noise in radon domain and analyze the root cause of streak artifacts; in Section III, we propose a novel ABF approach to reduce the streak artifacts while preserve edges, specify the adaptive filtering parameters selection; in Section IV, phantom and clinical results are presented to demonstrate the effectiveness and robustness of our approach; in Sec V, we will conclude the paper.

II. Analysis of Streak Artifacts

A. Noise Model in Radon Domain

The noise in CT system is primary composed of electronic noise and Poisson noise [7]. The electronic noise comes from the data acquisition system (DAS) that converts the x-ray photons received by detector to digital signals that can be processed in the computer. The DAS is composed of detector, current to voltage convertor, pre-amplifier, analog integrator and analog to digital convertor. For similarity, the electronic noise is considered as a constant noise floor independent of the input signal. The Poisson noise is the x-ray photon noise that is well known to follow a Poisson distribution in that the variance equals its mean. Considering a single detector channel and ignoring the variations in detector characteristics and data acquisition system, the variance in a single detector channel can be expressed as:

$$\sigma^2 = \sigma_e^2 + \sigma_q^2, \tag{1}$$

where σ_e^2 represents the electronic noise floor of DAS, σ_q^2 represents the x-ray photon quantum noise. There is a scale factor also called gain factor k relates the measured output current signal s in detector to the number of x-ray photons. By applying the properties of Poisson distribution, Eq. (1) can be rewritten as:

$$\sigma^2 = \sigma_e^2 + ks. \tag{2}$$

Chuang Miao, Ph.D. is formerly with FMI Medical Systems, Inc., 29001 Solon Road, Unit A, Solon, OH, USA

Hongbin Guo, Ph.D. and Aziz Ikhlef, Ph.D. are with FMI Medical Systems, Inc., 29001 Solon Road, Unit A, Solon, OH, USA

A logarithm operation needs to be performed to convert the measured photon data to the projection data so that the line integral of each ray path can be obtained:

$$p = f(s) = log(s). \tag{3}$$

To a second-order approximation [8], the relationship between the noise in projection and the measured photon signal can be expressed as:

$$\sigma_p^2 \cong [f'(s)]^2\sigma^2 + \frac{[f''(s)]^2}{4}\sigma^4 = \frac{k}{s} + \frac{4\sigma_e^2+k^2}{4s^2} + \frac{k\sigma_e^2}{2s^3} + \frac{\sigma_e^4}{4s^4}. \tag{4}$$

B. Streak Artifacts Analysis

The streak artifacts in the reconstructed images are caused by the inconsistency / fluctuations in the projection data. Due to the backprojection process, high fluctuations in the projection will be mapped onto a group of streaks in the final reconstructed images. Based on Eq. (4), it is obvious that the projection noise increases rapidly as the measured photon signal s decreases. Therefore, the detector channels that receive a small number of x-ray photons (photon starvation) will exhibit high fluctuations. These channels are the culprit of the streak artifacts in the final reconstructed images. In contrast, when the measured signal is large, the variance of the projection is very small. Therefore, these detector channels have very little channel to channel fluctuations and will not contribute to the streak artifacts. The objective of the streak artifacts correction is to reduce the fluctuations corresponding to the detector channels that are lack of x-ray photons and to make the noise even across all detector channels.

III. Adaptive Edge Preserve Filter

Based on the previous discussion, the filtering operation for streaks reduction should meet the following criteria. First, the filtering process should not be performed to the detector channels corresponding to large measured signals since these channels do not contribute to the streak artifacts. Applying filtering operation to these channels can result in degraded spatial resolution. Second, the filtering operation should be applied to detector channels lack of photons. According to Eq. (4), the variance in projection increases quickly as the signal becomes small. Hence, the amount of the filtering applied to each channel should be inversely proportional to the signal intensity.

To satisfy these conditions, we proposed an adaptive edge preserve bilateral filter. The bilateral filter [9] is well known for its robust performance on filtering while preserve edges. Mathematically, the bilateral filter can be expressed as

$$I^f(\boldsymbol{x}) = \frac{1}{W}\sum_{\boldsymbol{y}\in\Omega} I(\boldsymbol{x}) e^{\frac{\|x-y\|_2^2}{\sigma_S^2}} e^{\frac{(I(x)-I(y))^2}{\sigma_r^2}}, \tag{5}$$

where Ω is a window centered in $\boldsymbol{x}$. σ_s is the filtering parameters in spatial domain. The amount of smoothing increases as σ_s increases. σ_r is the filtering parameter in range domain. As σ_S gradually increases, the bilateral filter approaches to the Gaussian filter and stronger edges will be blurred. The normalization term is defined as

$$W = \sum_{\boldsymbol{y}\in\Omega} e^{\frac{\|x-y\|_2^2}{\sigma_S^2}} e^{\frac{(I(x)-I(y))^2}{\sigma_r^2}}. \tag{6}$$

The bilateral filter can be fully specified by the selection window Ω, σ_s and σ_r. Based on Eq. (4), the dynamic range of the noise variance is very large, so we adaptively select all the three filtering parameters Ω, σ_s and σ_r. For this paper, we limited the filter to one dimension and applied only to the channels in the same view. It is easy to extend the filtering to two dimensional to apply to both detector channel and detector row, or both detector channel and views. The filtering length L is determined based on the measured signal s by the following equation:

$$L = \frac{M \cdot r(\lambda - s)}{\lambda + \alpha \cdot [r(s-\delta)]^\beta} \tag{7}$$

where the rectifier function

$$r(z) = z^+ = \max(0, z). \tag{8}$$

In Eq. (7), M represents the maximum filter length when the measured signal x approaches to zero. λ is a threshold defines the cutoff point. When the measured signal s is greater than λ, no filtering will be performed. α and β are two hyper parameters control the rate of change of filter length. δ is a estimation of electronic noise. When the measured signal s is smaller than δ, the change of filter length is not affected by α and β and the filter length is almost constant.

The σ_s and σ_r are determined by the following equations:

$$\sigma_s = \sigma_{s0}\left(\frac{L}{M}\right)^{\gamma_S} \tag{9}$$

$$\sigma_r = \sigma_{r0}\left(\frac{L}{M}\right)^{\gamma_r} \tag{10}$$

Where σ_{s0} and σ_{r0} are the maximum σ for spatial domain and range domain, respectively. The filtering parameters decrease as the filtering length decrease. The decreasing rates for σ_s and σ_r are controlled by γ_s and γ_r, respectively. When γ_s and γ_r equal to zeros, the filtering parameters become constant and the filtering is only adaptive for window size. The institution to select the filtering parameters is the following. Smaller signal s results in stronger streaks so that greater filtering parameters should be selected to suppress streaks while preserving strong edges. Greater signal s results in weaker or no streaks so that smaller filtering parameters need to be selected to suppress streaks while preserving fine edges. By combining Eq. (7) and Eqs. (9) & (10), the following equations are obtained

$$\sigma_s = \sigma_{s0} \cdot \left(\frac{r(\lambda - s)}{\lambda + \alpha \cdot r^\beta(s-\delta)}\right)^{\gamma_S} \tag{11}$$

$$\sigma_r = \sigma_{r0} \cdot \left(\frac{r(\lambda - s)}{\lambda + \alpha \cdot r^\beta(s-\delta)}\right)^{\gamma_r} \tag{12}$$

Hence, the parameters $\lambda, \alpha, \beta, \delta$ together affect the filtering parameters selection.

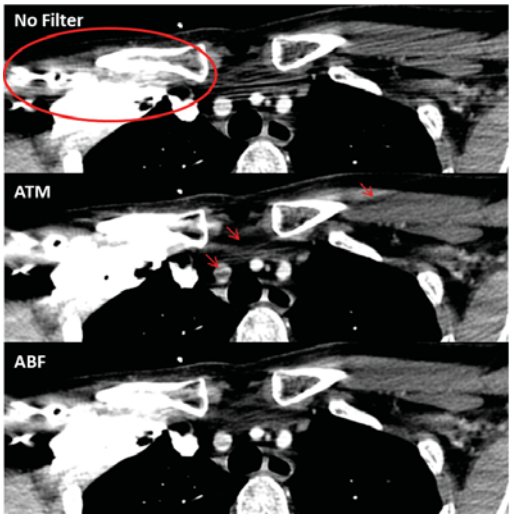


Fig. 1. Images from contrast scan. From top to bottom, images correspond to reconstruction with no streak reduction filtering, ATM and ABF, respectively. Display window level = 40HU, display window width = 300 HU.

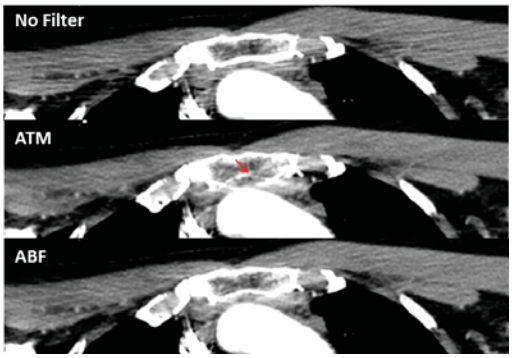


Fig. 2. Same as Fig. 1 but different slices.

IV. Experiments

To evaluate our proposed approach, a patient scan data was employed. In most CT applications, the streak artifacts are prominent when scanning shoulder area. However, the streakartifacts can exist in the low dose chest scan as well. In this experiment, the chest area was scanned by a Scinticare 16 slices CT scanner (Minfound Medical Systems Co., Ltd, Zhejiang, China) with 1 cm collimation, 120kV, 1 second rotational speed. A non-contrast scan was first performed followed by a scan with contrast agent. Images were reconstructed with 512 by 512 matrix size. Streak artifacts were observed in the reconstructed images from both datasets. The objective of this study is to evaluate the performance of the

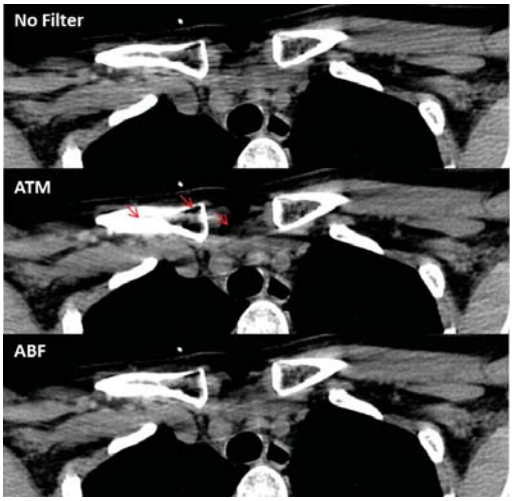


Fig. 3. Same slice as Fig. 1 but without contrast.

proposed ABF approach under different situations. We want to make sure the selected parameters works for both non-contrast scan and contrast scan with different level of streak artifacts. Suppressing the streak artifacts in a contrast scan is more challenging since the soft tissue regions are easily corrupted by the filtering operation. For the purpose of comparison, we implemented the ATM proposed in [7].

The parameters selection was performed as follows. Since the clinical specialist cares more about the streaks in contrast agent dataset for this case. We applied ATM to the contrast agent scan with a set of different filtering parameters. The whole reconstructed image volumes were reviewed by clinical specialist to select the best parameters. Similar procedure was performed for ABF to select the optimal filtering parameters. In practice, once the optimal filtering parameters tuned, the same parameters are applied to all different scans. So the ATM and ABF were then applied to the non-contrast scan with the same selected parameters. Fig. 1 shows the results for contrast scan. It can be clearly observed that the streak artifacts are significant without filtering operation. While both ATM and ABF suppress the streak artifacts to clinical acceptable level, the ATM tends to introduce artifacts in the tissue regions as indicated by the red arrows. Within the red circle area, the ATM completely blurred out the tissues between bone and contrast agents. Fig. 2 shows the results from a different slice. As indicated by the red arrow, the ATM blurred out the interface between high attenuation material (bone and contrast agent) and the soft tissue. The ABF suppressed the artifacts in the soft tissue while preserving the edges. In Fig. 3, the ATM and ABF were applied to the non-contrast agent scan with same parameters used in the contrast scan. The streaks in the non-contrast scan are much less than that of contrast scan as expected. Similar conclusions can be made, as indicated by red arrows, that the ATM tends to blur out edges and introduces artifacts in the soft tissue regions. It should be noted that the ATM does not introduce artifacts in every slice. Because same

filtering parameters were applied to the whole data set, we picked the slices with artifacts introduced to show.

To further investigate the streak artifacts elimination, the lung portion of CT images is displayed in Fig. 4 at the appropriate display window to show the lung nodules. It is clear that the proposed ABF method removes the streak artifacts in the right lobe of the lung without corrupting lung nodules.

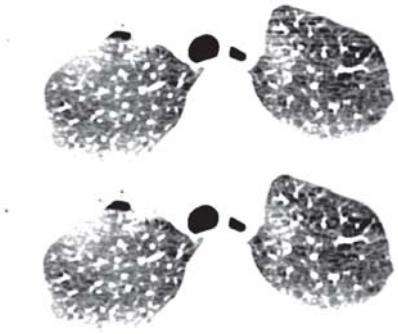


Fig. 4. Lung portion images of the contrast scan. Top: images reconstructed without streak artifacts reduction; Bottom: images reconstructed with proposed ABF method. Display window = 200, display level = -700.

V. Conclusion

In this paper, we developed an adaptive edge preserve bilateral filter to suppress streak artifacts due to photon starvation while preserving edges and spatial resolution. The filtering parameters such as window size, σ_s and σ_r are adaptively selected to adapt the signal intensity. Clinical studies have been performed to demonstrate the advantages and robustness of the proposed ABF method. Compare to the ATM approach, the proposed ABF method has a better edge preservation ability and less likely to introduce artifacts to soft tissues. The amount of streak reduction depends on the measured signal intensity due to the adaptive nature of the ABF method.

Due to the limited scope of this paper, we do not present phantom experiments in various conditions and do not perform detailed qualitative and quantitative analysis about the proposed adaptive formulas, the spatial resolution and noise characteristics of the proposed ABF. It is worth to point out that the bilateral filter used in this paper is just one option. Many other edge preserve filters such as non-local mean (NLM) [10] or even directional filters can be employed to replace the bilateral filter. The other format of adaptive formulas (Eqs. 7, 9&10) for filtering parameters selection is open to be investigated following the similar adaptive rules. For the purpose of simplicity, the filter used in this paper is one dimensional. It is easy to be extended to two dimensional (e.g. views direction) to consider more information to suppress the streak artifacts. We will present a more detailed analysis in a follow up journal paper.

References

[1] J. F. Barrett and N. Keat, "Artifacts in CT: Recognition and Avoidance," *RadioGraphics*, vol. 24, no. 6, pp. 1679–1691, Nov. 2004.

[2] B. D. Man, J. Nuyts, P. Dupont, G. Marchal, and P. Suetens, "Reduction of metal streak artifacts in X-ray computed tomography using a transmission maximum a posteriori algorithm," *IEEE Trans. Nucl. Sci.*, vol. 47, no. 3, pp. 977–981, Jun. 2000.

[3] G. Wang, D. L. Snyder, J. A. O'Sullivan, and M. W. Vannier, "Iterative deblurring for CT metal artifact reduction," *IEEE Trans. Med. Imaging*, vol. 15, no. 5, pp. 657–664, Oct. 1996.

[4] J. Hsieh, *Computed Tomography: Principles, Design, Artifacts, and Recent Advances*. SPIE Press, 2003.

[5] S. Leng *et al.*, "Streaking artifacts reduction in four-dimensional cone-beam computed tomography," *Med. Phys.*, vol. 35, no. 10, pp. 4649–4659, Oct. 2008.

[6] A. Sisniega *et al.*, "High-fidelity artifact correction for cone-beam CT imaging of the brain," *Phys. Med. Biol.*, vol. 60, no. 4, pp. 1415–1439, Feb. 2015.

[7] J. Hsieh, "Adaptive streak artifact reduction in computed tomography resulting from excessive x-ray photon noise," *Med. Phys.*, vol. 25, no. 11, pp. 2139–2147, Nov. 1998.

[8] "Papoulis: Probability, Random Variables and Stochastic Processes, Fourth Edition." [Online]. Available: http://www.mhhe.com/engcs/electrical/papoulis/. [Accessed: 20-Dec-2017].

[9] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*, 1998, pp. 839–846.

[10] A. Buades, B. Coll, and J. M. Morel, "A non-local algorithm for image denoising," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 2005, vol. 2, pp. 60–65 vol. 2.

Advancements in computed tomography for musculoskeletal imaging

Kai Mei, Felix K. Kopp, Johannes Hammel, Benedikt J. Schwaiger, Alexandra S. Gersing, Thomas Baum, Ernst J. Rummeny and Peter B. Noël

***Abstract*—Recent advancements in the field of computed tomography (CT) showed promising results for the application in musculoskeletal (MSK) imaging. Iterative reconstruction techniques and the resulting capability of dose reduction proofed to be eligible for osteoporosis diagnosis and (therapy) monitoring in clinical routine. Spectral CT allows for phantomless assessment of BMD and advanced tissue decomposition, predicting fracture risk in non-dedicated routine examinations or recognizing acute from old vertebral fractures. The potential introduction of sparse sampling acquisition schemes could even enhance the quality of the mentioned techniques for low dose MSK imaging. In this work, we present an overview of our recent CT developments for MSK imaging and provide an outlook over promising future technologies.**

***Index Terms*—CT imaging, spectral CT, sparse sampling, musculoskeletal imaging**

I. Introduction

The field of musculoskeletal (MSK) imaging is well suited for modalities involving ionizing radiation because of the high x-ray absorption of bone. Bone tumors and fractures are usually imaged by conventional radiography, while osteoporosis diagnosis and therapy monitoring is mostly performed by dual energy x-ray absorptiometry (DXA). However, computed tomography (CT) provides more precise information and improved diagnostics—compared to planar methods like DXA—because of additional three-dimensional information of the bone microstructure.

On the one hand, as a recently introduced technology, multi energy CT is now able to provide spectral based images in clinical studies. Material specific imaging leads to better quantification of calcium and thus better diagnostic performance [1]. These new detector technologies, based on dual-layer or photon-counting detectors, obtain energy dependent information directly in the sinogram space and can substantially influence and further improve iterative reconstruction algorithms, reducing artefacts.

On the other hand, over the last decade radiation dose reduction in CT has been extensively investigated because of the adverse effects of ionizing radiation. With the help of advanced reconstruction algorithms such as model-based iterative reconstruction, trabecular bone microstructure can be reconstructed from ultra-low dose CT acquisitions.

Sparse sampling, as an approach to preserve tube current but reduce the overall radiation exposure by obtaining fewer projections per rotation, was investigated for MSK imaging applications. With sparse sampling and iterative reconstruction, a significant radiation dose reduction can be achieved without compromising the diagnostic image quality [2] and with reduced influence of electronic detector noise [3].

In this work, we present an overview of our recent CT developments for MSK imaging and provide an outlook over promising future technologies.

II. Methods

A. Spectral CT for phantomless BMD assessment

Phantomless dual-layer spectral CT scans were performed for ex-vivo human vertebrae (n=13) and a phantom containing different known hydroxyapatite (HA) concentrations in a semi-anthropomorphic abdomen phantom, Fig. 1. Different degrees of obesity were simulated by adding extension rings of different size to the anthropomorphic phantom. Different tube current settings were also investigated (500, 250, 125 and 50 mAs) [4].

The calcium content of the vertebral samples was characterized and quantitated with reference values derived

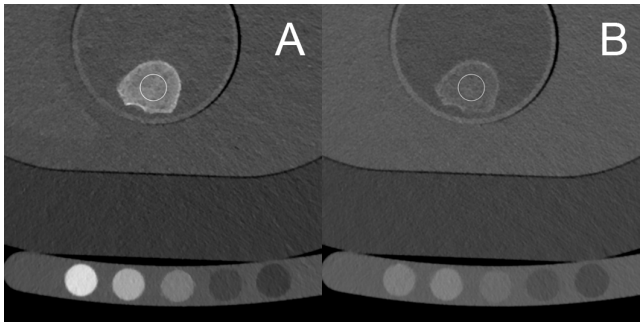


Fig. 1. Virtual monoenergetic images of bone imaging can be used for bone mineral density measurements without additional calibration phantoms. Fig. 1A shows the vertebral sample at virtual monochromatic energy at 50 keV. Fig. 1B shows the corresponding image at 200 keV. With conspicuous difference of the calcium attenuation, bone mineral density can be determined. (Window level 300 HU, width 1500 HU).

K. Mei*, F. Kopp, B. Schwaiger, A. Gersing and P. Noël are with the Department of Radiology, Technische Universität München, Germany.

T. Baum is with the Department of Neuroradiology, Technische Universität München, Germany.

P. Noël is with the Chair of Biomedical Physics, Physics Department, Technische Universität München, Germany.

*Email: kai.mei@tum.de

from spectral-based virtual monoenergetic images at 50 keV and 200 keV of the HA phantom. HA-specific BMD was derived. Values were compared to the conventional quantitative CT (qCT) measurements using the standard reference phantom (Mindways Osteoporosis Phantom, Austin, TX, USA).

B. *Spectral CT and material decomposition of bone marrow for fracture classification*

Sixty subjects with acute (n=41) and old (n=19) vertebral fractures were recruited and scanned in a dual-layer spectral CT. Virtual monoenergetic images at 50 keV (μ^{50}) and 200 keV (μ^{200}) were used to quantitate the composition of calcium (f_1), fat- (f_2) and water-based (f_3) materials inside bone marrow. The reconstruction algorithm assumes volume conservation of the three compositions.

$$\mu^{50} = f_1\,\mu_1{}^{50} + f_2\,\mu_2{}^{50} + f_3\,\mu_3{}^{50}$$

$$\mu^{200} = f_1\,\mu_1{}^{200} + f_2\,\mu_2{}^{200} + f_3\,\mu_3{}^{200}$$

$$f_1 + f_2 + f_3 = 1$$

Additional fat- and water-based spine images were used to distinguish between acute (water dominated) and chronic (fat dominated) spine fractures. The result was compared to the standard diagnostic procedure from magnetic resonance images.

C. *Sparse sampling with iterative reconstruction for osteoporosis diagnosis*

Twelve subjects with osteoporotic vertebral fractures and twelve age- and gender-matched controls were included in the study investigating the spine (average effective dose: 10 mSv). Similarly, twenty osteoporotic patients and twenty control subjects were included for the evaluation at the femoral bone.

Sparse sampling acquisition schemes were simulated on the raw projection data of a clinical CT by taking projections only at every 2nd, 4th and 10th position of the original angular position. Correspondingly, lower radiation doses were simulated by virtual lower tube current of 50%, 25%, and 10% of the original current.

Image reconstruction was performed with filtered backprojection (FBP) and maximum-likelihood based statistical iterative reconstruction (SIR) with ordered-subset paraboloidal surrogate and proper regularizations [5, 6]. BMD and trabecular bone parameters were extracted in T10 to L5 of the spine images. DXA-equivalent BMD values and T-scores were assessed in the femoral head and neck. The results between the different dose levels, numbers of projections, and image reconstructions were compared.

III. Results

A. *Spectral CT for phantomless BMD assessment*

Above 125 mAs, phantomless BMD assessment based on spectral information had an error ranged between -1.3% to 4.8%. In vertebral specimens, high correlations were found between values assessed with spectral CT and conventional qCT (r ranging between 0.96 and 0.99; p<0.001 for all) with different extension rings, Fig. 2A. And a high agreement was found in Bland Altman plots, Fig. 2B. Different degrees of obesity did not have a significant influence on measurements (p>0.05 for all).

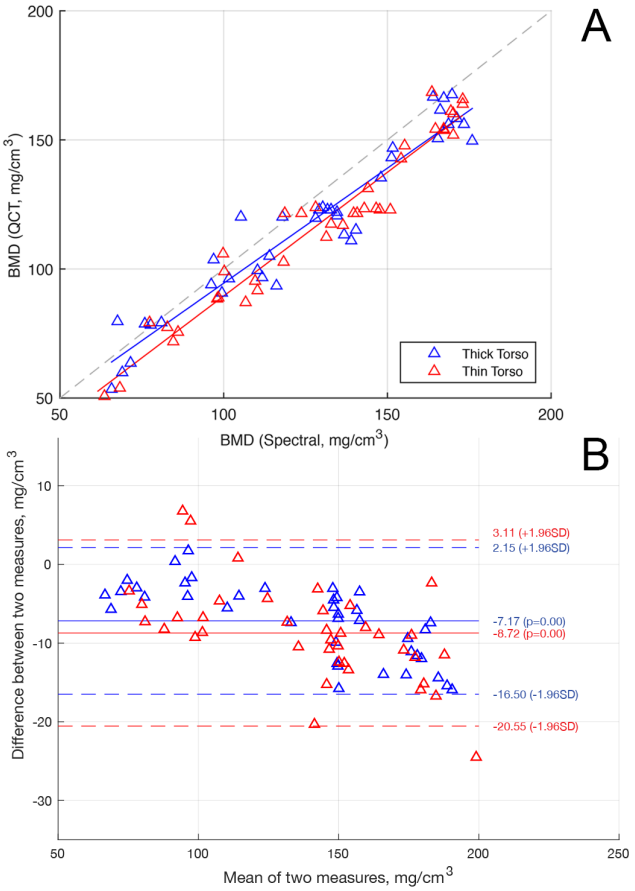


Fig. 2. Comparison of phantomless bone mineral density estimation in spectral CT versus traditional quantitative CT measurement. Fig. 2A shows the correlation plot for the measurement. Fig. 2B shows the corresponding Bland Altman plot. Both plots proved the possibility of calcium quantification in spectral CT.

B. *Spectral CT and material decomposition for fracture classification at the spine*

Fig. 3 shows a conventional CT image of an acute fracture and the spectral images. Images were surveyed and rated by radiologists. For the identification of acute fractures, conventional CT images showed a sensitivity of 63% and specificity of 84%, whereas with the material information derived from spectral CT the sensitivity (95%) and specificity (90%) were substantially higher. Accuracy for conventional images was 70% and increased to 93% with spectral CT. Readers considered MRI necessary for fracture age determination in 60% (based on conventional images) and 28% (based on spectral CT images), respectively. Interreader agreement was high for the classification of acute fractures based on conventional CT images (intraclass correlation coefficient (ICC): 0.79, 95% confidence interval (CI): [0.68–0.87]) and spectral CT images (ICC: 0.95, 95% CI: [0.92–0.97]).

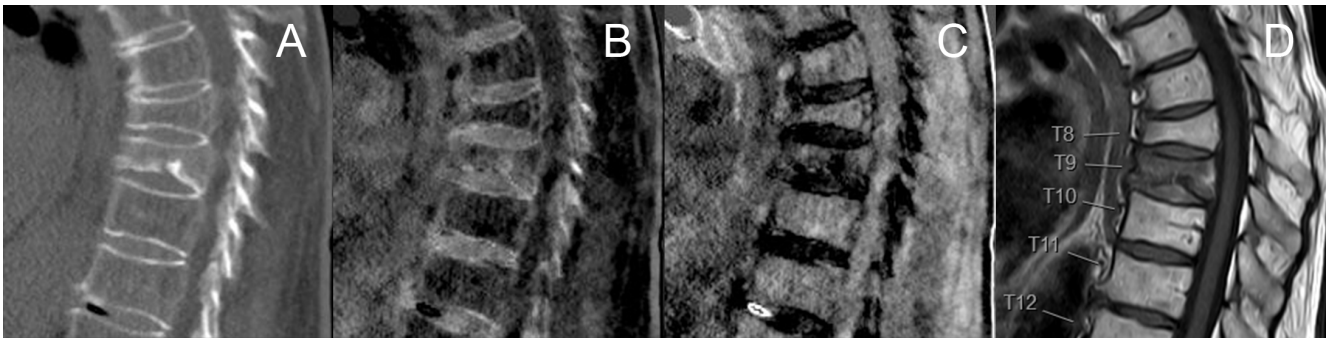


Fig. 3. Images of a patient's spine with accute fracture in thoracic vertebra T9. Fig. 3A shows the fracture in conventional CT. Fig. 3B and 3C show the corresponding water- and fat-based images. Fig. 3D shows the magnetic resonance image (MR). In the case of acute vertebral fractures, fat-dominated marrow is replaced by the water-dominated marrow, which is not visible in conventional CT; these patients are likely subject to additional MR scans.

C. *Sparse sampling CT*

At the spine, images from sparse sampling simulations had a better visual impression than the corresponding images with virtual lower tube current (Fig. 4) [7]. Absolute values of all trabecular parameters derived from ultra-low dose data were significantly different from those derived from original dose images ($p<0.05$). BMD, bone fraction and trabecular thickness were still consistently lower in subjects with than in those without fractures in the ultra-low dose data ($p<0.05$). BMD measurements in sparse sampling data were not significantly different from those derived from original dose images.

At the femur, virtual tube current lowering resulted in a statistically significant increase in DXA-equivalent T-scores for both FBP and SIR, whereas an opposite, but less markedly effect, was observed for sparse sampling. Images show increased noise for reduced radiation doses for virtual lower tube current compared to sparse sampling (Fig. 5). For both FBP and SIR, the application of sparse sampling with a reduction of up to 90% of projections did not lead to statistically significant proportions of misclassification when considering the World Health Organization (WHO) osteoporosis scheme, in contrast to findings with conventional dose reduction / virtually reduced tube currents.

IV. Discussion

In this work, we presented an overview of our results / developments in CT for MSK imaging. Firstly, we evaluated the application of spectral CT in osteoporosis diagnostics and for the classification of fractures at the spine. Secondly, we evaluated the effect of radiation dose reduction for MSK imaging, where we investigated two different approaches for radiation dose reduction—lowering the x-ray current and sparse sampling acquisition schemes—in combination with iterative reconstruction.

Spectral CT showed promising results for phantomless BMD assessment. With this technique, non-dedicated routine CT scans can retrospectively be used to determine the fracture risk. Compared to conventional CT, where a calibration phantom is needed during the scan, no additional examination is necessary if data without phantom is available. Moreover, advanced material decomposition allows the classification of fractures at the spine into acute and chronic fractures. Today, the standard modality for that classification task is MRI, while the detection of the fracture is performed with CT. Therefore, using spectral CT for the classification could spare the patient the discomfort of an additional examination.

The evaluation of the different radiation dose reduction techniques indicate that sparse sampling allows a further dose reduction than lowering the tube current. Therefore, sparse sampling acquisition schemes may allow ultra-low dose examination without loss of diagnostic value in the clinical routine.

V. Conclusion

Advancements in CT can significantly improve diagnostics in the field of MSK imaging. We showed that with spectral CT and advanced material decomposition, improved assessment of bone mineral density is feasible and the identification of acute fractures can be enhanced. Moreover, the presented results indicate that the potential future introduction of sparse sampling acquisition schemes allows a significant radiation dose reduction without compromising the diagnostic value.

Acknowledgment

We acknowledge support through NVIDIA with a hardware grant, the German Department of Education and Research (BMBF) under grant IMEDO (13GW0072C) and by the German Research Foundation (DFG) within the Research Training Group GRK 2274.

References

[1] S. Ehn, T. Sellerer, D. Muenzel, et al. "Assessment of quantification accuracy and image quality of a full-body dual-layer spectral CT system." *Journal of applied clinical medical physics,* 19(1), pp. 204-217, 2018.

[2] F. K. Kopp, R. Bippus, A. Sauter, et al. "Diagnostic value of sparse sampling computed tomography for radiation dose reduction: initial results", in *Medical Imaging 2018: Physics of Medical Imaging*, SPIE 2018.

[3] Z. Zhao, G. J. Gang and J. H. Siewerdsen, "Noise, sampling, and the number of projections in cone-beam CT with a flat-panel detector.", *Med. Phys.*, vol. 41, p. 061909, June 2014.

[4] K. Mei, B. J. Schwaiger, F. K. Kopp, et al. "Bone mineral density measurements in vertebral specimens and phantoms using dual-layer spectral computed tomography", *Scientific Reports*, *7*(1), p.17519, 2017.

[5] J. A. Fessler. "Statistical image reconstruction methods for transmission tomography", *Handbook of medical imaging* 2, p. 1-70, 2000.

[6] D. Kim, S. Ramani and J. A. Fessler, "Combining ordered subsets and momentum for accelerated X-ray CT image reconstruction", *IEEE transactions on medical imaging*, *34*(1), pp.167-178, 2015.

[7] K. Mei, F. K. Kopp, R. Bippus, et al. "Is multidetector CT-based bone mineral density and quantitative bone microstructure assessment at the spine still feasible using ultra-low tube current and sparse sampling?", *European radiology*, *27*(12), pp.5261-5271, 2017.

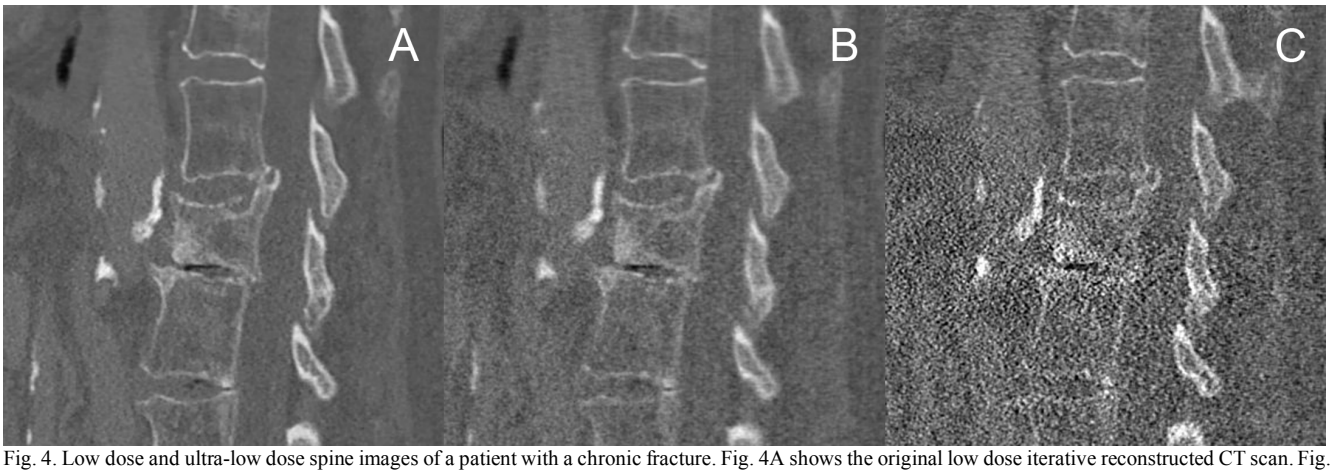


Fig. 4. Low dose and ultra-low dose spine images of a patient with a chronic fracture. Fig. 4A shows the original low dose iterative reconstructed CT scan. Fig. 4B shows the corresponding 10% sparse sampling simulation scan and Fig. 4C the equivalent 10% tube current. Combined with iterative reconstruction, sparse sampling provides better results when further lowering the radiation exposure for musculoskeletal CT imaging. (Window level 300 HU, width 1500 HU.)

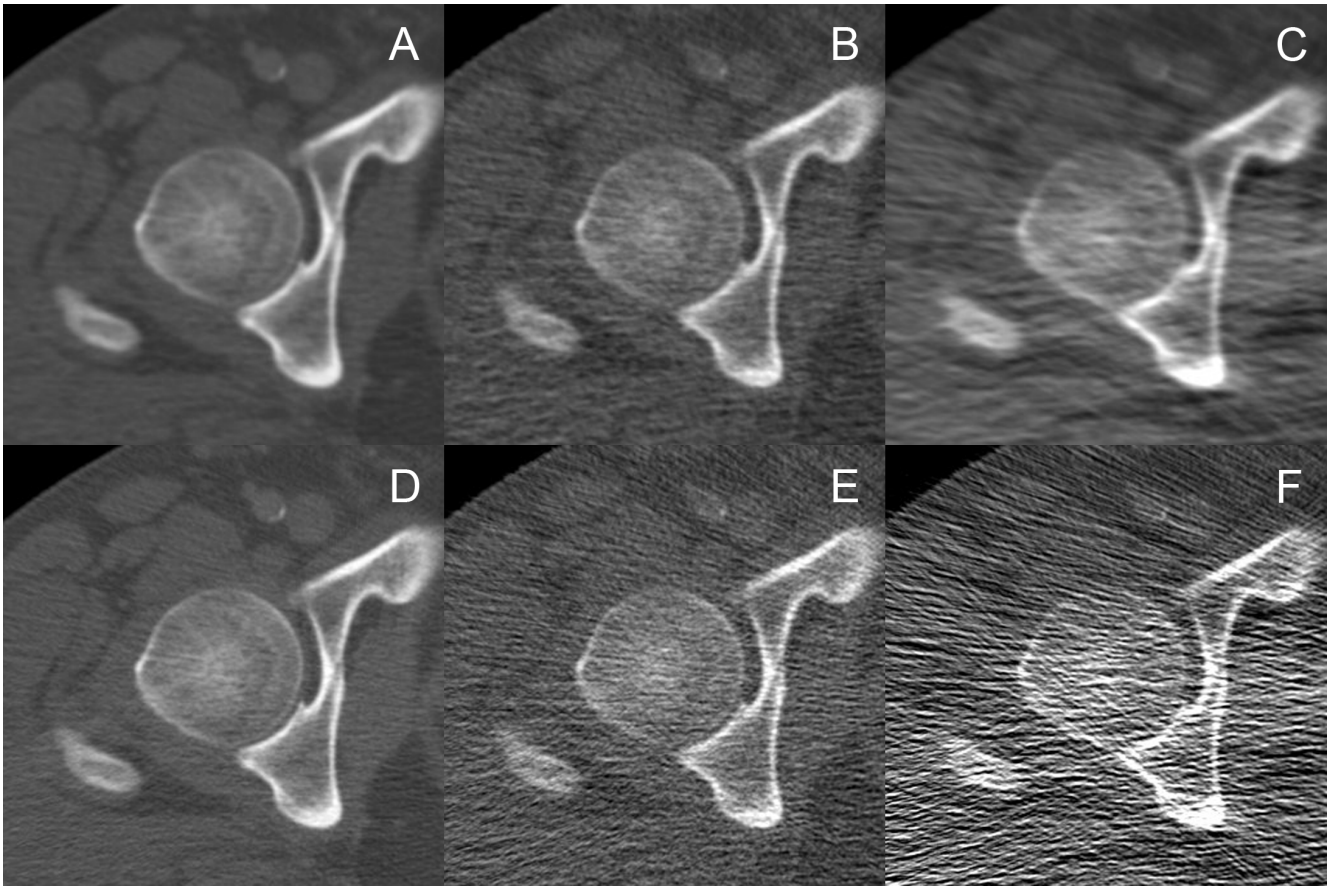


Fig. 5. Low dose and ultra-low dose images of a patient's right femur head. A, B and C (1st row) are reconstructed with iterative reconstruction (SIR) and D, E and F (2nd row) are reconstructed with filtered backprojection (FBP). A and D (left column) are the original low dose images. B and E (middle column) are the 10% sparse sample images, while C and F (right column) are the corresponding images with 10% tube current. SIR and sparse sampling have great potential to mitigate artefacts and to quantitate bone density in ultra-low dose musculoskeletal CT imaging. (Window level 300 HU, width 1500 HU.)

New GPU implementation of Separable Footprint (SF) Projector and Backprojector : first results

Camille Chapdelaine, Nicolas Gac, Ali Mohammad-Djafari and Estelle Parra

***Abstract*—Model-based iterative reconstruction methods enable to improve the quality of reconstruction in 3D X-ray Computed Tomography (CT). The main computational burden of these methods lies in successive projection and backprojection operations. Among existing pairs of projector and backprojector, Separable Footprint (SF) pair combines computational efficiency and accurate modelling of X-rays passing through the volume to image. In order to accelerate these operators, implementations on Graphical Processor Units (GPUs) for parallel-computing have been proposed for SF pair. Due to a CPU-loop, these implementations involve many memory transfers between CPU and GPU which are known to be the main bottleneck for GPU computing. In this paper, we investigate a new GPU implementation of SF projector and backprojector in order to minimize these memory transfers. Our proposed GPU SF projector and backprojector have no CPU-loop, and use two ray-driven kernels for the projection and one voxel-driven kernel for the backprojection. After having described their implementations, we study these operators as single modules and validate it in a MBIR method. Perspectives for this work are GPU optimizations and comparisons with the other existing implementations of SF pair.**

***Index Terms*—Computed Tomography, Separable Footprint, Graphical Processor Unit, iterative reconstruction methods**

I. Introduction

Compared with conventional filtered backprojection (FBP) methods, model-based iterative reconstruction (MBIR) methods have shown their advantages in terms of robustness and reconstruction quality [1]. Their major drawback is that they are highly computationally-costly due to successive projection and backprojection operations. The projection operator models the linear process of X-rays passing through the volume. The backprojector is defined as the mathematical adjoint of the projector.

Particularly in 3D applications, unmatched pairs of projector and backprojector (P/BP pairs) [2] have been widely used in order to alleviate the overall computational cost of iterative reconstruction methods. Although it has been found sufficient conditions for not having troubles when using an unmatched P/BP pair [2], these conditions have been derived for very simple iterative reconstruction methods which minimize the Euclidean distance between theoretical and measured projections by a gradient descent. For regularized least-squares using more sophisticated optimization algorithms, it has been highlighted in [3] that an unmatched P/BP pair can lead to suboptimal solutions or to divergence of the reconstruction algorithm. In order to avoid this mathematical approximation, computationally-efficient matched pairs have been recently proposed. The separable footprint (SF) pair [1] approximates the footprint of a voxel on the detector as a separable function which is trapezoidal in transaxial direction and rectangular in axial direction. The distance-driven (DD) pair [4] does the same kind of approximation, but models the footprint as rectangular both in transaxial and axial directions, which makes it less accurate than the SF pair [1].

To the best of our knowledge, two GPU implementations of SF P/BP pair have been proposed in [5], [6], the one of [6] having been shown faster than the one of [5]. Nevertheless, both these implementations have a CPU-loop on projection angles. This implies many memory transfers between CPU and GPU which are known to be the main bottleneck for accelerating computations on GPU. In order to minimize these memory transfers, in this paper, we investigate a new GPU implementation of SF P/BP pair and present first results. Our GPU implementation of SF projector is ray-driven and runs two independent kernels. Each of these kernels projects rays if the source is closer to x- or y-axis respectively. Our GPU SF backprojector is voxel-driven and runs only one kernel. These implementations have no CPU-loop.

In the following, we first present our SF projector and backprojector on GPU. Next, we analyze these operators as single modules, and validate it in simulation in a full MBIR method we detailed in [7]. Compared with an unmatched pair, we show the reconstruction converges in less global iterations with better convergence properties.

II. SF pair on GPU

A. SF projection

We consider a volume $\boldsymbol{f} = \{f(x_e, y_e, z_e)\}$ discretized in cubic voxels of side δ. Centers of voxels have normalized coordinates (x_e, y_e, z_e). On the detector, each cell's center has (u, v)-coordinates $(u_e\delta u, v_e\delta v)$. The volume is put between the detector and a source modeled as a point from which X-rays are sent in a conic beam. To acquire several perspectives, the source and the detector are rotated by an angle ϕ around z-axis. The rotation center has normalized (x, y, z)-coordinates $(x_{0_e}, y_{0_e}, z_{0_e})$. After X-rays have crossed the volume and

C. Chapdelaine is with Laboratoire des Signaux et des Systèmes (L2S) (CNRS–CentraleSupélec–Université Paris-Saclay) and SAFRAN SA, Safran Tech, Pôle Technologie du Signal et de l'Information (e-mail : camille.chapdelaine@l2s.centralesupelec.fr).

A. Mohammad-Djafari and N. Gac are with Laboratoire des Signaux et des Systèmes (L2S) (CNRS–CentraleSupélec–Université Paris-Saclay).

E. Parra is with SAFRAN SA, Safran Tech, Pôle Technologie du Signal et de l'Information.

reached the detector, the SF projection on a cell at angle ϕ reads

$$g(u_e, v_e, \phi) = l_{\theta_c}(u_e, v_e) \sum_{x_e, y_e} l_{\psi_v}(\phi; x_e, y_e) F_t(u_e, \phi; x_e, y_e) \times \sum_{z_e} F_a(v_e, \phi; x_e, y_e, z_e) f(x_e, y_e, z_e) \quad (1)$$

where $F_t(u_e, \phi; x_e, y_e)$ is the transaxial footprint and $F_a(v_e, \phi; x_e, y_e, z_e)$ the axial footprint of voxel (x_e, y_e, z_e) on cell (u_e, v_e) with projection angle ϕ [1]. Amplitude functions $l_{\theta_c}(u_e, v_e)$ and $l_{\psi_v}(\phi; x_e, y_e)$ are given by the A2 method in [1].

In order to avoid writing conflicts between threads, our GPU implementation of SF projector is ray-driven, i.e. one thread computes the SF projection of one ray defined by (u_e, v_e, ϕ). We follow each ray according to its primary direction, which is x-axis if the source is closer to x-axis, and y-axis otherwise [6]. Next, along this primary direction, we compute the voxels for which transaxial footprint and axial footprint are both non-zero. To ensure local memory accesses, the volume to project is copied on texture memory. Furthermore, variables related to the geometry of the acquisition are copied in constant memory.

Considering a ray with primary direction x or y leads to different calculations we need to do separately. That is why our SF projector runs two kernels. First kernel $proj\_x\_ker$ projects rays with primary direction x, while second kernel $proj\_y\_ker$ handles rays with primary direction y. These two kernels are independent since they compute disjoint sets of projections. In this paper, they are executed successively on one GPU but their parallel applications on two different GPUs can be considered.

For a ray with primary direction x, the corresponding thread performs a loop on x_e, $1 \leq x_e \leq N_x$. For each x_e, it computes the intersecting location $(x_e, y_e(x_e))$ with the ray, similarly to Joseph's method [8] :

$$y_e(x_e) = \frac{1}{\delta}\left(y_s(\phi) + \frac{y(u_e, \phi) - y_s(\phi)}{x(u_e, \phi) - x_s(\phi)}(x_e\delta - x_s(\phi))\right) \quad (2)$$

where $(x_s(\phi), y_s(\phi))$ is the (x, y)-location of the source and $(x(u_e, \phi), y(u_e, \phi))$ is the (x, y)-location of the center of cell (u_e, v_e) at projection angle ϕ. Then, for current x_e, the thread looks the voxels of which projections onto the median plane are between $(x_e, y_e(x_e) - 1)$ and $(x_e, y_e(x_e) + 1)$, i.e. pixels (x_e, y_e) of which left side is before $(y_e(x_e) + 1)$ and right side after $(y_e(x_e) - 1)$. Hence, the thread looks each y_e, such that

$$y_{e_{min}} \leq y_e \leq y_{e_{max}}, \begin{cases} y_{e_{min}} = \lfloor y_e(x_e) - 1.5 \rfloor \\ y_{e_{max}} = \lceil y_e(x_e) + 1.5 \rceil \end{cases} . \quad (3)$$

This makes a loop on y_e for each thread of kernel $proj\_x\_ker$ which is very small (typically size 4 or 5). For rays with primary direction y, the calculations are the same by reverting roles of x and y : the main loop is on y_e and we have a second loop on x_e, $x_{e_{min}} \leq x_e \leq x_{e_{max}}$. Here, we see the interest of dealing with rays with different primary directions in different kernels, in order to avoid divergence between threads.

In the double loop over (x_e, y_e) (with y_e or x_e varying in a very little set depending on the executed kernel), for each considered (x_e, y_e), we calculate the scaled transaxial footprint

$$F'_t(u_e, \phi; x_e, y_e) = l_{\psi_v}(\phi; x_e, y_e) F_t(u_e, \phi; x_e, y_e) \quad (4)$$

as described in [1]. Next, we find indices z_e for which $F_a(v_e, \phi; x_e, y_e, z_e) \neq 0$. Thanks to the chosen rectangular shape of the axial footprint, these indices are very simple to compute :

$$z_{e_{min}} \leq z_e \leq z_{e_{max}} \quad (5)$$

where

$$\begin{cases} z_{e_{min}} = \lfloor z_{0_e} - 0.5 + \frac{x_{\phi_e}\delta v}{D}(v_e - v_{c_e} - 0.5) \rfloor \\ z_{e_{max}} = \lceil z_{0_e} + 0.5 + \frac{x_{\phi_e}\delta v}{D}(v_e - v_{c_e} + 0.5) \rceil \end{cases} , \quad (6)$$

where D is the source-to-detector distance, and

$$x_{\phi_e} = \frac{R}{\delta} + (x_e - x_{0_e})\cos\phi + (y_e - y_{0_e})\sin\phi \quad (7)$$

where R is the source-to-rotation center distance. Knowing the bounds for z_e, threads run a loop on z_e to compute

$$F'_a(v_e, \phi; x_e, y_e) = \sum_{z_e = z_{e_{min}}}^{z_{e_{max}}} F_a(v_e, \phi; x_e, y_e, z_e) f(x_e, y_e, z_e). \quad (8)$$

This loop is very small and is typically size 3. Iteratively, through the double loop over x_e and y_e, threads calculate the sum

$$g'(u_e, v_e, \phi) = \sum_{x_e} \sum_{y_e = y_{e_{min}}}^{y_{e_{max}}} F'_t(u_e, \phi; x_e, y_e) F'_a(v_e, \phi; x_e, y_e) \quad (9)$$

in kernel $proj\_x\_ker$, and

$$g'(u_e, v_e, \phi) = \sum_{y_e} \sum_{x_e = x_{e_{min}}}^{x_{e_{max}}} F'_t(u_e, \phi; x_e, y_e) F'_a(v_e, \phi; x_e, y_e) \quad (10)$$

in kernel $proj\_y\_ker$. Finally, the thread handling ray (u_e, v_e, ϕ) computes the final value for the projection

$$g(u_e, v_e, \phi) = l_{\theta_c}(u_e, v_e) g'(u_e, v_e, \phi), \quad (11)$$

which is stored from GPU to CPU.

B. SF backprojection

Because SF projector and backprojector are matched, for a voxel (x_e, y_e, z_e), the SF backprojection is

$$b(x_e, y_e, z_e) = \sum_{\phi} \sum_{u_e} F_t(u_e, \phi; x_e, y_e) l_{\psi_v}(\phi; x_e, y_e) \times \sum_{v_e} F_a(v_e, \phi; x_e, y_e, z_e) l_{\theta_c}(u_e, v_e) g(u_e, v_e, \phi). \quad (12)$$

To prevent from writing conflicts between threads, we compute volume $\boldsymbol{b}$ by running a kernel $back\_ker$ which is voxel-driven : one thread calculates the backprojection of one voxel (x_e, y_e, z_e). Kernel $back\_ker$ has a main loop on projection angles ϕ. For each projection angle, a thread finds cells (u_e, v_e) overlapped by transaxial and axial footprints of voxel (x_e, y_e, z_e). u_e-coordinates of these cells are given by ordering

the projections of the four corners of pixel (x_e, y_e) in the middle plane, $\tau_0 \leq \tau_1 \leq \tau_2 \leq \tau_3$ [1] :

$$u_{e_{min}} \leq u_e \leq u_{e_{max}}, \begin{cases} u_{e_{min}} = \lfloor u_{c_e} - 0.5 + \frac{\tau_0}{\delta u} \rfloor \\ u_{e_{max}} = \lceil u_{c_e} + 0.5 + \frac{\tau_3}{\delta u} \rceil \end{cases} . \quad (13)$$

The set of v_e for which axial footprint $F_a(v_e, \phi; x_e, y_e, z_e) \neq 0$ is

$$v_{e_{min}} \leq v_e \leq v_{e_{max}}, \begin{cases} v_{e_{min}} = \lfloor v_{c_e} - 0.5 + \frac{\chi_0}{\delta v} \rfloor \\ v_{e_{max}} = \lceil v_{c_e} + 0.5 + \frac{\chi_1}{\delta v} \rceil \end{cases} \quad (14)$$

where $\chi_0 \leq \chi_1$ are the projections of voxels $(x_e, y_e, z_e - 0.5)$ and $(x_e, y_e, z_e + 0.5)$ respectively [1].

Next, double-loop on u_e and v_e is performed to compute the backprojection of voxel (x_e, y_e, z_e). The size of this double-loop is approximately the same for each voxel [5]. Projections $\boldsymbol{g}$ are copied on texture memory, so local memory accesses are ensured. Because of the separation of the voxel's footprint between transaxial and axial directions, the double-loop can be done on u_e then v_e or on v_e then u_e indifferently. Here, we chose the loop on u_e as main loop. For each u_e between $u_{e_{min}}$ and $u_{e_{max}}$, we run a loop on v_e, $v_{e_{min}} \leq v_e \leq v_{e_{max}}$. Like for z_e-loop in SF projector, this loop is typically size 3 and calculates each axial footprint to return

$$b_a(u_e, \phi; x_e, y_e, z_e) = \sum_{v_e=v_{e_{min}}}^{v_{e_{max}}} F_a(v_e, \phi; x_e, y_e, z_e) l_{\theta_c}(u_e, v_e) g(u_e, v_e, \phi). \quad (15)$$

Hence, the double-loop performs the summation

$$b_\phi(\phi; x_e, y_e, z_e) = \sum_{u_e=u_{e_{min}}}^{u_{e_{max}}} F_t(u_e, \phi; x_e, y_e) \times l_{\psi_v}(\phi; x_e, y_e) b_a(u_e, \phi; x_e, y_e, z_e) \quad (16)$$

Then, the backprojection can be updated

$$b(x_e, y_e, z_e) += b_\phi(\phi; x_e, y_e, z_e) \quad (17)$$

until all the projection angles have been considered.

III. First results on simulated data

In this section, we present first results of our GPU implementation of SF pair on a volume with 256^3 voxels. We use 64 projections of this volume uniformly distributed over $[0, 2\pi]$. Each projection has 256^2 pixels. For our tests, we use only one GPU, which is a NVIDIA's GeForce GTX TITAN X. We analyze our GPU SF P/BP as single modules and in a full MBIR method.

A. GPU SF projector and backprojector as single modules

Like [6], we show in table I the normalized root mean square error (NRMSE) with respect to our CPU version, which has been implemented following [1]. The NRMSE is

$$NRMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(\frac{y_i^{(GPU)} - y_i^{(CPU)}}{y_i^{(CPU)}} \right)^2}, \quad (18)$$

where y_i denotes projection or backprojection for ray or voxel i respectively. We see NRMSE is very low for both projector and backprojector, so our GPU implementation implies no deviation with respect to our CPU version. Furthermore, the coupling degree, introduced in [3], of our GPU SF pair is 1.0005, i.e. very close to 1, which means that our GPU SF P/BP pair is very well matched [3].

In table I, we also measure the number of used registers per thread. Because each of our kernels do many calculations, this number is quite high. Because the number of registers per block is limited, dimensioning thread blocks of the GPU must be done very carefully. The size of each block is $16 \times 16 \times 1$ for our SF projector, as for the backprojector. The size of the grid is $16 \times 16 \times 64$ for the projector and $16 \times 16 \times 256$ for the backprojector.

Lastly, we give in table I the computation times for our GPU SF projector and backprojector. We may underline that the presented version is the first one. Further optimizations, such as using multiple GPUs or finding a way to merge kernels $proj\_x\_ker$ and $proj\_y\_ker$ for the projection, are still needed and are undergoing works.

Operator	Computation time		NRMSE (%)	Registers/thread
	CPU	GPU		
Projector	143.9 s	8.3 s	1.2×10^{-4}	84
Backprojector	98.7 s	4.4 s	3.2×10^{-5}	63

TABLE I: Proposed GPU SF projector and backprojector as single modules

B. GPU SF pair in full MBIR algorithm

In order to fully validate our GPU implementation, we now test our GPU SF projector and backprojector in a full MBIR algorithm we presented in [7] with an unmatched pair. This algorithm performs alternate reconstruction and segmentation until convergence, and maximizes the joint posterior distribution of volume $\boldsymbol{f}$ and hyperparameters $\boldsymbol{\theta}$ of the prior models defined for the volume and the uncertainties on the projections. These prior models are detailed in [7]. Hyperparameters $\boldsymbol{\theta}$ are estimated jointly with the volume. According to Bayes's rule and by removing constant terms from the log-joint posterior distribution of $\boldsymbol{f}$ and $\boldsymbol{\theta}$, the stop criterion to maximize is

$$\mathcal{L}(\boldsymbol{f}, \boldsymbol{\theta}) = \ln(p(\boldsymbol{g}|\boldsymbol{f}, \boldsymbol{\theta})) + \ln(p(\boldsymbol{f}|\boldsymbol{\theta})) + \ln(p(\boldsymbol{\theta})). \quad (19)$$

We run the algorithm with our matched SF pair on GPU and compare the obtained results with the unmatched pair we used in [7]. This unmatched pair is described in [7], [9]. In order to compare the two considered P/BP pairs, for both experiments with matched and unmatched pair, we use same dataset, initialization and parameters, given in [7]. The algorithm has a maximum number of global iterations fixed to 20 and can be stopped before if the criterion (19) does not change by more than 10^{-6} % between two global iterations.

The results are shown in figure 1. We see the quality of reconstruction is good for both P/BP pairs : this validates our GPU implementation of SF P/BP pair. The obtained reconstructions with matched and unmatched pairs look visually the same : the SSIM [10] computed by MATLAB is approximately 1. Furthermore, the $\mathcal{L}_2$-relative error with respect to the original phantom, in table II, is approximately the same for both pairs. Nevertheless, as shown in table II,

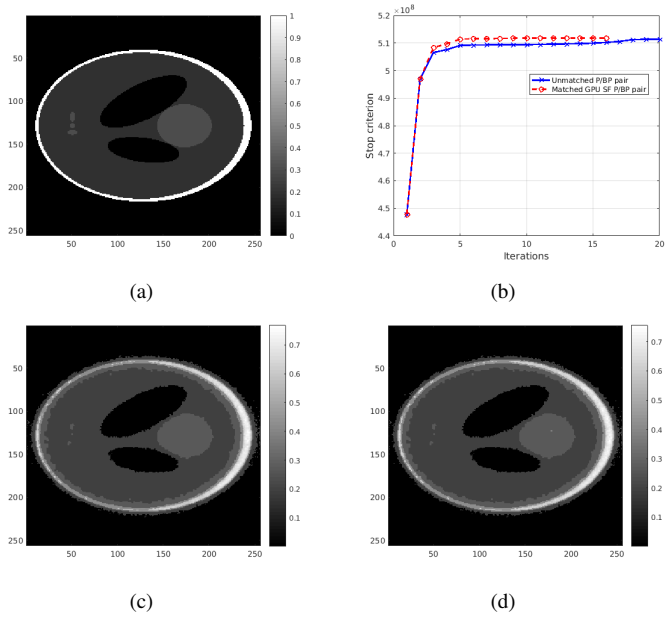


Fig. 1: Reconstruction of Shepp-Logan phantom (a) by the same MBIR method [7], with unmatched P/BP pair (c) and proposed matched GPU SF pair (d). Convergences are shown in (b).

the computation time is much longer when using matched SF pair, since it requires more calculations.

To deepen the comparison, we also analyze the convergence of the MBIR method. In table II and figure 1, we see the algorithm with matched SF pair converges in 16 global iterations, while with unmatched P/BP pair, it reaches the maximum number of 20. Moreover, in figure 1, the convergence of the algorithm is shown to be much smoother when using matched SF pair. In table II, the comparison of the final value of the stop criterion (19) shows that it is greater when using matched SF pair : that means the reconstruction obtained with an unmatched pair is suboptimal, as it has been already noticed in [3].

To conclude this section, we may emphasize that our GPU implementation of SF pair has been fully validated by its use in a MBIR method we developed with another P/BP pair [7]. The comparison between these two pairs has shown that using matched SF pair is more computationally intensive than using an unmatched pair, but ensures better convergence properties.

Used P/BP pair	Computation Time	$\mathcal{L}_2$-relative error	Number of global iterations	Final value of the criterion ($\times 10^8$)
Unmatched	**629.3 s**	**18.5 %**	20	5.1136
GPU SF	6517.0 s	18.7 %	**16**	**5.1183**

TABLE II: Comparison of the results for a MBIR method [7] with unmatched P/BP pair and proposed GPU SF pair

IV. Conclusion and perspectives

In this work, we have investigated a new GPU implementation of Separable Footprint (SF) projector and backprojector (P/BP) which minimizes memory transfers between CPU and GPU. We have presented a ray-driven GPU SF projector with two independent kernels which handle rays depending on their primary direction. Concerning SF backprojector, our GPU implementation is voxel-driven and uses only one kernel which takes advantage of the separability of the footprint in SF pair. Both our GPU SF projector and backprojector have no CPU-loop, so memory transfers are minimized.

As first results, we have fully validated the proposed GPU implementation. By computing the coupling degree [3], we have shown that our GPU SF implementation is well-matched. Our proposed implementation of SF pair has been shown to obtain very good results in a Model-Based Iterative Reconstruction (MBIR) algorithm. Compared with an unmatched P/BP pair, we have emphasized that the use of matched SF pair provides better convergence properties and accelerates the convergence in terms of global iterations.

Nevertheless, we have also seen the computational cost when using a matched pair is much higher, and results in a much longer computation time, as highlighted in table II. To reduce this computation time, further GPU optimizations are still necessary, such as, for instance, executing the two kernels of our GPU SF projector on two different GPUs, which is possible since these two kernels are independent. Moreover, our experiments have been done on relatively small volumes. Consequently, scaling our implementation for much larger volumes, as the aforementioned optimizations, is a perspective for this work. After having proceeded with these optimizations, comparisons with previous GPU implementations of SF P/BP pair [5], [6] will remain to be done.

References

[1] Y. Long, J. A. Fessler, and J. M. Balter, "3D forward and back-projection for X-ray CT using separable footprints," *IEEE transactions on medical imaging*, vol. 29, no. 11, pp. 1839–1850, 2010.

[2] G. L. Zeng and G. T. Gullberg, "Unmatched projector/backprojector pairs in an iterative reconstruction algorithm," *IEEE transactions on medical imaging*, vol. 19, no. 5, pp. 548–555, 2000.

[3] F. Arcadu, M. Stampanoni, and F. Marone, "On the crucial impact of the coupling projector-backprojector in iterative tomographic reconstruction," *arXiv preprint arXiv:1612.05515*, 2016.

[4] B. De Man and S. Basu, "Distance-driven projection and backprojection in three dimensions," *Physics in medicine and biology*, vol. 49, no. 11, p. 2463, 2004.

[5] M. Wu and J. A. Fessler, "GPU acceleration of 3D forward and backward projection using separable footprints for X-ray CT image reconstruction," in *Proc. Intl. Mtg. on Fully 3D Image Recon. in Rad. and Nuc. Med*, vol. 6. Citeseer, 2011, p. 021911.

[6] X. Xie, M. G. McGaffin, Y. Long, J. A. Fessler, M. Wen, and J. Lin, "Accelerating separable footprint (SF) forward and back projection on GPU," in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2017, pp. 101 322S–101 322S.

[7] C. Chapdelaine, A. Mohammad-Djafari, N. Gac, and E. Parra, "A 3D Bayesian Computed Tomography Reconstruction Algorithm with Gauss-Markov-Potts Prior Model and its Application to Real Data," *Fundamenta Informaticae*, vol. 155, no. 4, pp. 373–405, 2017.

[8] P. M. Joseph, "An improved algorithm for reprojecting rays through pixel images," *IEEE transactions on medical imaging*, vol. 1, no. 3, pp. 192–196, 1982.

[9] N. Gac, A. Vabre, A. Mohammad-Djafari, A. Rabanal, and F. Buyens, "GPU implementation of a 3D Bayesian CT algorithm and its application on real foam reconstruction," in *The First International Conference on Image Formation in X-Ray Computed Tomography*, 2010, pp. 151–155.

[10] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

High-Fidelity Modeling of Detector Lag and Gantry Motion in CT Reconstruction

Steven Tilley II, Alejandro Sisniega, Jeffrey H. Siewerdsen, J. Webster Stayman

***Abstract*—Detector lag and gantry motion during x-ray exposure and integration both result in azimuthal blurring in CT reconstructions. These effects can degrade image quality both for high-resolution features as well as low-contrast details. In this work we consider a forward model for model-based iterative reconstruction (MBIR) that is sufficiently general to accommodate both of these physical effects. We integrate this forward model in a penalized, weighted, nonlinear least-square style objective function for joint reconstruction and correction of these blur effects. We show that modeling detector lag can reduce/remove the characteristic lag artifacts in head imaging in both a simulation study and physical experiments. Similarly, we show that azimuthal blur ordinarily introduced by gantry motion can be mitigated with proper reconstruction models. In particular, we find the largest image quality improvement at the periphery of the field-of-view where gantry motion artifacts are most pronounced. These experiments illustrate the generality of the underlying forward model, suggesting the potential application in modeling a number of physical effects that are traditionally ignored or mitigated through pre-corrections to measurement data.**

I. INTRODUCTION

The need for high-resolution, quantitatively accurate CT reconstructions has increased with the rise of application-specific systems. For example, Cone-Beam CT (CBCT) mammography [1] and extremities systems [2] require high resolution to detect microcalcifications and visualize fine trabecular structure, respectively. Point-of-care CBCT head imaging [3] similarly requires highly accurate reconstruction of relative attenuation values to detect low contrast bleeds. Such dedicated imaging systems often use flat-panel detectors, which are selected for their high-resolution capability and ease of integration into compact systems. However, a number of physical effects including scintillator blur and detector lag can degrade measurement data, challenging the above applications. Similar examples of hardware limitations challenging particular applications can be found in traditional Multi-Detector CT (MDCT). For example, cardiac and emergency room scanning place high demands on lowering the scan time. The high rotation rates in such applications can result in significant blurring effects due to gantry motion during the integration time of the detector.

Previous work has suggested that such hardware limitations can be compensated through explicit modeling and incorporation into a Model-Based Iterative Reconstruction (MBIR) algorithm. In particular, we have found that scintillator blur and focal-spot blur in flat-panel systems can be modeled for potential resolution recovery [4]. The forward model used in that work is very general and permits incorporation of a wide range of physical effects. In this work we adopt the same mathematical form for the underlying forward model and apply MBIR to detector lag and gantry motion.

Detector lag results from the detector trapping and later releasing charge, causing a fraction of the signal from previously acquired projections to be added (temporally blurred) into subsequent projections [5], [6]. Detector lag effects are usually low contrast and extend across large areas of the reconstruction, originating near high contrast objects. A classic example of lag artifacts are the low contrast trails arcing off the skull into the brain in flat-panel-based head imaging. Traditionally, detector lag corrections are applied through preprocessing the measurements prior to reconstruction [7], [8]. To our knowledge, this work is the first attempt to correct for lag within the forward model of an MBIR approach.

Gantry motion blur shares some similarity with lag in that there is an effective blurring over angle. However, this blur occurs within a single measurement - effectively integrating an arc of projection images based on how far the source and detector have rotated during an integration period. Such blur exhibits as an azimuthal smearing of the CT volume and is most pronounced toward the edge of the field of view. Gantry motion effects have been addressed in hardware (e.g., collecting data with a step-and-shoot protocol or more complicated methods [9]) and in software (e.g., incorporating a blur model into a linearized forward model for MBIR [10]).

In this paper, we introduce specific models of detector lag and gantry motion, and integrate those models into the general form in [4]. Simulation studies are conducted for both blur scenarios. Image reconstructions are performed using both traditional (unmodeled blur) and the proposed high-fidelity models, and the resulting images are compared. Preliminary physical-experiment results using a head phantom and a CBCT test bench are also shown to illustrate application in a real system.

Department of Biomedical Engineering, Johns Hopkins University. email: web.stayman@jhu.edu.

TABLE I
BLUR KERNEL PARAMETERS

	0	1	2	3
a	—	0.998	0.0991	0.0152
b	0.965	0.0165	0.000572	4.51e-05

Fig. 1. A portion of the digital phantom for motion blur studies. The left most circle in this figure is at the center of the phantom. The circles are separated by $20\,\mathrm{mm}$.

II. METHODS

Both detector lag and gantry motion blur scenarios can use the same general forward model presented in [4]:

$$\bar{\boldsymbol{y}} = \mathbf{B}\exp\left(-\mathbf{A}\boldsymbol{\mu}\right) \tag{1a}$$

$$y \sim \mathcal{N}\left(\bar{\boldsymbol{y}}, \mathbf{K}\right) \tag{1b}$$

where $\mathbf{B}$, $\mathbf{A}$, and $\mathbf{K}$ are matrices, $\boldsymbol{\mu}$ is a vector of attenuation values, and $\boldsymbol{y}$ is a vector of measurements. The corresponding penalized likelihood objective function is

$$\|(\boldsymbol{y} - \mathbf{B}\exp\left(-\mathbf{A}\boldsymbol{\mu}\right))\|^2_{\mathbf{K}^{-1}} + \beta\,\mathrm{R}(\boldsymbol{\mu}) \tag{2}$$

where R is a penalty function and β is the penalty strength. The $\boldsymbol{\mu}$ that minimizes (2) is the reconstruction. A traditional forward model has $\mathbf{A}$ as the system matrix, $\mathbf{B}$ as a diagonal matrix which scales measurements by a gain factor (e.g., photon flux, etc.), and $\mathbf{K}$ as a diagonal matrix of measurement variances. However, the reconstruction method in [4] which minimizes the objective function (2) makes few assumptions about these matrices, allowing the forward model (1) to incorporate many physical properties. This reconstruction method may utilize ordered subsets and Nesterov momentum acceleration [11], [12].

A. Detector lag

Detector lag may be modeled as a convolution blur where the blur kernel is a sum of exponentials [8]:

$$h[k] = \begin{cases} b_0\delta[k] + \sum_{i=1}^{3} b_i \exp\left(-k a_i\right) & \text{if } 0 \le k < K \\ 0 & \text{otherwise} \end{cases}. \tag{3}$$

The K parameter in (3) determines the length of the blur kernel (i.e., the number of nonzero terms). This convolution is incorporated into $\mathbf{B}$ in (1a). Specifically, each row of $\mathbf{B}$ weights and combines a series of unblurred measurement data to form a measurment with lag. Physical blur kernel parameters for our test bench system were estimated from the falling edge of a bare-beam scan [8]. The estimated parameters used throughout this work are shown in Table I. Because $\mathbf{B}$ is no longer block diagonal with regards to projection number (i.e., $\mathbf{B}$ blurs among projections), we cannot trivially apply ordered subsets to speed convergence [13].

A simulation study was conducted with an ellipsoidal "head" phantom of fat surrounded by bone. Data were generated from a phantom with $0.25\,\mathrm{mm} \times 0.25\,\mathrm{mm}$ voxels on a system with $580.0\,\mathrm{mm}$ Source-Isocenter Distance (SID) and $800\,\mathrm{mm}$ Source-Detector Distance (SDD). Data were projected onto a detector with $0.278\,\mathrm{mm}$ pixels with $0.5 \times 10^5\,\mathrm{photons\,pixel}^{-1}$ over $360°$ in $1°$ increments. Poisson noise was added and data were binned by a factor of two, resulting in $0.556\,\mathrm{mm}$ pixels with $10^6\,\mathrm{photons\,pixel}^{-1}$. We then blurred the data by the calculated blur kernel with a length of $K = 359$ and added readout noise ($\sigma_{ro} = 7.12\,\mathrm{photons}$). Blurring the data after adding Poisson noise correlates the noise as in real systems [6].

Data were reconstructed with $0.5\,\mathrm{mm} \times 0.5\,\mathrm{mm}$ voxels, a quadratic regularizer, and the separable footprints projector [14]. Two reconstruction methods were used: identity blur modeling (i.e., no blur modeling) and detector lag blur modeling (with a kernel length of $K = 101$). In this work we assume uncorrelated noise for simplicity, specifically

$$\mathbf{K} = \mathrm{D}\{\boldsymbol{y}\} + \sigma^2_{ro}, \tag{4}$$

where $\mathrm{D}\{\cdot\}$ is a diagonal matrix with its argument on the diagonal). We used 5000 iterations and Nesterov acceleration. Reconstructions were noise matched by varying β and taking the standard deviation of the attenuation values in the center of the image.

Additionally, we scanned a physical head phantom on a CBCT test bench with parameters similar to those in the simulation study, except projection data were acquired in half angle increments. In order to focus on only detector lag in this preliminary study, we corrected the data for beam hardening due to water, scatter, and glare, as described in [15]. Data were reconstructed with $0.5\,\mathrm{mm} \times 0.5\,\mathrm{mm} \times 0.5\,\mathrm{mm}$ voxels using the same blur models as the simulation study (the blur model used a lag kernel length of $K = 201$). Nesterov acceleration was used with 4000 iterations. We used a quadratic regularizer, and the same regularization strength for both reconstructions.

B. Gantry motion

Gantry motion blur is the result of a continuous integration over angle, and may be modeled as

$$\bar{\boldsymbol{y}}_i = \mathbf{B}_2 \int_{\psi=\theta_i-\Delta\theta/2}^{\theta_i+\Delta\theta/2} \exp(-\mathbf{A}_\psi\boldsymbol{\mu})\mathrm{d}\psi \tag{5}$$

where $\bar{\boldsymbol{y}}_i$ is the mean measurement vector at projection i and gantry angle θ_i, $\Delta\theta$ is the angular distance over which data is collected for projection i, and $\mathbf{A}_\psi$ is the projection matrix at angle ψ. A discrete approximation is achieved by

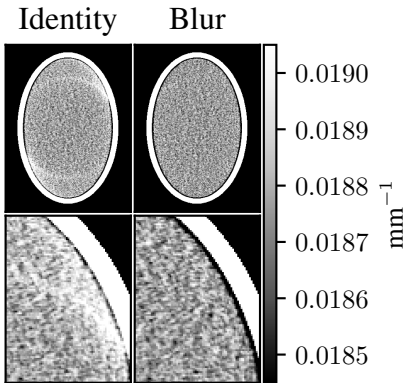


Fig. 2. Simulation phantom reconstructions with the identity model (left) and the detector lag model (right). The second row of images shows a smaller portion of the phantom for better visualization.

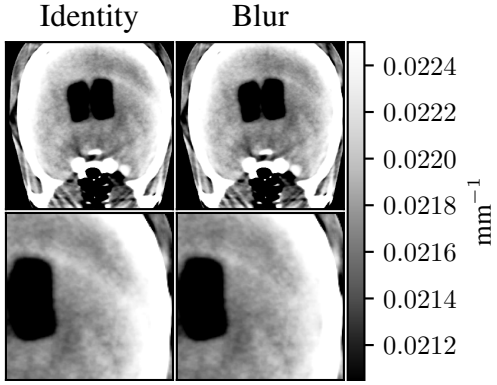


Fig. 3. Head phantom bench reconstructions with the identity model (left) and the detector lag model (right). The second row shows a smaller portion of the phantom to better visualize the detector lag effect.

oversampling in projection angle and summing the results to obtain the measurement sampling:

$$\bar{\boldsymbol{y}}_i = \mathbf{B}_2 J^{-1} \sum_{j=0}^{J} \exp\left(-\mathbf{A}_{\psi}\boldsymbol{\mu}\right) \tag{6}$$

$$\psi(j) = \theta_i + \Delta\theta\,(j/J - 1/2) \tag{7}$$

where J is the angular oversampling factor. $\mathbf{B}$ from (1) contains $\mathbf{B}_2$ and the summation term in (6), and $\mathbf{A}$ contains all the $\mathbf{A}_{\psi}$ used in (6). For example, if the measurement data contains 360 projections and J is 3, then $\mathbf{A}$ results in 1080 projections, and every three consecutive projections are summed together as part of $\mathbf{B}$.

A circular simulation phantom with a diameter of $25\,\mathrm{cm}$ and multiple round ROIs at different distances from the center of rotation was used to evaluate the proposed algorithm. A subset of this phantom is shown in Fig. 1. A continuous motion system was simulated with $500\,\mathrm{mm}$ SID, $1000\,\mathrm{mm}$ SDD, and 1000 projections per rotation. This geometry was chosen to approximate high-resolution MDCT systems. Data were generated from a phantom with $0.1\,\mathrm{mm} \times 0.1\,\mathrm{mm}$ voxels and a detector with $0.125\,\mathrm{mm}$ pixel pitch. We projected at $51\,000$ equally spaced angles over a 360° rotation. Poisson noise was added prior to binning to 1000 projections and spatially binning to $0.25\,\mathrm{mm}$ pixels. The photon flux after binning was $10^5\,\mathrm{photons\,pixel}^{-1}$. Finally, readout noise was added to the data ($\sigma_{ro} = 7.12\,\mathrm{photons}$).

Data were reconstructed with $0.2\,\mathrm{mm} \times 0.2\,\mathrm{mm}$ voxels. We used two blur models: an identity blur model (no blur, $\mathbf{A}$ produces 1000 projections), and a gantry motion blur model with an angular oversampling factor of $J = 5$ ($\mathbf{A}$ produces 5000 projections). We used an uncorrelated noise model (4), the Huber penalty ($\delta = 10^{-3}$) [16], and the separable footprints projector [14]. Nesterov acceleration was used with 1000 iterations and 10 subsets. Bias/noise measurements were calculated for each ROI. Bias was the Root Mean Squared Error (RMSE) between a noiseless reconstruction and truth at the ROI, and noise was the RMSE between a noisy reconstruction and a noiseless reconstruction in a nearby region. Bias and noise were calculated for multiple penalty strengths to obtain a bias/noise curve for each method. Data were also reconstructed with a quadratic penalty and $J = 5$ to compare this penalty to the Huber penalty.

III. RESULTS

A. Detector lag

The detector lag digital phantom reconstructions are shown in Fig. 2. The reconstructions are approximately noise matched — $7.96 \times 10^{-5}\,\mathrm{mm}^{-1}$ (identity) and $7.91 \times 10^{-5}\,\mathrm{mm}^{-1}$ (blur). When no blur model is used, detector lag causes a bright trail atrifact arcing off the skull and into the interior of the head. When blur modeling is used, this effect is eliminated. When lag modeling was applied to bench data, the bright trail off the skull was dramatically reduced (Fig. 3). The fact that the trail was not completely removed may be due to an insufficient number of iterations (non-converged estimate) or an inaccurate estimate of the lag blur kernel.

B. Gantry motion

Gantry motion results are summarized in Fig. 4. The bias/noise tradeoff is shown for each ROI at varying distances from the center of rotation. The identity model suffers from increased bias at large distances from the center of rotation, while the blur model bias is relatively unchanged (suggesting a recovery of spatial resolution). The identity model appears to outperform the blur model at $20\,\mathrm{mm}$ to $60\,\mathrm{mm}$ from the center of rotation, although the difference is small. These results are confirmed in the reconstructions in Fig. 4. These reconstructions were approximately noise matched at the ROI furthest from the center of rotation by altering penalty strength (noise is $3.790 \times 10^{-4}\,\mathrm{mm}$ for the identity model and $3.407 \times 10^{-4}\,\mathrm{mm}$ for the blur model). The circles in the identity model reconstruction get blurrier along the direction of rotation as distance from the center increases. However, with the blur model the circles are accurately reconstructed. Additionally, the blur model's bias

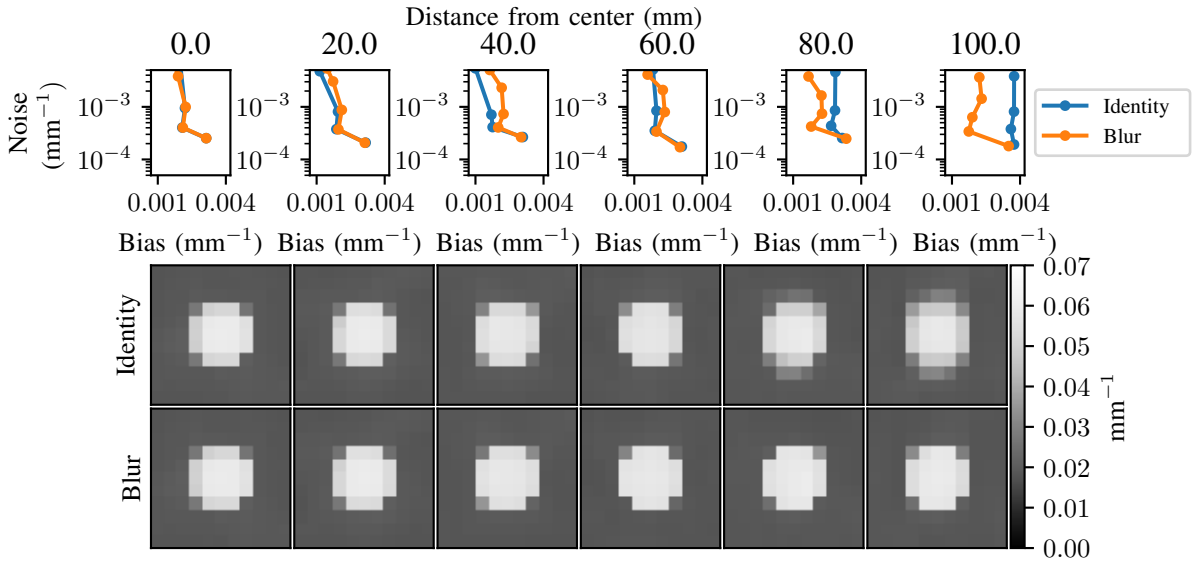


Fig. 4. Bias/noise curves (top) and reconstructions (bottom) for each ROI in Fig. 1. Each column corresponds to a distance from the center of rotation. The top row reconstructions use the identity model and the bottom row reconstructions use the gantry motion blur model.

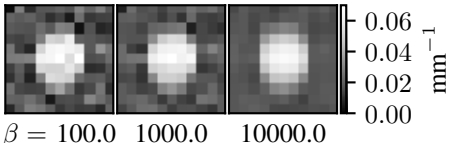


Fig. 5. Quadratic penalty reconstructions of the 100 mm ROI with blur modeling.

improvement in the 20 mm to 60 mm range is difficult to visualize.

Fig. 5 shows the 100 mm ROI reconstructed with the blur model and the quadratic penalty at three different penalty strengths. With this penalty the blur model is unable to deblur the circle without a substantial increase in noise.

IV. DISCUSSION

We have shown that the general reconstruction method presented previously [4] is capable of reducing effects due to detector lag and gantry motion blur. The methods presented here could trivially be extended to model detector lag with other forms (i.e., not sum of exponentials) or more complicated forms of gantry motion (e.g., when data acquisition only occurs during a fraction of the rotation). Additionally, these models may be combined with each other or other forms of blur, such as focal spot blur and scintillator blur, to further improve image quality.

A major limitation of modeling detector lag is the inability to use ordered subsets to speed convergence. In practice, one may initialize with a reconstruction without a lag model and with ordered subsets to get a relatively accurate estimate, and then reconstruct with the lag model for a handful of iterations. Additionally, a more accurate initialization may be obtained by lag correcting the projection data prior to simple MBIR (i.e., without lag modeling), and then the final reconstruction obtained with a few iterations with the full lag model and the original, uncorrected measurement data.

While this work is still preliminary, we note that the edge preserving Huber penalty plays an important role in the gantry motion reconstructions. We believe the quadratic penalty's tendency to enforce smooth edges prevents the fidelity term from deblurring the gantry motion effects. In contrast, the Huber penalty doesn't penalize sharp edges to the same degree, and allows the fidelity term to deblur the image. Ongoing work will further explore these issues by analyzing different penalties (e.g., sweeping the δ parameter) and using more complicated image quality targets.

High-fidelity system modeling with MBIR can improve image quality by overcoming hardware limitations such as detector lag and gantry motion. However, application specific systems may have different limitations and constraints. The forward model and MBIR algorithm used in this work are sufficiently general to accommodate many physical effects, and may therefore be used to improve image quality and quantitative accuracy in a wide range of clinical scenarios.

V. ACKNOWLEDGMENTS

This work was supported in part by NIH grant F31 EB023783. The bench data used in this work was acquired with support of an academic-industry partnership with Carestream (Rochester, NY). The authors would like to thank Ali Uneri and Yoshi Otake for GPU software contributions.

REFERENCES

[1] J. M. Boone and K. K. Lindfors, "Breast CT: Potential for breast cancer screening and diagnosis." *Future oncology (London, England)*, vol. 2, no. 3, pp. 351–356, 2006.

[2] E. Marinetto, M. Brehler, A. Sisniega, Q. Cao, J. W. Stayman, J. Yorkston, J. H. Siewerdsen, and W. Zbijewski, "Quantification of bone microarchitecture in ultrahigh resolution extremities conebeam CT with a CMOS detector and compensation of patient motion," in *Computer Assisted Radiology 30th International Congress and Exhibition*, Heidelberg, Germany, Jun. 2016, pp. S20–S21.
[3] J. Xu, A. Sisniega, W. Zbijewski, H. Dang, J. W. Stayman, M. Mow, X. Wang, D. H. Foos, V. E. Koliatsos, N. Aygun, and J. H. Siewerdsen, "Technical assessment of a prototype cone-beam CT system for imaging of acute intracranial hemorrhage," *Medical Physics*, vol. 43, no. 10, pp. 5745–5757, Oct. 2016.
[4] S. Tilley, M. Jacobson, Q. Cao, M. Brehler, A. Sisniega, W. Zbijewski, and J. W. Stayman, "Penalized-Likelihood Reconstruction with High-Fidelity Measurement Models for High-Resolution Cone-Beam Imaging," *IEEE Transactions on Medical Imaging*, vol. PP, no. 99, pp. 1–1, 2017.
[5] L. E. Antonuk, Y. El-Mohri, J. H. Siewerdsen, J. Yorkston, W. Huang, V. E. Scarpine, and R. A. Street, "Empirical investigation of the signal performance of a high-resolution, indirect detection, active matrix flat-panel imager (AMFPI) for fluoroscopic and radiographic operation," *Medical Physics*, vol. 24, no. 1, pp. 51–70, Jan. 1997.
[6] J. H. Siewerdsen, L. E. Antonuk, Y. el-Mohri, J. Yorkston, W. Huang, J. M. Boudry, and I. A. Cunningham, "Empirical and theoretical investigation of the noise performance of indirect detection, active matrix flat-panel imagers (AMFPIs) for diagnostic radiology," *Medical Physics*, vol. 24, no. 1, pp. 71–89, Jan. 1997.
[7] J. Starman, J. Star-Lack, G. Virshup, E. Shapiro, and R. Fahrig, "A nonlinear lag correction algorithm for a-Si flat-panel x-ray detectors," *Medical Physics*, vol. 39, no. 10, pp. 6035–6047, Oct. 2012.
[8] ——, "Investigation into the optimal linear time-invariant lag correction for radar artifact removal," *Medical Physics*, vol. 38, no. 5, pp. 2398–2411, May 2011.
[9] T. Nowak, M. Hupfer, F. Althoff, R. Brauweiler, F. Eisa, C. Steiding, and W. A. Kalender, "Time-delayed summation as a means of improving resolution on fast rotating computed tomography systems," *Medical Physics*, vol. 39, no. 4, pp. 2249–2260, Apr. 2012.
[10] J. Cant, W. J. Palenstijn, G. Behiels, and J. Sijbers, "Modeling blurring effects due to continuous gantry rotation: Application to region of interest tomography," *Medical Physics*, vol. 42, no. 5, pp. 2709–2717, May 2015.
[11] Y. Nesterov, "Smooth minimization of non-smooth functions," *Mathematical Programming Journal, Series A*, vol. 103, pp. 127–152, 2005.
[12] D. Kim, S. Ramani, and J. A. Fessler, "Combining Ordered Subsets and Momentum for Accelerated X-Ray CT Image Reconstruction," *IEEE Transactions on Medical Imaging*, vol. 34, no. 1, pp. 167–178, 2015.
[13] J. Nuyts, B. De Man, J. A. Fessler, W. Zbijewski, and F. J. Beekman, "Modelling the physics in iterative reconstruction for transmission computed tomography," *Physics in medicine and biology*, vol. 58, no. 12, pp. R63–R96, Jun. 2013.
[14] Y. Long, J. A. Fessler, and J. M. Balter, "3D forward and back-projection for X-ray CT using separable footprints." *IEEE Transactions on Medical Imaging*, vol. 29, no. 11, pp. 1839–50, Nov. 2010.
[15] A. Sisniega, W. Zbijewski, J. Xu, H. Dang, J. W. Stayman, J. Yorkston, N. Aygun, V. Koliatsos, and J. H. Siewerdsen, "High-fidelity artifact correction for cone-beam CT imaging of the brain," *Physics in Medicine and Biology*, vol. 60, no. 4, p. 1415, 2015.
[16] P. J. Huber, *Robust Statistics*. New York: Wiley, 1981.

Iterative Image Reconstruction for CT by Emphasizing Local Image Quality

Jianmei Cai, Jiayu Duan, Yongyi Shi and Xuanqin Mou

***Abstract*—Iterative reconstruction with an appropriate regularization term can well handle noisy data reconstruction for low-dose X-ray CT scanning, which will effectively alleviate the X-ray radiation exposed to the patient. To acquire the best image quality from the reconstruction, the selection of the regularization parameter is crucial to balance the noise and resolution properties of the reconstructed image. There are two types of image quality assessment (IQA) metrics involved in the selection of the parameter, i.e., general IQA and task-specific IQA. The ultimate goal of CT reconstruction is to provide adequate pathological information for diagnosis, which can be measured by a task-specific IQA model. However, the pathological information is insufficient during the iterative reconstruction procedure, thus the task-specific discrimination metric is not capable to guide reconstruction optimization. In contrast, general IQA metric can be used for the guidance but it always results in inconsistent selection result with the task-specific IQA. By recognizing that the task-specific IQA indeed focuses on local image quality property, in this study we proposed to explore how to improve the local image quality while maintaining the global quality in the iterative reconstruction. Inspired by an advanced IQA metrics PAMSE, which introduces a Gaussian smoothed filter to MSE, we propose a new constraint to the regularization term in dictionary learning based iterative reconstruction. Preliminary experiment results both on simulation data and real data demonstrated that the proposed method can elevate the local image quality on CT images when using general IQA in optimization of regularization parameter and put the selection being closer to the result of task-specific IQA guidance.**

Keywords—computed tomography; PAMSE; MSE; general IQA; task-specific IQA

I. INTRODUCTION

Nowadays, X-ray computed tomography (CT) imaging has been widely used for diagnosis and intervention in hospitals and clinicals. As is well known to all, X-ray radiation can cause potential damage to human body which may induce genetic, cancerous and other diseases. Therefore, the ALARA principle (As Low as Reasonably Achievable) is put forward to avoid the excessive radiation, which dictates that restriction on radiation dose is extremely essential during any radiation procedure by all radiologists, whilst at the same time maintaining the diagnostic quality of medical images. As the radiation dose decreases, the medical image quality may degrade. Consequently, considering other conditions, image quality assessment (IQA) has got an enormous amount of attention from researchers.

This work was supported by the National Key Research and Development Program of China (No. 2016YFA0202003) and National Natural Science Foundation of China (NSFC) (No. 61571359).

JM Cai, JY Duan, YY Shi and XQ Mou are with the Institute of Image processing and Pattern recognition, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China.

For our primary interest in this paper, i.e., medical X-ray CT imaging, IQA is related with every aspect through the whole medical care. As mentioned in [2], medical care is composed of four layers, namely, physical layer, algorithm layer, diagnosis layer and retrieval layer, in which different parameters settings will influence on image quality, for instance, scan protocols, methods to reconstruct, the extent of pathological information and postoperative recovery for patients. As for the engineering application of X-ray CT imaging, we pay more attention on algorithm layer and diagnosis layer.

Iterative reconstruction has been validated to be efficient to reconstruct image with acceptable quality from noisy data in low-dose CT scanning. Image quality associated with iterative reconstruction will be tuned by the selection of involved regularization parameter, which is a crucial issue in this type of reconstruction algorithms. In this study, we focus on the IQA issues related in the regularization parameter selection in iterative CT reconstruction. More specifically, we investigate how an IQA task affects the selection of the regularization parameter and find a new way to tune the selection toward to producing high quality image for diagnosis.

During the algorithm and diagnosis layers when a CT image goes through medical care process, there are two types of IQA metrics involved, i.e., general IQA and task-specific IQA. Traditionally, the former mainly focuses on the general image quality conducted by various algorithms and parameters based on the obtained raw dataset, which can be evaluated by a general IQA metric, such as SSIM[3], GMSD[4], etc. The latter concerns the capability of recognizing a specific pathological signal from its local background in the image, which is usually assessed by a task-specific IQA metric. The most typical IQA model for this purpose is observer model. Up to now, these two classes of IQA models have been utilized to improve or optimize various algorithms in literature. Bai et al tried to select regularization parameter tuned by blind image quality assessment (BIQA) in iterative CT reconstruction[5]. Wunderlich and Noo tried to evaluate the effect of tube current modulation on lesion detectability in CT images[6]. In practice, the ultimate goal of CT reconstruction is to provide enough pathological information for diagnosis by clinicians[7], which hints that the task-specific shall be used to dominate the optimization. However, in the algorithm layer, the task-specific IQA model cannot be applied because the task-specific IQA is a problem of signal known exactly (SKE) and background known exactly/statistically (BKE/BKS), and obviously, the pathological signal and its background are unknown before the image is well reconstructed. In contrast, applying general IQA in optimizing image reconstruction cannot always produce the best image quality recognized by the task-specific IQA, which fact has been validated by our

previous work[2]. Hence, investigating a new iterative reconstruction algorithm that produces consistent image quality in terms of general IQA and task-specific IQA becomes an interesting research topic.

Observer model was built based on some low-level image feature decompositions. With the decompositions, image signals are represented by removing mutual dependencies as much as possible and hence the signal and the background can be distinguished efficiently. Meanwhile, general IQA models also sense the low-level image features to evaluate how much extent these features are similar to the reference ones. Both the two types of the IQA models all adopt the low-level features to perceive quality property of the image. This hints us to emphasize low-level features in image reconstruction process so that the two kinds of IQA would steer toward consistent outputs by the emphasized features.

To this end, in this study we investigate to emphasize low-level feature representation in a dictionary learning based iterative CT reconstruction algorithm to enhance the local image quality appearance while maintaining adequate global quality. More specifically, we amend the sparse representation term in the dictionary learning based algorithm by adding a low-level feature term. Inspired by our previous work [8], the proposed amended term has a simple form of valid distance metric and exhibits excellent IQA performance. With the amended term, we propose a new CT reconstruction algorithm, named FAIR. By experiments both on simulation data and real data, we found that the proposed FAIR can reconstruct images with more consistent global and local image qualities than traditional iterative algorithm.

II. METHODOLOGY

Primarily, the classical iterative reconstruction can be formulated as follows:

$$\hat{\mathrm{x}}_{\beta} = \arg\min_{x}\{\boldsymbol{\Phi}(\mathrm{x}) + \beta\boldsymbol{\Psi}(\mathrm{x})\} \tag{1}$$

where $\boldsymbol{\Phi}(\mathrm{x})$ and $\boldsymbol{\Psi}(\mathrm{x})$ represent the fidelity term and regularization term. And the scale β is the regularization parameter which adjusts the fidelity and penalty of reconstruction images. The inverse problem $\boldsymbol{\Phi}(\mathrm{x})$ can be demonstrated as $\|\mathbf{A}\,\mathrm{x}-\mathbf{b}\|_2^2$, where $\mathbf{A}\in\mathbf{R}^{M\times N}$ represents a bound and linear operator, and $\mathbf{b}\in\mathbf{R}^{M}$ is denoted as the measure data which is composed of noise $\varepsilon\in\mathbf{R}^{N}$ and the exact solution $\mathrm{x}\in\mathbf{R}^{N}$. $\boldsymbol{\Psi}(\mathrm{x})$ represents prior information of the reconstructed images, such as dictionary learning. In accordance with reference [1], the sparsity constraint as the regularization term can be written as

$$\boldsymbol{\Psi}(\mathrm{x}) = \sum_{s}\|\mathbf{E}_s\mathrm{x}-\mathbf{D}\boldsymbol{\alpha}_s\|_2^2 + \sum_{s}\upsilon_s\|\boldsymbol{\alpha}_s\|_0 \tag{2}$$

where $\mathbf{E}_s$ is an operator to extract patches from the image. $\mathbf{D}$ is denoted as the overcomplete dictionary, and $\boldsymbol{\alpha}_s$ is a sparse representation which has few nonzero entries with the l_0-norm. And then the goal of image reconstruction becomes to solve the following minimization problem:

$$\min_{\mathrm{x},\boldsymbol{\alpha}_s}\|\mathbf{A}\,\mathrm{x}-\mathbf{b}\|_2^2 + \beta\Big(\sum_{s}\|\mathbf{E}_s\mathrm{x}-\mathbf{D}\boldsymbol{\alpha}_s\|_2^2 + \sum_{s}\upsilon_s\|\boldsymbol{\alpha}_s\|_0\Big) \tag{3}$$

In this study, we use this reconstruction framework in exploring how to emphasize local image quality measure. Obviously, the regularization term in (3) is constrained by $\|\bullet\|_2$ being the l_2-norm, i.e., a measure of MSE. Despite of excellent mathematical properties, such as simplicity, convexity and differentiability, MSE is not consistent with human perceptual vision[3,4,8]. Therefore, the optimal solution of (3) can merely improve the general CT image quality. In order to improve visibility of pathological signal in the image, we add a structural MSE (SMSE) measure to emphasize local image quality by introducing a linear structure extractor $\mathbf{S}$ to construct an MSE-like l_2-norm metric. Combing the linear structure extractor and regularization term, we can get a new constraint norm for regularization term as follows:

$$\begin{aligned}\boldsymbol{\Psi}(\mathrm{x}) &= \sum_{s}\Big(\|\mathbf{E}_s\mathrm{x}-\mathbf{D}\boldsymbol{\alpha}_s\|_2^2 + \alpha\|\mathbf{S}(\mathbf{E}_s\mathrm{x}-\mathbf{D}\boldsymbol{\alpha}_s)\|_2^2\Big) + \sum_{s}\upsilon_s\|\boldsymbol{\alpha}_s\|_0 \\ &= \sum_{s}(\mathbf{E}_s\mathrm{x}-\mathbf{D}\boldsymbol{\alpha}_s)^T\left(\mathbf{I}+\alpha\mathbf{S}^T\mathbf{S}\right)(\mathbf{E}_s\mathrm{x}-\mathbf{D}\boldsymbol{\alpha}_s) + \sum_{s}\upsilon_s\|\boldsymbol{\alpha}_s\|_0\end{aligned} \tag{4}$$

where α is a constant. Let $\mathbf{M}=\mathbf{I}+\alpha\mathbf{S}^T\mathbf{S}$. As analyzed in [8], for the validity of (4), $\mathbf{M}$ should be positive semi-definite (PSD). Then $\mathbf{M}$ can be decomposed into $\mathbf{M}=\mathbf{P}^T\mathbf{P}$. And (4) can be re-written as the following formula:

$$\boldsymbol{\Psi}(\mathrm{x}) = \sum_{s}\|\mathbf{P}(\mathbf{E}_s\mathrm{x}-\mathbf{D}\boldsymbol{\alpha}_s)\|_2^2 + \sum_{s}\upsilon_s\|\boldsymbol{\alpha}_s\|_0 \tag{5}$$

Generally, $\mathbf{P}$ can be considered as an image smoothing operator. Moreover, it be inferred to be a Gaussian smoothed filter by using two structure extractors, namely the gradient operator and the Laplacian operator [8]. As a result, the minimization problem for (3) can be re-written as follows:

$$\min_{\mathrm{x},\boldsymbol{\alpha}_s}\|\mathbf{A}\,\mathrm{x}-\mathbf{b}\|_2^2 + \beta\Big(\sum_{s}\|h\otimes(\mathbf{E}_s\mathrm{x}-\mathbf{D}\boldsymbol{\alpha}_s)\|_2^2 + \sum_{s}\upsilon_s\|\boldsymbol{\alpha}_s\|_0\Big) \tag{6}$$

where h is Gaussian smoothed filter in default setting as reference [8]. For more detailed analysis, readers are suggested to consult reference [8]. Equation (6) shows a new CT reconstruction algorithm that is called as feature-aware iterative reconstruction, abbreviated as FAIR. In the following sections of this paper, we will experimentally examine the FAIR in terms of general IQA and task-specific IQA and check if the inconsistency between the two IQA evaluations can be improved.

III. MATERIALS AND EXPERIMENTS

A. Materials

In this paper, we conduct two experiments to test the proposed FAIR. The associated experimental data sets are from simulation data and real CT scanning respectively.

The simulation noisy data were generated from a simulation phantom of abdomen added by 5000 Poisson photons, which was produced by the DXCAT2 program, as

showed in Fig. 1(a). The simulation lesion was a uniform disk signal as indicated by the red rectangle, whose radius was 6 pixels and contrast was 0.0719. The real data were obtained from the quarter dose of patient L067 of the training database provided by Low Dose Grand Challenge from Mayo Clinic, which was showed in Fig. 1(b). The sinogram views for all data were 672×1160 . We reconstructed noisy images by the proposed FAIR and traditional dictionary learning based iterative reconstruction (DL_IR). The regularization parameters were ranging form $1e(-5)\times1.2$ to $1e(-5)\times1.2^{50}$ and $1e(-4)\times1.2^{-5}\times1.2$ to $1e(-4)\times1.2^{-5}\times1.2^{45}$ respectively. The dictionary was trained with 256 atoms whose size was 8. The coding loss was 3e-5.

(a) (b) (c)

Fig. 1. Experimental data. (a) Simulation phantom circled with simulation lesion. The display window is [0.007 0.02]. (b) Real data from Mayo Clinic. The display wimdow is [0.0168 0.0235]. (c) The dictionry.

B. Description of evaluation metrics

In order to evaluate the local and global image quality on CT reconstruction images, two kinds of IQA metrics were used. Considering the availability of ideal reference image, the general IQA metrics were different for simulation data and real data. In the case of phantom simulation experiment, a full reference IQA metric, GMSD[4], was used because the ground truth is available. While in real data experiment, a blind IQA (BIQA) metric [10] was used since corresponding clean reconstruction image without any noise contaminated is unavailable. A brief introduction to the used general IQA metrics will be described in late paragraph. The task-specific IQA metric was a channelized Hotelling observer model that was used for all experimental data.

We briefly introduced the general IQA metrics, GMSD and BIQA. GMSD is short for gradient magnitude similarity deviation. It is a pooling strategy after obtaining the pixel-wise gradient magnitude similarity (GMS) between reference image and distorted image [4]. As for simulation data, the abdomen phantom with simulation lesion was denoted as the reference image and the reconstructed images were denoted as the distorted images. The parameters were in default settings as reference [4]. And the value of GMSD score approaching 0 meant better image quality. The BIQA model utilized the joint statistics of gradient magnitude and the Laplacian of Gaussian operator to extract features and trained regression model learning with subjective human scores as stated in reference [10]. Due to lack of effective CT image databases, we used natural images from LIVE database [11] to train BIQA model in default settings as reference [10]. Differential mean opinion score (DMOS) was regarded as subjective scores, ranging from 0 to 100. Thus the BIQA score approaching 0 meant better image quality.

As for task-specific IQA metric, the channelized Hotelling observer model was conducted to measure the performance of a specific signal from its background in visual perception [9]. It can be regarded as a two-classification problem. The discriminant function for this binary task was a scalar test statistic $\lambda(\mathrm{x})$ compared to a threshold value λ_t for signal detection. In this paper, we set channels as Gabor channels as described in reference [6]. The matched filter and pre-whitening filter were trained from training database of Mayo clinic. According to the radius of lesion, the size of image patches was 32 and 48 for simulation data and real data, respectively. After training, we set 0 as the threshold value for all experiments. If the $\lambda(\mathrm{x})$ larger than 0, the image patch was absent with lesion. Otherwise, that was present with lesion.

IV. RESULTS

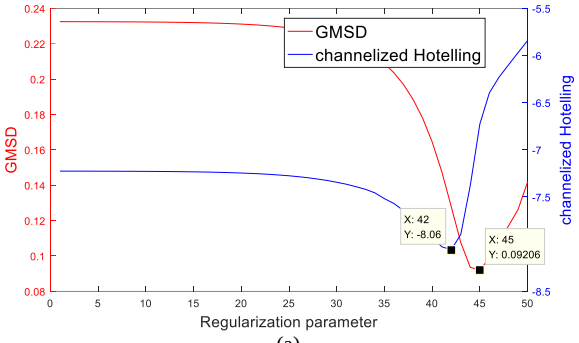
In this section, experimental results are presented to validate our proposed method. Firstly we demonstrate the results for simulation data in Fig. 2 and Fig. 3, which were reconstructed by FAIR and traditional DL_IR respectively. The curves in both Fig. 2(a) and Fig. 3(a) show the regularization parameters β versus GMSD scores (red line) and the scores of channelized Hotelling observer model (blue line) separately. In both Fig. 2 and Fig. 3, the images (b) and (c) are the optimal ones selected by general IQA and task-specific IQA respectively. The lesion is pointed by a red arrow. By comparing Fig. 2(b) with Fig. 3(b), we can find that the optimal image reconstructed is smoother than that reconstructed by DL_IR. As for Fig. 2(c) and Fig. 3(c), Fig. 3(c) contain more noise and block structure than Fig. 2(c) as well. As observed from the curves in Fig. 2(a) and Fig. 3(a), the selections of two optimal images reconstructed by FAIR are closer than those reconstructed by DL_IR.

As for real data, the results are demonstrated in Fig. 4 and Fig. 5, which were reconstructed by FAIR and DL_IR respectively. The curves in both Fig. 4(a) and Fig. 5(a) are IQA scores versus the regularization parameters β . And in both Fig. 4 and Fig. 5, the images (b) and (c) are the optimal images selected by general IQA and task-specific IQA respectively. The lesion is pointed by a red arrow. The optimal image in Fig. 4(b) is smoother than that in Fig. 5(b), same as Fig. 4(c) and Fig. 5(c). In Fig. 4 and Fig. 5, similar as the results for simulation data, the two optimal images reconstructed by FAIR have more consistency than those reconstructed by DL_IR.

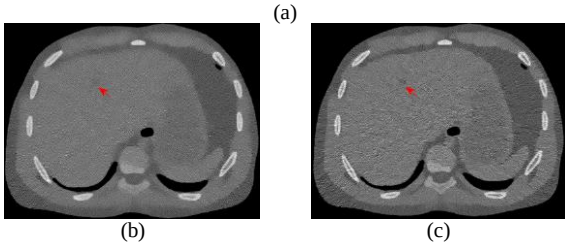
V. DISCUSSIONS AND CONCLUSIONS

In this paper, we propose a new CT reconstruction algorithm, named FAIR, by introducing a Gaussian smoothed filter to regularization term. According to experimental results, the two optimal images reconstructed by FAIR and selected by the two kinds of IQA metrics exhibit more consistent than that reconstructed by traditional iterative reconstruction on all experimental data. As a conclusion, the proposed FAIR can improve the local image quality while maintaining general image quality to some extent. However, the experimental dataset was too small for validation. It is

necessary to test on more real clinical data. And the reconstructed images should be evaluated by professional clinicals for subjective scores. This research is still a fairly challenging mission on CT field.

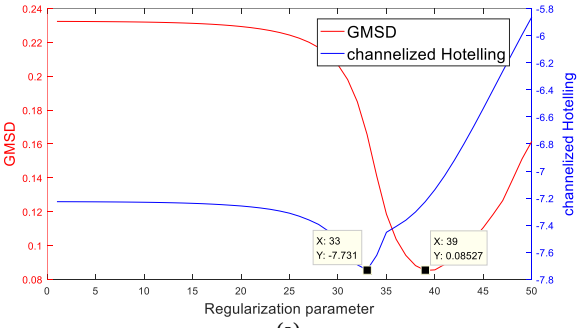


(a)

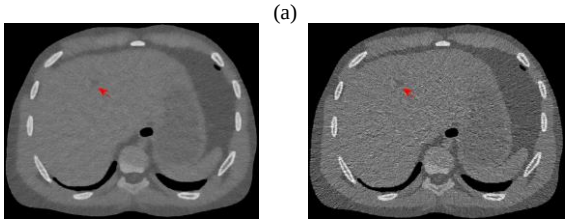


(b) (c)

Fig. 2. Evaluation on simulation data reconstructed by proposed FAIR. (a) Curves of regularization parameters β vs GMSD scores (red) and the scores of channelized Hotelling observer model (blue) respectively. (b) The best reconstruction image on general IQA. (c) The best reconstruction image on task-specific IQA. The display window is [0.007 0.02].

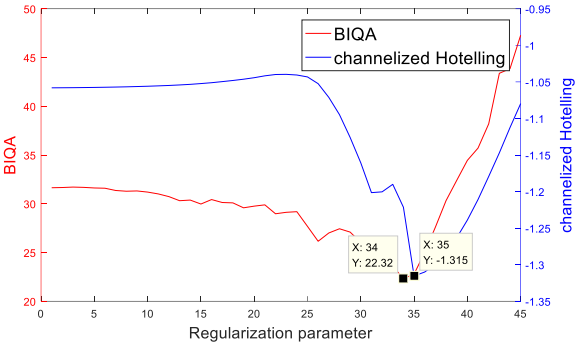


(a)

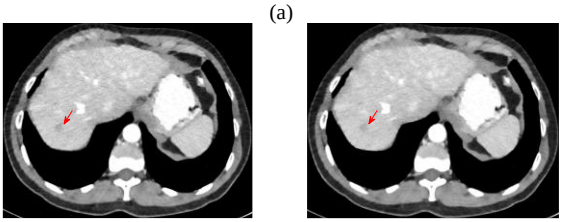


(b) (c)

Fig. 3. Evaluation on simulation data reconstructed by DL_IR. (a) Curves of regularization parameters β vs GMSD scores (red) and the scores of channelized Hotelling observer model (blue) resperctively. (b) The best reconstruction image on general IQA. (c) The best reconstruction image on task-specific IQA. The display window is [0.007 0.02].

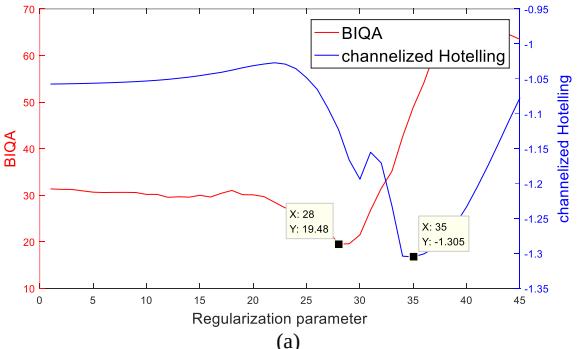


(a)

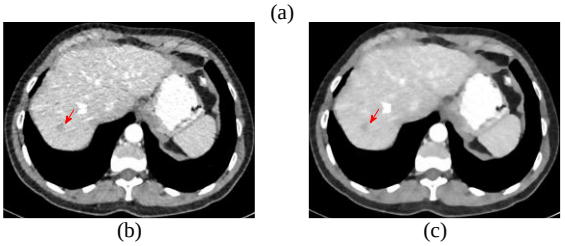


(b) (c)

Fig. 4. Evaluation on real data reconstructed by proposed FAIR. (a) Curves of regularization parameters β vs BIQA scores (red) and the scores of channelized Hotelling observer model (blue) respectively. (b) The best reconstruction image on general IQA. (c) The best reconstruction image on task-specific IQA. The display wimdow is [0.0168 0.0235].



(a)



(b) (c)

Fig. 5. Evaluation on real data reconstructed by DL_IR. (a) Curves of regularization parameters β vs BIQA scores (red) and he scores of channelized Hotelling observer model (blue). (b) The best reconstruction image on general IQA. (c) The best reconstruction image on task-specific IQA. The display wimdow is [0.0168 0.0235].

ACKNOWLEDGMENT

The real data of the experiment was based on the train database of Low Dose CT Grand Challenge organized by Mayo Clinics, so the authors would like to acknowledge Dr. Cynthia McCollough, the Mayo Clinic, the American Association of Physicists in Medicine, and grants EB017095 and EB017185 from the national Institute of Biomedical Imaging Bioengineering.

REFERENCES

[1] Q. Xu, H. Yu, X. Mou, L. Zhang, J. Hsieh and G. Wang, "Low-dose x-ray ct reconstruction via dictionary learning," Medical Imaging, IEEE Transactions on, vol. 31, no. 9, pp. 1682-1697, 2012.

[2] J. Cai, X. Chen, W. Huang and X. Mou, "Image Quality Assessment on CT Reconstruction Images: Task-specific vs. General Quality Assessment," 14th Annual Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine, pp. 814-819, 2017.

[3] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," Image Processing, IEEE Transactions on, vol. 13, no. 4, pp. 600-612, 2004.

[4] W. Xue, L. Zhang, X. Mou and A. C. Bovik, "Gradient magnitude similarity deviation: a highly efficient perceptual image quality index," Image Processing, IEEE Transactions on, vol. 23, no. 2, pp. 684-95, 2013.

[5] T. Bai, X. Mou, W. Xue, H. Yan and S. B. Jiang, "Iterative CT Reconstruction with Regularization Parameter Tuned by Blind Image Quality Assessment," 13th Annual Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine, pp. 124-127, 2015.

[6] A. Wunderlich and F. Noo, "Evaluation of the impact of tube current modulation on lesion detectability using model observers," 30th Annual International IEEE EMBS Conference, vol.2008, pp.2705-2708, 2008.

[7] C. J. Martin, P. F. Sharp and D. G. Sutton, "Measurement of image quality in diagnostic radiology," Applied Radiation & Isotopes Including Data Instrumentation & Methods for Use in Agriculture Industry & Medicine, vol. 50, no. 1, pp. 21-38, 1999.

[8] W. Xue, X. Mou, L. Zhang and X. Feng, "Perceptual Fidelity Aware Mean Squared Error," IEEE International Conference on Computer Vision, vol. 2013, pp. 705-712, 2013.

[9] H. H. Barrett, "Predicting human performance by a channelized Hotelling observer model," Proceedings of SPIE - The International Society for Optical Engineering, vol. 1768, pp. 161-168, 1992.

[10] W. Xue, X. Mou, L. Zhang, A. C. Bovik and X. Feng, "Blind image quality assessment using joint statistics of gradient magnitude and laplacian features," Image Processing, IEEE Transactions on, vol. 23, no. 11, pp. 4850-4862, 2014.

[11] H. R. Sheikh, M. F. Sabir, A. C. Bovik, "A Statistical Evaluation of Recent Full Reference Image Quality Assessment Algorithms," Image Processing, IEEE Transactions on, vol. 15, no. 11, pp. 3440-51, 2006.

Analysis on Optimal Selection of Regularization Parameters Based on CT Image Statistics

Jiayu Duan, Jianmei Cai and Xuanqin Mou

***Abstract*—Regularization parameter selection is pivotal in optimizing reconstructed images which controls a balance between fidelity and penalty term. Images reconstructed with the optimal regularization parameter will keep the detail preserved and the noise restrained at the same time. Our previous works have used the second order derivates of image variance (Soda-curve) to select the optimal regularization parameter which exhibits well in multiple regularizers. However, we have not yet given a justification for this hypothesis. In this work, we proposed a generalized model between variance and regularization parameters based on medical CT image statistics to explain the effectiveness of Soda-curve method. The reconstructed images with different β were decomposed into structure and noise parts with the dual dictionary learning method. And variances of the decomposed results vs β were modeled by polynomial functions. Calculating the Soda-curve with modeled functions, we found that the presence of maximum in Soda-curve doesn't rely on the parameters range in functions nor employed sparse operator in different cases which provides a solid basis to support the efficiency of Soda-curve method.**

Keywords—regularization parameter selection, image statistics, dual dictionaries, Soda-curve

I. INTRODUCTION

Inverse problem exists in many imaging computation cases, especially in computed tomography reconstruction. Computed tomography needs to transform projections gathered by detectors into the attenuation coefficient of the reconstructed area. Besides, in accordance with the ALARA (as low as reasonably achievable) principle, low dose CT scan is applied to avoid excessive radiation by adjusting the operating current, the operating potential and exposure time of an X-ray tube, which leads to the noisy projections. In general, this kind of problem is ill-posed. When settling this kind of problem, Bayesian approaches, or maximum a posteriori (MAP) methods, are effective in providing solutions [1]. Usually, by constructing the Lagrangian function as Eq. (1) to minimum this MAP problem, an optimal solution can be found.

$$\hat{x}_\beta = \arg\min_x \{\phi(x) + \beta\varphi(x)\} \quad (1)$$

In the Lagrangian function, the non-negative functional φ and ϕ are known as the fidelity term and regularization term. The selection of regularization term is always based on some kind of prior, such as total variance (TV) [2] and sparse representation based on an over-complete dictionary [3]. The regularization parameter β is used to balance the strength of the fidelity and penalty terms to make good tradeoff between resolution and noise suppression. In practice, a larger value of β will result in over smooth reconstructed image, while a smaller one will bring about noisy result. Hence the selection of the parameter β plays a crucial role in producing a high-quality reconstruction. Currently, many fashions have been proposed to seek optimal parameters in regularization term. As for the discrepancy principle, it requires an estimate of the noise level, which is not always available in practice and the existence of a solution is not guaranteed for some non-smooth functional [4]. And the GCV (Generalized Cross Validation) method estimates the mean square error (MSE), but the minimization of the objective functions is nontrivial because of their flat property over a broad scale. Moreover, the L-curve method is totally based on data, which is sensitive to curvature estimation and the correlation with β is implicit [5]. We have investigated this issue by introducing CT image statistics to estimate the correlation function between a proposed sparse feature and the parameter β and hence to find the resolution based the correlation. In [5], we analyzed the image by applying a sparse operator. As a hypothesis, the analyzed coefficient consists of signal and noise components, each of which subjects to a generalized Gaussian distribution. And we discovered that the second order derivation of the variance curve of the coefficient with regarding to β, in this study we call it as Soda-curve, can be used to locate the optimal selection of β by maximizing the Soda-curve. Meanwhile, we experimentally validated this method by using LOG (Laplacian of Gaussian) as the sparse operator. Then in [11], we used the pairwise products of neighboring LOG signals as the new sparse feature to obtain more accuracy selection. It was confirmed that the new feature is more efficient to remove structural redundancy than the LOG signal. In [6], we constructed a 'Z-curve' that is originated from L-curve by applying scale transformations to analyze the relationship between dictionary learning based sparse representation coefficients and β. Although the best parameter selection strategy is different from that in [5, 6], they also follow similar principle. In any case, above works based on the same hypothesis, i.e., the optimal β is associated with the maximum point of the Soda-curve. However, we have not yet given a justification for this hypothesis. In this work, we will conduct an analysis on CT image statistics and propose a general model to explain how the Soda-curve method works based on the generalized model.

The rest of the paper is organized as follows: in Section 2, we introduce the proposed methods. Section 3 gives the experiment setup and results, and analysis about the results. The conclusion and discussion can be found in Section 4.

JY Duan, JM Cai and XQ Mou are with the Institute of Image processing and Pattern recognition, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China.

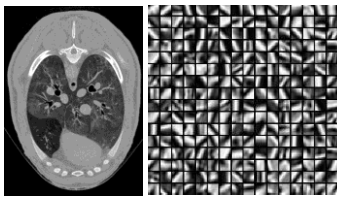
This work was supported by National Natural Science Foundation of China (NSFC) (No. 61571359). And the National Key Research and Development Program of China (No. 2016YFA0202003).

II. METHODS

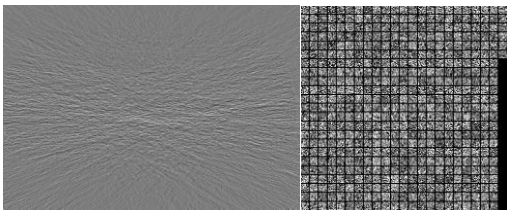
Let $x(\beta) \in \mathbb{R}^{N\times N}$ denote a serial of reconstructed images with a number of parameter β. The values of β are selected to ensure that the reconstructed images vary from noisy ones to over-smooth ones which guarantees the optimal image is exist in this range. $\{x^i\}_{i=1}^m$ denotes the patches extracted from the image x. For every patch, there consists of two parts: structure and noise although the structure can be over-smooth when the β is too big. The patches can be generated by randomly sampling in the image [12] or extracting blocks with (non-) overlap interval [13]. Nowadays, with the emergence of dual dictionaries learning technics (DDL), it offers a chance to solve many matters such as super-resolution image reconstruction via DDL [8], removal the limited angle artifacts [9,10] etc. We apply the DDL method to separate structure and noise from the reconstructed images with the structure dictionary and noise dictionary. Therefore, the $x^i(\beta)$ can be represented as follow:

$$x^i(\beta)=A\alpha^i(\beta)+B\gamma^i(\beta) \tag{2}$$

where A and B are the structure dictionary and noise dictionary separately and β is the regularization parameter.



(a) structure dictionary with training image



(b) noise dictionary with training image

Fig. 1. structure and noise dictionary with training images

In addition, we consider the sparse coefficients distribution of samples patches extracted from the image as the image sparse coefficients distribution. In sparse coding, the OMP (Orthogonal Matching Pursuit) algorithm is employed to chase for the optimal representation of the input signal. Note that the OMP algorithm uses MSE as a principle to minimum the coding residue. Theoretically, there is the rule that OMP only selects the same atom once which means the structure atoms in the structure dictionary cannot be selected when the coding process begins to represent the noise part of the image [14]. Nevertheless, the OMP seeks the optimal combination in the atoms span which means the combination of this atoms can be regarded as a "new" atom. Besides, we experimentally find that all the sparse coefficients produced by different sparse operators from structure and noise dictionary are independent, which accords with the spirit of sparse representation.

As a consequence, we mimic the "structure atom" in noise dictionary depicted as Eq. (3)

$$A=CB \tag{3}$$

Thus the $x^i(\beta)$ can be described as follow:

$$x^i(\beta)=A\alpha^i(\beta)+CB\eta^i(\beta) \tag{4}$$

As can be seen, the patches can be represented by the same sparse operator or same atom after Eq. (4). And in [5], we model noise and structure sparse coefficients as two independent generalized Gaussian distribution. Hence, the image distribution can be formularized as Eq. (5).

$$f^i(\beta)=\{\alpha^i(\beta),\eta^i(\beta)\} \tag{5}$$

Note that the sparse coefficients distributions in structure and noise parts are statistical independent, which means the variance of the image sparse coefficients can be seen as Eq. (6)

$$\mathrm{var}(f(\beta))=\mathrm{var}(\alpha(\beta))+\mathrm{var}(\eta(\beta)) \tag{6}$$

By our experimental observations based on a large number of resultant images resolved by (1), no matter for CT images or other natural image processing cases, and for any one of the structure atoms, var(α(β)) and var(η(β)) can be generally fitted by a fourth-order polynomial. Fig 2 shows a sample. Considering the noise level will be nearly to zero when the regularization parameter is enough large, we introduce a step function to limit the domain of the noise component. Then the variances can be formularized as Eq. (7).

$$\begin{aligned}&\mathrm{var}(\alpha(\beta))=p_1^\alpha\beta^4+p_2^\alpha\beta^3+p_3^\alpha\beta^2+p_4^\alpha\beta+p_5^\alpha\\&\mathrm{var}(\eta(\beta))=(p_1^\eta\beta^4+p_2^\eta\beta^3+p_3^\eta\beta^2+p_4^\eta\beta+p_5^\alpha)u(\beta_0-\beta)\end{aligned} \tag{7}$$

where $p_i^\alpha, i=1,2,3,4,5$, $p_j^\eta, j=1,2,3,4,5$ are the polynomial coefficients of structure and noise, respectively and $u(\beta_0-\beta)$ is the step function.

The solution of Soda-curve can be generated with the general fitting polynomial as Eq. (8)

$$\begin{aligned}\frac{\partial^2\mathrm{var}(f(\beta))}{\partial\beta^2}&=\frac{\partial^2\mathrm{var}(\alpha(\beta))}{\partial\beta^2}+\frac{\partial^2\mathrm{var}(\eta(\beta))}{\partial\beta^2}\\&=(p_1^\alpha\beta^4+p_2^\alpha\beta^3+p_3^\alpha\beta^2+p_4^\alpha\beta+p_5^\alpha)^{''}\\&\quad+[(p_1^\eta\beta^4+p_2^\eta\beta^3+p_3^\eta\beta^2+p_4^\eta\beta+p_5^\alpha)u(\beta_0-\beta)]^{''}\\&=\begin{cases}12(p_1^\alpha+p_1^\eta)\beta^2+6(p_2^\alpha+p_2^\eta)\beta+2(p_3^\alpha+p_3^\eta) & (\beta_0>\beta)\\ 12p_1^\alpha\beta^2+6p_2^\alpha\beta+2p_3^\alpha & (\beta_0<\beta)\end{cases}\end{aligned} \tag{8}$$

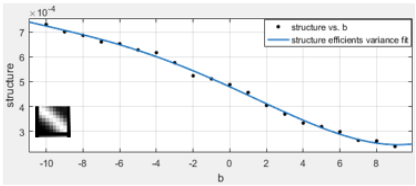
According to [5], we analyzed that the structure variance will gradually increase with decay of regularization parameter. Because of the non-sparsity of noise, the variance of the noise coefficients is always bigger than variance of the structure coefficients. With the regularization parameter gradually increasing, the reconstructed image will become more and more clear and we assumed that there exists a point β_0 where the reconstructed image preserves detail structures while the noise does not or just appear. We assume that the optimal regularization parameter shows up around

β_0. In order to simulate this phenomenon, we add a step function in Eq. (8). Thus, the optimal point always locates around maximum point at Eq. (8). Eventually, the generalized model is proposed based on medical image statistics.

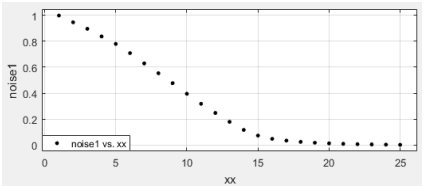
In this work, we will use generalized model to explain how the Soda-curve method works based on the generalized model. In addition, visual and quantitative evaluation are validated to verity the ability of this method.

III. Experiments and Analysis

In this section; we designed two type of validation experiments to demonstrate our hypothesis. First, we experimentally validated that the maximum of Eq. (8) just locates around the point β_0. In the fitting process, we had value ranges about these variables. And we observed that the maximum point always near the optimal regularization parameters no matter what value we choose in the range. Here for present preference, we randomly selected several values in the range value and showed in Fig.3. As can be seen, maximum in all curves shows up around β_0 which indicates the optimal point always shows up around β_0. The relation between β_0 and optimal point will be one of our further research directions.



(a)structure



(b)noise

Fig. 2. variance tendency in structure and noise with regularization parameter increasing with a specified atom

Second, we demonstrated that the optimal point doesn't rely on the sparse operator. The data is the real data from mayo clinic which was utilized to numerically generate a 672×1160 sinogram. All were added by 1E4 incident photons to synthesize low dose noisy data. And we used two different regularization methods: total variation (TV) and dictionary learning (DL) in the reconstruction algorithm. In TV regularization algorithm, we reconstructed 25 images with regularization parameter varying from $(1.2)^1$ to $(1.2)^{25}$. And in DL, the regularization parameter was set from $0.01\times(1.2)^{-20}$ to $0.01\times(1.2)^4$ generating 25 images and the coding length was set 8 and coding loss was set 5e-3.

In dual dictionaries learning process, we trained structure dictionary in clean sheep reconstruction with 256 atoms and noise dictionary is trained by the residue images with 512 atoms because the noise image is denser than clean image the dictionaries and training image are showed in Fig. 1. The whole workflow about the experiment can be described as Fig. 4.

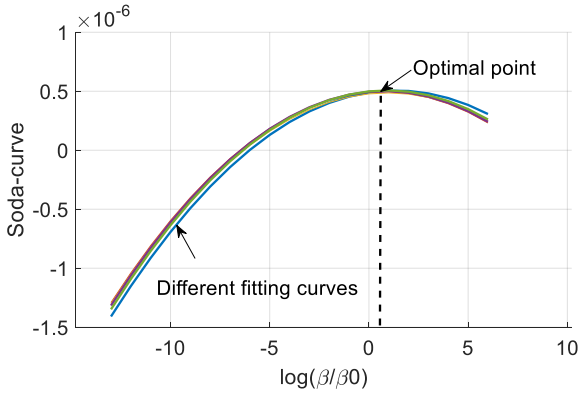


Fig. 3. Maximum point at vertical axis

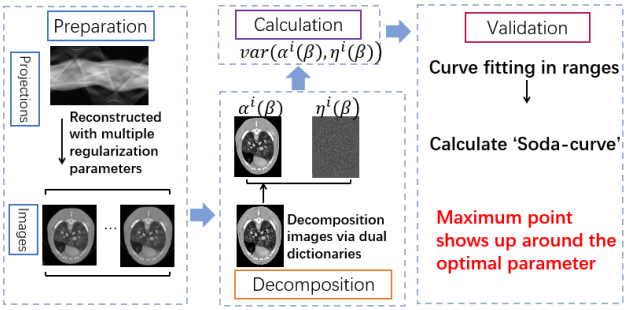
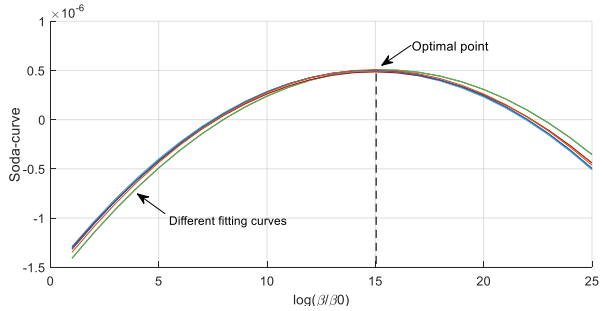
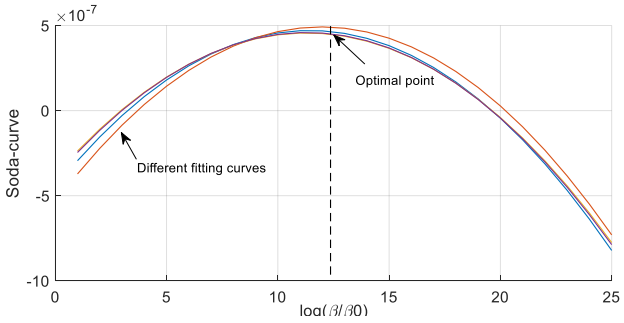


Fig. 4. the work flow of the validation



(a) different sparse operators fitting curves in TV as regularizer

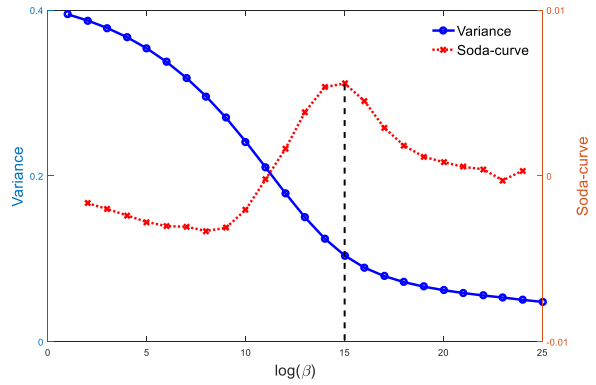


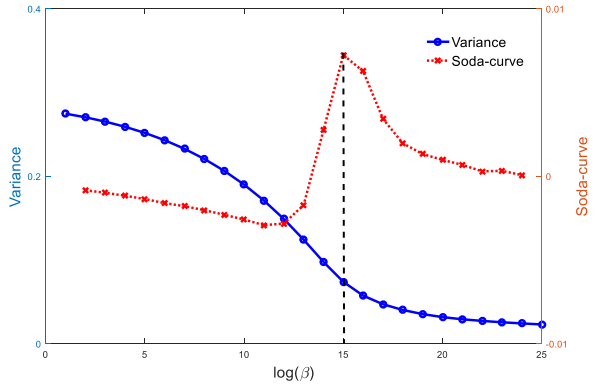
(b) different sparse operators fitting curves in DL as regularizer

Fig. 5. Different fitting curves when using different sparse operators

(a1) (b1) (c1)

(a2) (b2) (c2)





(d)

Fig. 6. First and second row are the human abdomen reconstructions in TV and DL regularization algorithms while reconstructions from the 1st to the 3rd column correspond to different β. Curves of image variance vs regularization parameter β are illustrated in d. And the reconstructed window is set as [0, 0.035], b(1) and b(2) are optimal reconstructed images.

Fig 5 showed the results about different sparse operators There is no obvious difference in Soda-curve when choose TV or DL regularizers. The reconstruction ranges from very noisy to oversmoothed which guaranteed excellent trade-off versions existence. Results are depicted in Fig.5. it can be seen that all the fitted curves with different atoms had almost same maximum value which means the selection of sparse operator doesn't influence the result of the optimal point of Soda-curve.

Table 1
MSE AND SSIM VALUES OF SELECTED OPTIMAL RECONSTRUCTED IMAGE IN TV AND DL

	TV		DL	
MSE	L-curve	5.696e-7	L-curve	8.7126e-7
	Soda-curve	**5.5559e-7**	**Soda-curve**	**6.5804e-7**
SSIM	L-curve	0.8635	L-curve	0.7912
	Soda-curve	**0.8696**	**Soda-curve**	**0.8273**

It is arbitrary to claim Soda-curve is superior than other selection method without comparison. Under this circumstance, we designed a comparison experiment between Soda-curve and L-curve. In Soda-curve selection for convenience, we used LoG as sparse operator and the data is also from mayo clinic same as second demonstration experiment. The result is showed in Fig. 6.

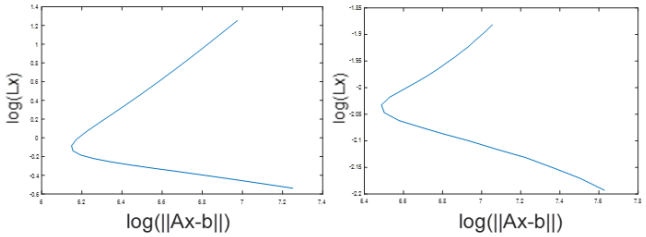
L-curve illustrates the relationship between fidelity and penalty terms. In L-curve, more horizontal part corresponds to the solutions where the regularization parameter is too large and vertical parts corresponds to the solution where regularization parameter is too small. The optimal point locates in the maximum of the curvature. The L-curve of different reconstruction algorithm is depicted in Fig. 7. And the optimal reconstruction results selected by L-curve and Soda-curve is showed in Fig. 8. From the perceptual visual evaluation, Soda-curve shows potential to select the optimal reconstruction in better edges and low noise than L-curve. Apart from perceptual visual evaluation, we conducted quantitative evaluation by calculating MSE (mean square error) and SSIM (structural similarity index) values and compared our method with L-curve method. And the quantitative evaluation results are presented in Table. 1.

From the quantitative results, the Soda-curve exhibits well in both preserving structure (higher SSIM value) and suppressing noise (lower MSE value) than L-curve which shows the ability to select the optimal regularization parameters.

IV. Discussion and Conclusion

In this study, in order to explain the mechanism of Soda-curve. We investigated the statistical changes vs the regularization parameters and found out that all the tendencies are similar with different sparse operators. Based on experimental analysis, we modeled them as polynomial functions as the generalized model. it is valuable to mention that the generalized model we brought up in this paper demonstrates that the optimal regularization parameter point shows up without changes in sparse operators and range of value in the regularization parameters which explains the efficiency of the Soda-curve. What's more, visual and quantitative results also illustrate the potential of the Soda-curve.

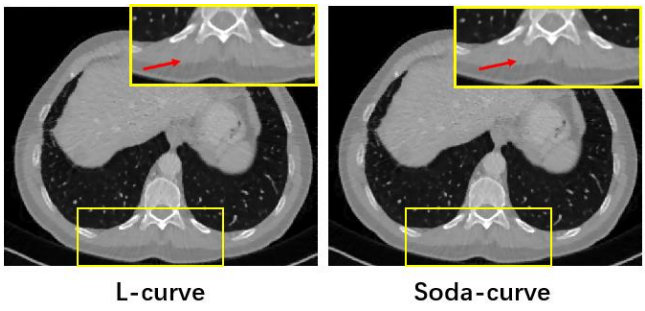
One of the barriers of the present work is that calculating the model with dual dictionaries can be time consuming which means the demand of creating a simplified manner is keen. In practice, adjacent LOG multiplication can be regarded as the line atom in structure, employing this king of operator can save a lot of time [11]. In other word, this study will provide a solid basis to support that the maximum point of the 'Soda-curve' of the adjacent LOG signals can be located as the optimal selection regularization parameter.



(a) TV as regularizer (b) DL as regularizer

Fig. 7. is the L-curve about TV and DL reconstruction method with different regularization parameters. The Y axis is the logarithm of the penalty term and X axis is the logarithm of the fidelity term. The optimal regularization parameter is located by maximizing the curvature of L-curve.

(a)L-curve and Soda-curve selection result comparison in TV



(b)L-curve and Soda-curve selection result comparison in DL

Fig. 8. is depicted the result about the optimal reconstruction results selected by L-curve and Soda-curve method. The top right corner is the magnified ROIs marked in the yellow rectangles which shows that the Soda-curve exhibits well in noise suppression.

One of the barriers of the present work is that calculating the model with dual dictionaries can be time consuming which means the demand of creating a simplified manner is keen. In practice, adjacent LOG multiplication can be regarded as the line atom in structure, employing this king of operator can save a lot of time [11]. In other word, this study will provide a solid basis to support that the maximum point of the 'Soda-curve' of the adjacent LOG signals can be located as the optimal selection regularization parameter.

As can be seen, in this paper, we first used the dual dictionary to separate noise and structure components in series of regularization parameters. And based on former statistical hypothesis, we explained how the selection works in 'Soda-curve' which had nothing to do with the change in regularization parameter value and the selection in sparse operators.

ACKNOWLEDGMENT

The real data is adopted from the train database of Low Dose CT Grand Challenge organized by Mayo Clinics. The authors would like to acknowledge Dr. Cynthia McCollough, the Mayo Clinic, the American Association of Physicists in Medicine, and grants EB017095 and EB017185 from the national Institute of Biomedical Imaging Bioengineering.

REFERENCES

[1] Chen Y, Ma J, Feng Q, et al. Nonlocal Prior Bayesian Tomographic Reconstruction[J]. Journal of Mathematical Imaging & Vision, 2008, 30(2):133-146.

[2] Sidky E, Pan X. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization[J]. Physics in Medicine & Biology, 2008, 53(17):4777.

[3] Xu Q, Yu H, Mou X, et al. Low-dose X-ray CT reconstruction via dictionary learning.[J]. IEEE Transactions on Medical Imaging, 2012, 31(9):1682.

[4] Ito, Kazufumi, B. Jin, and T. Takeuchi. "A Regularization Parameter for Nonsmooth Tikhonov Regularization." Siam Journal on Scientific Computing 33.3(2011):1415-1438.

[5] Xuanqin Mou, Ti Bai, Xi Chen, Hengyong Yu, Qingsong Yang and Ge Wang. "Optimal Selection for Regularization Parameter in Iterative CT Reconstruction Based on the Property of Natural Image Statistics."Fully 3D,2015

[6] Ti Bai, Hao Yan, Xun Jia, Steve Jiang, Ge Wang, and Xuanqin Mou, Z-Index Parameterization for Volumetric CT Image Reconstruction via 3-D Dictionary Learning, IEEE Transactions on Medical Imaging, 36(12): 2466-2478, 2017.

[7] D. L. Donoho and M. Elad, "Maximal sparsity representation via l1 minimization," Proc. Nat. Aca. Sci., vol. 100, pp. 2197–2202, 2003.

[8] H.Zhang, Y.Zhang, and T.S.Huang, "Efficient sparse representation based image super resolution via dual dictionary learning,"in Proceedings of the 12th IEEE International Conference on Multimedia and Expo(ICME'11),July2011.

[9] M.Cao and Y.X.Xing, "Limited angle reconstruction with two dictionaries," in IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC '13), pp. 1–4, Seoul, Republic of Korea,October2013.

[10] Z.Chen,X. Jin,L.Li,and G.Wang, "A limited angle CT reconstruction method based on an isotropic TV minimization," Physics in Medicine and Biology,vol.58,no.7,pp.2119–2141,2013

[11] J. Duan, X. Chen, B. Shen and X. Mou, "An Optimal Metric for Regularization Parameter Selection in Iterative Reconstruction for CT Image," Proc. Fully3D, pp. 820-824, 2017.

[12] Lewicki M S, Olshausen B A. Probabilistic framework for the adaptation and comparison of image codes[J]. Journal of the Optical Society of America A, 1999, 16(7):1587-1601.

[13] Aharon M, Elad M, Bruckstein A. -SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation[M]. IEEE Press, 2006.

[14] Tropp, Joel A. Greed is good: algorithmic results for sparse approximation[J]. IEEE Transactions on Information Theory, 2004, 50(10):2231-2242.

Model based image reconstruction with concomitant scale estimation

Jingyan Xu and Frédéric Noo

***Abstract*—Model base image reconstruction (MBIR) often includes an edge-preserving regularizer to stabilize the ill-posed inverse problem. Edge preservation can be achieved by adjusting the penalty weights of the regularizer: the weights should be small if they are near edges, the weights can be large if they are in a homogeneous region. In this work, the adjustment of the penalty weights is accomplished by concomitant scale estimation, a reconstruction framework that simultaneously estimate the image intensity and the scale parameters that are inversely proportional to the penalty weights. If the original problem is convex and the regularizer is a convex function, then the problem of image reconstruction with concomitant scale estimation is jointly convex (and nonsmooth) in image intensities and the scale parameters. Such problems can be solved by a number of primal-dual splitting algorithms from convex programming; we use the alternating direction method of multipliers (ADMM) in this work. We discuss the applicability and versatility of the general framework using CT reconstruction examples.**

I. Introduction

The objective function for MBIR is often written as

$$\Phi(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{y} - A\boldsymbol{x}\|^2 + \beta \sum_i h([D\boldsymbol{x}]_i; \delta_i) \qquad (1)$$

where $\boldsymbol{y}$ is the acquired projection data, and A is the forward projection operator, D is the finite difference operator. For simplicity, we assume $[D\boldsymbol{x}]_i = [x_i - x_{i,h}; x_i - x_{i,v}]$ is a 2-vector consisting of horizontal (h) and vertical (v) differences at pixel location i in a 2-D setting. The penalty function h, which itself may have parameters $\{\delta_i\}$, is assumed to be convex, penalizing large intensity differences based on the underlying assumption that CT images should be piecewise constant. The hyper-parameter $\beta > 0$ balances between data fidelity and the prior information. The parameter values, β, $\{\delta_i\}$, directly affect image quality. However, parameter tuning is a thorny issue with arguably no satisfying solutions. Without a better handle on a solution, their selections often resort to visual inspection, semi-exhaustive search or more sophisticated optimization using image quality metrics.

To introduce the concept of concomitant scale estimation, we first consider a special case of (1) where h is a quadratic regularizer, *i.e.*, $h(\boldsymbol{s}) \equiv q(\boldsymbol{s}) = \|\boldsymbol{s}\|^2/2$. The hyper-parameter β can be split into two parts, one overall constant $\beta \geq 0$, the other $\delta_i > 0$ that is position-dependent.

$$\Phi(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{y} - A\boldsymbol{x}\|^2 + \beta \sum_i \frac{1}{\delta_i} q([D\boldsymbol{x}]_i) \qquad (2)$$

J Xu is with Department of Radiology, Johns Hopkins University. F Noo is with Department of Radiology and Imaging Sciences, the University of Utah. This project was supported by the National Cancer Institute of the National Institutes of Health under R21CA211035.

The objective function (2) abstracts some recent approaches of using edge-preserving quadratic regularizers. The key is to adjust δ_i, hence the overall penalty weight (β/δ_i) at each location i, based on the image content. In a homogeneous region where the image gradient is small, δ_i can be small as well thus increase the overall weighting; near region boundaries where the image gradient is large, δ_i should be large so that overall less smoothing is applied. This idea is the principle of all edge-preserving image filters and regularizers.

The quadratic function satisfies the following identity

$$q(\boldsymbol{s})/\delta = q(\boldsymbol{s}/\delta)\delta,$$

the right-hand-side (RHS) of which is the perspective transform of q. We define $Q(\boldsymbol{s}, \delta) \triangleq q(\boldsymbol{s}/\delta)\delta$, then it is well known from convex analysis that $Q(\boldsymbol{s}, \delta)$ is jointly convex in $\boldsymbol{s}$ and δ. Thus potentially the RHS of (2) can be jointly optimized in image intensity $\boldsymbol{x}$ and the scale parameters $\{\delta_i\}$. This is what we refer to as concomitant scale estimation.

We used a quadratic regularizer as an example to explain the concept. Indeed, many popular potential functions used in MBIR for CT, such as the Huber and the Fair potential, can be written as

$$H(\boldsymbol{s}, \delta) \triangleq h(\boldsymbol{s}; \delta) = h(\boldsymbol{s}/\delta; 1)\delta, \qquad (3)$$

so that H is jointly convex in $\boldsymbol{s}$ and δ. Plugging (3) into (1) and jointly optimize image intensity $\boldsymbol{x}$ and the scale parameters $\{\delta_i\}$ comprises the basic formulation of our generic framework.

To our knowledge, the term *concomitant scale estimation* was used in [1, Chapter 7.7] to describe the same concept as we have just presented. It was also mentioned there that the function $H(\boldsymbol{s}, \delta)$ was jointly convex in $\boldsymbol{s}, \delta$ if $h(\boldsymbol{s}; \delta)$ is convex. In terms of computation, an alternating minimization algorithm was proposed to update image intensities (the location variables in [1]) and the scale parameter δ. However, the function $H(\boldsymbol{s}, \delta)$, though convex, is not smooth at $\delta = 0$. Thus, an alternating minimization algorithm is not strictly applicable.

To deal with the singularity at $\delta = 0$, we rewrite $H(\boldsymbol{s}, \delta)$ using its conjugate function. As $H(\boldsymbol{s}, \delta)$ is a homogeneous function, its conjugate function is the indicator function of a convex set. This reformulation makes a number of primal-dual splitting type algorithms applicable for the concomitant estimation problem. As an example, we use the alternating direction method of multipliers (ADMM) in this work.

The proposed framework requires some background from convex analysis. These are discussed next for self-containment. Readers should consult standard textbooks on convex analysis, *e.g.*, [2], [3], for more details.

II. Background on convex analysis

A. Conjugate functions of convex functions

The conjugate function (Fenchel-Legendre transform) of a closed proper convex function $f(\boldsymbol{x}) : \Omega \subset R^n \to R \cup \{\infty\}$, denoted by $f^*(\boldsymbol{p})$, is defined as

$$f^*(\boldsymbol{p}) = \sup_{\boldsymbol{x}\in\Omega} \boldsymbol{p}^t\boldsymbol{x} - f(\boldsymbol{x})$$

If $f(\boldsymbol{x})$ is convex, we further have $f(\boldsymbol{x}) = f^{**}(\boldsymbol{x})$.

B. Perspective transform of convex functions

The perspective transform H of a closed, proper, convex function $h(\boldsymbol{s})$ is defined as

$$H(\boldsymbol{s},\delta) = \begin{cases} \delta h(\boldsymbol{s}/\delta) & \delta > 0 \\ h^\infty(\boldsymbol{s}) & \delta = 0 \\ +\infty & \text{elsewhere} \end{cases}$$

where h^∞ is the horizon function defined by $h^\infty(\boldsymbol{s}) := \lim_{\lambda\to\infty}(h(\boldsymbol{w} + \lambda\boldsymbol{s}) - h(\boldsymbol{w}))/\lambda$. Perspective transform is one of the known transforms that preserves convexity. That is, $H(\boldsymbol{s},\delta)$ is jointly convex in $\boldsymbol{s},\delta$. See *e.g.*, [4, page 89]. When h has superlinear growth, *e.g.*, when h is a quadratic function, $h^\infty(\boldsymbol{s}) = \infty$ unless $\boldsymbol{s} = 0$. We present a formal derivation of the conjugate function of H.

$$\begin{aligned} H^*(\boldsymbol{p},q) &= \sup_{(\boldsymbol{s},\delta)} (\boldsymbol{p},q)(\boldsymbol{s},\delta) - H(\boldsymbol{s},\delta), \\ &= \max \begin{cases} \sup_{(\boldsymbol{s},\delta>0)}\{(\boldsymbol{p},q)(\boldsymbol{s},\delta) - \delta h(\boldsymbol{s}/\delta)\}, \\ \sup_{\boldsymbol{s}}\{(\boldsymbol{p},\boldsymbol{s}) - h^\infty(\boldsymbol{s})\}, \end{cases} \\ &= \max \begin{cases} \sup_{(\boldsymbol{s}',\delta>0)}\{(\boldsymbol{p}\cdot\boldsymbol{s}' + q)\delta - \delta h(\boldsymbol{s}')\}, \\ \sup_{\boldsymbol{s}}\{\boldsymbol{p}\cdot\boldsymbol{s} - h^\infty(\boldsymbol{s})\}, \end{cases} \\ &= \max \begin{cases} \sup_{\delta>0}\{\delta(q + h^*(\boldsymbol{p}))\}, \\ \iota_{\boldsymbol{p}\in\text{dom}(h^*)} \qquad (*) \end{cases} \\ &= \iota_K, \text{where} \\ &\quad K = \{(\boldsymbol{p},q) | q + h^*(\boldsymbol{p}) \le 0, \boldsymbol{p} \in \text{dom}(h^*)\} \quad (4) \end{aligned}$$

where h^* is the convex conjugate of h. In (*), the horizon function h^∞ and the indicator function of the domain of h^* are conjugate to each other, see, *e.g.*, [5, Theorem 11.5]. Note that H^* is the indicator function of the convex set K specified by (4). For convenience we often write simply $K = \{(\boldsymbol{p},q) | q + h^*(\boldsymbol{p}) \le 0\}$ and omit the domain specification. An example of the set K is illustrated in Fig. 1 for the Huber potential h (6). In this case, its conjugate function $h^*(\boldsymbol{p}) = \|\boldsymbol{p}\|^2/2$, for $\|\boldsymbol{p}\| \le 1$. The convex set K is the flipped epigraph of h^*. See [6], [7] for other examples of the set K and more details about computing the projection onto it.

III. Method

The objective function for MBIR with concomitant scale estimation is written as

$$\begin{aligned} \min_{\boldsymbol{x},\boldsymbol{\delta}} \Phi(\boldsymbol{x},\boldsymbol{\delta}) = \{&\frac{1}{2}\|\boldsymbol{y} - A\boldsymbol{x}\|^2 \\ &+ \beta_1 \sum_i H([D\boldsymbol{x}]_i, \delta_i) + \beta_2 R(\boldsymbol{\delta},\boldsymbol{x})\} \quad (5) \end{aligned}$$

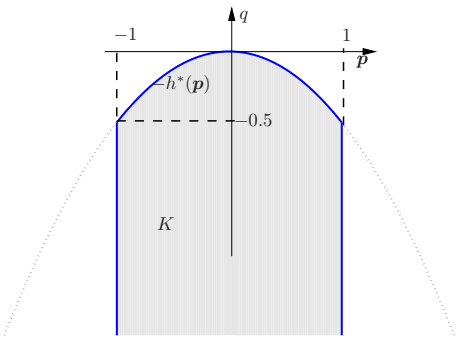


Fig. 1: The conjugate function H^* of the perspective transform $H(\boldsymbol{s},\delta)$ is the indicator function ι_K of the convex set K. Computing the proximal mapping of ι_K is simply projection onto the convex set.

Compared with (1), we rewrite $h(\cdot;\delta_i)$ as $H(\cdot,\delta_i)$ [cf. (3)] to emphasize that we optimize with respect to $\boldsymbol{\delta} := \{\delta_i\}$ as well. We have also added an additional regularizer $R(\cdot,\cdot)$ to penalize large values of δ_i. As discussed in [1], the main purpose of introducing R is to ensure a well-defined (coercive) objective function. Otherwise, for any fixed $\boldsymbol{x}$, we can always increase $\boldsymbol{\delta}$ to decrease the objective. In addition to ensure coerciveness, we may encourage a more structured image formation by penalizing large values of $\boldsymbol{\delta}$. As later our examples show, we also allow $\boldsymbol{\delta}$ to be dependent on the unknown $\boldsymbol{x}$. This is expressed as the joint penalty $R(\boldsymbol{\delta},\boldsymbol{x})$ in (5). We assume $R(\boldsymbol{\delta},\boldsymbol{x})$ is jointly convex in $\boldsymbol{\delta}$ and $\boldsymbol{x}$.

Due to the perspective transform, the objective function (5) is convex but nonsmooth. Using the conjugacy relation (4), a number of primal-dual splitting algorithms can be applied to minimize $\Phi(\boldsymbol{x},\boldsymbol{\delta})$. The key issue in computation becomes that of calculating the proximal mapping of $H(\boldsymbol{s},\delta_i)$ [8], which, upon invoking the Moreau identity, can be obtained from the proximal mapping of its conjugate function $H^*(\boldsymbol{p},q)$. As H^* is the indicator function of the convex set K, its proximal mapping is nothing but projection onto K, *i.e.*, the epigraph of h^* [9], [10], which can be related to the proximal mapping of h^* in an implicit form. See Appendix A.

Some familiar potential functions can be written as perspective transforms. Below we list some of them. Some of these examples also appear in the recent paper [8], which focuses on the proximal mapping and the calculus of perspective transform and its applications in machine learning.

A. Examples of potential functions

- The Huber function

$$h_1(\boldsymbol{s};\delta) = \begin{cases} \frac{\|\boldsymbol{s}\|^2}{2\delta}, & \|\boldsymbol{s}\| \le \delta \\ \|\boldsymbol{s}\| - \frac{\delta}{2} & \|\boldsymbol{s}\| > \delta \end{cases} \quad (6)$$

  It is easy to see that $h_1(\boldsymbol{s};\delta) = \delta h_1(\boldsymbol{s}/\delta;1)$.

- The Berhu function

$$h_2(\boldsymbol{s},\delta) = \begin{cases} \|\boldsymbol{s}\| & \|\boldsymbol{s}\| \le \delta \\ \frac{\|\boldsymbol{s}\|^2}{\delta} + \delta = \delta\left[\left(\frac{\|\boldsymbol{s}\|}{\delta}\right)^2 + 1\right] & \|\boldsymbol{s}\| > \delta \end{cases}$$

- The Fair potential

$$h_3(\boldsymbol{s};\delta) = \|\boldsymbol{s}\| - \delta \log\left(1 + \frac{\|\boldsymbol{s}\|}{\delta}\right) = \delta h_3(\boldsymbol{s}/\delta; 1) \quad (7)$$

- The relative difference prior [11][1]

$$h_4(x_1, x_2) = \frac{(x_2 - x_1)^2}{x_1 + x_2 + \gamma|x_2 - x_1|}, \quad \gamma > 0 \quad (8)$$

where x_1 and x_2 are neighboring voxel intensity values. To relate (8) to perspective transform, we first consider the following convex function

$$r(s) = \frac{s^2}{1 + \gamma|s|} \quad (9)$$

which is a special case of q-GGMRF [12]. The perspective transform of (9) is

$$r(s/\delta)\delta = \frac{s^2}{\delta + \gamma|s|} \quad (10)$$

Now that $r(s/\delta)\delta$ is jointly convex in s, δ, for $\delta \geq 0$, we further obtain (8) from (10) by the linear transformation of the variables $(s, \delta) \to (s-\delta, s+\delta)$. The non-negativity of image intensity values ensure that (8) is convex in x_1 and x_2. A similar construction may be applied to other potential functions, *e.g.*, the Fair or the Huber potential.

B. Application of the ADMM algorithm

In the following, we use the (linearized) ADMM algorithm combined with a generic $R(\boldsymbol{\delta}, \boldsymbol{x})$ and give the algorithm steps. We first introduce the following auxiliary variables for (5),

$$\boldsymbol{s} = D\boldsymbol{x}, \quad \boldsymbol{\delta} = \boldsymbol{\eta} \quad (11)$$

Minimizing (5) can be rewritten as a constrained problem

$$\min_{\boldsymbol{x},\boldsymbol{\eta},\boldsymbol{s},\boldsymbol{\delta}} \ \frac{1}{2}\|\boldsymbol{y} - A\boldsymbol{x}\|^2 + \beta_1 \sum_i H(\boldsymbol{s}_i, \delta_i) + \beta_2 R(\boldsymbol{\eta}, \boldsymbol{x})$$
$$\text{subject to} \quad (11)$$

The augmented Lagrangian of the constrained problem is

$$\frac{1}{2}\|\boldsymbol{y} - A\boldsymbol{x}\|^2 + \beta_1 \sum_i H(\boldsymbol{s}_i, \delta_i) + \beta_2 R(\boldsymbol{\eta}, \boldsymbol{x}) + \frac{\mu}{2}\|D\boldsymbol{x} - \boldsymbol{s} + \frac{\boldsymbol{\lambda}_1}{\mu}\|^2 + \frac{\mu}{2}\|\boldsymbol{\eta} - \boldsymbol{\delta} + \frac{\boldsymbol{\lambda}_2}{\mu}\|^2 \quad (12)$$

where $\boldsymbol{\lambda}_1$, $\boldsymbol{\lambda}_2$ are the Lagrange multipliers, $\mu > 0$. The update equations of the linearized ADMM are given by

- Update $\{\boldsymbol{x}, \boldsymbol{\eta}\}$

$$\{\boldsymbol{x}^{k+1}, \boldsymbol{\eta}^{k+1}\} = \arg\min_{\boldsymbol{x},\boldsymbol{\eta}} \{\frac{1}{2}\|\boldsymbol{y} - A\boldsymbol{x}\|^2 + \beta_2 R(\boldsymbol{\eta}, \boldsymbol{x}) + \frac{\tau}{2}\|\boldsymbol{x} - \left[\boldsymbol{x}^k - \frac{\mu}{\tau} D^t(D\boldsymbol{x}^k - \boldsymbol{s}^k + \boldsymbol{\lambda}_1^k/\mu)\right]\|^2 + \frac{\mu}{2}\|\boldsymbol{\eta} - \boldsymbol{\delta}^k + \boldsymbol{\lambda}_2^k/\mu\|^2\} \quad (13)$$

Convergence requires that $\tau > \mu\|D^t D\|$.

[1]The prior was used in a negated form in a maximum a posteriori reconstruction problem, therefore in [11] it was a concave prior.

- Update $\{\boldsymbol{s}, \boldsymbol{\delta}\}$

$$\{\boldsymbol{s}^{k+1}, \boldsymbol{\delta}^{k+1}\} = \arg\min_{\boldsymbol{s},\boldsymbol{\delta}} \{\beta_1 \sum_i H(\boldsymbol{s}_i, \delta_i) + \frac{\mu}{2}\|D\boldsymbol{x}^{k+1} - \boldsymbol{s} + \boldsymbol{\lambda}_1^k/\mu\|^2 + \frac{\mu}{2}\|\boldsymbol{\eta}^{k+1} - \boldsymbol{\delta} + \boldsymbol{\lambda}_2^k/\mu\|^2\} \quad (14)$$

- Update Lagrange multipliers

$$\boldsymbol{\lambda}_1^{k+1} \leftarrow \boldsymbol{\lambda}_1^k + \mu(D\boldsymbol{x}^{k+1} - \boldsymbol{s}^{k+1})$$
$$\boldsymbol{\lambda}_2^{k+1} \leftarrow \boldsymbol{\lambda}_2^k + \mu(\boldsymbol{\eta}^{k+1} - \boldsymbol{\delta}^{k+1})$$

Now we discuss two examples in image denoising ($A = \mathrm{Id}$), corresponding to two choices of $R(\boldsymbol{\delta}, \boldsymbol{x})$, to see the detailed steps involved in the above update equations.

C. Example 1: $R(\delta, \boldsymbol{x}) = 1_{\{\delta \leq \gamma\bar{\boldsymbol{x}}\}}$

In this example, we require that the single value δ is at most γ times the mean image intensity $\bar{\boldsymbol{x}}$, which is unknown. The objective function is given by

$$\Phi_1(\boldsymbol{x}, \delta) = \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{x}\|^2 + 1_{\{\delta \leq \gamma\bar{\boldsymbol{x}}\}} + \beta_1 \sum_i H([D\boldsymbol{x}]_i, \delta) \quad (15)$$

In (15), we have removed the weighting coefficient β_2 as it becomes irrelevant for an indicator function. The constraints in (11) translates to $\boldsymbol{\delta} = \eta\mathbf{1}$, *i.e.*,

$$\eta = \delta_i, i = 1, \cdots, N$$

From (13), define

$$\eta^{k.5} = \sum_i (\delta_i^k - \lambda_{2,i}^k/\mu)/N \quad (16a)$$

$$\boldsymbol{\chi}^k = \boldsymbol{x}^k - \frac{\mu}{\tau} D^t(D\boldsymbol{x}^k - \boldsymbol{s}^k + \boldsymbol{\lambda}_1^k/\mu) \quad (16b)$$

$$\boldsymbol{x}^{k.5} = (\boldsymbol{y} + \tau\boldsymbol{\chi}^k)/(1 + \tau) \quad (16c)$$

- Update $\{\boldsymbol{x}, \eta\}$. The update equation (13) becomes

$$\{\boldsymbol{x}^{k+1}, \eta^{k+1}\} = \arg\min_{\boldsymbol{x},\eta} \{1_{\{\eta \leq \gamma\bar{\boldsymbol{x}}\}} + \frac{\tau + 1}{2}\|\boldsymbol{x} - \boldsymbol{x}^{k.5}\|^2 + \frac{N\mu}{2}\|\eta - \eta^{k.5}\|^2\}$$

where $\bar{\boldsymbol{x}} = (x_1 + \cdots + x_N)/N$. This subproblem can be conceptualized as projection onto the half-space $\{\eta \leq \gamma\bar{\boldsymbol{x}}\}$ with (unequal) weighting. The solution is given by

$$r^k = \frac{\max\{\eta^{k.5} - \gamma\bar{\boldsymbol{x}}^{k.5}, 0\}}{1 + \gamma^2/(Nw)}$$
$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^{k.5} + r^k\gamma/(Nw)$$
$$\eta^{k+1} = \eta^{k.5} - r^k, \quad \text{where } w = (\tau + 1)/(N\mu)$$

and $\bar{\boldsymbol{x}}^{k.5} = (x_1^{k.5} + \cdots + x_N^{k.5})/N$.

- Update $\boldsymbol{s}, \boldsymbol{\delta}$. Again we define short-hand notation

$$\boldsymbol{s}^{k.5} = D\boldsymbol{x}^{k+1} + \boldsymbol{\lambda}_1^k/\mu$$
$$\boldsymbol{\delta}^{k.5} = \eta^{k+1} + \boldsymbol{\lambda}_2^k/\mu$$

the update equation (14) can be decomposed componentwise for each voxel i:

$$\{\boldsymbol{s}_i^{k+1}, \delta_i^{k+1}\} = \arg\min_{\boldsymbol{s}_i, \delta_i} \{\beta_1 H(\boldsymbol{s}_i, \delta_i) + \frac{\mu}{2}\|\boldsymbol{s}_i - \boldsymbol{s}_i^{k.5}\|^2 + \frac{\mu}{2}\|\delta_i - \delta_i^{k.5}\|^2\} \quad (17)$$

We invoke Moreau's identity that relates the proximal mapping of H with the proximal mapping of its conjugate function H^*. Let $\boldsymbol{z} = \{\boldsymbol{s}, \delta\}$, we have

$$\boldsymbol{z} = \text{prox}_{\kappa H}(\boldsymbol{z}) + \kappa \text{prox}_{\kappa^{-1} H^*}(\boldsymbol{z}/\kappa) \quad (18)$$

Let $\kappa = \beta_1/\mu$, then the problem is converted to calculating the proximal mapping of $H^* = \iota_K$ [cf. (4)], which amounts to projection onto the convex set K, or the epigraph of h^*.

Although we have not explored this, it may speed up convergence when the two quadratic terms in (14) or (17) have unequal weighting (different μs). In this case, we would need a slight generalization of the Moreau identity with weighting. This is derived in Appendix B.

D. Example 2: $R(\boldsymbol{\delta}) = \sum_i \|\boldsymbol{\delta}_i\| = \sum_i \sqrt{\delta_{i,h}^2 + \delta_{i,v}^2}$

The objective function is given by

$$\Phi_2(\boldsymbol{x}, \boldsymbol{\delta}) = \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{x}\|^2 + \beta_2 \sum_i \|\boldsymbol{\delta}_i\| + \beta_1 \sum_i H_a([D\boldsymbol{x}]_i, \boldsymbol{\delta}_i)$$

where H_a is the anisotropic version of the Huber potential, *i.e.*, the scalar Huber potential function applied to each component of the finite difference $[D\boldsymbol{x}]_i$ at voxel i.

- Update $\boldsymbol{x}, \boldsymbol{\eta}$. Using the definitions (16), the update equation (13) becomes

$$\{\boldsymbol{x}^{k+1}, \boldsymbol{\eta}^{k+1}\} = \arg\min_{\boldsymbol{x}, \boldsymbol{\eta}} \{\beta_2 \sum_i \|\boldsymbol{\eta}_i\| + \frac{\tau+1}{2}\|\boldsymbol{x} - \boldsymbol{x}^{k.5}\|^2 + \frac{\mu}{2}\|\boldsymbol{\eta} - \boldsymbol{\eta}^{k.5}\|^2\}$$

  therefore, $\boldsymbol{x}^{k+1} = \boldsymbol{x}^{k.5}$, and $\boldsymbol{\eta}^{k+1} = \text{prox}_{\beta_2/\mu\|\cdot\|}(\boldsymbol{\eta}^{k.5})$ is the soft block-thresholding of $\boldsymbol{\eta}^{k.5}$.
- Update $\boldsymbol{s}, \boldsymbol{\delta}$. The update equations are almost identical to (17), but for an anisotropic Huber function. In other words, not only the update equation decomposes voxelwise, it also decomposes direction-wise for each element of the 2-vectors $\boldsymbol{s}_i$ and $\boldsymbol{\delta}_i$.

E. Generalization to image reconstruction

The objective function can be written as

$$\min_{\boldsymbol{x}, \boldsymbol{\eta}, \boldsymbol{s}, \boldsymbol{\delta}, \bar{\boldsymbol{y}}} \frac{1}{2}\|\boldsymbol{y} - \bar{\boldsymbol{y}}\|^2 + \beta_1 \sum_i H(\boldsymbol{s}_i, \delta_i) + \beta_2 R(\boldsymbol{\eta}, \boldsymbol{x}) \quad (19)$$

$$\text{subject to} \quad \begin{bmatrix} D & 0 \\ 0 & \text{Id} \\ A & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{\eta} \end{bmatrix} = \begin{bmatrix} \boldsymbol{s} \\ \boldsymbol{\delta} \\ \bar{\boldsymbol{y}} \end{bmatrix} \quad (20)$$

If we define $\boldsymbol{\xi} = \{\boldsymbol{x}, \boldsymbol{\eta}\}$, $\boldsymbol{\zeta} = \{\bar{\boldsymbol{y}}, \boldsymbol{s}, \boldsymbol{\delta}\}$ we may further rewrite the linear constraint (20) as $\mathbb{A}\boldsymbol{\xi} = \boldsymbol{\zeta}$. Applying the linearized ADMM algorithm, we obtain the following update equations

- Update $\boldsymbol{\xi} = \{\boldsymbol{x}, \boldsymbol{\eta}\}$

$$\{\boldsymbol{x}^{k+1}, \boldsymbol{\eta}^{k+1}\} = \arg\min_{\boldsymbol{x}, \boldsymbol{\eta}} \{\beta_2 R(\boldsymbol{\eta}, \boldsymbol{x}) + \frac{\tau}{2}\|\boldsymbol{x} - \boldsymbol{x}^{k.5}\|^2 + \frac{\mu}{2}\|\boldsymbol{\eta} - \boldsymbol{\delta}^k + \boldsymbol{\lambda}_2^k/\mu_1\|^2\}$$

$$\boldsymbol{x}^{k.5} = \boldsymbol{x}^k - \frac{\mu}{\tau}[A^t(A\boldsymbol{x}^k - \bar{\boldsymbol{y}}^k + \frac{\boldsymbol{\lambda}_3^k}{\mu}) + D^t(D\boldsymbol{x}^k - \boldsymbol{s}^k + \frac{\boldsymbol{\lambda}_1^k}{\mu})]$$

  For convergence we require $\tau > \mu\|A^tA + D^tD\|$.
- Update $\boldsymbol{\zeta} = \{\bar{\boldsymbol{y}}, \boldsymbol{s}, \boldsymbol{\delta}\}$

$$\{\bar{\boldsymbol{y}}^{k+1}, \boldsymbol{s}^{k+1}, \boldsymbol{\delta}^{k+1}\} = \arg\min\{\frac{1}{2}\|\boldsymbol{y} - \bar{\boldsymbol{y}}\|^2 + \beta_1 \sum_i H(\boldsymbol{s}_i, \delta_i) + \frac{\mu}{2}\|D\boldsymbol{x}^{k+1} - \boldsymbol{s} + \boldsymbol{\lambda}_1^k/\mu_1\|^2 + \frac{\mu}{2}\|\boldsymbol{\eta}^{k+1} - \boldsymbol{\delta} + \boldsymbol{\lambda}_2^k/\mu_1\|^2 + \frac{\mu}{2}\|A\boldsymbol{x}^{k+1} - \bar{\boldsymbol{y}} + \boldsymbol{\lambda}_3^k/\mu\|^2\}$$

  This subproblem further splits between $\bar{\boldsymbol{y}}$ and $\boldsymbol{s}, \boldsymbol{\delta}$: $\bar{\boldsymbol{y}}^{k+1}$ can be obtained by completion of squares, and $\boldsymbol{s}, \boldsymbol{\delta}$ can be obtained using projection onto epigraphs similar to the denoising problem.
- Update Lagrange multipliers $\boldsymbol{\lambda}_1$, $\boldsymbol{\lambda}_2$, $\boldsymbol{\lambda}_3$

$$\boldsymbol{\lambda}_1^{k+1} \leftarrow \boldsymbol{\lambda}_1^k + \mu(D\boldsymbol{x}^{k+1} - \boldsymbol{s}^{k+1})$$
$$\boldsymbol{\lambda}_2^{k+1} \leftarrow \boldsymbol{\lambda}_2^k + \mu(\boldsymbol{\eta}^{k+1} - \boldsymbol{\delta}^{k+1})$$
$$\boldsymbol{\lambda}_3^{k+1} \leftarrow \boldsymbol{\lambda}_3^k + \mu(A\boldsymbol{x}^{k+1} - \bar{\boldsymbol{y}}^{k+1})$$

IV. Extension to joint multiphase image reconstruction

Our development was partially motivated by multiphase CT reconstruction. During a multiphase CT exam, a sequence of two, three, or even four scans are acquired at different time points following contrast injection. These multiple independent acquisitions share a huge amount of information, the utilization of which presents opportunities to reduce radiation dose while maintaining image quality.

An example of extending (5) to joint multiphase reconstruction can be the following:

$$\min \Phi(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_P, \Delta_{\{i\}}) = \{\sum_p \frac{1}{2}\|\boldsymbol{y}^p - A^p\boldsymbol{x}^p\|^2 + \beta_1 \sum_i J(T_i(\boldsymbol{x}), \Delta_i) + \beta_2 \sum_i R(\Delta_i)\} \quad (21)$$

where $\boldsymbol{y}^p$, A^p are the projection data and the forward projection matrix, $\boldsymbol{x}^p$ is the reconstructed image, at each phase $p = 1, \cdots, P$. T_i is the (multiphase) image gradient tensor at voxel i, $T_i = \{D\boldsymbol{x}_i^1, \cdots, D\boldsymbol{x}_i^P\}$ and $\Delta_i \succeq 0$ is a positive semi-definite matrix playing a role similar to $\delta \geq 0$ for single image reconstruction. The joint penalty term J is the matrix analogue to the quadratic penalty $q(\boldsymbol{s}) = \|\boldsymbol{s}\|^2/(2\delta)$.

$$J(T, \Delta) = \sum_i \text{Tr}\{T_i^t \Delta_i^{-1} T_i\}$$

Note that the mapping $(T, \Delta) \rightarrow T^t\Delta^{-1}T$ is (operator) convex [13] for $\Delta \succ 0$. We may use a variety of matrix norms for the penalty $R(\Delta)$, *e.g.*, the trace norm or the Frobenius

norm. For computation, we will need a matrix version of the conjugacy relationship in (4).

The formulation (21) may also be applicable to single image reconstruction. If we compose T_i using the gradient vectors within a small neighborhood of the voxel i, and letting $R(\Delta_i) = \mathrm{Tr}\,\Delta_i$, the regularizer in (21) resembles the structure tensor TV [14].

V. Discussion

Our contribution in this work can be interpreted in different ways. One is that the combination of $H(\boldsymbol{s}, \boldsymbol{\delta})$ and $R(\boldsymbol{\delta}, \boldsymbol{x})$, upon minimizing with respect to $\boldsymbol{\delta}$, generates new regularizers on image intensity values. On the other hand, it also provides a systematic approach for exploring the large hyper-parameter space for some well known regularizers, including the Huber and the Fair potential. Moreover, the concept of concomitant scale estimation can be applied to the data-fitting terms [8] that are constructed as perspective transforms, thus allowing novel, non-least-squares types of data-fitting terms to be handled.

In [1], $R(\boldsymbol{\delta}, \boldsymbol{x}) = \boldsymbol{\delta}$ was used as an example to penalize large values of $\boldsymbol{\delta}$. If this choice is combined with the quadratic regularizer, *i.e.*, $q(\boldsymbol{s}, \delta) = \|\boldsymbol{s}\|^2/(2\delta)$, then as

$$\min_{\delta_i \geq 0} \quad \frac{\|\boldsymbol{s}_i\|^2}{2\delta_i} + \beta_2' \delta_i = \|\boldsymbol{s}_i\| \sqrt{\beta_2'/2} \tag{22}$$

we simply recover the conventional TV. Indeed, (22) is known as the half quadratic formulation (in the multiplicative form) of TV. This example, and the example of structure tensor TV we discussed in Sec IV, indicate that not all combinations of $H(\boldsymbol{s}, \boldsymbol{\delta})$ and $R(\boldsymbol{\delta}, \boldsymbol{x})$ produces new regularizers. In this sense, our development provides an alternative perspective and a computational procedure that unifies the existing approaches. However, new regularizers will arise if we incorporate nonlocal or image intensity information in constructing $R(\boldsymbol{\delta}, \boldsymbol{x})$, as is done in the relative difference prior [11].

Our numerical examples mostly consider using the Huber function as $H(\boldsymbol{s}, \boldsymbol{\delta})$. Note that the computational procedure is general, and can be adapted to other priors as well. The only difference would be calculating the projection onto the convex set K [cf. Fig. 1]. For the Huber or the quadratic prior, the convenience is that the projection can be calculated in analytical form [7], which may not exist for other functions, such as the Fair potential.

VI. Conclusions

We investigated a general framework of concomitant scale estimation, that is, a framework that estimates simultaneously image intensity values and the hyper-parameters in the regularizer. We also proposed a primal-dual splitting type computational procedure for concomitant scale estimation. We used three examples to demonstrate the wide applicability and the versatility of concomitant scale estimation.

We are at an early stage of this investigation. Many questions remain unanswered, for example, how to choose the penalty for the hyper-parameters $R(\boldsymbol{\delta}, \boldsymbol{x})$ given a set of desirable imaging properties. But the general framework of concomitant scale estimation provides a systematic approach of exploring the large hyper-parameter space, and holds promise for new regularizer designs and image quality improvement.

Appendix A
Epigraphical projection

It is possible to relate the projection onto the epigraph of a generic function h to its proximal mapping, prox_h. The projection (prox operator) onto the epigraph is defined as the mapping $(\boldsymbol{x}, y) \to (\boldsymbol{u}, v)$ such that

$$(\boldsymbol{u}, v) = \arg\min_{\boldsymbol{x}', y'} \ \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{x}'\|^2 + \frac{1}{2}(y - y')^2 + \iota_{(\boldsymbol{x}', y') \in \mathrm{epi}(h)} \tag{23}$$

When $(\boldsymbol{x}, y) \in \mathrm{epi}(h)$, *i.e.*, $y \geq h(\boldsymbol{x})$, we have $(\boldsymbol{u}, v) = (\boldsymbol{x}, y)$. Otherwise, $y' = h(\boldsymbol{x}')$, and we seek $\boldsymbol{u} = \boldsymbol{x}'$ such that

$$\min_{\boldsymbol{x}'} \ \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{x}'\|^2 + \frac{1}{2}(y - h(\boldsymbol{x}'))^2 + \{\iota_{(\boldsymbol{x}', y') \in \mathrm{epi}(h)}\} \tag{24}$$

we can ignore the constraint on the epigraph in (24) as it is satisfied. Taking derivative with respect to $\boldsymbol{x}'$ of the objective function in (24), we have

$$\boldsymbol{x} \in \boldsymbol{x}' + (h(\boldsymbol{x}') - y)\partial h(\boldsymbol{x}') \tag{25}$$

which is a characterization of the solution to (23) when $(\boldsymbol{x}, y) \notin \mathrm{epi}(h)$ [3, Proposition 28.28]. Letting $h(\boldsymbol{x}') - y = \rho > 0$ (unknown for now), then

$$0 \in \boldsymbol{x}' - \boldsymbol{x} + \rho \partial h(\boldsymbol{x}') \to \boldsymbol{x}' = \mathrm{prox}_{\rho h}(\boldsymbol{x})$$

Appendix B
A generalization of Moreau identity

We consider the weighted proximal mapping

$$\mathrm{prox}_{\kappa F}^{W}(\boldsymbol{z}) \stackrel{\triangle}{=} \arg\min_{x} \ \kappa F(\boldsymbol{x}) + \frac{1}{2}\|\boldsymbol{z} - \boldsymbol{x}\|_w^2$$

where $F(\boldsymbol{x})$ is a closed, proper, convex function, $\|\boldsymbol{z}\|_W^2 = \boldsymbol{z}^t W \boldsymbol{z}$ denotes the weighted least squares, and the weighting matrix $W \succ 0$ is positive definite. Let $F^*(\boldsymbol{p}) = \sup_x(\boldsymbol{p}^t \boldsymbol{x} - F(\boldsymbol{x}))$ be the conjugate function of $F(\boldsymbol{x})$. Assume the generic saddle problem

$$\min_{\boldsymbol{x}} \max_{\boldsymbol{p}} \quad \Phi(\boldsymbol{x}, \boldsymbol{p}) = \kappa \boldsymbol{p}^t \boldsymbol{x} - \kappa F^*(\boldsymbol{p}) + \frac{1}{2}\|\boldsymbol{z} - \boldsymbol{x}\|_W^2$$

has at least one solution $\boldsymbol{x}_*, \boldsymbol{p}_*$, such that $\Phi(\boldsymbol{x}_*, \boldsymbol{p}) \leq \Phi(\boldsymbol{x}_*, \boldsymbol{p}_*) \leq \Phi(\boldsymbol{x}, \boldsymbol{p}_*), \quad \forall \boldsymbol{x}, \boldsymbol{p}$. The solution $\boldsymbol{p}^*$ satisfies

$$\boldsymbol{p}_* = \arg\min_{\boldsymbol{p}} \kappa^{-1} F^*(\boldsymbol{p}) + \frac{1}{2}\|\boldsymbol{p} - \kappa^{-1} W \boldsymbol{z}\|_{W^{-1}}^2$$

and $\boldsymbol{x}_*$ and $\boldsymbol{p}_*$ are related by $\boldsymbol{x}_* + \kappa W^{-1} \boldsymbol{p}_* = \boldsymbol{z}$ which can be rewritten as

$$\boldsymbol{z} = \mathrm{prox}_{\kappa F}^{W}(\boldsymbol{z}) + \kappa W^{-1} \mathrm{prox}_{\kappa^{-1} F^*}^{W^{-1}}(\kappa^{-1} W \boldsymbol{z}) \tag{26}$$

Obviously, (26) coincides with (18) when $W = \mathrm{Id}$.

References

[1] P. J. Huber and E. M. Ronchetti, *Robust Statistics*. Hoboken, New Jersey: Wiley, 2 ed., 2009.

[2] R. T. Rockafellar, *Convex analysis*. Princeton University Press, 2015.

[3] H. H. Bauschke and P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*, vol. 408. Springer, 2011.

[4] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[5] R. T. Rockafellar and R. J.-B. Wets, *Variational analysis*, vol. 317. Springer Science & Business Media, 2009.

[6] T. Pock, D. Cremers, H. Bischof, and A. Chambolle, "Global solutions of variational models with convex regularization," *SIAM Journal on Imaging Sciences*, vol. 3, no. 4, pp. 1122–1145, 2010.

[7] E. Strekalovskiy, A. Chambolle, and D. Cremers, "Convex relaxation of vectorial problems with coupled regularization," *SIAM Journal on Imaging Sciences*, vol. 7, no. 1, pp. 294–336, 2014.

[8] P. L. Combettes and C. L. Müller, "Perspective functions: Proximal calculus and applications in high-dimensional statistics," *Journal of Mathematical Analysis and Applications*, 2016.

[9] G. Chierchia, N. Pustelnik, B. Pesquet-Popescu, and J.-C. Pesquet, "A nonlocal structure tensor-based approach for multicomponent image recovery problems," *IEEE Transactions on Image Processing*, vol. 23, no. 12, pp. 5531–5544, 2014.

[10] G. Chierchia, N. Pustelnik, J.-C. Pesquet, and B. Pesquet-Popescu, "Epigraphical projection and proximal tools for solving constrained convex optimization problems," *Signal, Image and Video Processing*, vol. 9, no. 8, pp. 1737–1749, 2015.

[11] J. Nuyts, D. Beque, P. Dupont, and L. Mortelmans, "A concave prior penalizing relative differences for maximum-a-posteriori reconstruction in emission tomography," *IEEE Transactions on nuclear science*, vol. 49, no. 1, pp. 56–60, 2002.

[12] J.-B. Thibault, K. D. Sauer, C. A. Bouman, and J. Hsieh, "A three-dimensional statistical approach to improved image quality for multislice helical ct," *Medical physics*, vol. 34, no. 11, pp. 4526–4544, 2007.

[13] E. H. Lieb and M. B. Ruskai, "Some operator inequalities of the schwarz type," *Advances in Mathematics*, vol. 12, no. 2, pp. 269–273, 1974.

[14] S. Lefkimmiatis, A. Roussos, P. Maragos, and M. Unser, "Structure tensor total variation," *SIAM Journal on Imaging Sciences*, vol. 8, no. 2, pp. 1090–1122, 2015.

Impact of the Non-Negativity Constraint in Model-Based Iterative Reconstruction from CT Data

Viktor Haase, Katharina Hahn, Harald Schöndube, Karl Stierstorfer, and Frédéric Noo

***Abstract*—Model-based iterative reconstruction is a promising approach to achieve dose reduction without affecting image quality in diagnostic X-ray computed tomography. In the problem formulation, it is common to enforce non-negative values, which is motivated by physics but narrows down the choice of optimization algorithm. In this work, we report on experiments assessing the impact of the non-negativity constraint on image quality and reconstruction speed. The assessment is performed under eight scenarios that challenge the usefulness of the constraint. These include reconstructions from full and sparse view sampling, with quadratic or edge-preserving regularization, for two different objects. Our results show that improvements due to the non-negativity constraint are strongly scenario-dependent, and likely negligible for conventional full view CT imaging. This implies that for specific reconstructions, the non-negativity constraint could be disregarded to simplify the optimization problem.**

I. Introduction

A lot of the research in clinical computed tomography (CT) is driven by the aim to reduce radiation dose while maintaining image quality. One promising way to achieve this goal is model-based iterative reconstruction (MBIR). Its potential for diagnostic CT imaging was shown in recent studies [1]–[4]. A popular MBIR formulation is penalized least squares reconstruction [5], which includes two key components: (i) the data fidelity term, which is characterized by the choice of a forward projection model and the option of a statistical weighting of the projections; (ii) the penalty term, which defines a regularization process with a potential function and additional incorporation of a priori knowledge, such as the non-negativity constraint.

To get the most out of the MBIR approach, it is valuable to understand the impact of each component and its subparts. For that reason, Thibault et al. have examined different potential functions for the regularization when they first introduced the concept of MBIR [5]. As another example, Hahn et al. have compared linear interpolation models for iterative CT reconstruction in various imaging scenarios [6]. In our previous work, we focused on the effect of statistical weights, which was analyzed using a lesion detection study with human observers [7].

The purpose of this paper is to look at the impact of the non-negativity constraint. Since the attenuation coefficient of X-rays is known to be positive, this constraint appears very natural, and its use is reinforced by knowledge that reconstruction without it leads to non-physical negative attenuation values, likely due to inaccuracy of the forward projection model and noise in the projection data. Our primary question of interest is about quantifying the impact of the constraint in regions of the image that are within the object. To our knowledge, this question has not been thoroughly addressed in the context of MBIR from real CT data.

V. Haase, K. Hahn, H. Schöndube, and K. Stierstorfer are with Siemens Healthcare GmbH, Forchheim, Germany. V. Haase is also with the Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany. F. Noo is with the Department of Radiology and Imaging Sciences, University of Utah, Salt Lake City, Utah, USA.

II. Background

MBIR from CT data is formulated as optimization of a convex objective function. The optimization requires a dedicated algorithm that is consistent with the characteristics of this function. These two aspects are covered in the following subsections.

A. Objective Function

Let $\vec{x} \in \mathbb{R}^N$ be a discrete vector for the 3-D reconstruction volume and $\vec{p} \in \mathbb{R}^M$ a discrete vector for the measured projection data. Our objective function F consists of three parts: the data fidelity term f, the regularization term g with its hyper parameter $\beta \geq 0$, and the indicator function $\iota_{\mathbb{R}_+}$:

$$F(\vec{x}) = f(\vec{x}) + \beta\, g(\vec{x}) + \iota_{\mathbb{R}_+}(\vec{x}). \tag{1}$$

To ensure the data fidelity, we use the squared residual between the forward projected reconstruction and the projection data:

$$f(\vec{x}) = \frac{1}{2}\|A\vec{x} - \vec{p}\|_2^2. \tag{2}$$

The forward projection is modeled using Joseph's method [8]. It is a ray driven approach that provides a good compromise between accuracy and computational cost [6]. The forward projection process is symbolized by matrix A. The matching back projection operator is written as its transpose, A^T.

To reduce the noise in the reconstruction, a regularization term is used that is defined as

$$g(\vec{x}) = \sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}\psi(x_i - x_j), \tag{3}$$

with $w_{ij} = 1$ for the neighbors of each voxel found in the three Cartesian directions and $w_{ij} = 0$ otherwise. ψ is the potential function that assigns a cost to the difference between each voxel and one of its neighbors. We use two different potential functions for two different reconstruction scenarios, namely a quadratic and an edge-preserving potential function that can both be fine-tuned with $\delta > 0$:

$$\begin{aligned} &\text{(i)} \quad \psi(t,\delta) = \frac{t^2}{2\delta}, \\ &\text{(ii)} \quad \psi(t,\delta) = \sqrt{t^2+\delta^2} - \delta. \end{aligned} \tag{4}$$

Parameter δ controls the importance given to differences between neighboring voxel values. Both choices for ψ are convex and differentiable, which implies that they are easy to handle for an optimization algorithm.

When we enforce non-negative voxels for our reconstruction result, the indicator function $\iota_{\mathbb{R}_+}$ is part of the objective function. This non-negativity constraint is defined as

$$\iota_{\mathbb{R}_+}(\vec{x}) = \begin{cases} 0 & \text{if } x_i \in \mathbb{R}_+ \\ +\infty & \text{otherwise} \end{cases} . \tag{5}$$

B. Optimization Algorithm

Because the indicator function is non-smooth, it has to be processed via its proximal operator. This means the optimization algorithm needs to be able to minimize a convex objective function that consists of a smooth and a proximable part. The fast iterative shrinkage-thresholding algorithm, also known as FISTA, meets these demands [9]. It requires only one gradient evaluation per iteration. The pseudocode of FISTA applied to our reconstruction problem is shown in Algorithm 1. We chose the version of FISTA with a fixed step size λ. The convergence is guaranteed for $\lambda < 1/L$, where L is the Lipschitz constant for the gradient of the smooth part, $f + \beta\, g$, of our objective function. The non-negativity constraint is enforced in line 3 of the algorithm. Because the proximal operator of an indicator function of a given set is the orthogonal projection operator onto the same set,

$$\text{prox}_{\iota_{\mathbb{R}_+}}(\vec{x}) = \begin{cases} x_i & \text{if } x_i \in \mathbb{R}_+ \\ 0 & \text{otherwise} \end{cases} . \tag{6}$$

In the pseudocode we use the shorter symbol $(\cdot)_+$ instead of $\text{prox}_{\iota_{\mathbb{R}_+}}(\cdot)$.

Algorithm 1: FISTA with constant step size

Input: Parameters $\beta \geq 0, \lambda > 0$ and initial image $\vec{x}_0$.

1 $\vec{y}_1 = \vec{x}_0$, $t_1 = 1$

2 **for** $k = 1, 2, \ldots$ **do**

3 $\quad \vec{x}_k = (\vec{y}_k - \lambda(A^T(A\vec{y}_k - \vec{p}) + \beta(\nabla g)(\vec{x})))_+$

4 $\quad t_{k+1} = \dfrac{1 + \sqrt{1 + 4t_k^2}}{2}$

5 $\quad \vec{y}_{k+1} = \vec{x}_k + (\dfrac{t_k - 1}{t_{k+1}})(\vec{x}_k - \vec{x}_{k-1})$

If we do not apply the non-negativity constraint, $\iota_{\mathbb{R}_+}$ is not used and our objective function is purely smooth. In this case, the algorithm simplifies itself to Nesterov's accelerated gradient descent [10]. The only change in the presented pseudocode is that the operation $(\cdot)_+$ disappears from line 3. The convergence condition for the step size is the same as in FISTA.

III. Experimental Setup

To evaluate our reconstruction method, experiments were carried out on a state-of-the-art clinical CT system. An overview of the scanner geometry can be found in Table I. The X-ray tube was operated at $80\,\text{kV}$ and $500\,\text{mAs}$. We have used a circular trajectory scan with a flying-focal-spot (FFS) in the x,y-plane that records 2304 projection images distributed over $360°$. We refer to this setting as full view sampling. To create a second scenario that is more challenging for the reconstruction algorithm, we also consider using only every 4th projection image taken from one of the two focus positions. This results in $^1/_8$th of the original projection data and represents a sparse view sampling.

TABLE I
Parameters of scanner geometry

Source to detector distance	$108.56\,\text{cm}$
Source trajectory radius	$59.5\,\text{cm}$
Anode angle	$7°$
Number of detector channels	736
Angular detector width	$0.067\,864°$
Number of detector rows	8
Detector row height at isocenter	$0.06\,\text{cm}$
Number of projections	2304 (full view sampling) 288 (sparse view sampling)

The ACR CT accreditation phantom (model 464, Gammex-RMI, Middleton, WI, USA) was scanned as a test object. It has a cylindrical shape with a $20\,\text{cm}$-diameter and a length of $16\,\text{cm}$. The phantom is divided into four different modules of which we have looked at two for our study. The first one, called module A, has five cylinders representing the attenuation behavior of bone, polyethylene, water, acrylic, and air, respectively. Also two ramps are included that consists of small bars which are visible in $0.5\,\text{mm}$ z-axis increments. The module can be used to assess positioning and CT number accuracy. The second one, called module D, contains eight aluminum bar patterns with up to $12\,\text{lp/cm}$. The bar patterns provide very high contrast relative to the background and are used to assess the spatial resolution for high contrast objects. During the scan, each module was centered on the rotation axis, and the plane of the source trajectory passed through the middle of the module.

All reconstructions were done on a grid of $512 \times 512 \times 16$ voxels centered at the isocenter of the scanner. The voxel size was $0.1\,\text{cm}$ in x and y, and $0.06\,\text{cm}$ in z. The radius for the field of view (FOV) in the x, y-plane was set to $25\,\text{cm}$. 5000 iterations of the reconstruction algorithm were calculated and after every 25th iteration the intermediate result was saved. As initial reconstruction volume we used $\vec{x}_0 = \vec{0}$. The chosen step size λ was based on the Lipschitz constant of the data fidelity term, $L(f)$. This can be calculated as the largest eigenvalue of $A^T A$ using the power method [11]. To account for g, we used $\lambda = 0.95/L(f)$ and the knowledge that $L(g) \ll L(f)$. For the full view sampling this means $\lambda = 0.000065$, and for the sparse view sampling, it is $\lambda = 0.00052$. The hyper parameter β is also linked to the size of the projection data. For the full view sampling we used $\beta = 0.1$; for the sparse sampling, which has $^1/_8$th of the full projection data, we used $\beta = 0.0125$ so that the amount of applied regularization is comparable for the different data sets. The value for the parameter of the potential function was empirically chosen as $\delta = 0.0005$

for the quadratic regularization and $\delta = 0.001$ for the edge-preserving regularization

To summarize, we have four different projection data sets that differ in the scanned object and the number of projections, and for reconstruction we use two different regularizers. This results in eight reconstruction scenarios for which we compare the result with and without the non-negativity constraint.

To assess the image quality, difference images between the result with and without non-negativity constraint were calculated. Where needed, profile plots through the difference image were created. Note that the result of the reconstruction is a 3-D volume, of which we only analyze one of the central slices for simplification. As figure of merit for convergence, we used the root-mean-square error (RMSE). To ignore the irrelevant structures outside of the phantom, a binary mask was applied with a radius of $10.5\,\mathrm{cm}$ that only contains the ACR module including its edges. The RMSE was calculated within the binary mask for the central six z-slices of the reconstruction. The final result after 5000 iterations was used as ground truth.

IV. Results

The results are separated into two parts. We first present the outcome of the experiments for the projection data with full view sampling, and then the results for the sparse view projection data. The results are focused on difference images and convergence speed. Fig. 1 shows how some reconstructions appear prior to computing differences.

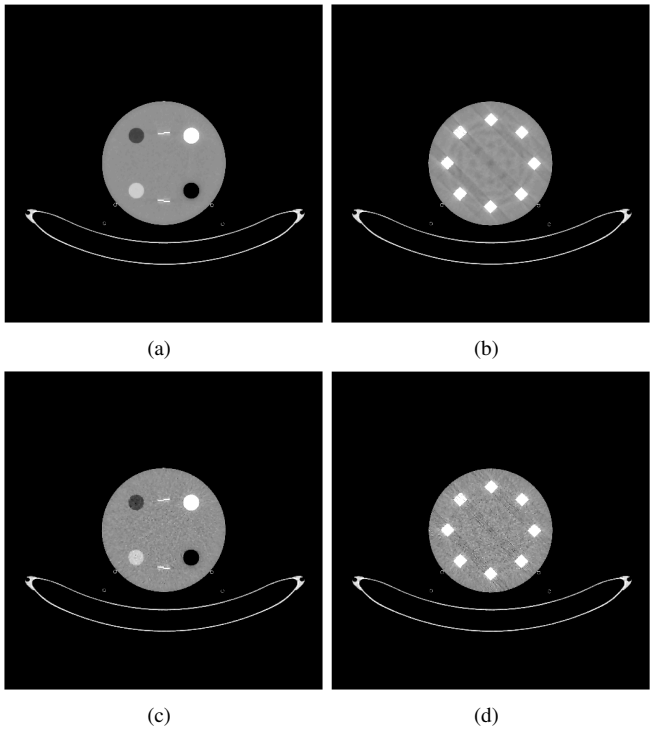


Fig. 1. Ground truth for reconstruction with edge-preserving regularization and non-negativity constraint after 5000 iterations. (Top row) full view sampling, (bottom row) sparse view sampling. Module A (left) and module D (right) are both displayed with a grayscale of $[-200,\ 200]$ HU.

A. Full View Sampling of Projection Data

Fig. 2 shows the difference images between the reconstruction with non-negativity constraint and without. On the basis of the displayed grayscale window of $[-2,\ 2]$ HU, no significant differences are observed inside the phantom, for both phantom modules and both potential functions. The summary values given in Table II confirm this visual impression.

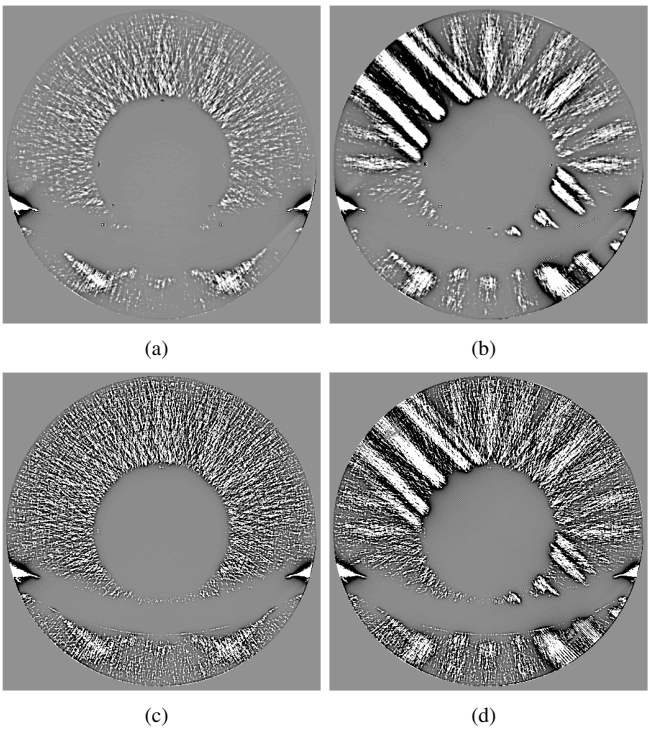


Fig. 2. Difference images for reconstruction with and without non-negativity constraint. These images are for *full view sampling* of module A (left) and module D (right). The applied potential function is either edge-preserving (top) or quadratic (bottom). Grayscale window: $[-2, 2]$ HU.

Convergence according to the RMSE value is presented in Fig. 3. The results for the different scanned modules are similar. For the first 100 iterations, there is no difference in the convergence behavior with and without non-negativity. In the later iterations, differences in the RMSE can be observed when using the quadratic potential, whereas the curves essentially remain the same when using the edge-preserving potential. For the quadratic potential, we thus see a gain in convergence speed when enforcing non-negativity; this gain is observed for a small improvement in an already small RMSE.

B. Sparse View Sampling of Projection Data

The difference images for the sparse view sampling are shown in Fig. 4. For the edge-preserving regularization, in Figs. 4(a) and (b), the difference inside the phantom is a noisy pattern with little to no structural information. For the quadratic regularization, in Figs. 4(c) and (d), the differences inside the phantom show a noisy pattern with enhancement of the edges with the sharp objects. Compared to the results with full view sampling, a larger grayscale window was required for display. This visual impression is confirmed by the summary values in Table II. The profiles through the edge of the air

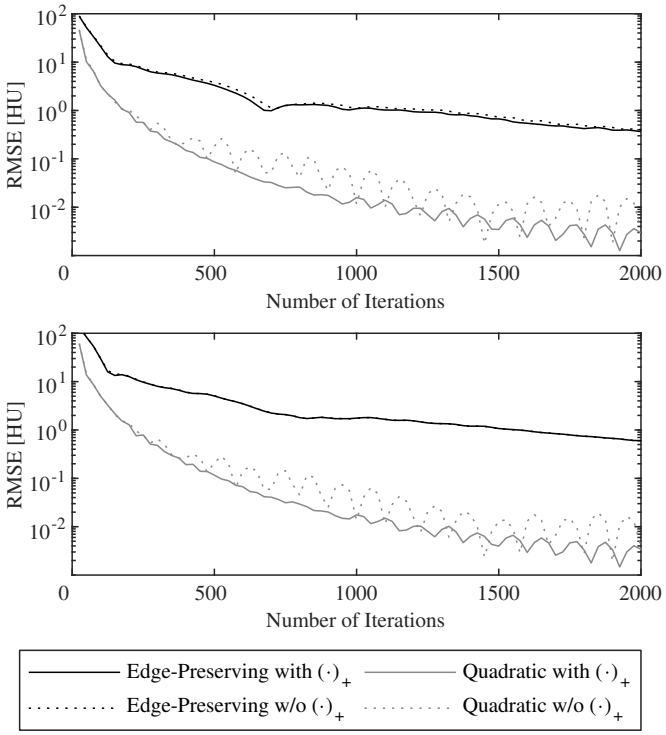


Fig. 3. Convergence behavior for reconstruction with *full view sampling* of module A (top) and module D (bottom). A solid (resp. dotted) line is used for reconstruction with and without the non-negativity constraint (resp.). The potential function is either edge-preserving (black) or quadratic (gray).

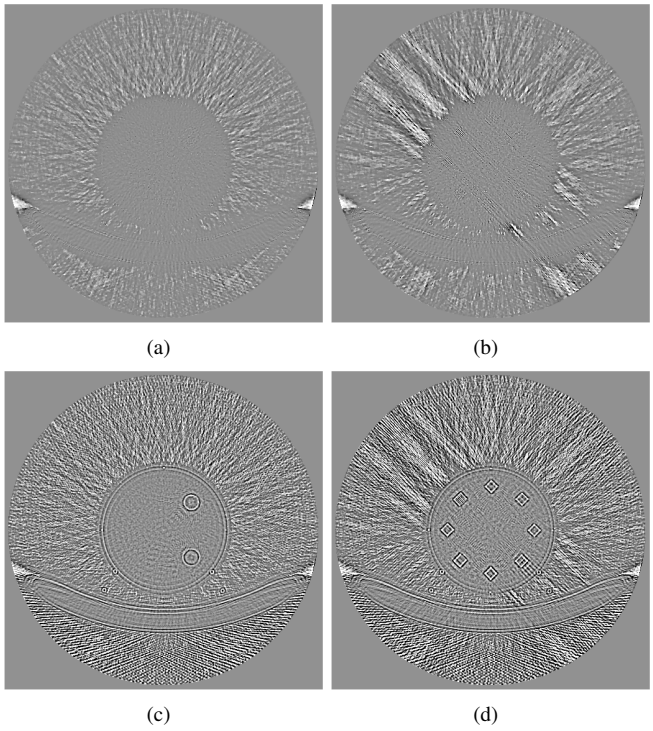


Fig. 4. Difference images for reconstruction with and without non-negativity constraint. These images are for *sparse view sampling* of module A (left) and module D (right). The applied potential function is either edge-preserving (top) or quadratic (bottom). Grayscale window: $[-20, 20]$ HU.

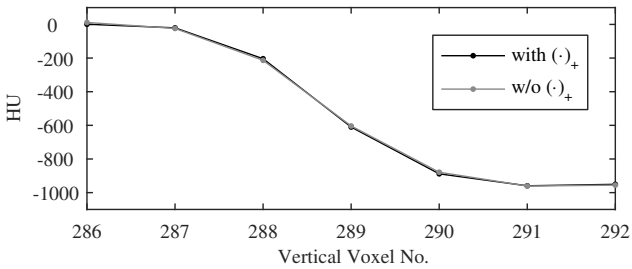


Fig. 5. Profiles through the edge of the air cylinder in module A for reconstruction with quadratic regularization and sparse view sampling. The profiles show the voxel value for the edge crossing from water to air, in HU.

cylinder in module A are compared in Fig. 5, where we see that the non-negativity constraint leads to a slightly sharper profile, though the difference is tiny and likely not significant for a human observer.

The convergence behavior is depicted in Fig. 6. It has the same characteristics as seen in Fig. 3, showing that the non-negativity constraint does not impact convergence speed for the edge-preserving regularization, but does impact it, at a similar rate, for the quadratic regularization.

TABLE II
MINIMUM, MAXIMUM, MEAN, AND STANDARD DEVIATION OF THE DIFFERENCE IMAGES. ALL VALUES ARE IN HU AND MEASURED INSIDE THE PHANTOM.

	Min.	Max.	Mean	SD
Full view sampling				
Module A, edge-preserving reg.	-1.50	1.85	-0.02	0.03
Module A, quadratic reg.	-0.20	0.05	-0.05	0.03
Module D, edge-preserving reg.	-1.61	1.36	-0.09	0.09
Module D, quadratic reg.	-0.79	0.35	-0.12	0.10
Sparse view sampling				
Module A, edge-preserving reg.	-13.02	11.29	-0.08	1.09
Module A, quadratic reg.	-29.17	26.45	-0.14	3.77
Module D, edge-preserving reg.	-42.90	38.98	-0.16	3.56
Module D, quadratic reg.	-47.39	54.89	-0.26	7.01

V. DISCUSSION AND CONCLUSIONS

In this work, we have reported results assessing the impact of the non-negativity constraint on image appearance and convergence speed under eight different scenarios based on real CT data. These scenarios were selected to challenge the potential usefulness of the non-negativity constraint.

Our observations in terms of image quality were as follows. For experiments with full view sampling, we observed that the use of the non-negativity constraint has only minimal impact on the image appearance inside the object. This was the case for both quadratic and edge-preserving regularization. The differences were on the order of 1-2 HU, and thus would likely have no effect on human observer performance.

For the more challenging problem of reconstruction from sparse view sampling, we observed more important differences: (i) enhancement of sharp edges, indicating a slight difference in resolution for sharp-contrast features, (ii) a noisy pattern inside the object. The difference in resolution is not likely to be significant for human observer performance but the noisy pattern might; this aspect requires further investigation.

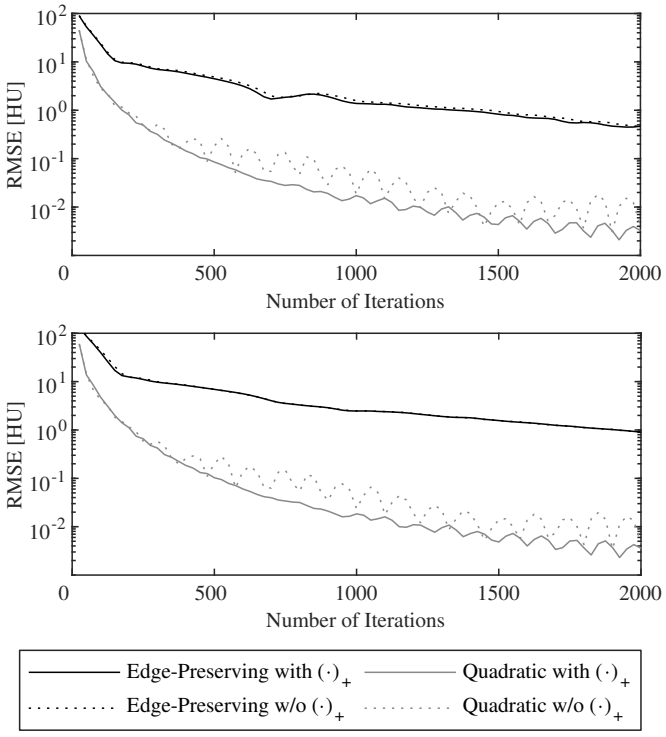


Fig. 6. Convergence behavior for reconstruction with *sparse view sampling* of module A (top) and module D (bottom). A solid (resp. dotted) line is used for reconstruction with and without the non-negativity constraint (resp.). The potential function is either edge-preserving (black) or quadratic (gray).

Also of importance is the observation that the differences are of a fairly smaller magnitude for reconstruction with edge-preserving regularization.

In terms of convergence speed, we only identified benefits in using the non-negativity constraint for reconstruction with the quadratic penalty. In this case, the same benefit was observed for both full and sparse view sampling. However, this benefit plays a role only for improving the convergence when the RMSE is already small (below 1 HU).

Overall, we conclude that the non-negativity constraint may not offer any benefit for conventional diagnostic CT imaging, but could possibly slightly help for reconstruction under challenging conditions like sparse view sampling. To further verify this conclusion, a wider range of objective functions and reconstruction algorithms should be examined. Also, the outcome of our experiments could be affected by the choice of the phantom. A more complex anthropomorphic object with several air cavities (e.g., as encountered in lung imaging of patients with COPD) should be the basis for further investigations.

From an algorithm viewpoint, it is important to know that there are situations where the non-negativity constraint has little impact on image quality, because, for these situations, the choice of the optimization algorithm would not be restricted to those that can handle the non-smooth indicator function for non-negative values. This can have an impact on reconstruction speed.

Disclaimer

The concepts and information presented in this paper are based on research and are not commercially available.

Acknowledgment

This project was partly supported by Siemens Healthcare GmbH and partly by the National Cancer Institute of the National Institutes of Health under R21CA211035.

References

[1] Vardhanabhuti, V., Riordan, R. D., Mitchell, G. R., Hyde, C., and Roobottom, C. A., "Image comparative assessment using iterative reconstructions: clinical comparison of low-dose abdominal/pelvic computed tomography between adaptive statistical, model-based iterative reconstructions and traditional filtered back projection in 65 patients," *Investigative radiology* **49**(4), 209–216 (2014).

[2] Ichikawa, Y., Kitagawa, K., Nagasawa, N., Murashima, S., and Sakuma, H., "CT of the chest with model-based, fully iterative reconstruction: comparison with adaptive statistical iterative reconstruction," *BMC medical imaging* **13**(1), 27 (2013).

[3] Desai, G. S., Uppot, R. N., Elaine, W. Y., Kambadakone, A. R., and Sahani, D. V., "Impact of iterative reconstruction on image quality and radiation dose in multidetector CT of large body size adults," *European radiology* **22**(8), 1631–1640 (2012).

[4] Pickhardt, P. J., Lubner, M. G., Kim, D. H., Tang, J., Ruma, J. A., del Rio, A. M., and Chen, G.-H., "Abdominal ct with model-based iterative reconstruction (MBIR): initial results of a prospective trial comparing ultralow-dose with standard-dose imaging," *American journal of roentgenology* **199**(6), 1266–1274 (2012).

[5] Thibault, J.-B., Sauer, K. D., Bouman, C. A., and Hsieh, J., "A three-dimensional statistical approach to improved image quality for multislice helical CT," *Medical physics* **34**(11), 4526–4544 (2007).

[6] Hahn, K., Schöndube, H., Stierstorfer, K., Hornegger, J., and Noo, F., "A comparison of linear interpolation models for iterative CT reconstruction," *Medical physics* **43**(12), 6455–6473 (2016).

[7] Haase, V., Griffith, A., Guo, Z., Oktay, M., Hahn, K., Schöndube, H., Stierstorfer, K., and Noo, F., "Penalized least-square CT reconstruction without and with statistical weights: effect on lesion detection performance with human observers," in [*Proceedings of the 14th International Meeting on Fully Three-dimensional Image Reconstruction in Radiology and Nuclear Medicine*], 831–835 (2017).

[8] Joseph, P. M., "An improved algorithm for reprojecting rays through pixel images," *IEEE transactions on medical imaging* **1**(3), 192–196 (1982).

[9] Beck, A. and Teboulle, M., "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM journal on imaging sciences* **2**(1), 183–202 (2009).

[10] Nesterov, Y., "A method for unconstrained convex minimization problem with the rate of convergence $\mathcal{O}(1/k^2)$," in [*Doklady AN USSR*], **269**, 543–547 (1983). (In Russian).

[11] Sidky, E. Y., Jørgensen, J. H., and Pan, X., "Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle–Pock algorithm," *Physics in Medicine & Biology* **57**(10), 3065 (2012).

Sparse-view Spectral CT Reconstruction Using Image Gradient ℓ_0–norm and Tensor Dictionary

Weiwen Wu, Yanbo Zhang, Qian Wang, Fenglin Liu, Peijun Chen and Hengyong Yu

***Abstract*—Spectral computed tomography (CT) has a great superiority in lesion detection, tissue characterization and material decomposition. To further explore its potential in clinical applications, we propose an improved tensor dictionary learning method for sparse-view low-dose spectral CT reconstruction with a constraint of image gradient ℓ_0-norm, which is named as ℓ_0TDL. The ℓ_0TDL method inherits the advantages of tensor dictionary learning (TDL) by employing the similarity of spectral CT images. On the other hand, by introducing an ℓ_0-norm constraint in gradient image domain, the proposed method emphasizes the spatial sparsity to overcome the weakness of TDL on preserving edge information. The results show that the proposed ℓ_0TDL method outperforms other competing methods.**

***Index Terms*— spectral computed tomography, image reconstruction, sparse-view, ℓ_0-norm of image gradient, tensor dictionary.**

I. Introduction

The state-of-the-art spectral CT scanner employs photon-counting detectors (PCDs) to record the energy of each individually incoming x-ray photon [1]. Thus, we can obtain the material decomposition maps from multiple projection datasets in different energy bins after a series of post-processing steps. The spectral CT has obtained tremendous successes in low-dose CT, contrast media imaging and K-edge imaging [2].

The spectral CT image reconstruction has made a great progress in recent years. Xu *et al*. considered spectral projection data from each channel as an independent traditional dataset and employed the total variation (TV) penalty to reconstruct interior spectral images [3]. Then, the dictionary learning technique was developed for conventional low-dose CT image reconstruction and spectral CT reconstruction [4]. To suppress the noise of global intensity and protect the image edge, a PRISM (prior rank, intensity, and sparsity model) was utilized for multi-energy CT reconstruction [5]. To exploit the correlations among all the dimensions simultaneously and edge-preserving/enhancement, Semerci *et al*. combined a tensor nuclear norm (TNN) with TV for multi-energy reconstruction [6]. Recently, a tensor dictionary learning (TDL) method was proposed by considering the similarity of spectral CT images of different energy channels [7]. Although the TDL algorithm has a relatively better performance in preserving fine structures, it is not good at preserving edge information. Besides, this method is not good at suppressing noise and reducing artifacts in the reconstructed images, either.

To overcome the aforementioned limitations of the TDL method, we combine the image gradient ℓ_0-norm minimization and the TDL technique to generate an ℓ_0TDL algorithm. The proposed ℓ_0TDL method has the following advantages: (a) introducing the image gradient ℓ_0-norm to encourage TDL-based method to recover fine structures and edge information; and (b) improving sparse ability on image spatial domain to suppress noise and reduce image artifacts.

II. Mathematic Model

1) ℓ_0TDL mathematic model

Because the image edges and fine structures are corrupted by severe noise in low-dose spectral CT, the TDL may fail to recover high-quality edge information from such an undersampling dataset. To achieve a better image with more accurate edge information and less noise, it is natural for us to combine the image gradient ℓ_0–norm minimization and TDL technology. As a result, we formulate an image reconstruction framework:

$$\arg\min_{\mathbf{X},\boldsymbol{\alpha}_r,\mathbf{m}_r}\left(\begin{array}{l}\sum_{s=1}^{S}\left\|\mathbf{A}\mathbf{x}_s-\mathbf{y}_s\right\|_2^2+\mu\sum_{s=1}^{S}\left\|\nabla\mathbf{x}_s\right\|_0\\+\lambda\left(\sum_r\left\|\mathbb{Z}_r\left(\mathbf{X}\right)-\mathbf{D}_m\times_4\mathbf{m}_r-\mathbf{D}\times_4\boldsymbol{\alpha}_r\right\|_F^2+\sum_r\kappa_r\left\|\boldsymbol{\alpha}_r\right\|_0\right)\end{array}\right). \quad (1)$$

Weiwen Wu is with the Department of Electrical and Computer Engineering (ECE), UMass Lowell (UML), Lowell, MA 01854 USA (email: weiwenwu12@gmail.com). He is also with the Key Lab of Optoelectronic Technology and Systems, Chongqing University, Chongqing, 400044 China.

Yanbo Zhang, Qian Wang and Hengyong Yu are with the Department of ECE, UML, Lowell, MA 01854 USA (email: yanbo_zhang@uml.edu;qian_wang@student.uml.edu;hengyong-yu@ieee.org).

Fenling Liu is with the Key Lab of Optoelectronic Technology and Systems. He is also with the Engineering Research Center of Industrial Computed Tomography Nondestructive Testing, Ministry of Education, Chongqing University, Chongqing, 400044 China (email: liufl@cqu.edu.cn).

Peijun Chen is with the Key Lab of Optoelectronic Technology and Systems, Chongqing University, Chongqing, 400044 China (email: pjchen52@gmail.com).

where $\mathbf{X}\in\Re^{I_1\times I_2\times S}$ and $\mathbf{Y}\in\Re^{J_1\times J_2\times S}$ are respectively the 3rd-order reconstructed image and projection data tensors, I_1 and I_2 are width and height of the reconstructed image, J_1 and J_2 present the numbers of detector and projection view, S is a number of energy channels, $\mathbf{x}_s$ and $\mathbf{y}_s$ are respectively the vectorized s^{th} image and projection, $\mathbf{A}$ is the system matrix which depends on the system scanning structure and calculation method, $\mathbf{m}_r$ presents the mean vector of each channel, the operator $\mathbb{Z}_r$ is used to extract r^{th} small tensor block ($N\times N\times S$) from $\mathbf{X}$ and $\boldsymbol{\alpha}_r\in\Re^K$ is the sparse representation coefficient of r^{th} tensor block. $\mathbf{D}=\left\{\mathbf{D}^{(k)}\right\}\in\Re^{N\times N\times S\times K}$ is the trained tensor dictionary and $\mathbf{D}_m=\left\{\mathbf{D}_m^{(k)}\right\}\in\Re^{N\times N\times S\times S}$ represents the mean removal process [4]. The parameter κ_r is a factor to modulate representation precision and the sparsity level, and λ is designed to balance the data fidelity term and the sparse representation regularization. The system matrix $\mathbf{A}$ from specified scanning configuration has a large impact on the parameter λ. $\|\nabla\mathbf{x}_s\|_0$ represents the ℓ_0-norm of image gradient and it can be denoted as

$$\|\nabla\mathbf{x}_s\|_0=\sum\nolimits_p\#\left\{p\,\middle|\,\left|\partial_x\mathbf{x}_s^p\right|+\left|\partial_y\mathbf{x}_s^p\right|\neq 0\right\}\ , \tag{2}$$

where $\#\{\cdot\}$ is a counting operator and p (p=1, 2, …, $I_1\times I_2$) index the location of $(i_1\ ,i_2)^{\text{th}}$ element in the image. The $\partial_x\mathbf{x}_s^p$ and $\partial_y\mathbf{x}_s^p$ represent $\left(\mathbf{x}_s(i_1,i_2)-\mathbf{x}_s(i_1-1,i_2)\right)$ and $\left(\mathbf{x}_s(i_1,i_2)-\mathbf{x}_s(i_1,i_2-1)\right)$, respectively.

2) Solution of the ℓ_0TDL

Because Eq. (1) contains three searched-for variables, we further divide it into three sub-problems:

$$\mathbf{X}^{n+1}=\arg\min_{\mathbf{X}}\left(\sum_{s=1}^{S}\|\mathbf{A}\mathbf{x}_s-\mathbf{y}_s\|_2^2+\mu\sum_{s=1}^{S}\|\nabla\mathbf{x}_s\|_0 +\lambda\sum_r\left\|\mathbb{Z}_r(\mathbf{X})-\mathbf{D}_m\times_4\mathbf{m}_r^n-\mathbf{D}\times_4\boldsymbol{\alpha}_r^n\right\|_F^2\right), \tag{3a}$$

$$\mathbf{m}_r^{n+1}=\arg\min_{\mathbf{m}_r}\left\|\mathbb{Z}_r\left(\mathbf{X}^{n+1}\right)-\mathbf{D}_m\times_4\mathbf{m}_r-\mathbf{D}\times_4\boldsymbol{\alpha}_r^n\right\|_F^2\ , \tag{3b}$$

$$\boldsymbol{\alpha}_r^{n+1}=\arg\min_{\boldsymbol{\alpha}_r}\ \left\|\mathbb{Z}_r\left(\mathbf{X}^{n+1}\right)-\mathbf{D}_m\times_4\mathbf{m}_r^{n+1}-\mathbf{D}\times_4\boldsymbol{\alpha}_r\right\|_F^2+\kappa_r\|\boldsymbol{\alpha}_r\|_0\ . \tag{3c}$$

Eqs. (3b) and (3c) can be easily solved by following the same steps in [7]. Eq. (3a) contains the ℓ_0-norm of image gradient and tensor dictionary based sparse representation, which is a non-convex and non-deterministic polynomial hard (NP-hard) problem. To solve this optimization problem effectively, we employ an alternating direction minimization method (ADMM). First, we introduce an auxiliary variable. $\mathbf{u}_S$. Eq. (3a) can be re-expressed as a constrained optimization model:

$$\arg\min_{\mathbf{X}}\left(\sum_{s=1}^{S}\|\mathbf{A}\mathbf{x}_s-\mathbf{y}_s\|_2^2+\lambda\sum_r\left\|\mathbb{Z}_r(\mathbf{X})-\mathbf{D}_m\times_4\mathbf{m}_r^n-\mathbf{D}\times_4\boldsymbol{\alpha}_r^n\right\|_F^2 +\mu\sum_{s=1}^{S}\|\nabla\mathbf{u}_s\|_0\ \right),\ \text{s.t. }\mathbf{u}_s=\mathbf{x}_s \tag{4}$$

where $\mathbf{u}_S$ is an auxiliary matrix in $\Re^{I_1\times I_2}$ for the s^{th} energy channel which is an element of tensor $\mathtt{U}$ in $\Re^{I_1\times I_2\times S}$. Thus, the scaled augmented Lagrangian function of problem (4) [8] can be written as

$$\arg\min_{\mathbf{X},\mathbf{U},\mathbf{T}}\left(\sum_{s=1}^{S}\|\mathbf{A}\mathbf{x}_s-\mathbf{y}_s\|_2^2+\lambda\sum_r\left\|\mathbb{Z}_r(\mathbf{X})-\mathbf{D}_m\times_4\mathbf{m}_r^n-\mathbf{D}\times_4\boldsymbol{\alpha}_r^n\right\|_F^2 +\mu\sum_{s=1}^{S}\|\nabla\mathbf{u}_s\|_0+\beta\sum_{s=1}^{S}\|\mathbf{x}_s-\mathbf{u}_s-\mathbf{t}_s\|_F^2\right), \tag{5}$$

where $\mathbf{t}_s$ is an auxiliary variable in $\Re^{I_1\times I_2}$ for the s^{th} energy channel which is a cell of tensor $\mathtt{T}$ in $\Re^{I_1\times I_2\times S}$. In fact, the ADMM method is utilized to iteratively and alternately solve Eq. (5) with respect to $\mathbf{X}$, $\mathtt{U}$ and $\mathbf{T}$. β is the Lagrangian multiplier for all energy channels. The ADMM algorithm of Eq. (5) contains the following three steps:

$$\mathbf{X}^{n+1}=\arg\min_{\mathbf{X}}\left(\begin{array}{l}\sum_{s=1}^{S}\|\mathbf{A}\mathbf{x}_s-\mathbf{y}_s\|_2^2++\beta\sum_{s=1}^{S}\|\mathbf{x}_s-\mathbf{u}_s^n-\mathbf{t}_s^n\|_F^2\\ +\lambda\sum_r\left\|\mathbb{Z}_r(\mathbf{X})-\mathbf{D}_m\times_4\mathbf{m}_r^n-\mathbf{D}\times_4\boldsymbol{\alpha}_r^n\right\|_F^2\end{array}\right), \tag{6}$$

$$\mathbf{U}^{n+1}=\arg\min_{\mathbf{U}}\mu\sum_{s=1}^{S}\|\nabla\mathbf{u}_s\|_0+\beta\sum_{s=1}^{S}\left\|\mathbf{x}_s^{n+1}-\mathbf{u}_s-\mathbf{t}_s^n\right\|_F^2\ , \tag{7}$$

$$\mathbf{T}^{n+1}=\mathbf{T}^n+\mathbf{U}^{n+1}-\mathbf{X}^{n+1}\ . \tag{8}$$

In this work, the solution of Eq. (6) is given by a separable surrogate method and its form can be expressed as:

$$\mathbf{X}_{i_1i_2s}^{n+1}=\mathbf{X}_{i_1i_2s}^{n}-\frac{\left[\mathbf{A}^{\mathrm{T}}\left(\mathbf{A}\mathbf{x}_s^n-\mathbf{y}_s\right)\right]_{i_1i_2}}{\left[\mathbf{A}^{\mathrm{T}}\mathbf{A}\right]_{i_1i_2}+\lambda\left[\sum_r\mathbb{Z}_r^{\mathrm{T}}\mathbb{Z}_r\right]_{i_1i_2s}+\beta}- \frac{\lambda\sum_r\mathbb{Z}_r^{\mathrm{T}}\left(\mathbb{Z}_r\left(\mathbf{X}^n\right)-\mathbf{D}_m\times_4\mathbf{m}_r^n-\mathbf{D}\times_4\boldsymbol{\alpha}_r^n\right)_{i_1i_2s}+\beta\left[\left(\mathbf{x}_s^n-\mathbf{u}_s^n-\mathbf{t}_s^n\right)\right]_{i_1i_2}}{\left[\mathbf{A}^{\mathrm{T}}\mathbf{A}\right]_{i_1i_2}+\lambda\left[\sum_r\mathbb{Z}_r^{\mathrm{T}}\mathbb{Z}_r\right]_{i_1i_2s}+\beta} \tag{9}$$

Eq. (7) includes the ℓ_0-norm minimization of image gradient, resulting in a non-convex and NP-hard problem. Fortunately, an approximate method was proposed in [9] to solve this problem. For the approximate method, another two auxiliary variables $\left(h_s^p,v_s^p\right)$ corresponding to the gradients $\left(\partial_x\mathbf{u}_s^p,\partial_y\mathbf{u}_s^p\right)$ are introduced. Therefore, Eq. (7) can be converted into the following problem

$$\mathbf{u}_s^{n+1}=\arg\min_{\mathbf{u}_s,h_s,v_s}\left(\sum\nolimits_p\left\|\left(\mathbf{x}_s^p\right)^{n+1}-\mathbf{u}_s^p-\left(\mathbf{t}_s^p\right)^n\right\|_F^2 +\lambda^*\Gamma\left(\mathbf{h}_s,\mathbf{v}_s\right)+\tau\left(\left(\partial x\left(\mathbf{u}_s^p\right)-h_s^p\right)^2+\left(\partial y\left(\mathbf{u}_s^p\right)-v_s^p\right)^2\right)\right), \tag{10}$$

where $\lambda^*=\mu/\beta$, $\Gamma\left(h_s,v_s\right)=\left\{p\left|h_s^p\right|+\left|v_s^p\right|\neq 0\right\}$, $\left(\mathbf{x}_s^p\right)^{n+1}$ and $\left(\mathbf{t}_s^p\right)^n$ are components of the p^{th} pixel in (n+1) and n iterations, and τ aims to balance the similarity

between $(\mathbf{h}_s, \mathbf{v}_s)$ and $(\partial_x u_s^p, \partial_y u_s^p)$. $\Gamma(\mathbf{h}_s, \mathbf{v}_s)$ can be expressed as:

$$\Gamma(\mathbf{h}_s, \mathbf{v}_s) = \sum_p \Lambda\left(\left|h_s^p\right| + \left|v_s^p\right|\right), \tag{11}$$

where

$$\Lambda\left(\left|h_s^p\right| + \left|v_s^p\right|\right) = \begin{cases} 1 & if \ \left|h_s^p\right| + \left|v_s^p\right| \neq 0 \\ 0 & \text{otherwise} \end{cases}. \tag{12}$$

Because Eq. (10) contains three variables, we also employ the ADMM scheme. Thus, the optimization problem of Eq. (12) can be divided into the following two steps:

$$\left\{\mathbf{h}_s^{m+1}, \mathbf{v}_s^{m+1}\right\} = \min_{\mathbf{h}_s, \mathbf{v}_s} \left(\begin{array}{l} \frac{\lambda^*}{\tau}\Gamma(\mathbf{h}_s, \mathbf{v}_s) + \\ \sum_p \left(\left(\partial x\left(\mathbf{u}_s^p\right)^n - h_s^p\right)^2 + \left(\partial y\left(\mathbf{u}_s^p\right)^n - v_s^p\right)^2 \right) \end{array} \right) \tag{13}$$

$$\mathbf{u}_s^{n+1} = \arg\min_{\mathbf{u}_s} \left(\begin{array}{l} \sum_p \left\| \left(\mathbf{x}_s^p\right)^{n+1} - \mathbf{u}_s^p + \left(\mathbf{t}_s^p\right)^n \right\|_F^2 + \\ \tau\left(\left(\partial x\left(\mathbf{u}_s^p\right) - \left(h_s^p\right)^{m+1}\right)^2 + \left(\partial_y\left(\mathbf{u}_s^p\right) - \left(v_s^p\right)^{m+1}\right)^2 \right) \end{array} \right) \tag{14}$$

Substituting Eq. (11) into (13), we have:

$$\left\{\mathbf{h}_s^{m+1}, \mathbf{v}_s^{m+1}\right\} = \min_{\mathbf{h}_s, \mathbf{v}_s} \sum_p \left(\begin{array}{l} \left(\partial x\left(\mathbf{u}_s^p\right)^n - h_s^p\right)^2 + \left(\partial y\left(\mathbf{u}_s^p\right)^n - v_s^p\right)^2 \\ + \frac{\lambda^*}{\tau}\Lambda\left(\left|h_s^p\right| + \left|v_s^p\right|\right) \end{array} \right). \tag{15}$$

Because each pixel $\mathbf{u}_s^p$ is considered independently in the iterative process of Eq. (15), we can separate each pixel so that it can be addressed independently.

For Eq. (15), the energy function can easily reach its minimum with the optimization criteria as follow:

$$\left(\left(h_s^p\right)^{m+1}, \left(v_s^p\right)^{m+1}\right) = \begin{cases} 1 & \left(\partial x\left(\mathbf{u}_s^p\right)^n\right)^2 + \left(\partial y\left(\mathbf{u}_s^p\right)^n\right)^2 \le \frac{\lambda^*}{\tau} \\ 0 & otherwise \end{cases}. \tag{16}$$

Now, let's consider Eq. (14). Because the function is quadratic, even a gradient descent method can make it shrink to a global minimum solution. Alternatively, we employ a fast analytic technique [10, 11] which integrates diagonalization derivative operators after Fast Fourier Transform (FFT). Therefore

$$\mathbf{u}_s^{n+1} = \mathbb{F}^{-1}\left\{ \frac{\mathbb{F}\left(\mathbf{x}_s^{n+1} - \mathbf{t}_s^n\right) + \tau\left(\left(\mathbb{F}^*(\partial x)\right)\left(\mathbb{F}^* h_s^{m+1}\right) + \left(\mathbb{F}^* \partial y\right)\left(\mathbb{F}^* v_s^{m+1}\right)\right)}{\mathbb{F}(1) + \tau\left(\left(\mathbb{F}^*(\partial x)\right)\left(\mathbb{F}^* \partial x\right) + \left(\mathbb{F}^*(\partial y)\right)\left(\mathbb{F}^* \partial y\right)\right)} \right\}$$

where $\mathbb{F}$ and $\mathbb{F}^*$ represent Fourier transform and conjugate Fourier transform, respectively.

III. EXPERIMENTS

A. Numerical simulations

We employ a simulated mouse thorax phantom, and 1.2% iodine contrast agent is injected to the blood circulation (Fig. 1) [12]. A 50KVp x-ray spectrum is utilized, and it is divided into 8 energy channels: [16, 22) keV, [22, 25) keV, [25, 28) keV, [28, 31) keV, [31, 34) keV, [34, 37) keV, [37, 41) keV, and [41, 50) keV [13]. The geometrical protocol is as follows: the distances from x-ray source to the PCD and rotation center are 180*mm* and 132*mm*, respectively; the PCD consists of 512 units and each of which is 0.1*mm*; the size of reconstructed channel image is 256×256×8 and each pixel covers an area of 0.15×0.15mm^2. We collect 640 projections for each full scan, and the default photon number is 5000 for each x-ray beam. For all the simulations, Poisson noises are superimposed on the obtained projections.

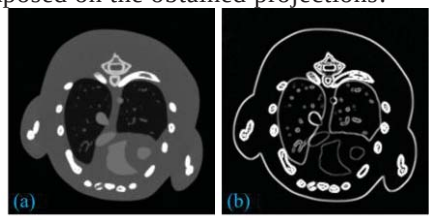


Fig 1. The mouse thorax phantom (left) and the gradient image (right).

To demonstrate the advantages of our proposed algorithm in recovering high-quality images from sparse-view projections, reconstructed images from 160 projections are shown in Figs. 2, along with the counterparts from other competing algorithms. To save space, we only show the reconstructed images for the first energy channel.

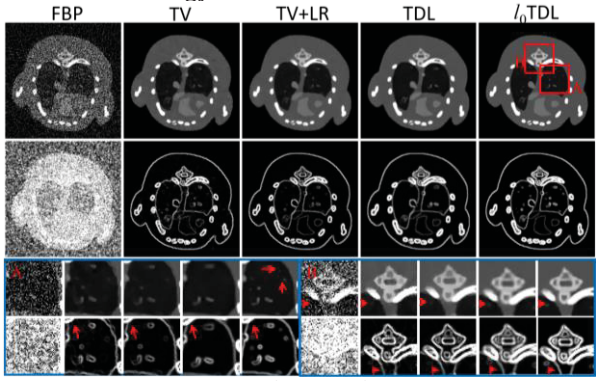


Fig. 2. Reconstruction results of the modified mouse thorax phantom. The first two rows are the reconstructed and gradient images from 160 projections and the last two rows are the magnified images of ROIs A and B. The display window of the reconstructed images is [0 3] cm-1 and the gradient images are in [0 0.8] cm-1.

In order to investigate the convergences of the proposed ℓ_0TDL and other competing methods, the convergence curves in terms of averaged RMSEs vs. iteration number are given in Fig. 3. Compared with other competing algorithms, the ℓ_0TDL method can converge to an optimized solution quickly with a smaller RMSE. From Fig. 3, for the ℓ_0TDL method, one can see that the RMSE decreases rapidly at first and then it is subsequently stable after 40 iterations.

B. Real data experiments

To demonstrate the advantages of the proposed ℓ_0TDL algorithm for low-dose spectral reconstruction in practical applications, an injected gold nanoparticles (GNP) mouse is scanned by a micro-CT system including one x-ray source and one photon counting detector[7]. The distance from the x-ray source to the PCD is set as 255 mm, the distance between the x-ray source and rotation axis is 158 mm, and 371 projections are uniformly acquired over a full scan. The energy spectrum of x-ray source is divided

into 13 channels via multiple scans. The PCD consists of 512 elements, each of which covers a length of 0.11 mm. The detector offset for this datasets was 1.0 mm. Fig. 4 shows some representative channel images reconstructed by the FBP algorithm from full projections, where each channel image is a matrix of 512×512 covering an area of 18.41×18.41mm^2. To investigate the sparse view reconstruction for this real dataset, Fig. 5 shows the results from 120 views. From the two magnified ROIs A and B, we can observe the proposed method has the best ability to protect image edges. From the magnified ROI B, we can see the x-ray CT image artifacts can be further reduced by our method. The image gradient ℓ_0–norm can penalize the sparsity in image gradient domain, which may result in reduced image artifacts.

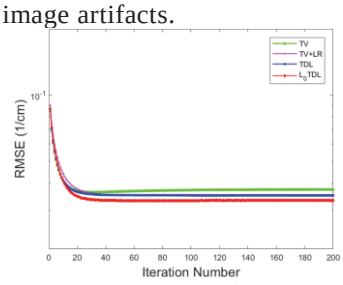


Fig. 3.The convergence curves in terms of average RMSE *vs.* iteration number.

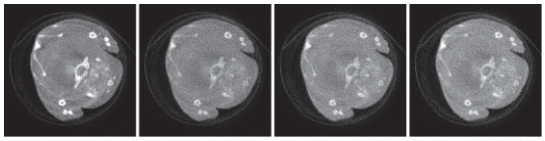


Fig.4. From the left to right columns, images are reconstructed for the 1st, 4th, 9th and 13th channels and the display window is [0, 0.8] cm^{-1}.

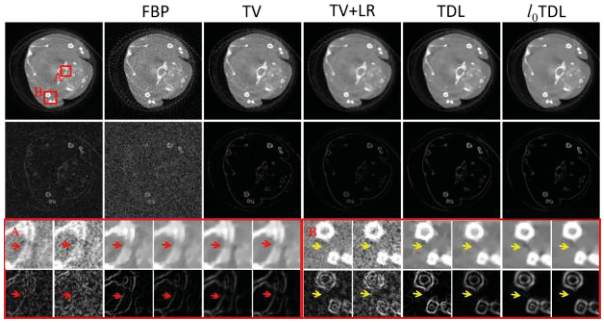


Fig. 5. Reconstructed images from the 120 projections. The display window of reconstructed and gradient images is [0, 0.8] and [0, 0.4].

IV. CONCLUSIONS

To penalize the image gradient to preserve image edge information from each channel and improve the anti-noising capability of the TDL method, we develop an ℓ_0TDL algorithm for sparse-view spectral CT reconstruction. By incorporating the image gradient ℓ_0–norm into the TDL based reconstruction framework, the image quality of channel reconstructions is dramatically improved, especially in the cases of sparse-view low-dose reconstruction. Both numerical simulations and realistic preclinical mouse study confirm that the proposed ℓ_0TDL algorithm outperforms the TV, TV+LR, and TDL methods.

In conclusion, we propose an ℓ_0TDL algorithm based on a global tensor dictionary and image gradient ℓ_0-norm for sparse-view low-dose spectral CT reconstruction. The developed ℓ_0TDL method can not only well maintain fine structures and image edges, but also reduce image artifacts especially in the areas of bone. This will be extremely meaningful for sparse-view spectral CT reconstruction.

ACKNOWLEDGEMENT

This work was supported in part by the National Natural Science Foundation of China (No. 61471070), National Instrumentation Program of China (No. 2013YQ030629), NIH/NIBIB U01 grant (EB017140) and China Scholarship Council (No. 201706050070). The authors would like to thank the MARS team in New Zealand for the realistic mouse datasets.

REFERENCES

[1] K. Taguchi and J. S. Iwanczyk, "Vision 20/20: Single photon counting x‐ray detectors in medical imaging," *Medical physics*, vol. 40, no. 10, 2013.

[2] W. Wuest, M. May, M. Saake, M. Brand, M. Uder, and M. Lell, "Low-dose CT of the paranasal sinuses: minimizing X-ray exposure with spectral shaping," *European radiology*, vol. 26, no. 11, pp. 4155-4161, 2016.

[3] Q. Xu *et al.*, "Image reconstruction for hybrid true-color micro-CT," *IEEE transactions on biomedical engineering*, vol. 59, no. 6, pp. 1711-1719, 2012.

[4] Q. Xu, H. Yu, X. Mou, L. Zhang, J. Hsieh, and G. Wang, "Low-dose X-ray CT reconstruction via dictionary learning," *IEEE Transactions on Medical Imaging*, vol. 31, no. 9, p. 1682, 2012.

[5] H. Gao, H. Yu, S. Osher, and G. Wang, "Multi-energy CT based on a prior rank, intensity and sparsity model (PRISM)," *Inverse problems*, vol. 27, no. 11, p. 115012, 2011.

[6] O. Semerci, N. Hao, M. E. Kilmer, and E. L. Miller, "Tensor-based formulation and nuclear norm regularization for multienergy computed tomography," *IEEE Transactions on Image Processing*, vol. 23, no. 4, pp. 1678-1693, 2014.

[7] Y. Zhang, X. Mou, G. Wang, and H. Yu, "Tensor-based dictionary learning for spectral CT reconstruction," *IEEE transactions on medical imaging*, vol. 36, no. 1, pp. 142-154, 2017.

[8] Q. Liu, D. Liang, Y. Song, J. Luo, Y. Zhu, and W. Li, "Augmented Lagrangian-Based Sparse Representation Method with Dictionary Updating for Image Deblurring," *Siam Journal on Imaging Science*, vol. 6, no. 3, pp. 1689-1718, 2013.

[9] L. Xu, C. Lu, Y. Xu, and J. Jia, "Image smoothing via L 0 gradient minimization," *Acm Transactions on Graphics*, vol. 30, no. 6, p. 174, 2011.

[10] M. Tao and J. Yang, "Alternating direction algorithms for total variation deconvolution in image reconstruction," *TR0918*, 2009.

[11] W. Yu, C. Wang, and M. Huang, "Edge-preserving reconstruction from sparse projections of limited-angle computed tomography using ℓ 0-regularized gradient prior," *Review of Scientific Instruments*, vol. 88, no. 4, p. 043703, 2017.

[12] W. P. Segars, B. M. Tsui, E. C. Frey, G. A. Johnson, and S. S. Berr, "Development of a 4-D digital mouse phantom for molecular imaging research," Molecular Imaging & Biology Mib the Official Publication of the Academy of Molecular Imaging, vol. 6, no. 3, p. 149, 2004.

[13] Y. Zhang, X. Mou, G. Wang, and H. Yu, "Tensor-based Dictionary Learning for Spectral CT Reconstruction," *IEEE Transactions on Medical Imaging*, vol. 36, no. 1, pp. 142-154, 2016.

Adaptive Non-Local Means Method for Denoising Basis Material Images from Dual-energy CT

Yuan Yuan, Yanbo Zhang, and Hengyong Yu

***Abstract*—We present an adaptive non-local means (ANLM) approach for image-domain material decomposition in low-dose dual-energy micro-CT. The key idea is to create a decomposition error distribution map and assign a smooth weight for a given pixel. This method is directly applied to the decomposed images for three basis materials: bone, soft tissue and gold in our application. We assume that bone and gold cannot co-exist in the same pixel and regroup these basis materials into two categories. For soft tissue, the proposed algorithm is implemented in a non-iterative version. For bone and gold, an iterative version is used and followed by a post-iteration process. Our results show that the proposed adaptive NLM outperforms other two common state-of-the-art denoising algorithms.**

***Keywords*— Image-domain material decomposition; Adaptive non-local mean; Edge detection.**

I. INTRODUCTION

Dual-energy CT has been widely used in various clinical applications, especially in contrast agent quantification and visualization [1]. Compared with standard single-energy CT, the dual-energy CT has a pair of linear attenuation coefficients at certain energies, which can be applied to decompose basis material images. This study will focus on the decomposed basis materials in image-domain. The basic idea of image-domain material decomposition is that the linear attenuation coefficients at the same position derived from the reconstructed images at low- and high-energy scans can be expressed as a linear combination of the pixel values of two basis materials. The direct matrix inversion method is easy for implementation but with amplified noise, and it is necessary to incorporate denoising techniques. There are several methods available for this purpose. Clark *et al.* [2] applied joint a bilateral filtration to reconstructed images before decomposition. Niu *et al.* [3] designed an iterative algorithm to denoise in the process of decomposition. Zhao *et al.* [1] established a new framework which can be directly applied to the decomposed images. However, for the low dose dual-energy material decomposition images, the aforementioned denoising methods still cannot provide satisfied results. To address this issue, we will develop a new adaptive non-local means (ANLM) method that can preserve details in material specific images and meanwhile suppress severe noises. The main idea is from a fact that the decomposed material images have additional edges compared to the reconstructed images. We demonstrate the advantages of the proposed approach in numerical simulations. The rest of this paper is organized as follow. The next section describes our methods. The third section presents the results. The last section discusses some related issues and concludes the paper.

II. MATHEMATIC MODEL

A. Dual-Energy Material Decomposition

We employ a conventional direct matrix inversion method to obtain the material specific images from dual energy micro-CT images. Material specific images are quantitative maps of the fractions of each material. In each material specific image, its pixel values are solutions of the following equations:

$$\begin{bmatrix} \mu_{1,H} & \mu_{2,H} & \mu_{3,H} \\ \mu_{1,L} & \mu_{2,L} & \mu_{3,L} \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix} = \begin{bmatrix} \mu_H \\ \mu_L \\ 1 \end{bmatrix}, \qquad (1)$$

where μ is the effective linear attenuation coefficients with the unit cm^{-1}, the subscripts 1-3 represent the sequence number of three basis materials, the subscripts H and L represent the high and low scanning energies, respectively. and f_1, f_2 and f_3 represent the corresponding unknown volumetric fractions. We assume that each pixel contains only basis materials and the sum of volumetric fraction is equal to 1. That is, we have 3 equations and 3 unknowns in Eq. (1). In fact, the direct matrix inversion method for material decomposition can magnify noise and degrade the accuracy. Here, we introduce a new method to reduce the noise and preserve basis material features.

* Corresponding Author (hengyong-yu@ieee.org).
1 Department of Physics and Applied Physics, University of Massachusetts Lowell, Lowell, MA 01854, USA.
2 Department of Electrical and Computer Engineering, University of Massachusetts Lowell, Lowell, MA 01854, USA.

B. Numerical Simulation Configuration

A realistic mouse thorax phantom generated by the Moby software (Fig.1 (a)) is used for numerical simulations. The diameter of this phantom is 30.0 mm, and 16mg/ml gold-based contrast agent is superimposed to the blood circulation system (Fig.1(d)). The qualities of decomposed images are significantly affected by scanning energies. In our previous study, the optimal decomposition results were obtained when the scanning energies are 40kVp and 65 kVp in dual-energy micro CT. In this study, an extrinsic Aluminum (Al) filtration is 1mm. A 2D fan-beam CT geometry is assumed with circular scan. The distance from source to rotation center is 351mm, the distance from source to detector is 541 mm, and 720 projections are uniformly acquired over a full scan range. The detector includes 1024 elements and each element is 0.078 mm. Poisson noise is superimposed into the projections. The total number of photons is 5×10^4 for each detector element and the ratio of the number of photons with high- and low- energy spectrum is 1.8. A conventional filtered back-projection algorithm is employed for image reconstruction.

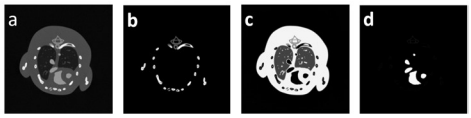


Fig.1. Moby mouse phantom for study. (a) is the 2D mouse thorax phantom. (b), (c) and (d) are the reference images of bone, soft tissue and gold, respectively.

C. Adaptive Non-Local Means

In the conventional non-local means (NLM) filter [4], the filtered value at a pixel i is calculated as a weighted average of all the pixels across the whole image. Given a 2D noisy image x={x(i)| $i \in I$ }, the NLM can be written as:

$$\mathrm{X}_{NLM}\ (i) = \sum_{j\in\Omega_i} \omega(i,j) * x(j), \qquad (2)$$

where Ω_i is referred to as a search window centered at the pixel i and it usually is a square neighborhood with a fixed size. The weight $\omega(i,j)$ represents the similarity between the pixels i and j, and it satisfies the constraint conditions $0\leq \omega(i,j) \leq 1$ and $\sum_j \omega(i,j) = 1$. The weight is constructed as follow:

$$\omega(i,j) = \frac{1}{Z(i)} exp\left(-\frac{\|x(\mathrm{N}_i)-x(\mathrm{N}_j)\|_2^2}{h^2}\right), \qquad (3)$$

where N_i is a square neighborhood of fixed size and centered at a pixel i . It is usually referred to as the similarity window. The similarity of two pixels i and j measured by Gaussian weighted Euclidean distance of pixel gray values $x(\mathrm{N}_i)$ and $x(\mathrm{N}_j)$ between the two neighborhoods N_i and N_j within the searching window Ω_i. $Z(i)$ is the normalizing factor to make sure $\sum_j \omega(i,j) = 1$. The parameter h controls the degree of filtering by adjusting the weight decay of the exponential function. In general, h directly affects the image quality in the non-local mean denoising. In our proposed adaptive NLM (ANLM) filter h is calculated by

$$h = h_{whole} * (1 + r * M_{error}), \qquad (4)$$

Here, h_{whole} is a smoothing parameter across the whole image. It is a small number to preserve as many detailed basis materials features as possible. M_{error} represents a decomposition error distribution map to indicate the error region where the filtering strength is enhanced. The method to construct decomposition error distribution map is shown in next section. r is the weight of M_{error} , and a proper value of r can balance the whole image smoothing and basis material feature enhancement. The ANLM filter can be reduced to the conventional NLM filter when r is equal to zero. In this study, we will apply the new ANLM filter directly to the original decomposed images. Based on the basis material structures, we label them as bone-like, soft-tissue-like and gold-like images. We assume that the most non-zero pixels in bone-like image have a high likelihood to belong to bone. It is the same for soft-tissue-like and gold-like images.

D. Decomposition Error Distribution Map

As the aforementioned, a decomposition error distribution map is used to adjust the smooth parameter for each pixel. It is called error map for short below. We create the error map from an original basis material image. Each pixel in the error map represents the probability of decomposition error of the corresponding pixel in the material images. We will first explain the method to construct error map for a soft-tissue-like image, then move to more complicated bone-like and gold-like images.

Soft tissue

To construct a proper error map for a soft-tissue-like image (Fig.2 (a)), we calculate a binarized image (Fig.2 (b))

$$x_{BW}(i) = \begin{cases} 0, & x(i) < \lambda \\ 1, & x(i) \geq \lambda' \end{cases} \qquad (5)$$

Here, we choose a threshold value $\lambda = 0.1$ to select the pixels where the fraction of soft tissue is above 0.1 and ignore the pixels whose values are below 0.1.

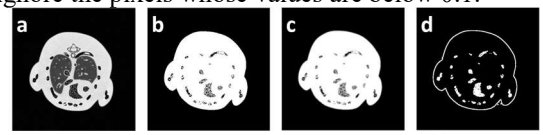


Fig. 2. Example if error map. (a) is soft-tissue-like image; (b) is binarized image; (c) is pixel density map; and (d) is decomposition error distribution map.

In a binarized image, the distribution of 'bright' pixel is visually nonuniform. We assume that the regions with high concentrated level of 'bright' pixels have high likelihood to belong to the basis material and the regions with low concentrated level have high likelihood to be errors. The concentrated level is defined as pixel density (Fig.2 (c)). The pixel density P(i) at a given pixel i is computed as the ratio between the pixel number in a certain category to the total number in a small region centered at the pixel i:

$$\mathrm{P}(i) = \left(\frac{\Delta \mathrm{N}}{\mathrm{N}}\right)_{Ai}, \qquad (6)$$

where N is the total number of pixels in a given small area, A_i centered at i and ΔN is the number of pixels in A_i that the value of the pixel is one. The radius of A_i is 2.

To obtain a qualified classification result, a proper threshold is needed to identify pixels likely belong to the decomposition error. Namely, a pixel likely belongs to the decomposition error when its corresponding pixel density value is below the threshold. Here, the threshold is empirically set as 0.7. The distribution of non-zero pixels in the error map (Fig.2 (d)). represent the region where decomposition error located. The value of each pixel shows the strengths of error. Last, a small mean filter window is used to smooth background without reducing the spatial resolution. The parameters used in the ANLM are listed as follow: 1) the radius of search window is 5; 2) the radius of similarity window is 2; 3) h_{whole}=0.02; 4) r=1.

Bone and Gold

Similarly, pixel density (X_d) maps for bone-like and gold-like images are obtained and shown in Fig 3 (c) and

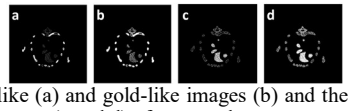


Fig.3. Bone-like (a) and gold-like images (b) and the corresponding pixel density maps (c and d) of mouse phantom.

However, it is difficult to set a threshold to distinguish the basis material features and the decomposed noises based on the pixel density. Here, we incorporate the Canny edge detection to construct the error map. The Canny edge detection is one of the most strictly defined methods by searching for a local maximum in its neighborhood in the direction of gradient at each pixel in a Gaussian filter smoothed image. As demonstrated in Fig. 4, the edge of Moby phantom CT image is obtained by detecting the edge of a composite image that are generated by averaging the high- and low-energy images with equal weighting.

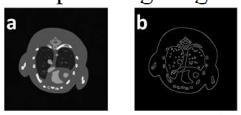


Fig.4 The composite image (a) and its edge (b).

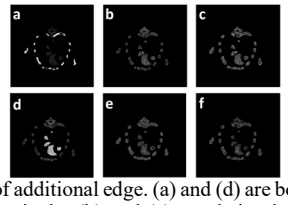


Fig.5 Illustration of additional edge. (a) and (d) are bone-like and gold-like images, respectively; (b) and (e) are their edge images; and (c) and (f) are additional edge maps.

Since all the edge detection results are easily affected by image noise, the non-preprocessed decomposition images, such as bone-like image and gold-like images in Fig. 5, have additional edges compared to the composite image. We assume that these additional edges are caused by noise. Although noise is randomly distributed, the region of additional edges could be identified. The additional edges of bone in Fig 5(c) are obtained by subtracting the edge of the composite image Fig. 4 (b) from the edge of non-preprocessed decomposed images Fig. 5 (b). The additional edges of gold Fig. 5 (f) are obtained using a similar method. The additional edge probability (Fig. 6 (a) and Fig. 6 (d)) are derived from the additional edges by applying Eq. (6). The noise contaminated probability (Fig. 6 (b) and Fig. 6 (e)) are calculated by $X_{nc} = X_a(1 - X_d)$, where X_{nc} represents the noise contaminated probability; X_a represents the additional edges probability; X_d represents pixel density.

To preserve fine details in the post-processed image, in Fig. 6 (b) and (e), some regions with low pixels density (red arrow) are not treated as noise contaminated ones. Fig. 6 (c) and (f) show the error maps which will be used in our ANLM filtration.

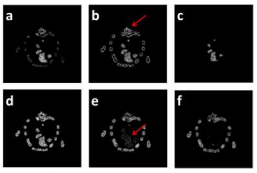


Fig. 6. Illustration of additional edge probability. (a) and (d) are for bone and gold. (b) and (e) are noise maps; and (c) and (f) are error maps.

In this section, the parameters used for ANLM filtration are listed as follow: 1) the radius of search window is 5; 2) the radius of similarity window is 2; 3) h_{whole}=0.01; 4) r=2.

As to the post-iterations images, the isolated bright pixels are removed since pixels that belong to basis material are also spatially connected. A quantitative pharmacokinetics study [5] showed that after gold nanoparticle injection, the highest tissue gold concentration was in the kidney followed by tumor, liver and muscle. That is, gold cannot propagate to bone. Therefore, we assume that bone and gold cannot co-exist in the same pixel. Here, we add an additional constraint:

1) M_i is a square neighborhood of fixed size and centered at a pixel i . $L_{B(G)}$ is the number of pixels in M_i which are greater than a certain threshold r. The subscript B (G) represent bone or gold.
2)

$$P_i = \begin{cases} belongs\ to\ bone & if\ L_B > L_G \\ belongs\ to\ gold & if\ L_G > L_B \end{cases}, \quad (8)$$

where P_i is the pixel i. Here, the threshold r=0.25 and the radius of M_i =11.

E. Comparison

For comparison, we denoise the same decomposed images using the conventional NLM and TV minimization methods. In the conventional NLM, the same parameters are used as ANLM In the TV minimization method, we select the optimal parameters. Further, the PSNRs (Peak signal-to-noise ratio) are used

to quantitatively evaluate the quality of denoised images.

III. Results

Fig.7 shows the denoised soft tissue images by different methods. The first image is reference. One can see that the ANLM outperforms other two methods. For the bone-like and gold-like images, to further suppress noise, the ANLM is implemented in an iterative fashion: the ANLM denoised material images are set as the input images of the next ANLM iteration. Here, the iteration number is 3.

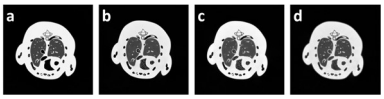


Fig.7 Denoised soft tissue images by different algorithms: NLM (b), TV (c) and ANLM (d) and (a) is reference.

Fig.8 shows the effect of post-iteration process. The two rows represent the images before and after the process. The last two columns show the magnified details in the red square. For each basis material, the display windows are the same. Fig.9 shows the post-processed bone and gold images by different methods. The iteration number of the NLM is also 3. Obviously, the ANLM significantly improves the decomposition accuracy of basis material images, especially in the red square region magnified in Fig.10. The quantitative PSNR (Fig.11) results confirm that the proposed ANLM outperforms the competing methods. Among the basis materials results, the PSNR of NNLM are always higher than NLM and TV.

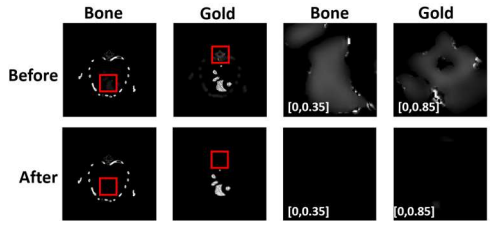


Fig.8 The bone and gold images before/after post-iteration process and the corresponding magnified details.

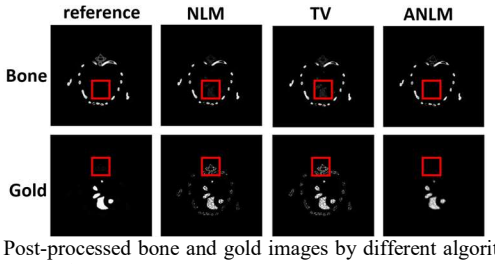


Fig.9 Post-processed bone and gold images by different algorithms: NLM, TV and ANLM.

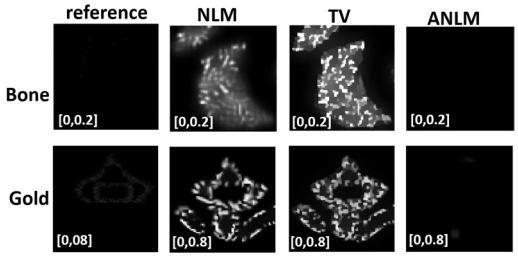


Fig.10. Magnified bone and gold images in the red squares in Fig.9.

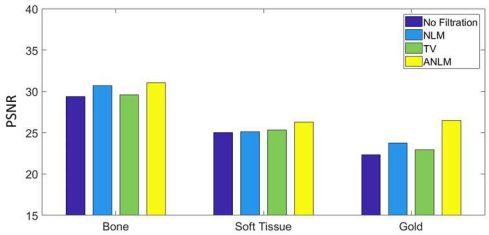


Fig. 11. The PSNR results of basis materials from different denoising methods.

IV. Discussion and Conclusion

This study introduces an ANLM algorithm for noise reduction of image-domain dual-energy material decomposition. It can be applied directly to the original decomposed images. The proposed method is intended to preserve basis material-like features and remove noise-like features for low SNR images. In this filter, the smooth factors include two parts: the first one is a constant for the whole image and the second one is a decomposition error distribution map which allows us to greatly denoise the selected regions. The decomposition error distribution map is especially created for enhancing weak material-like pixels and greatly suppressed noise-like pixels at the same time. In addition, we divide three decomposed images into two categories. For soft tissue-like image, a non-iterative version is used and for the other two decomposed images, an iterative version is applied to further denoise since the noise levels of these two basis material images are higher and the decomposition error distribution maps are more complicated. The quantitative PSNR values confirm that the proposed ANLM outperforms other algorithms, and it can significantly improve the decomposition accuracies for bone, soft tissue and gold. In the future, we will continue to perform a comprehensive evaluation for the ANLM including *in vivo* small animal experiments and stability characterization with respect to different noise levels.

Acknowledgment

This work was partly supported by an NIH/NIBIB grant (EB017140).

References

[1] W. Zhao *et al.*, "Using edge-preserving algorithm with non-local mean for significantly improved image-domain material decomposition in dual-energy CT," *Phys. Med. Biol.*, p. 1332.

[2] D. P. Clark, K. Ghaghada, E. J. Moding, D. G. Kirsch, and C. T. Badea, "In vivo characterization of tumor vasculature using iodine and gold nanoparticles and dual energy micro-CT," *Phys. Med. Biol*., vol. 58, no. 6, p. 1683, 2013.

[3] T. Niu, X. Dong, M. Petrongolo, and L. Zhu, "Iterative image-domain decomposition for dual-energy CT," *Med. Phys*., vol. 41, no. 4, p. 41901, 2014.

[4] Buades, Antoni Coll, Bartomeu Morel, Jean-Michel J.-M. , "A non-local algorithm for image denoising," *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol.2, p. 60-65.

[5] J. F. Hainfeld, D. N. Slatkin, T. M. Focella, and H. M. Smilowitz, "Gold nanoparticles: a new X-ray contrast agent," *Br. J. Radiol*., vol. 79, no. 939, pp.248-253,2006. pp.248–253, 2006.

DIRA-3D - a Model-based Dual-Energy Iterative Algorithm for Quantitative 3D Helical CT

Maria Magnusson, Magnus Björnfot, Åsa Carlsson Tedgren and Alexandr Malusek

***Abstract*—Quantitative dual-energy computed tomography may improve the accuracy of treatment planning in radiation therapy. Of special interest are algorithms that can estimate material composition of the imaged object. One example of such an algorithm is the 2D-based model-based iterative reconstruction algorithm DIRA. The aim of this work is to extend this algorithm to 3D so that it can be used with cone-beams and helical scanning. The new algorithm was implemented using the PI-method, which includes filtered backprojection. Its performance was tested using a mathematical phantom consisting of six ellipsoids. The algorithm substantially reduced the beam hardening and cone-beam artifacts after four iterations. If desired, the dual energy X-ray sources may be placed orthogonally and on different helices. The results indicate that the concepts used in DIRA can be extended to 3D geometries. More work is needed to fully test this algorithm.**

I. Introduction

Dual-energy computed tomography (DECT) may improve the accuracy of radiation treatment planning [1]. The field may especially benefit from algorithms providing quantitative information about CT numbers or material composition of the imaged object. The latter is provided by projection-based basis material decomposition (PBBMD) and image-based basis material decomposition (IBBMD) methods [2]. PBBMD methods, such as the well-known Alvarez-Macovski's base material decomposition (AMBMD) [3], require geometrically consistent projections and only two base materials can be assigned for the whole object. IBBMD methods can use any number of tissue-specific material doublets or triplets, nevertheless the methods suffer from beam hardening artefacts produced by classical reconstruction methods like the filtered backprojection (FBP). Better results are achieved by applying AMBDM before IBBMD [4] or by combining IBBMD methods with model-based iterative reconstruction algorithms, for example the Monoenergetic Plus algorithm [5] or DIRA [6] developed by the authors. The original DIRA algorithm uses two- and three-material decomposition to tissue-specific, user-defined material doublets and triplets for the characterization of the imaged object. FBP is used in the iterative loop for image reconstruction. For the sake of simplicity, the algorithm was implemented for 2D geometries only.

Clinical scanners shorten acquisition times by using helical scanning and multi-row detectors. Reconstruction is performed by FPB or iterative algorithms, where backprojection is part of the inner loop algorithm [7]. Examples of FPB algorithms for helical scanning are Siemens' Weighted FPB (WFBP) [8] and the PI-method [9], which are approximate only, and Katsevich's FBP [10], which is exact. Both the PI and Katsevich's methods discard projection data outside the Tam window, which makes them less favorable in clinical environment compared to the WFBP method.

The aim of this paper is to extend the 2D algorithm DIRA to 3D Helical geometry.

Maria Magnusson[1,2,3] (e-mail: maria.magnusson@liu.se), Åsa Carlsson Tedgren[2,3,4] and Alexandr Malusek[2,3] are with [1]Dept. of Electrical Engineering, [2]Dept. of Medical and Health Sciences, Radiation Physics, [3]Center for Medical Image Science and Visualization (CMIV), Linköping University, SE-581 83 Linköping, Sweden and [4]Dept. of Medical Radiation Physics and Nuclear Medicine, Karolinska University Hospital, SE-171 76 Stockholm, Sweden.

II. Theory

A. Material Decomposition

Two-material decomposition (2MD) assumes that a mixture is composed of two materials. 2MD determines mass fractions, w_1 and w_2, of the two base materials and the mass density, ρ, of the mixture. Three-material decomposition (3MD) assumes that a mixture consists of three base materials. 3MD determines mass fractions, w_1, w_2 and w_3, and the mass density of the mixture, which is calculated as $\rho^{-1} = \sum_{k=1}^{3} w_k/\rho_k$, where ρ_k and w_k are the mass density and mass fraction, respectively, of the kth material. In both methods, the mass fractions are normalized so that $\sum_k w_k = 1$. More information on the resulting systems of linear equations is in [6].

B. Forward projection generation

The logarithm of attenuation, here referred to as the polyenergetic projection P, is calculated as

$$P = \ln \frac{I_0}{I}, \tag{1}$$

where I and I_0 are the detector responses with and without, respectively, the imaged object. The intensity I_0, is calculated for an ideal energy integrating detector as

$$I_0 = \int_0^{E_{max}} EN(E)dE, \tag{2}$$

where E is the photon energy and $N(E)$ is the energy spectrum of photons emitted from the x-ray tube. The intensity I is calculated as

$$I = \int_0^{E_{max}} EN(E) \exp\left[-\int_L \mu(x,y,z,E)dl\right] dE, \tag{3}$$

where $\mu(x,y,z,E)$ is the LAC of pixel (x,y,z) at energy E and $\int_L dl$ is a line integral through the object. This calculation is time consuming since the line integrals must be calculated for all energies in the energy spectrum. The calculation of

projections change slightly when material decomposition is introduced. The line integrals are calculated through volume fractions of the different base materials. The intensity I is then

$$I = \int_0^{E_{max}} EN(E) \exp\left[-\sum_k \mu_k(E) l_k\right] dE, \quad (4)$$

where μ_k is the LAC of the kth base material and l_k is computed as

$$l_k = \rho_k^{-1} \int_L \rho(x,y,z) w_k(x,y,z) dl, \quad (5)$$

where ρ_k is the tabulated density of the kth base material, $\rho(x,y,z)$ is the calculated density in voxel (x,y,z) and $w_k(x,y,z)$ is the mass fraction of the kth material in voxel (x,y,z). The density ρ and the mass fractions w_k are obtained from the two-material or three-material decomposition.

A monoenergetic projection P_{E_i}, where E_i, $i = 1, 2$ is a specific energy, is calculated as

$$P_{E_i} = \sum_k \mu_k(E_i) l_k. \quad (6)$$

The line integral $\int_L dl$ can be calculated using Josephs method [11].

C. The PI-method, an FBP method for helical CT

1) Helical cone-beam geometry: Third generation CT-scanners typically use cone-beam projections and helical trajectory of the source, see Fig. 1, left. In this figure, κ is the cone-angle, γ is the fan-angle and P represents the pitch of the helix (the height of one complete helix turn, measured parallel to the axis of the helix). The maximum value of the cone-angle is called the cone-beam angle and is denoted κ_{max} and the maximum value of the fan-angle is called the fan-beam angle and is denoted γ_{max}. The s-axis starts at the point C, the cross-section between the central ray and the z-axis, and is aligned with the z-axis. The x-ray source is visualized as a dot which rotates around the z-axis in a helical trajectory. The detector is cylindrical and has its centre in the x-ray source.

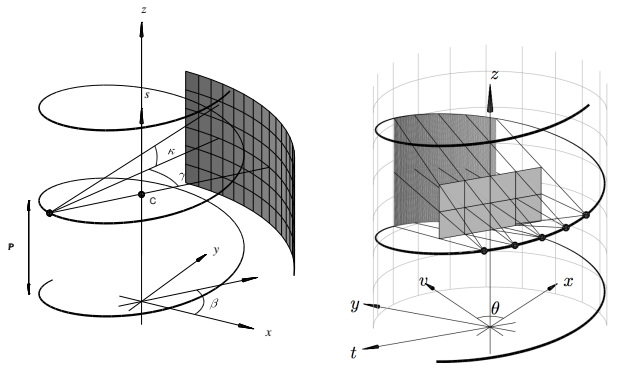


Fig. 1. Left: Helical CT geometry. Right: The PI-method geometry. The Tam window is located between two turns of the helix. The PI-detector is shown as the rectangular (t,s)-plane. Source: [9].

2) The algorithm of the PI-method: In the PI-method [9], cone-beam projections are rebinned to semi-parallel projections, see figure 1, right. The mathematical relation between a cone-beam projection, $p(\beta,\gamma,s)$, and the corresponding semi-parallel projection, $p^P(\theta,t,s)$, is

$$p^P(\theta,t,s) = p(\beta,\gamma,s), \quad (7)$$

where the relations between the parameters (θ,t,s) and (β,γ,s) are

$$\begin{aligned} \theta &= \beta + \gamma \\ t &= R\sin\gamma \end{aligned} \quad (8)$$

In (8), R is the radius of the helix. Projections to the (x,y) plane from the rays generating the semi-parallel projections are parallel line segments, see figure 1. In the (v,z) plane the rays generating the semi-parallel projections are fan-beam shaped.

The PI-method assumes that a virtual detector is positioned between the two turns of the helix (Tam window), see figure 1. Then in the (t,s) plane the rays define a rectangular area called the PI-detector. The s-axis has its origin at the center of the PI-detector and is aligned with the z-axis. It has been shown that the PI-detector gives complete and non-redundant projection data, i.e. no redundant data are present during reconstruction [9]. The PI-detector restricts the illumination interval of a voxel (x,y,z) to exactly 180°. A basic outline for the simplest version of the PI-method is:

1) Obtain cone-beam projections from the CT-scanner.
2) Perform rebinning, i.e. calculate semi-parallel projections from the cone-beam projections using (7) and (8).
3) Throw away all projection data outside the Tam window.
4) Pre-weight with $\cos\kappa$.
5) Perform rampfiltering along the rows of the virtual PI-detector.
6) Perform three-dimensional backprojection along the semi-parallel rays.

The PI-method is not an exact method and small artefacts can be seen in the reconstructed images. In general, the larger the cone-beam angle, the larger artefacts are observed [12].

D. DIRA-3D

DIRA-3D is an extension of the 2D algorithm DIRA presented in [6]. The algorithm is illustrated in Fig. 2 and performs the following steps:

1) Obtain measured semi-parallel polyenergetic projections, P_{M,U_i}, for two different tube voltages, $U_i, i = 1, 2$, giving P_{M,U_1} and P_{M,U_2}.
2) Reconstruct these projections so that the reconstructed LAC at energy E_i is $\mu_i = \Delta\mu_i + \mu_{P_i}$, where μ_{P_i} is the reconstructed LAC from mono-energetic calculated parallel projections at energy E_i and $\Delta\mu_i$ is the result of reconstructing $P_{M,U_i} - P_{U_i}$, where P_{U_i} are the calculated semi-parallel projections. For the first iteration, $P_{U_i} = P_{E_i} = 0$ and thus μ_i is the reconstruction of the P_{M,U_i} only.
3) Perform automatic threshold segmentation on μ_1 and μ_2.

4) Classify tissues using the material decomposition methods (section II-A).
5) Calculate polyenergetic projections P_{U_i} for semi-parallel geometry, see equations (1), (2) and (4).
6) Calculate monoenergetic projections P_{E_i} for parallel geometry, see equation (6).

Points 2-6 are repeated a predefined number of times.

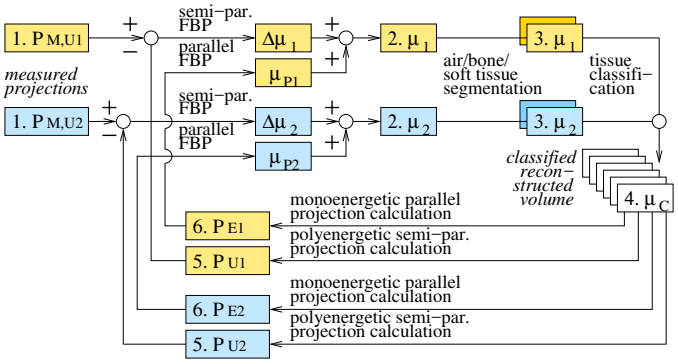


Fig. 2. A flowchart of the DIRA-3D algorithm.

To formulate DIRA in mathematical terms, set

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \mathbf{P}_{\mathbf{M,U}} = \begin{pmatrix} P_{M,U_1} \\ P_{M,U_2} \end{pmatrix}, \tag{9}$$

for the reconstructed images and the measured projections, respectively. Furthermore, denote the filtered backprojection operators $\mathcal{B}_{\mathrm{PI}}$ and $\mathcal{B}_{\mathrm{II}}$; the former represents the PI-method working with semi-parallel projections, the latter represents ordinary parallel FBP. Set

$$\mathcal{P}_U = \begin{pmatrix} \mathcal{P}_{U_1} \\ \mathcal{P}_{U_2} \end{pmatrix}, \quad \mathcal{P}_E = \begin{pmatrix} \mathcal{P}_{E_1} \\ \mathcal{P}_{E_2} \end{pmatrix}, \tag{10}$$

for the projection operator for polyenergetic projections and monoenergetic projections, respectively. These projection operators include the automatic tissue segmentation and classification. The linear attenuation coefficient $\boldsymbol{\mu}^{(i+1)}$ obtained at the $(i+1)$th iteration is

$$\boldsymbol{\mu}^{(i+1)} = \mathcal{B}_{\mathrm{PI}}(\mathbf{P}_{\mathbf{M,U}}) - \mathcal{B}_{\mathrm{PI}}\mathcal{P}_U(\boldsymbol{\mu}^{(i)}) + \mathcal{B}_{\mathrm{II}}\mathcal{P}_E(\boldsymbol{\mu}^{(i)}). \tag{11}$$

Ideally, the calculated polyenergetic projections $\mathcal{P}_U(\boldsymbol{\mu}^{(i)})$ converge towards the measured projection $\mathbf{P}_{\mathbf{M,U}}$. The $(i+1)$th iteration then gives $\boldsymbol{\mu}^{(i+1)} \approx \mathcal{B}_{\mathrm{II}}\mathcal{P}_E(\boldsymbol{\mu}^{(i)})$, which is the filtered backprojection result of the monoenergetic projections. The generation of monoenergetic projections followed by backprojection in DIRA serves as a regularization [6].

III. Methods

The new DIRA-3D algorithm was implemented according to the description in sections II-D and III-B. It was tested in the geometry described in section III-A.

A. Mathematical phantom and projection geometry

The phantom consisted of six ellipsoids, see figure 3. Two ellipsoids consisting of protein and water had their centers located at slice 24. Four ellipsoids consisting of adipose tissue, lipid, protein and compact bone had their centers located at slice 10. The scanning geometry for the semi-parallel projection generation is shown in figure 3. The parameters were: number of projection angles = 800 $(0-799)$, number of helix turns = 2, helical pitch = 32 voxels, cone beam angle = $\pm 3.138074°$, voxelsize: $\Delta x = \Delta y = 2.76$ mm, detector size: $N_t \times N_s = 192 \times 32$ pixels, detector pixel size of $\Delta t \times \Delta s = \Delta x/1.5 \times \Delta x/2$, resolution of the voxel array of $128 \times 128 \times 48$ voxels.

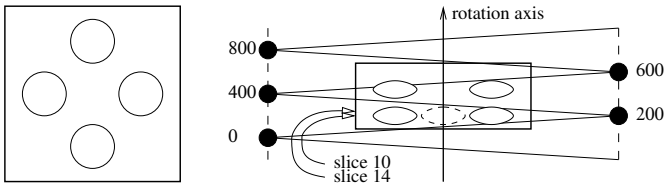


Fig. 3. Left: Phantom slice at $z = 10$. Right: The scanning geometry of the semi-parallel projection generation through the phantom showing projection number $0, 200, 400, 600$ and 800 and highlighting slice positions at $z = 10$ and $z = 14$.

B. Implementation details

The rebinning step in the PI-method was omitted for simplicity. The "measured" semi-parallel projections were simulated using the line integrals in (4) with $l_k = \int_L m_k(x,y,z)dl$, where m_k are the masks for the different ellipsoids of different materials. Energy spectra for the x-ray tube voltages of 80 and 140 kV with Sn filtration were used.

At each iteration, reconstructed volumes μ_1 and μ_2 were threshold segmentented into air, soft tissue and bone regions. Air was then decomposed into a (lipid, water) doublet, soft tissue was decomposed into a (lipid, protein, water) triplet and bone was decomposed into a (compact bone, bone marrow) doublet. As a consequence, the ellipsoid containing bone was decomposed into the (compact bone, bone marrow) doublet and the other ellipsoids were decomposed into the (lipid, protein, water) triplet.

The calculated semi-parallel polyenergetic projections were computed using equations (1), (2), (4) and the calculated parallel monoenergetic projections were were computed using (6). The generation of parallel projections with succeeding reconstruction was done slice by slice. The number of parallel detector elements was 128 and the number of projection angles was 400, with the angular interval $[0°, 180°)$.

C. Error calculation

The relative error of reconstructed LAC was estimated as $\delta(\bar{\mu}) = (\bar{\mu} - \mu_t)/\mu_t$, where μ_t is the tabulated value and $\bar{\mu}$ is the average of the calculated LAC in a spherical region of interest (ROI). The ROI was defined as a sphere with a radius of one third of the evaluated ellipsoid radius (in the x, y plane) and positioned in the center of the evaluated ellipsoid.

IV. Results

Figure 4 shows color-maps of reconstructed LAC in slice $z = 14$ for iterations 1, 2, 3 and 4 of DIRA-3D. The cone-beam artifact was decreased with the increasing number of

TABLE I

CALCULATED MASS FRACTIONS (IN %) OF LIPID, PROTEIN AND WATER IN ROIS R_1 AND R_2 CORRESPONDING TO ADIPOSE TISSUE AND MUSCLE, RESPECTIVELY, FOR THE FIRST (w^1) AND TENTH (w^{10}) ITERATION. THE TRUE VALUES (w^t) ARE ALSO LISTED.

	lipid			protein			water		
R	w^1	w^{10}	w^t	w^1	w^{10}	w^t	w^1	w^{10}	w^t
R1	93	70.0	70.1	2	2.9	2.9	5	27.1	27.0
R2	20	-12.3	-12.8	22	13.0	12.8	58	99.2	100.0

iterations. The convergence was very fast; visual differences between iterations 4 and 25 (not shown) were small. Streaks in the reconstructed images are most likely caused by partial volume artefacts.

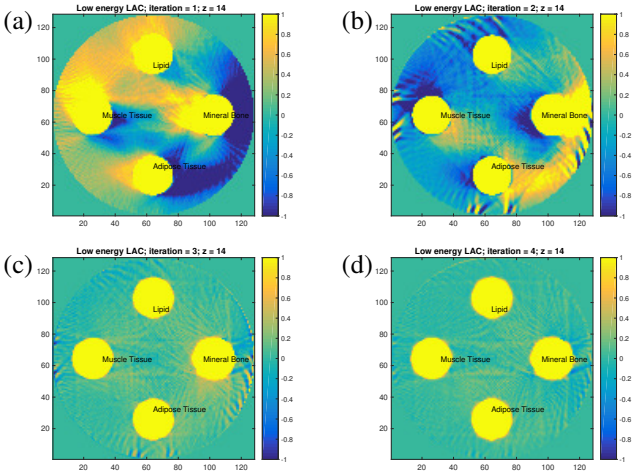


Fig. 4. Suppression of cone-beam artifacts in the surrounding air. Colormaps of reconstructed LAC (in m^{-1}) for iterations 1 (a), 2 (b), 3 (c) and 4 (d) of regular DIRA-3D reconstruction for the slice $z = 14$ and photon energy $E_1 = 50$ keV.

Figure 5 shows that the relative error of the LAC in the six different ellipsoids converged to a value close to zero already after 10 iterations (and even earlier for soft tissues).

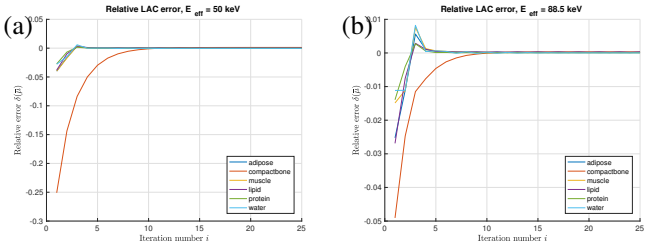


Fig. 5. The relative errors of the LAC in the six different ellipsoids as a function of the number of iterations for $E_1 = 50$ keV (a) and $E_2 = 88.5$ keV (b). The functions are plotted using a solid line for visual guidance only

Table I shows that was a good agreement between mass fractions of lipid, protein and water calculated in the ellipsoids containing adipose tissue (region R_1) and muscle (region R_2) and the corresponding true values when the algorithm converged. In the ellipsoid containing compact bone, the calculated mass fractions of the compact bone and bone marrow were $w^1 = 0.43$ and $w^1 = 0.57$, respectively, for the first iteration and $w^{10} = 1.0$ and $w^{10} = 0.0$ for the tenth iteration. The relative error of the density decreased from 0.0259 at iteration 1 to approximately 0 at iteration 10.

V. DISCUSSION AND CONCLUSION

We have presented DIRA-3D, a model-based iterative reconstruction algorithm that estimates material composition of the imaged object from DECT projections obtained in helical geometries. Specifically, the algorithm determines mass fractions of components of user defined material doublets and triplets. The current implementation was based on FBP and the PI-method, nevertheless the WFPB should also work.

The algorithm was evaluated using computer simulations with a simple phantom consisting of ellipsoids of different materials. In the studied geometry, the algorithm effectively removed beam-hardening and cone-beam artifacts and quickly converged. More work is needed to test the stability of the algorithm in the presence of quantum noise. The technique for handling the long object problem [13] was evaluated in a simulation experiment with good results (not presented here). Another successful experiment (not presented), was to place the two X-ray sources orthogonally and on different helices.

The presented result indicate that the concepts used in DIRA can be extended to cone-beams and helical scanning trajectories.

REFERENCES

[1] E. Bär, A. Lalonde, G. Royle, H.-M. Lu, and H. Bouchard. The potential of dual-energy CT to reduce proton beam range uncertainties. *Medical Physics*, 44(6):2332–2344, June 2017.

[2] B. J. Heismann, B. T. Schmidt, and T. Flohr. *Spectral Computed Tomography*. SPIE Press, Bellingham, Wash. (1000 20th St. Bellingham WA 98225-6705 USA), 2012.

[3] R. E. Alvarez and A. Macovski. Energy-selective reconstructions in X-ray computerised tomography. *Physics in Medicine and Biology*, 21:733–744, September 1976.

[4] P. R. S. Mendonca, P. Lamb, and D. V. Sahani. A Flexible Method for Multi-Material Decomposition of Dual-Energy CT Images. *IEEE Transactions on Medical Imaging*, 33(1):99–116, 2014.

[5] P. Wohlfahrt, C. Möhler, V. Hietschold, S. Menkel, S. Greilich, M. Krause, M. Baumann, W. Enghardt, and C. Richter. Clinical Implementation of Dual-energy CT for Proton Treatment Planning on Pseudo-monoenergetic CT scans. *International Journal of Radiation Oncology*Biology*Physics*, 97(2):427–434, February 2017.

[6] A. Malusek, M. Magnusson, M. Sandborg, and G. Alm Carlsson. A model-based iterative reconstruction algorithm DIRA using patient-specific tissue classification via DECT for improved quantitative CT in dose planning. *Medical Physics*, 44(6):2345–2357, 2017.

[7] M. Beister, D. Kolditz, and W. A. Kalender. Iterative reconstruction methods in X-ray CT. *Physica Medica*, 28(2):94–108, 2012.

[8] K. Stierstorfer, A. Rauscher, J. Boese, H. Bruder, S. Schaller, and T. Flohr. Weighted FBP—a simple approximate 3D FBP algorithm for multislice spiral CT with good dose usage for arbitrary pitch. *Physics in Medicine and Biology*, 49(11):2209–2218, 2004.

[9] H. Turbell. *Cone-Beam Reconstruction Using Filtered Backprojection*. Number 672 in Linköping studies in science and technology Dissertations. Univ, Linköping, 2001. OCLC: 248449210.

[10] A. Katsevich. Analysis of an exact inversion algorithm for spiral cone-beam CT. *Physics in medicine and biology*, 47(15):2583, 2002.

[11] P. M. Joseph. An improved algorithm for reprojecting rays through pixel images. *IEEE transactions on medical imaging*, 1(3):192–196, 1982.

[12] J. Sunnegårdh and P.-E. Danielsson. Regularized iterative weighted filtered backprojection for helical cone-beam CT: Regularized iterative filtered backprojection for cone-beam CT. *Medical Physics*, 35(9):4173–4185, 2008.

[13] M. Magnusson, P.-E. Danielsson, and J. Sunnegårdh. Handling of long objects in iterative improvement of nonexact reconstruction in helical cone-beam CT. *IEEE Transactions on Medical Imaging*, 25(7):935–940, 2006.

Physical Constraints for Beam Hardening Reduction using Polynomial Models

Tobias Würfl, Nicole Maaß, Frank Dennerlein, Anré Aichert and Andreas K. Maier

***Abstract*—Reconstruction algorithms for X-ray computed tomography typically assume a monochromatic X-ray beam and an energy independent attenuation coefficient of the materials along the ray. However, the attenuation coefficient of every material depends on energy, which leads to beam hardening artifacts in the reconstructed images. Recently reference-free algorithms for mono-material beam hardening artifact reduction based on the epipolar consistency condition have been introduced. These and reference-based algorithms apply a univariate polynomial model to the measured intensities prior to reconstruction. However consistency conditions reflect all sources of measurement errors. Other sources of inconsistency, notably truncation, may impact the model fitting and lead to low-quality reconstructions in spite of higher consistency. This work aims at avoiding such problems by imposing physically plausible constraints on the compensation functions. We introduce two necessary constraints on compensation functions namely monotonicity and convexity over the range of observations. Subsequently, we reformulate the optimization problem of polynomial models to yield only solutions obeying these constraints. Our formulation presents the advantage of being able to fit exactly all those functions, therefore not discarding plausible solutions. We show that this problem, despite being non-convex in the general case, is convex for the special case of polynomials of degree three. A measured data experiment is presented to demonstrate the effectiveness of our method.**

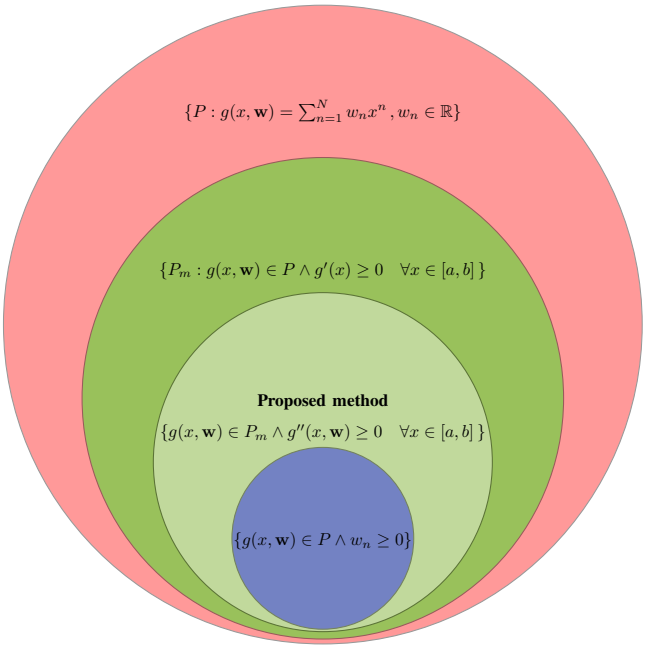


Figure 1: Venn diagramm of the different constraints. $[a, b]$ denotes the range of measured values of q.

I. Introduction

The combination of the polychromatic spectrum of X-ray tubes and the energy dependence of the linear attenuation coefficient causes a common problem in X-ray computed tomography known as beam hardening. This typically degrades image quality by introducing artifacts such as negative regions, cupping and streaks in reconstructions [1].

Conventionally the effect of beam hardening is compensated by a combination of software approaches which use reference measurements and methods optimizing the effective X-ray spectrum. An important distinction between software approaches is whether they assume a mono-material or a multi-material model. A computationally efficient mono-material method was presented by Kachelrieß et al. [2]. The approach introduces a polynomial compensation model and constructs a linear optimization problem to estimate its parameters using a reference measurement.

However, beam hardening introduces inconsistency in raw projection data which can be used to estimate parameters, even without any reference. New consistency conditions have been introduced for cone-beam data by Clackdoyle et al. [3], Lesaint et al. [4] and Debbeler et al. [5]. Aichert et al. provided an efficient flexible formulation of [5] in terms of epipolar geometry known as the epipolar consistency condition [6].

Recently, two reference-free beam hardening reduction algorithms based on the epipolar consistency condition and a univariate polynomal model have been presented by Abdurahman et al. [7] and Würfl et al. [8]. The approach by Würfl et al. uses the linearity of the Radon operator to speed up the algorithm significantly by reformulating the optimization problem on the Radon intermediate function. Additionally, the authors propose to improve robustness to other sources of inconsistency by requiring the coefficients of the polynomials to be non-negative. This non-negativity constraint on the coefficients is motivated by the observation that it is a sufficient but not necessary condition for a polynomial to be monotonously increasing.

We show that the requirement on the model functions to be monotonously increasing can be restricted further by considering the physics of X-ray attenuation by additionally requiring the functions to be convex.

This is implicitly satisfied by requiring non-negativity of

T. Würfl and A. K. Maier are with the Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany.
N. Maaß and F. Dennerlein are with Siemens Healthcare GmbH, Erlangen, Germany.

the coefficients. However non-negativity of the coefficients is too restrictive in the sense that it prohibits many physically plausible solutions by not being a necessary condition. We illustrate this situation in Fig. 1. In this work, we present a new parametrization of polynomial functions which is necessary and sufficient for monotonously increasing polynomials with a monotonously increasing derivative over the range of interest. This is also depicted in Fig. 1 as the set of our proposed formulation is equal to the set of our proposed constraint. Despite the fact that the original optimization problem is convex the domain of the new parametrization is in general not convex which renders optimization non-convex. However we are able to show convexity for the special, practically relevant case of a polynomial of degree three.

II. Methodology

In section II-A we present physical constraints on our model function. We construct a parametrization of polynomials obeying these constraints in section II-B. Subsequently we discuss optimization of this new parametrization in section II-C.

A. Physical constraints on beam hardening reduction models

The log attenuation along a line in X-ray imaging is given as:

$$q(L) = -\ln \int S(L,E) e^{-\int_0^\infty \mu(E,\mathbf{s}+\lambda\mathbf{l})d\lambda} dE\,, \qquad (1)$$

where $S(E)$ is the normalized spectrum over energy E on the line of integration L parametrized as $\mathbf{s}+\lambda\mathbf{l}$. Here $\mathbf{s}$ denotes the source position and $\mathbf{l}$ the direction, while $\mu(E,\mathbf{r})$ denotes the spatial distribution at position $\mathbf{r}$ of the energy-dependent attenuation values along a line L. It is common to assume we can decompose the energy dependence from the spatial dependence. This allows us to reformulate Eq. (1) to:

$$q(\mathbf{r}) = -\ln \int S(E) e^{-p(\mathbf{r})\psi(E)} dE\,, \qquad (2)$$

where $p(\mathbf{r})$ denotes the mono-chromatic line-integral at position $\mathbf{r}$ at some effective energy, while $\psi(E)$ denotes the energy dependence. In order to obtain p from measurements q, the task is now to find the inverse to this function, which we will denote as f. This is depicted in Fig. 2.

Physically plausible model functions have to be monotonous and convex giving rise to the requirement:

$$f'(q) > 0 \quad \wedge \quad f''(q) > 0 \quad \forall\, q \in [0, q_{max}]\,. \qquad (3)$$

B. Parametrization

We introduce a new parametrization of our polynomial by extending a recently presented monotonic parametrization by Murray et al. [9]. The goal of their method is to fit a polynomial:

$$g(x,\mathbf{w}) = w_0 + w_1 x + \cdots + w_d x^d\,, \qquad (4)$$

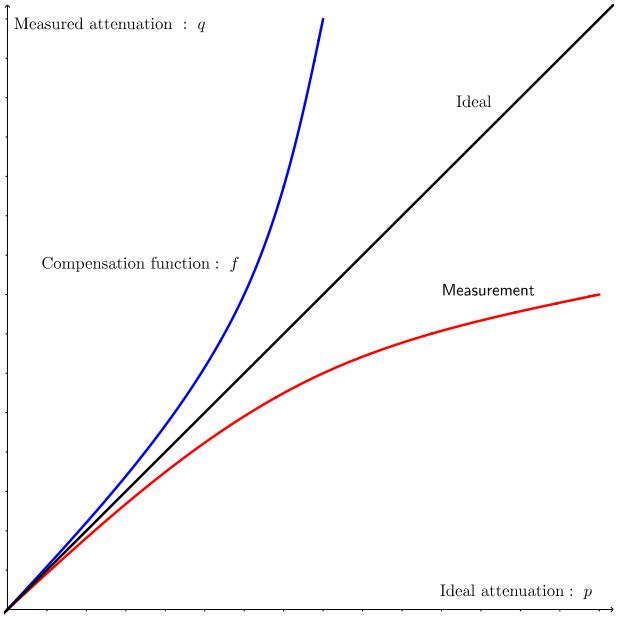


Figure 2: Illustration of the beam hardening effect.

where d denotes the degree of the polynomial, subject to the constraint of being monotonous over a range $[a,b]$. They use the fact that a polynomial of degree $d = 2k$ is positive on $[a,b]$ if and only if it can be written as

$$\hat{g}(x,\hat{\mathbf{w}}_1,\hat{\mathbf{w}}_2) = \hat{g}_1(x,\hat{\mathbf{w}}_1)^2 + (x-a)(b-x)\hat{g}_2(x,\hat{\mathbf{w}}_2)^2\,, \quad (5)$$

where k is a positive integer denoting the degrees of $\hat{g}_1$ and $\hat{g}_2$. If instead the degree is $d = 2k+1$ we have

$$\hat{g}(x,\hat{\mathbf{w}}_1,\hat{\mathbf{w}}_2) = (x-a)\hat{g}_1(x,\hat{\mathbf{w}}_1)^2 + (b-x)\hat{g}_2(x,\hat{\mathbf{w}}_2)^2\,. \quad (6)$$

Integration over this non-negative polynomial $\hat{g}(x,\hat{\mathbf{w}}_1,\hat{\mathbf{w}}_2)$ yields:

$$g(x,\hat{\mathbf{w}}) = \delta + \alpha \int_0^x \hat{g}(u,\hat{\mathbf{w}}_1,\hat{\mathbf{w}}_2)\mathrm{d}u\,. \qquad (7)$$

Note that all monotonic polynomials can be constructed in this manner. Their method interprets the parameters of polynomials $\hat{g}_1(x,\hat{\mathbf{w}}_1)$ and $\hat{g}_2(x,\hat{\mathbf{w}}_2)$ together with δ as a set of new parameters:

$$\hat{\mathbf{w}} = (\hat{\mathbf{w}}_1, \hat{\mathbf{w}}_2)\,,$$

making them a reparametrization of $g(x,\mathbf{w})$. By writing polynomial multiplication and addition as convolution and addition of coefficient vectors, an intermediate polynomial with coefficients $\tilde{w}$ representing a non-negative polynomial of degree $d-1$ from $\hat{\mathbf{w}}$ can be calculated. If we e.g. pick an uneven degree:

$$\tilde{\mathbf{w}} = (-a,1)^T * (\hat{\mathbf{w}}_1 * \hat{\mathbf{w}}_1) + (b,-1)^T * (\hat{\mathbf{w}}_2 * \hat{\mathbf{w}}_2)\,,$$

where $*$ denotes convolution. An integral over $\tilde{\mathbf{w}}$ can be computed according to:

$$\mathbf{w} = \left(\delta, \alpha\tilde{w}_0, \alpha\frac{\tilde{w}_1}{2}, \cdots, \alpha\frac{\tilde{w}_{d-1}}{d}\right)^T\,, \qquad (8)$$

where α controls if it is monotonously increasing or decreasing, i.e. $\alpha = \pm 1$.

We set $\alpha = 1$, since we only need increasing functions. Additionally we require the second derivative to be non-negative to satisfy Eq. (3). To this end we integrate a second time:

$$f(x) = \int_0^x \delta + \int_0^u \hat{g}(v, \hat{\mathbf{w}}) \mathrm{d}v \, \mathrm{d}u \,. \tag{9}$$

We set the constant coefficient arising from the second integration to zero because we expect zero attenuation for zero traversed material. This new parametrization restricts $f(x)$ to be convex but does not enforce monotonicity. Because the first integration yields a strictly positive polynomial over the range of interest we only need to constrain the parameter δ to be non-negative. This has a straightforward interpretation since δ is simply the slope at $x = 0$. A negative slope here cannot yield a sensible compensation polynomial. Because we started from a necessary and sufficient condition of non-negativity this new parametrization includes all polynomials which meet our physical requirements of Eq. (3).

C. Optimization of the new parametrization

The new parametrization can be incorporated into any scheme estimating a polynomial compensation model from measurements using a general non-linear optimizer. The parameters of the optimization are reparametrized and the cost function is evaluated using the intermediate weights.

An important distinction of optimization problems is whether they are convex and so any initial value will lead to the same unique global minimum. When applying the new parametrization this is in general not the case [9]. The problem stems from the fact that the domain of all monotonic polynomials $f(x)$ is not convex. Therefore, even though the unconstrained problem is convex, an optimizer can get stuck in local minima which are on the boundaries of the restricted domain.

In practice, a degree of $d = 3$ is often found to be sufficient for measured data. If we restrict our attention to this special case, we can investigate the function $\mathbf{m}_3(\delta, \hat{\mathbf{w}})$ which maps the optimization parameters $\delta, \hat{\mathbf{w}}_1, \hat{\mathbf{w}}_2$ to the parameters of our polynomial model. In this case $\hat{\mathbf{w}}_1 = \hat{w}_1$ and $\hat{\mathbf{w}}_2 = \hat{w}_2$ and thus:

$$\mathbf{w} = \mathbf{m}_3(\delta, \hat{\mathbf{w}}) = \begin{pmatrix} 0 \\ \delta \\ \frac{1}{2}\left(b\hat{w}_2^2 - a\hat{w}_1^2\right) \\ \frac{1}{6}\left(\hat{w}_1^2 - \hat{w}_2^2\right) \end{pmatrix} . \tag{10}$$

Since δ is constant, the shape of the domain depends only on the intermediate weights $\hat{w}_1$ and $\hat{w}_2$. We visualize this domain in Fig. 3. Examining the mapping in Eq. (10) we can state that the square-function maps every value in the four quadrants of $\hat{\mathbf{w}}$ to the same values in $\mathbf{w}$. In addition, we can determine that the boundaries of this domain are characterized by lines produced when $\hat{w}_1 = 0$ or $\hat{w}_2 = 0$ respectively. The lines are explicitly given as the left boundary: $\begin{pmatrix} \frac{1}{2} \\ -\frac{1}{6} \end{pmatrix} \hat{w}_2^2$ and the lower boundary: $\begin{pmatrix} -\frac{a}{2} \\ \frac{1}{6} \end{pmatrix} \hat{w}_1^2$. Because the expression of the boundaries are lines which can also be seen graphically in Fig. 3 the domain of optimization is actually convex for the special case of a polynomial of degree three. This implies that the whole optimization problem is convex.

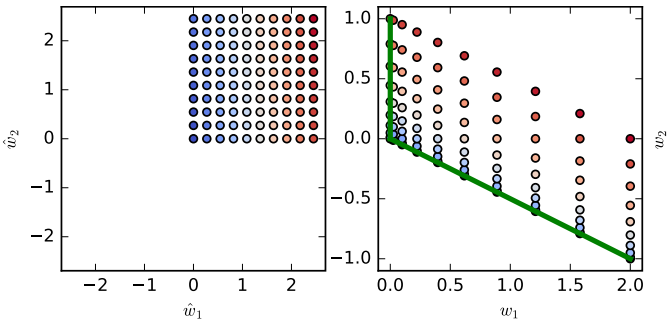


Figure 3: Visualization: $\mathbf{m}_3(\delta, \hat{\mathbf{w}})$ in relevant dimensions. The green lines form the treshold of the convex domain. $a = 0\,, b = \frac{4}{6}\,, \hat{w}_1 = \left[0, \sqrt{6}\right]\,, \hat{w}_2 = \left[0, \sqrt{6}\right]$

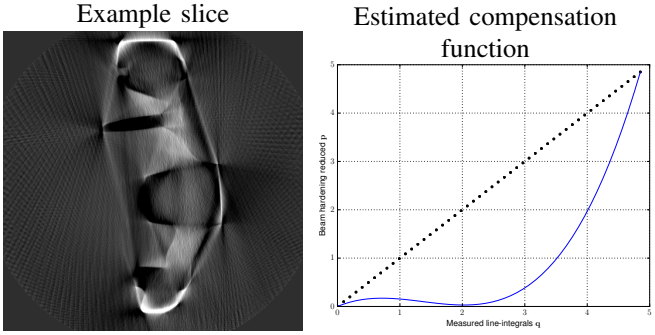


Figure 4: Results of the unconstrained ECC^2 algorithm using measured data of an aluminum part (cf. fig 5). The estimated compensation function shows undesired curvature and as a consequence, the reconstruction is flawed.

D. Application to the ECC^2 algorithm

The optimization problem for the reference-free beam hardening reduction algorithm of reference [8] is given as

$$\min\left(\|\mathbf{A}\mathbf{w}\|_2^2\right) \quad s.t. : \mathbf{w}^T\mathbf{b} = \beta; \quad w \geq 0 \quad \forall w \in \mathbf{w}\,, \tag{11}$$

Where $\mathbf{A}$ denotes a measurement matrix which is constructed using the epipolar consistency condition and $\mathbf{b}$ is a Vandermonde vector which fixes a point p to a value q to deal with the scale problem, inherent to the homogeneous least squares problem. We modify this algorithm using our new formulation to

$$\min\left(\|\mathbf{A}\,\mathbf{m}_d(\delta, \hat{\mathbf{w}})\,\|_2^2\right) \quad s.t. : \mathbf{m}_d(\delta, \hat{\mathbf{w}})^T\mathbf{b} = \beta\,, \delta > 0\,. \tag{12}$$

We can solve this problem using a standard solver for constrained convex optimization problems.

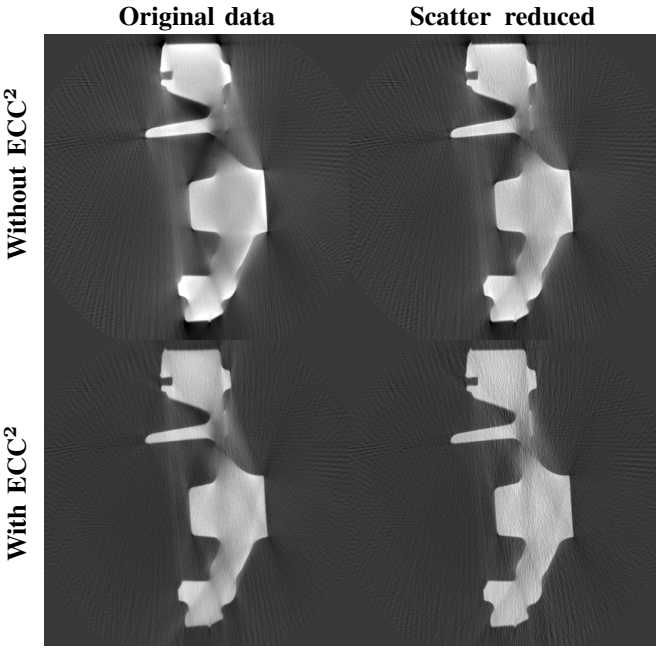


Figure 5: Comparison of the effect of scatter reduction on our modification of the ECC^2 algorithm. Because the ECC^2 algorithm does not preserve scale, we normalized every image to a mean of one (Grayscale window: C/W = 0.53/1.91).

III. Experiments

We present an experiment demonstrating our new method on a measured dataset of an aluminum object additionally affected by scatter.

The dataset shows severe scatter artifacts in addition to beam hardening. This is reflected in additional inconsistency. In Fig. 4 we present the result of the ECC^2 algorithm without a constraint on the polynomial:

The image-quality is severely impaired. The reason can be observed from the estimated polynomial which is neither monotonous nor convex.

We next compare the application of our proposed method to both the original data and scatter reduced data. The scatter reduction was performed using a beam stop array method. The results are presented in Fig. 5.

In Fig. 5 we demonstrate, that our proposed constraint makes the algorithm robust to additional sources of inconsistency. In addition we can see that the beam hardening related artifacts are removed in both cases, while the additional removal of scatter provides increased image quality independent of this.

IV. Conclusion and Outlook

We have shown a new parametrization of polynomial models which restricts the space of functions to a physically plausible subset. Specifically we improve previous approaches by precisely specifying the necessary conditions on these functions and providing a method restricting the results to all those functions which obey them. This can directly be used to improve a number of algorithms relying on such a model. Especially reference-free algorithms profit from our new formulation, as their functions of merit may actually reflect any other imaging problem in addition to beam hardening. We applied our method to the ECC^2 algorithm of Würfl et al.[8] and showed the effectiveness of our method in dealing with severe scatter conditions. Our algorithm is more complicated in terms of implementation and looses the advantage of presenting a convex optimization problem, if polynomials higher than degree three are considered. However we have not found this to be a practical limitation.

We will extend the physical constraints to multi-material methods in future research using similar techniques. This is more complicated since there is no unique definition of a convex functional of two coupled variables. We expect multi-material methods to benefit even more from such techniques because the problem has more degrees of freedom.

Additionally we are interested in applying our method to simultaneous multi-dimensional optimization of reference-free geometric and physical compensation methods promising to provide improved results for all those methods.

Acknowledgment

A. Aichert is supported by the German Research Foundation; DFG MA 4898/3-1.

Disclaimer

The concepts and information presented in this paper are based on research and are not commercially available.

References

[1] T. M. Buzug, Ed., *Computed Tomography*. Springer, 2008.
[2] M. Kachelrieß, K. Sourbelle, and W. A. Kalender, "Empirical cupping correction: A first-order raw data precorrection for cone-beam computed tomography," *Medical Physics*, vol. 33, no. 5, pp. 1269–1274, 2006.
[3] R. Clackdoyle, L. Desbat, J. Lesaint, and S. Rit, "Data consistency conditions for cone-beam projections on a circular trajectory," *IEEE Signal Processing Letters*, vol. 23, no. 12, pp. 1746–1750, 2016.
[4] J. Lesaint, R. Clackdoyle, S. Rit, and L. Desbat, "Two cone-beam consistency conditions for a circular trajectory," in *Proc. 4th Int. Conf. Image Formation X-Ray Comput. Tomography*, 2016, pp. 431–434.
[5] C. Debbeler, N. Maass, M. Elter, F. Dennerlein, and T. M. Buzug, "A new CT rawdata redundancy measure applied to automated misalignment correction," in *Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, 2013, pp. 264–267.
[6] A. Aichert, M. Berger, J. Wang, N. Maass, A. Doerfler, J. Hornegger, and A. K. Maier, "Epipolar consistency in transmission imaging," *IEEE Transactions on Medical Imaging*, vol. 34, no. 11, pp. 2205–2219, 2015.
[7] S. Abdurahman, R. Frysch, R. Bismarck, M. Friebe, and G. Rose, "Calibration free beam hardening correction using Grangeat-based consistency measure," in *Nuclear Science Symposium and Medical Imaging Conference*, 2016.
[8] T. Würfl, N. Maass, F. Dennerlein, X. Huang, and A. Maier, "Epipolar Consistency Guided Beam Hardening Reduction - ECC^2," in *Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, 2017, pp. 181–185.
[9] K. Murray, S. Müller, and B. Turlach, "Fast and flexible methods for monotone polynomial fitting," *Journal of Statistical Computation and Simulation*, vol. 86, no. 15, pp. 2946–2966, 2016.
[10] G. S. Fung, K. Stierstorfer, W. P. Segars, K. Taguchi, T. G. Flohr, and B. M. Tsui, "Xcat/drasim: a realistic CT/human-model simulation package," in *Proc. SPIE*, vol. 7961, 2011, p. 79613D.

Reducing partial volume artifacts with spectral CT

Mats Persson and Norbert J. Pelc

***Abstract*—Partial volume artifacts arise in computed tomography near sharp interfaces of highly attenuating material. These artifacts appear as shading and streaking and can degrade diagnostic performance. We propose a method of reducing these artifacts using energy-resolving photon-counting detectors, based on the observation that the partial volume artifacts have a unique spectral characteristic that distinguishes them from tissues obeying the low-dimensional approximation used in basis material decomposition. We simulate imaging of a simple mathematical phantom with an ideal photon-counting detector, and show that the level of artifacts in the image can be reduced substantially. While the artifact removal comes at the cost of increased noise, the basis function noise increase can be lowered to 4-7% using a frequency-splitting method. This method can be an important step towards using photon-counting spectral CT for accurate quantitative measurements in CT images.**

I. Introduction

The partial volume effect occurs in all imaging systems, including computed tomography (CT), when there are variations in the image on a length scale shorter than the resolution of the system. When the X-ray attenuation in a voxel-sized cell of the object contains a combination of two or more materials, the reconstructed voxel value will be an average of the true CT number over the voxel. Apart from this "linear" partial volume effect, however, there is also another effect called the nonlinear partial volume (NLPV) effect which can give rise to image artifacts. [1], [2] This phenomenon arises from the nonlinear form of X-ray attenuation (described by the Beer-Lambert law) in conjunction with the finite system resolution. As illustrated in Fig. 1, this effect means that the number of transmitted photons is higher for a beam passing an interface parallel to the beam direction (Fig. 1b) compared to a beam traversing a homogeneous mixture of the same materials (Fig. 1a). This leads to the X-ray attenuation being underestimated and causes streaks in the reconstructed image, called partial volume artifacts, that depend on the geometry of such structures in the object.

Different methods have been proposed for reducing these artifacts. [1], [3], [4] These methods build on making prior assumptions about the shape of the attenuating objects or the attenuation of the surrounding tissue. However, despite these developments and despite the fact that smaller detector pixels and thinner slices have decreased the severity of these artifacts over the years, partial volume artifacts can still be a problem e.g. when imaging dense bone structures in the skull. We propose a new method of correcting for partial volume artifacts using spectral CT, by exploiting the fact that the NLPV effect can be identified from its unique energy response, which does not agree with that of any material occurring in the human body. This fact has previously been exploited to improve the resolution of images. [5] In this work we use this insight to develop a new method for correcting partial volume artifacts and demonstrate it in a simulation study with a simple phantom.

M. Persson is with the departments of Bioengineering and Radiology, Stanford University, Stanford 94305, California, USA. e-mail: matspers@stanford.edu

N. J. Pelc is with the departments of Electrical Engineering, Bioengineering and Radiology, Stanford University, Stanford 94305, California, USA

II. Methods

We developed our artifact correction method within the framework of basis material decomposition, which is a method for quantifying material composition in spectral CT. [6]–[8] Basis material decomposition builds on the fact that the linear attenuation for any material occurring in the human body can be expressed as a linear combination of basis functions, e.g. $\mu(E) = a_w f_w(E) + a_b f_b(E)$ where we have chosen the linear attenuation coefficients of water $f_w(E)$ and bone $f_b(E)$ as basis functions. Integrating the basis coefficients (a_w, a_b) along the projection lines yields basis projections (A_w, A_b) which can be estimated from the measured data in each detector pixel. Since the spectral response of the NLPV effect cannot be expressed in this way, we model it by adding a third basis function:

$$f_{\mathrm{NLPV}}(E) = -\ln\left(\frac{1}{2}e^{-A_b^0 f_b(E)} + \frac{1}{2}e^{-A_w^0 f_w(E)}\right) - \frac{1}{2}A_w^0 f_w(E) - \frac{1}{2}A_b^0 f_b(E) \quad (1)$$

The rationale behind this definition is that it allows the transmission for a beam, half of which passes through a thickness A_w^0 of water and half of which passes through a thickness A_b^0 of bone, to be expressed as $\exp\left(-\frac{1}{2}A_w^0 f_w(E) - \frac{1}{2}A_w^0 f_w(E) - 1 \cdot f_{\mathrm{NLPV}}(E)\right)$. In this way, we have constructed a third basis function that measures the strength of nonlinear partial volume effect, exactly for the material combination for which it was constructed and approximately for other material combinations.

The proposed artifact reduction technique builds on performing a three-material decomposition using water, bone and the NLPV effect as basis functions. With the NPLV basis function constructed according to the above, the NLPV effect will be captured by the NLPV basis, so that the water and bone images will contain the average basis projections of these two materials, respectively. The water and bone images will thus be free from partial volume artifacts, at least if the measured material combination is similar enough to the case that was used to construct the basis function.

We simulated a 16 cm cylindrical water phantom with three cylindrical inserts made of cortical bone, each 15 mm in diameter. (Fig. 2.) The inserts are protruding halfway into the slice, two from below and one from above. This phantom was

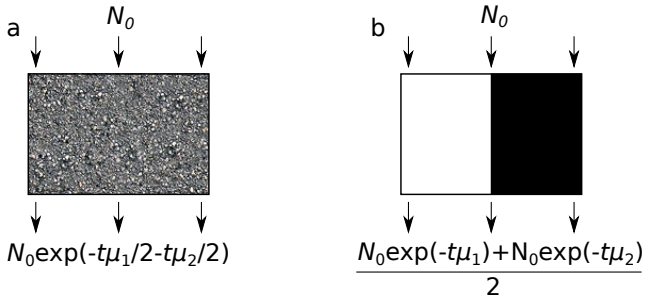


Fig. 1. Illustration of the nonlinear partial volume effect. A single voxel contains equal amounts of two materials (white and black) with attenuation coefficients $\mu_1(E)$ and $\mu_2(E)$ respectively. In case (a) the materials are mixed, and the effective attenuation coefficient is the linear average of μ_1 and μ_2. In case (b) the materials are separated by an interface parallel to the beam, and the effective attenuation coefficient is a nonlinear average of μ_1 and μ_2. This gives a higher number of transmitted photons compared to case (a). Here, t denotes the path length.

mathematically represented as an upper and a lower subslice, and the `fanbeam` function of MATLAB (The MathWorks Inc., Natick, MA, USA) was used to calculate sinograms of the projected amounts of water and bone for each of these subslices. The forward projection was made with 1000 view angles covering 360°, 500 mm source-to-isocenter distance and an arc-shaped array of detector pixels corresponding to 0.5 mm pixel pitch in isocenter. To model spectrum filtering by the bowtie and other filters in the X-ray tube housing, the prepatient spectrum was attenuated by an amount of teflon varying with the fan angle. The filter thickness was fitted to a X-ray dose profile measured on a GE VCT with medium field of view [9] with a half-value layer of 6.7 mm Al in the central ray. By adding the transmitted counts in the two subslices, we simulated the photon counts in each pixel for an energy-resolving photon-counting detector, with unity detection efficiency, perfect energy response and 1-keV wide energy bins with thresholds from 0.5 to 119.5 keV. We used a 12° tungsten anode X-ray spectrum from [10] and linear attenuation coefficients from [11], implemented in Spektr3. [12] We simulated Poisson noise with $4 \cdot 10^5$ prepatient photons per measurement (in total $4 \cdot 10^8$ photons per detector pixel during the entire rotation), corresponding to approximately 200 mAs.

We constructed an NLPV basis function according to (1) with $A_w^0 = A_b^0 = 15$ mm i.e. a common path length of 15 mm water and soft tissue, corresponding to a projection ray passing through one of the bone inserts at its thickest part. Maximum-likelihood basis material decomposition [7] was then used to generate three basis sinograms, for water, bone and NLPV, from the measured data. For reference, we also made a conventional two-basis decomposition into water and bone images. The basis images were reconstructed with `ifanbeam` in MATLAB using a pure ramp filter. In addition, a "pure" photon-counting image was reconstucted from a sinogram of log-normalized counts $-\ln(N/\overline{N}_0)$ where N is the total number of counts summed over all energy bins and $\overline{N}_0$ is the expected number of prepatient photons in the same projection line.

Since we observed that replacing the two-basis decomposition by three-basis decomposition decreased the level of artifacts at the expense of increased noise, we used a frequency-splitting method to merge the two-basis and three-basis images. Letting $a_{\alpha,2}(x,y)$ for $\alpha = \{b, w\}$ denote the reconstructed basis images resulting from the two-basis decomposition and letting $a_{\alpha,3}(x,y)$ be the corresponding basis images from the three-material decomposition, the difference images $\Delta a_\alpha(x,y) = a_{\alpha,3}(x,y) - a_{\alpha,2}(x,y)$ capture the NLPV artifacts. Since these artifacts have a predominantly low-frequency character, low-pass filtering the correction images can reduce noise without affecting the artifacts much. We therefore generated a corrected image as $a_{\alpha,2}(x,y) + \Delta a_\alpha(x,y) * K(x,y)$ where $*$ denotes convolution and K is a smoothing kernel, in our case a Gaussian with standard deviation 2 pixels. For each of the reconstructed basis image sets we also generated a synthetic monoenergetic image from the water and bone basis images as $\mu(E_{\text{mono}}) = a_w f_w(E_{\text{mono}}) + a_b f_b(E_{\text{mono}})$ with $E_{\text{mono}} = 60$ keV.

III. Results

Figure 2 shows reconstructed images of the phantom. Fig. 2(a-b) show the ground truth images with the regions of interest (ROIs) used for measuring noise standard deviation and mean water and bone image value, and the line along which the profiles in Fig. 3a are measured. Fig. 2(c) shows the pure photon-counting image. Fig. 2(d-l) show the reconstructed water and bone basis images and synthetic monoenergetic images for two-basis decomposition, three-basis decomposition and the merged image obtained by combining the two-basis and three-basis images using frequency splitting. Figure 3(a) shows a plot of the different reconstructed synthetic monoenergetic images along the line shown in Fig. 2(b). Figure 3(b) shows the average basis coefficient in the water and bone basis images, measured in the background ("Bg"), ROI in the most artifact-intense region between the inserts, and in an ROI positioned in one of the inserts. Figure 3(c) shows standard deviation in the water and bone basis images, measured in the square "Noise" ROI located in a more peripheral part of the phantom, where there are no visible artifacts.

IV. Discussion

As seen in in Fig. 2(c,f), both the pure photon-counting image and the synthetic monoenergetic image exhibit severe partial volume artifacts, primarily along the projection lines passing through two of the inserts. The appearance of the artifacts is different depending on which direction the inserts enter the slice from. Between the two inserts that protrude from the same direction, the linear attenuation coefficient is underestimated, giving a dark streak in the image. Between the upper insert and one of the lower inserts, the dark streaks are concentrated to both sides of the line through the center of the two inserts, i.e. to rays passing a much larger distance through one of the inserts than through the other. The artifacts are even more visible in the basis images resulting from two-basis decomposition (Fig. 2 (d-e)), with opposite polarity in the water and bone images, i.e. they cancel out to some extent when forming a synthetic monoenergetic image.

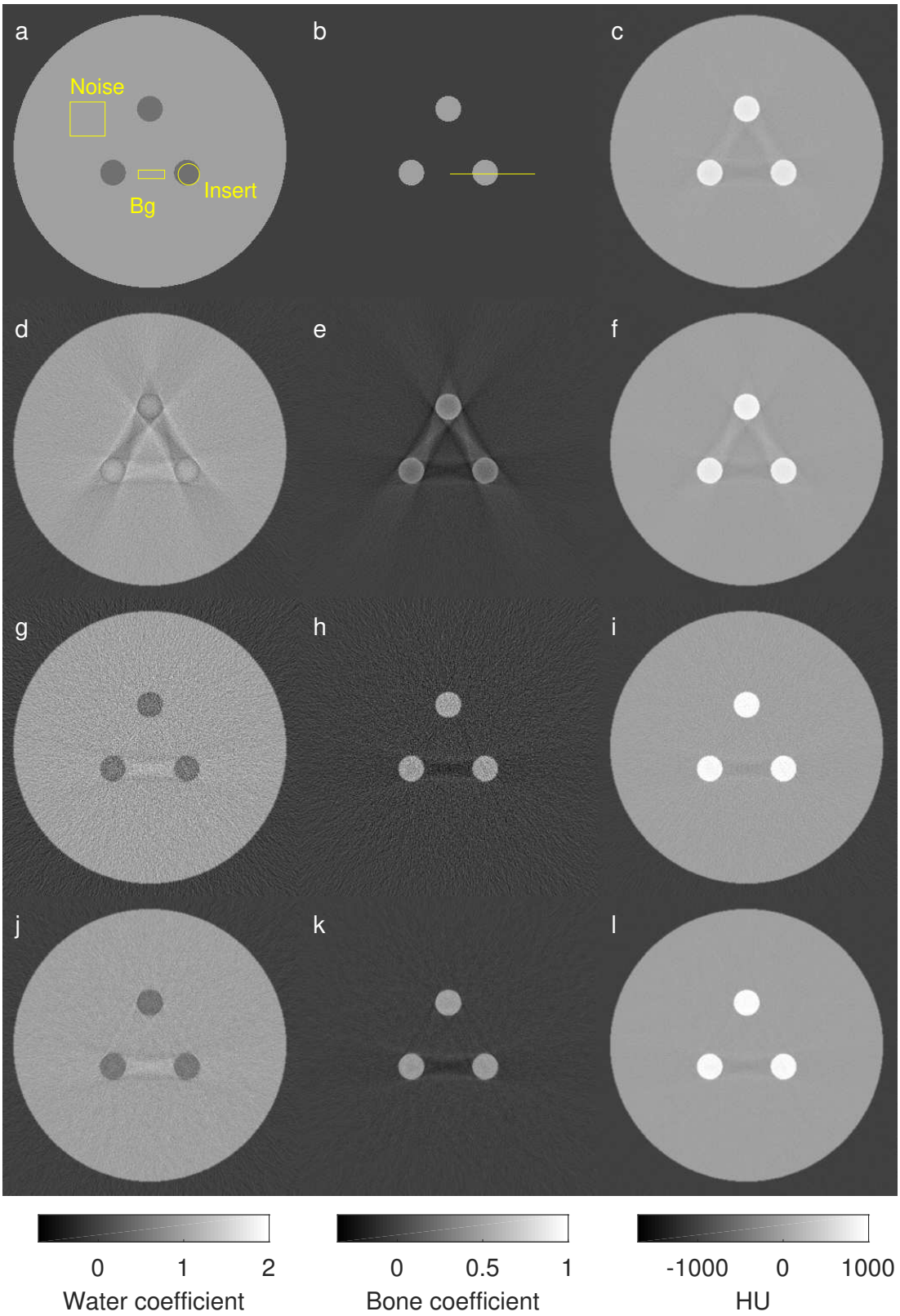


Fig. 2. Reconstructed images of a 16 cm water phantom with three 15 mm diameter bone inserts partially protruding into the slice. The upper insert protrudes from above and the bottom inserts from below. Left column: water basis images. Center column: Bone basis images. Right column: photon counting and synthetic monoenergetic images. (a-b) Water and bone basis images, ground truth. The ROIs for measuring mean and standard deviation are shown in (a) and the profile plotted in Fig. 3 is shown in (b). (c) Pure photon-counting image. (d-e) Water and bone basis images from two-basis decomposition. (f) Synthetic 60 keV monoenergetic image resulting from two-basis decomposition. (g-h) Water and bone basis images from three-basis decomposition. (i) synthetic monoenergetic image from three-basis decomposition. (j-k) Water and bone basis images created by merging images from two- and three-basis decomposition. (l) Synthetic 60 keV monoenergetic image generated from the merged images in (j-k).

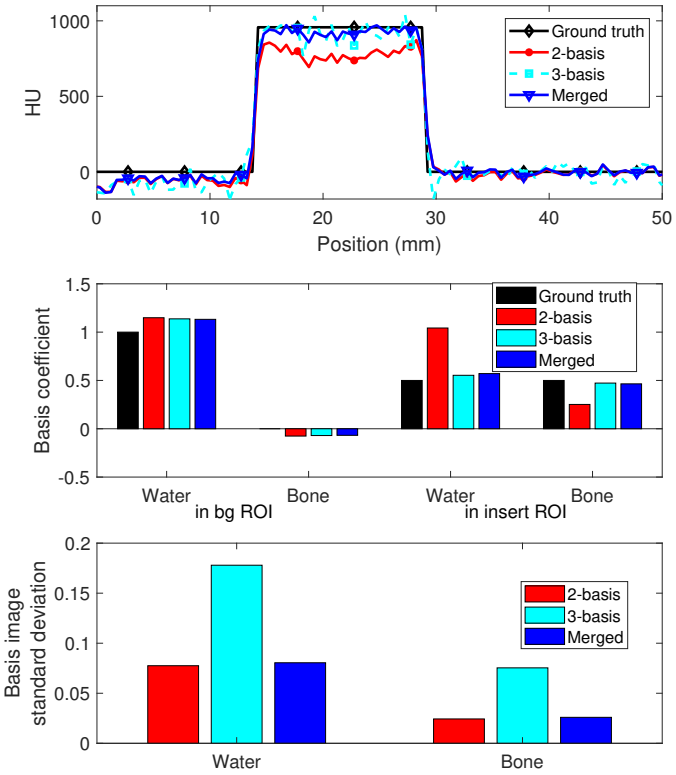


Fig. 3. (a) Horizontal profiles through the synthetic monoenergetic images, through the center of one of the bone inserts. (b) Mean basis coefficient measured in the background ("Bg") and in the "Insert" ROI. (c) Standard deviation in the basis images, measured in the square ROI denoted "Noise".

The water and bone images resulting from three-basis decomposition (Fig. 2 g-i) exhibit a much lower level of artifacts. The streaks between the upper and lower inserts are no longer visible and the artifact between the lower inserts is reduced in magnitude. The NLPV effect has thus been captured in the third basis projection instead of contaminating the two ordinary basis projections. Also note that the water and bone coefficients measured in the insert ROI are close to the ground truth value in 3-material and merged images (5-14% bias) compared to the 2-basis images which have a bias of 109% (water) and 50% (bone), respectively. Some residual artifacts are visible for lines passing a very large path length (up to 30 mm) of an interface between soft tissue and bone. This is sufficiently different from the case that the basis function was constructed for (15 mm common path length of soft tissue and bone) that the NLPV basis function does not capture the NLPV effect entirely.

Fig. 2(g-i) also show that the noise level increases when going from two-basis to three-basis decomposition, a common phenomenon in signal processing when the number of estimated parameters is increased. To mitigate this problem, we used the frequency-splitting technique to generate a merged image with low-frequency features from the three-basis decomposition and high-frequency features from the two-basis decomposition (Fig. 2(j-l)). As seen in the figure, these merged images have a low level of artifacts comparable to the three-basis decomposition but with a substantially reduced noise level. As seen in Fig. 3, the noise standard deviation in the merged water and bone images are only 4% and 7% higher than the images from the two-basis decomposition, respectively. This can be compared with the images from the three-basis decomposition which have 130% and 210% higher noise than the 2-basis image, for water and bone, respectively.

V. Conclusion

We have demonstrated that photon-counting energy-resolving CT can be used to reduce partial volume artifacts substantially, by performing basis material decomposition with an additional basis function. While this artifact reduction is accompanied by an noise increase in the basis image, this noise increase can be lowered to 4-7% using a frequency splitting method. In future work, we will investigate ways of further reducing the remaining artifacts.

Acknowledgments

This study was financially supported by by NIH Grant U01 EB01714003. M. Persson is stockholder of and consultant for Prismatic Sensors AB. Norbert J. Pelc is consultant for Prismatic Sensors AB.

References

[1] G. H. Glover and N. J. Pelc, "Nonlinear partial volume artifacts in x-ray computed tomography," *Med. Phys.*, vol. 7, no. 3, pp. 238–248, 1980.

[2] P. M. Joseph and R. D. Spital, "The exponential edge-gradient effect in X-ray computed tomography," *Physics in Medicine and Biology*, vol. 26, no. 3, p. 473, 1981.

[3] D. J. Goodenough, K. E. Weaver, H. Costaridou, H. Eerdmans, and P. Huysmans, "A new software correction approach to volume averaging artifacts in CT," *Computerized Radiology*, vol. 10, no. 2, pp. 87 – 98, 1986.

[4] J. Hsieh, "Partial volume artifact reduction in computed tomography," in *Proceedings of 3rd IEEE International Conference on Image Processing*, vol. 3, Sep 1996, pp. 567–570 vol.3.

[5] M. Persson, S. Holmin, S. Karlsson, H. Bornefalk, and M. Danielsson, "Subpixel x-ray imaging with an energy-resolving detector," *Journal of Medical Imaging*, vol. 5, p. 013507, 2018.

[6] R. E. Alvarez and A. Macovski, "Energy-selective reconstructions in x-ray computerised tomography," *Phys. Med. Biol.*, vol. 21, no. 5, pp. 733–744, Sept. 1976.

[7] E. Roessl and R. Proksa, "K-edge imaging in x-ray computed tomography using multi-bin photon counting detectors," *Phys. Med. Biol.*, vol. 52, no. 15, p. 4679, 2007.

[8] J. P. Schlomka, E. Roessl, R. Dorscheid, S. Dill, G. Martens, T. Istel, C. Bumer, C. Herrmann, R. Steadman, G. Zeitler, A. Livne, and R. Proksa, "Experimental feasibility of multi-energy photon-counting K-edge imaging in pre-clinical computed tomography," *Phys. Med. Biol.*, vol. 53, no. 15, pp. 4031–4047, Aug. 2008.

[9] M. Persson, R. Bujila, P. Nowik, H. Andersson, L. Kull, J. Andersson, H. Bornefalk, and M. Danielsson, "Upper limits of the photon fluence rate on ct detectors: Case study on a commercial scanner," *Medical Physics*, vol. 43, no. 7, pp. 4398–4411, 2016.

[10] A. M. Hernandez and J. M. Boone, "Tungsten anode spectral model using interpolating cubic splines: Unfiltered x-ray spectra from 20 kV to 640 kV," *Medical Physics*, vol. 41, no. 4, p. 042101, 2014.

[11] M. J. Berger, J. H. Hubbell, S. M. Seltzer, J. Chang, J. S. Coursey, R. Sukumar, D. S. Zucker, and K. Olsen, "XCOM: Photon Cross Section Database," http://physics.nist.gov/xcom. National Institute of Standards and Technology, Gaithersburg, MD, 2005.

[12] J. Punnoose, J. Xu, A. Sisniega, W. Zbijewski, and J. H. Siewerdsen, "Technical note: spektr 3.0-A computational tool for x-ray spectrum modeling and analysis," *Medical Physics*, vol. 43, no. 8Part1, pp. 4711–4717, 2016.

Locally Linear Transform based Gradient L_0-norm Minimization for Spectral CT Reconstruction

Qian Wang[†], Weiwen Wu[‡†], Hengyong Yu[*†]

***Abstract*—By extending the conventional single-energy computed tomography (SECT) along the energy dimension, spectral CT achieves superior energy resolution and material distinguishability. However, because the emitted photons with a fixed total number for one X-ray beam are divided into several different energy bins, the noise level is increased in the reconstructed image in each channel, and it further leads to an inaccurate material decomposition. To improve the reconstruction quality and decomposition accuracy, in this work, we first construct a global three-dimensional (3D) gradient sparsity of two-dimensional (2D) spectral CT images by employing a locally linear transform. Then, the sparse property is measured with an L_0-norm and incorporated into an optimization model. The corresponding iterative algorithm is also deduced. Experiment results with simulation data and real data demonstrate a superiority of the proposed method.**

***Index Terms*—Spectral CT, locally linear transform, three-dimensional gradient, L_0-norm minimization, material decomposition.**

I. Introduction

The matter attenuation properties of X-rays are energy- and material-dependent [1]. This allows deriving material properties by using multiple different X-ray energies for imaging. Spectral CT is proposed by extending the conventional single-energy computed tomography (SECT) along the energy dimension, and it can be realized by employing photon-counting detectors (PCDs) or other methods. The state-of-the-art PCD can divide the X-ray photons into 8 different energy channels, and then achieve multiple energy-dependent projection sets [2]. Thus, compared with the SECT, spectral CT is superior in energy resolution and material distinguishability. It has a great potential in both medical and industrial applications, such as bone mineral density and liver iron concentration measurements, beam-hardening correction and contrast enhancement of soft tissues, calculation of pseudo-monochromatic images, and so forth [3].

Similar to the conventional image reconstruction problems, the target problem of spectral CT is also ill-posed, i.e., the solution process is very sensitive to noise. Because the emitted photons with a fixed total number for one X-ray beam are divided into various energy bins, the noise level will increase in each channel. This degrades the reconstructed image quality and further reduces the decomposition accuracy. Therefore, one of the fundamental problems for spectral CT is how to reconstruct high-quality images from noisy projections.

To overcome the ill-posedness in spectral CT reconstruction, prior knowledge needs to be greatly concerned and effectively employed. One popular prior knowledge is based on the fact that the scanned specimens are usually piece-wise constant. This property can be employed to establish a sparse gradient constraint in the spatial domain. In [4], Xu *et al.* proposed a total variation (TV) based spectral CT method by individually considering the channel image reconstruction. Generally, if the image is compressible, the prior information can be developed in a transform domain in the framework of compressive sensing, such as Fourier domain, wavelet domain and so on. In [5], a tight-frame based iterative reconstruction (TFIR) approach was investigated for spectral breast CT. Another spatial property is the structural sparsity, which concerns the feature in high-dimensional space. One representative utilization is the dictionary learning (DL) based sparse representation, which is also applied to spectral CT [6]. These methods just consider the prior knowledge in the spatial domain or its transform domain, but overlook the correlation in the spectral dimension. By incorporating the similarity among channel images, all the aforementioned methods achieve further developments. For example, the TV-based method is combined with low-rank property [7], the DL-based method is extended to tensor dictionary [8], and so on.

Different from directly employing the correlation among channel images, we use a locally linear transform to establish a gradient sparsity in the spectral dimension. Combining the piece-wise constant prior knowledge in the spatial domain, a three-dimensional (3D) gradient sparsity is formed. This property is further measured by an L_0-norm and incorporated into an optimization model as a regularization term. We also develop the corresponding iterative algorithm.

The remainder of this paper is organized as follows. In section II, we present the mathematical model of spectral CT, and review the image-guided filtering and two-dimensional (2D) gradient L_0-norm minimization. In section III, the global 3D gradient sparsity of spectral CT images is established and measured by an L_0-norm. Correspondingly, we propose an optimization model and develop an iterative algorithm. In section IV, we perform both numerical simulations and real experiments to verify the effectiveness of the proposed method. In last section, we discuss some related issues and conclude this work.

* Corresponding Author (hengyong-yu@ieee.org).
[†]Department of Electrical and Computer Engineering, University of Massachusetts Lowell, Lowell, MA 01854, USA.
[‡]Key Lab of Optoelectronic Technology and Systems, Ministry of Education, Chongqing University, Chongqing 400044, China.

II. Theory

A. Mathematical Model of Spectral CT

In realistic CT applications, the emitted X-ray photons are not ideally monochromatic. They obey a spectral distribution, i.e., energy-dependent. However, the conventional CT scanners employ energy-integrating detectors, which do not have the capability to distinguish photon energies. Thus, the corresponding post-log projection P_0 is represented as

$$P_0 = \ln\Big[\frac{\int_E S(E)\,\mathrm{d}E}{\int_E S(E)\exp\big(-\mathcal{P}(\mu(E,\mathbf{x}))\big)\,\mathrm{d}E}\Big], \tag{1}$$

where $\mu(E,\mathbf{x})$ is the linear attenuation coefficient for energy E at position $\mathbf{x}$, $\mathcal{P}(\cdot)$ represents the ray transform, $S(E)$ is the emission spectrum.

For a photon-counting detector based spectral CT scanner, the transmitted photons can be discriminated and counted by different energy bins. In this case, the whole X-ray spectrum can be divided into several intervals with appropriate post-processing steps which lead to multiple projections along energy dimension. Assuming there are C energy channels and the energy interval is noted as $\omega_c\,(1 \le c \le C)$, the corresponding channel projection reads,

$$P_c = \ln\Big[\frac{\int_E S(E)W_c(E)\,\mathrm{d}E}{\int_E S(E)W_c(E)\exp\big(-\mathcal{P}(\mu(E,\mathbf{x}))\big)\,\mathrm{d}E}\Big], \tag{2}$$

where

$$W_c(E) = \begin{cases} 1, & E \in \omega_c, \\ 0, & \text{otherwise}, \end{cases}$$

is a window function of channel c. Obviously, (1) is a special case of (2) when the energy bin number is 1 and the energy window covers the whole spectrum.

By denoting the inverse operator of $\mathcal{P}(\cdot)$ as $\mathcal{P}^{-1}(\cdot)$, we can obtain the reconstructed channel image as follow,

$$\begin{aligned} f_c(\mathbf{x}) &= \mathcal{P}^{-1}(P_c) \\ &= \mathcal{P}^{-1}\Big\{\ln\Big[\frac{\int_E S(E)W_c(E)\,\mathrm{d}E}{\int_E S(E)W_c(E)\exp\big(-\mathcal{P}(\mu(E,\mathbf{x}))\big)\,\mathrm{d}E}\Big]\Big\}. \end{aligned} \tag{3}$$

Denoting the size of $f_c(1 \le c \le C)$ as $M \times N$ and heaping up all the channel images along the spectral dimension, we can obtain a 2D spectral CT image F with size of $M \times N \times C$. Denote the c^{th} channel image of spectral CT reconstruction volume F as F_c, i.e., $F_c = f_c(1 \le c \le C)$. Then, by performing the same operation on the channel projection data, we can obtain the spectral projection volume P.

B. Image-guided Filtering

Image-guided filtering is a reference image based edge-preserving smoothing technique [9]. As a neighborhood operation, it establishes a locally linear relationship between the reference image f and the filtering output g as follow,

$$g(\mathbf{x}) = a(\mathbf{k})f(\mathbf{x}) + b(\mathbf{k}), \forall \mathbf{x} \in \Omega_2(\mathbf{k}), \tag{4}$$

where $\Omega_2(\mathbf{k})$ represents an image patch centered at the position $\mathbf{k}$, and $\big(a(\mathbf{k}), b(\mathbf{k})\big)$ are a pair of constant coefficients depending on the patch $\Omega_2(\mathbf{k})$. $\big(a(\mathbf{k}), b(\mathbf{k})\big)$ can be determined by solving the following quadratic optimization model,

$$\min_{\left(a(\mathbf{k}),b(\mathbf{k})\right)} \sum_{\mathbf{x}\in\Omega_2(\mathbf{k})} \big[\big(a(\mathbf{k})f(\mathbf{x}) + b(\mathbf{k}) - g_0(\mathbf{x})\big)^2 + \epsilon a^2(\mathbf{k})\big]. \tag{5}$$

Here g_0 strands for the filtering input. Thus, the first quadratic term measures the difference between the filtering input and output, i.e., the data fidelity term. ϵ is a regularization parameter to penalize large $a(\mathbf{k})$. Further, the closed form solution of (5) can be written as,

$$a(\mathbf{k}) = \frac{\frac{1}{\#(\Omega_2)}\sum_{\mathbf{x}\in\Omega_2(\mathbf{k})} f(\mathbf{x})g_0(\mathbf{x}) - \bar{f}(\mathbf{k})\bar{g}_0(\mathbf{k})}{\sigma^2(\mathbf{k}) + \epsilon}, \tag{6a}$$

$$b(\mathbf{k}) = \bar{g}_0(\mathbf{k}) - a(\mathbf{k})\bar{f}(\mathbf{k}), \tag{6b}$$

where $\#(\cdot)$ stands for the counting operator to count the pixel number in an image path. $\bar{f}(\mathbf{k})$ and $\bar{g}_0(\mathbf{k})$ are respectively the mean of f and g_0 in $\Omega_2(\mathbf{k})$, and $\sigma^2(\mathbf{k})$ represents the variance of f in $\Omega_2(\mathbf{k})$.

Further, we can employ (4) to compute the filtering output g. According to (5), it is obvious that g is similar to the filtering input g_0. Meanwhile, by utilizing (4), g contains the features of the reference image f. This means any structure in f is transferred into g in terms of locally linear transform.

Considering each pixel $\mathbf{x}$ is covered by several patches, we adopt an averaging strategy. Thus, (4) is converted to

$$g(\mathbf{x}) = \bar{a}(\mathbf{x})f(\mathbf{x}) + \bar{b}(\mathbf{x}), \forall \mathbf{x} \in \Omega_2,$$

where $\big(\bar{a}(\mathbf{x}), \bar{b}(\mathbf{x})\big)$ is a pair of averaged coefficients in all the patches covering the pixel $\mathbf{x}$.

C. 2D Gradient L_0-norm Minimization

The gradient L_0-norm regularization based image denoising optimization model reads,

$$\min_g \Big\{\sum_{\mathbf{x}\in\Omega_2}\big(g(\mathbf{x}) - g_0(\mathbf{x})\big)^2 + \lambda\Phi(g)\Big\}, \tag{7}$$

where g_0 represents the obtained noisy image, g is the searched-for noise-free image, and Ω_2 stands for the image support.

$$\Phi(g) = \#\big\{\mathbf{x} \in \Omega_2 \big| |\partial_x g(\mathbf{x})| + |\partial_y g(\mathbf{x})| \neq 0\big\}$$

is the gradient L_0-norm regularization term, which counts the number of pixels satisfying $|\partial_x g(\cdot)| + |\partial_y g(\cdot)| \neq 0$. $\lambda > 0$ is a regularization parameter to balance the fidelity and regularization terms to obtain satisfactory results.

Comparing with the Tikhonov regularization (L_2-norm of image gradient) and TV regularization (L_1-norm of image gradient) based models, (7) has the strongest capability for sparse representation. However, the counting function $\Phi(\cdot)$ leads to an NP-hard problem, and it greatly increases the solution difficulty.

Recently, Xu *et al.* proposed an approximation method, which can fast and stably solve the gradient L_0-norm minimization problem [10]. To split the original optimization model into sub-problems that are easy to be solved, two auxiliary variables $h(\mathbf{x})$ and $v(\mathbf{x})$ are introduced to substitute

$\partial_x g(\mathbf{x})$ and $\partial_y g(\mathbf{x})$ respectively. Thus, (7) is converted to a constrained optimization model as follow,

$$\min_{g} \Big\{ \sum_{\mathbf{x}\in\Omega_2} \big(g(\mathbf{x}) - g_0(\mathbf{x})\big)^2 + \lambda\Phi(h,v)\Big\}, \tag{8}$$
$$\text{s.t. } h(\mathbf{x}) = \partial_x g(\mathbf{x}), v(\mathbf{x}) = \partial_y g(\mathbf{x}), \forall \mathbf{x}\in\Omega_2,$$

where

$$\Phi(h,v) = \#\big\{\mathbf{x}\in\Omega_2 \big| |h(\mathbf{x})| + |h(\mathbf{x})| \neq 0\big\}.$$

Then, by relaxing the constraints, (8) can be further converted into the following unconstrained version,

$$\min_{g,h,v} \Big\{ \sum_{\mathbf{x}\in\Omega_2} \big(g(\mathbf{x}) - g_0(\mathbf{x})\big)^2 + \lambda\Phi(h,v) \tag{9}$$
$$+ \beta \sum_{\mathbf{x}\in\Omega_2} \big[\big(\partial_x g(\mathbf{x}) - h(\mathbf{x})\big)^2 + \big(\partial_y g(\mathbf{x}) - v(\mathbf{x})\big)^2\big]\Big\},$$

where $\beta > 0$ is an automatically adapting parameter to control the relaxation degree. (9) can be further split into two sub-problems with respect to g and (h, v), i.e.,

$$\min_{g} \sum_{\mathbf{x}\in\Omega_2} \Big\{ \big(g(\mathbf{x}) - g_0(\mathbf{x})\big)^2 + \beta\big[\big(\partial_x g(\mathbf{x}) - h(\mathbf{x})\big)^2 + \big(\partial_y g(\mathbf{x}) - v(\mathbf{x})\big)^2\big]\Big\}, \tag{10a}$$

$$\min_{h,v} \Big\{\lambda C(h,v) + \beta \sum_{\mathbf{x}\in\Omega_2} \big[\big(\partial_x g(\mathbf{x}) - h(\mathbf{x})\big)^2 + \big(\partial_y g(\mathbf{x}) - v(\mathbf{x})\big)^2\big]\Big\}. \tag{10b}$$

The solution of (10a) has a closed-form and can be accelerated by Fast Fourier Transform (FFT). Moreover, by employing the characteristic of spatial decomposition, (10b) can also be quickly solved.

III. Method

A. Locally Linear Transform based 3D Gradient Sparsity Constraint

Comparing all the channel images, the correlation is conspicuous. For one fact, all the slices contain the same structures and textures, i.e., **structural similarity**. For the other fact, the gray value and contrast vary a lot, i.e., **quantitative diversity**. Because the scanned specimen is often piecewise-constant, the gradient image of each slice image F_c is sparse, which can be viewed as a 2D sparsity in the spatial domain. However, the sparsity along the spectral dimension is not obvious and direct.

Inspired by the idea of image-guided filtering, we first employ the locally linear transform to establish a gradient sparsity along the spectral direction. Assuming the current target channel is c, we fix F_c as the input image, and choose $F_i\,(1 \le i \le C)$ as a reference image. Then, by employing the method in II-B, we obtain a filtering output F_i^c, which satisfies

$$F_i^c(\mathbf{x}) = \bar{a}_i^c(\mathbf{x})F_i(\mathbf{x}) + \bar{b}_i^c(\mathbf{x}), \forall \mathbf{x}\in\Omega_2. \tag{11}$$

According to (5), F_i^c can be viewed as a copy of F_c. When the reference image traverses all the energy channels, we can heap up the corresponding filtering outputs along the spectral direction to form a volume F^c, of which the i^{th} channel image is $F_i^c\,(1 \le i \le C)$. Thus, we further represent (11) in a volume version as follow,

$$F^c(\mathbf{x}) = \bar{a}^c(\mathbf{x})F(\mathbf{x}) + \bar{b}^c(\mathbf{x}), \forall \mathbf{x}\in\Omega_3. \tag{12}$$

The corresponding filtering input volume is represented as V^c, which is a duplicate extension of F_c along the spectral dimension. It is emphasized while F_c represents a 2D channel image, F^c is the corresponding 3D extension along the spectral dimension. The transform parameters $\bar{a}^c(\mathbf{x})$ and $\bar{b}^c(\mathbf{x})$ are determined by using (6a) and (6b).

It is worth noting that the channel images of F^c successfully overcome the shortcoming of quantitative diversity, and well maintain the structural similarity at the same time. Thus, its 3D gradient volume is globally sparse, i.e., 2D spatial sparsity and 1D spectral sparsity.

B. Optimization Model and Iterative Algorithm

To describe the 3D gradient L_0-norm, we first extend the definition of counting function $\Phi(\cdot)$ as follow,

$$\Phi(F^c) = \#\big\{\mathbf{x}\in\Omega_3 \big| |\partial_x F^c(\mathbf{x})| + |\partial_y F^c(\mathbf{x})| + |\partial_z F^c(\mathbf{x})| \neq 0\big\},$$

where Ω_3 is the spectral volume range.

Considering the gradient sparsity of $F^c\,(1 \le c \le C)$, we employ it as a constraint and develop the corresponding regularization term in light of L_0-norm. Combining the data fidelity term, we propose the following optimization model,

$$\min_{F,F^c} \big\{\|\mathcal{P}(F) - P\|_{L_2}^2 + \lambda_c\Phi(F^c)\big\}, \tag{13}$$
$$\text{s.t. } F^c(\mathbf{x}) = \bar{a}^c(\mathbf{x})F(\mathbf{x}) + \bar{b}^c(\mathbf{x}), 1 \le c \le C.$$

Here F is the spectral CT reconstruction volume by heaping up all the channel images $F_c\,(1 \le c \le C)$ along the spectral dimension. F^c is a duplication volume of the searched-for c^{th} channel image along the spectral dimension. $\bar{a}^c(\mathbf{x})$ and $\bar{b}^c(\mathbf{x})$ are determined by the following quadratic optimization model,

$$\min_{a^c(\mathbf{x}),b^c(\mathbf{x})} \Big\{ \sum_{\mathbf{t}\in\Omega_2(\mathbf{x})} \Big[\big(a^c(\mathbf{x})F(\mathbf{t}) + b^c(\mathbf{x}) - V^c(\mathbf{t})\big)^2 + \epsilon\big(a^c(\mathbf{x})\big)^2\Big]\Big\}.$$

The solution process is the same as the method given in section II-B. By relaxing the constraint in 13, it is converted to the following unconstrained model,

$$\min_{F,F^c} \Big\{ \|\mathcal{P}(F) - P\|_{L_2}^2 + \lambda_c\Phi(F^c) \tag{14}$$
$$+ \tau_c \sum_{\mathbf{x}\in\Omega_3} \Big[F^c(\mathbf{x}) - \big(\bar{a}^c(\mathbf{x})F(\mathbf{x}) + \bar{b}^c(\mathbf{x})\big)\Big]^2\Big\},\ 1 \le c \le C,$$

where $\tau_c > 0\,(1 \le c \le C)$ is a free parameter to control the relaxation degree. For each target channel $c\,(1 \le c \le C)$, we split (14) into the following sub-problems,

$$\min_{F} \Big\{ \|\mathcal{P}(F) - P\|_{L_2}^2 + \tau_c \sum_{\mathbf{x}\in\Omega_3} \Big[F^c(\mathbf{x}) - \big(\bar{a}^c(\mathbf{x})F(\mathbf{x}) + \bar{b}^c(\mathbf{x})\big)\Big]^2\Big\}, \tag{15a}$$

$$\min_{F^c} \Big\{ \lambda_c\Phi(F^c) + \tau_c \sum_{\mathbf{x}\in\Omega_3} \Big[F^c(\mathbf{x}) - \big(\bar{a}^c(\mathbf{x})F(\mathbf{x}) + \bar{b}^c(\mathbf{x})\big)\Big]^2\Big\}. \tag{15b}$$

(15a) is a quadratic optimization problem, which can be iteratively solved by using POCS scheme. (15b) can be viewed as a 3D generalization of model (7), and the solution can be achieved by extending the 2D method in [10]. By averaging F^c along the spectral dimension, we can obtain the searched-for channel images. Further, we can employ the decomposition method [11] to get material percentage images.

IV. Results

A. Numerical Simulation

In the numerical simulations, a mouse thorax phantom with 1.2% iodine contrast agent injection is employed to simulate spectral CT scan. The spectrum is under 50 KV and it is split into 8 energy intervals. The emitted photons are fixed to 10^5 and 5×10^3 to simulate the normal-dose and low-dose cases, respectively. The scan geometrical protocol is as follows: the source-to-object distance (SOD) is 132 mm; the source-to-detector distance (SDD) is 180 mm; the detector consists of 512 cells and each has a length of 0.1 mm. With this configuration, projections are uniformly collected from 640 views with Poisson noise, and 512×512 channel images are reconstructed with a pixel size of $0.075 \times 0.075\,\mathrm{mm}^2$.

The conventional FBP, TV, TV with low rank (TVLR) and 2D L_0-norm (L_0) based methods are employed as comparisons. The decomposed material images and synthesized color images are illustrated in Figs. 1 and 2 respectively for normal-dose and low-dose cases. Moreover, the peak signal-to-noise ratio (PSNR), normal mean absolute deviation (NMAD) and structural similarity (SSIM) are employed to assess the decomposition quality and accuracy (seen tables I and II). Both visual and quantitative results demonstrate the superiority of the proposed method.

B. Real CT Data

In the real experiments, a mouse with gold nanoparticles (GNP) injection was scanned on a MARS (medipix all resolution system) micro-CT with a Medipix MXR CdTe layer detector. The spectrum channel number is 13 implemented by multiple scans. The SDD is 255 mm and the SOD is 158 mm. The preprocessed projections are obtained from 360 views by a detector consisting of 512 cells and each has a length of 0.11 mm. 512×512 channel images are reconstructed with a pixel size of $0.036 \times 0.036\,\mathrm{mm}^2$.

We also compare the proposed method with the conventional FBP, TV, TVLR and L_0 based methods, The corresponding decomposed material images and synthesized color images are shown in Fig. 3. From the magnified region-of-interests, it can be seen that the results with the proposed method can dramatically suppress the noise and meanwhile well maintain fine structures.

V. Conclusion

In this work, we first construct a gradient sparsity in spectral dimension by employing the locally linear transform. Then, we combine this property with the piece-wise constant prior knowledge in the spatial domain, and establish a 3D sparse gradient constraint of spectral channel reconstructions. By extending the 2D L_0-norm based measurement to a 3D version, we effectively incorporate the proposed constraint into an optimization model. Meanwhile, we develop the corresponding iterative solution strategy. Both numerical simulations and real experiments confirm the outstanding performance of the proposed method in noise suppression and structure maintenance.

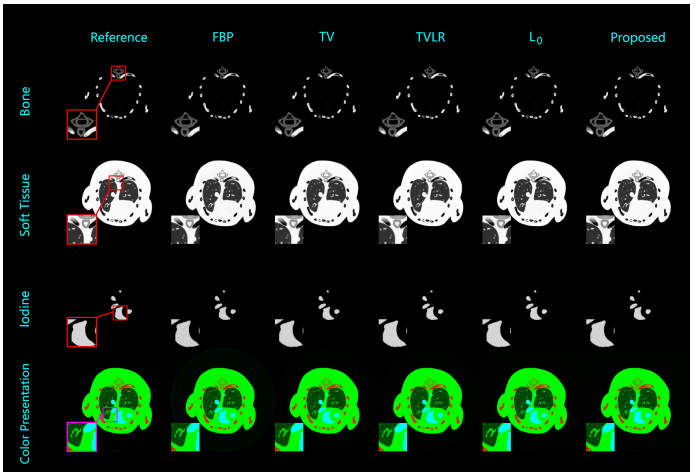


Fig. 1. Decomposed material images and synthesized color images with normal-dose projections. The display window for the first three rows are [0, 1], [0.1, 1], and [0, 0.013], respectively.

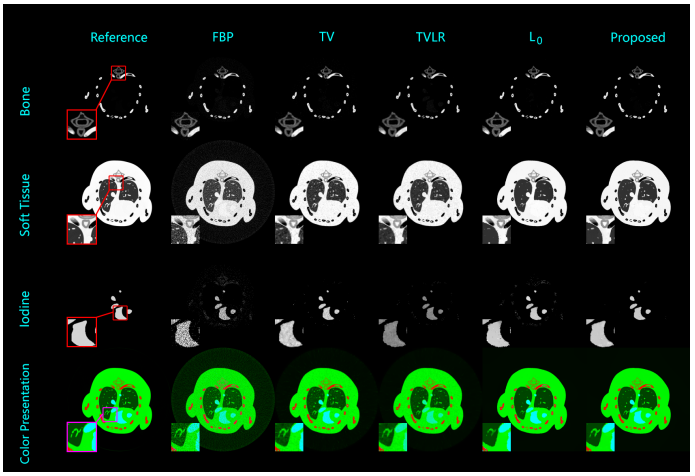


Fig. 2. Same as Fig. 1 but for low-dose projections.

TABLE I
QUANTITATIVE EVALUATION RESULTS OF THE MOUSE PHANTOM STUDIES WITH NORMAL-DOSE PROJECTIONS.

		PSNR	NMAD	SSIM
Bone	FBP	43.946	0.0986	0.9787
	TV	45.754	0.0536	0.9970
	TVLR	46.503	0.0498	0.9971
	L_0	48.005	0.0504	0.9927
	Proposed	**51.274**	**0.0320**	**0.9976**
Soft Tissue	FBP	33.847	0.0336	0.8963
	TV	36.813	0.0194	0.9694
	TVLR	37.423	0.0184	0.9698
	L_0	37.391	0.0194	0.9554
	Proposed	**39.839**	**0.0157**	**0.9702**
Iodine	FBP	40.708	0.0723	0.9927
	TV	41.501	0.0390	0.9959
	TVLR	43.993	0.0298	0.9976
	L_0	43.543	0.0313	0.9938
	Proposed	**45.641**	**0.0245**	**0.9991**

TABLE II
QUANTITATIVE EVALUATION RESULTS OF THE MOUSE PHANTOM STUDIES WITH LOW-DOSE PROJECTIONS.

		PSNR	NMAD	SSIM
Bone	FBP	33.800	0.3720	0.8223
	TV	39.443	0.1530	0.9609
	TVLR	37.566	0.1632	0.9809
	L_0	38.465	0.1226	0.9849
	Proposed	**39.903**	**0.0999**	**0.9928**
Soft Tissue	FBP	20.346	0.2074	0.2991
	TV	29.455	0.0510	0.8780
	TVLR	29.274	0.0554	0.9007
	L_0	28.795	0.0419	0.9468
	Proposed	**30.020**	**0.0417**	**0.9540**
Iodine	FBP	23.953	0.6896	0.8289
	TV	33.997	0.1342	0.9830
	TVLR	26.917	0.4494	0.9741
	L_0	29.076	0.2323	0.9578
	Proposed	**36.971**	**0.0924**	**0.9950**

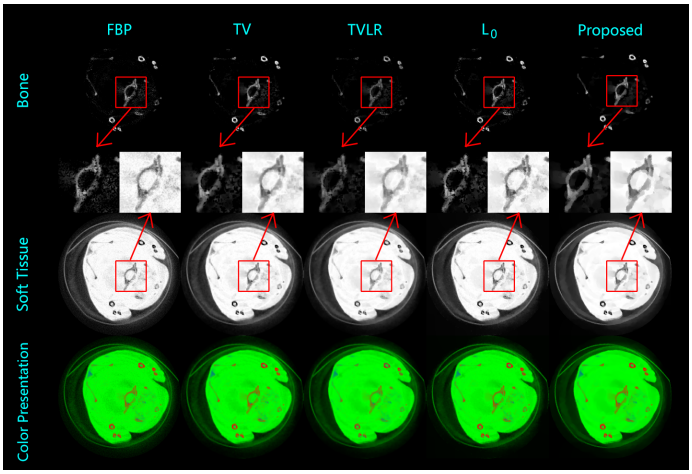


Fig. 3. Real experiments for material-based decomposition and the corresponding color images. The display window for material images is consistently [0, 1].

In the near future, we will combine noise model and high-order decomposition model into the proposed optimization framework.

ACKNOWLEDGMENT

This work was supported in part by the NIH/NIBIB U01 grant EB017140.

REFERENCES

[1] R. E. Alvarez and A. Macovski, "Energy-selective reconstructions in x-ray computerised tomography," *Physics in medicine and biology*, vol. 21, no. 5, p. 733, 1976.
[2] K. Taguchi and J. S. Iwanczyk, "Vision 20/20: Single photon counting x-ray detectors in medical imaging," *Medical physics*, vol. 40, no. 10, 2013.
[3] T. G. Schmidt, "Optimal image-based weighting for energy-resolved CT," *Medical physics*, vol. 36, no. 7, pp. 3018–3027, 2009.
[4] Q. Xu, H. Yu, J. Bennett, P. He, R. Zainon, R. Doesburg, A. Opie, M. Walsh, H. Shen, A. Butler *et al.*, "Image reconstruction for hybrid true-color micro-CT," *IEEE transactions on biomedical engineering*, vol. 59, no. 6, pp. 1711–1719, 2012.
[5] B. Zhao, H. Gao, H. Ding, and S. Molloi, "Tight-frame based iterative image reconstruction for spectral breast CT," *Medical physics*, vol. 40, no. 3, 2013.
[6] Q. Xu, H. Yu, X. Mou, L. Zhang, J. Hsieh, and G. Wang, "Low-dose X-ray CT reconstruction via dictionary learning," *IEEE Transactions on Medical Imaging*, vol. 31, no. 9, pp. 1682–1697, 2012.
[7] J. Chu, L. Li, Z. Chen, G. Wang, and H. Gao, "Multi-energy CT reconstruction based on low rank and sparsity with the split-bregman method (MLRSS)," pp. 2411–2414, 2012.
[8] Y. Zhang, X. Mou, G. Wang, and H. Yu, "Tensor-based dictionary learning for spectral CT reconstruction," *IEEE transactions on medical imaging*, vol. 36, no. 1, pp. 142–154, 2017.
[9] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 6, pp. 1397–1409, 2013.
[10] L. Xu, C. Lu, Y. Xu, and J. Jia, "Image smoothing via L0 gradient minimization," vol. 30, no. 6, p. 174, 2011.
[11] P. R. Mendonça, R. Bhotika, M. Maddah, B. Thomsen, S. Dutta, P. E. Licato, and M. C. Joshi, "Multi-material decomposition of spectral CT images," vol. 7622, p. 76221W, 2010.

Synthetic Energy Combinations for Material Decomposition in Photon Counting CT

Thomas O'Donnell, Ahmed Halaweish, Zahi Fayad, Venkatesh Mani

Abstract*— Current photon counting technologies facilitate the simultaneous acquisition of multiple energy image volumes through the specification of photon energy thresholds. Determination of the optimal choice of these energy thresholds can be challenging. For a photon counting CT machine capable of 70 threshold values and four simultaneous acquisitions, the number of possible threshold combinations is Choose(70, 4) = *916895*. In this paper we hypothesize: For the task of material decomposition, it is possible to "recycle" previously acquired (native) energy volumes to create a new synthetic threshold energy combination which has not actually been scanned. And that this synthetic scan is essentially equivalent to a native scan with the same thresholds. We test this hypothesis on four materials (Bismuth, Calcium, Iodine, and Gadolinium) using 8 native 4-threshold acquisitions. We limit ourselves to threshold (as opposed to bin) images and examine the image-based decomposition of 2-material mixtures only. We find that under certain circumstances synthetic energy combinations do an excellent job of standing in for native combinations.

Keywords—photon; counting; material; decomposition; optimal

I. Introduction

Current photon counting technologies facilitate the simultaneous acquisition of multiple energy image volumes [1]. These energy volumes may be created from the photons above some energy threshold (threshold images) or between two thresholds (bin images) [2, 3]. For the task of material decomposition, determination of the optimal choice of these energy thresholds can be challenging. To illustrate: for a photon counting CT machine capable of 70 threshold values (e.g., ranging from 20 to 90keV) and four simultaneous acquisitions, the number of possible threshold combinations is Choose(70, 4) = 916895. Clearly, many of these combinations can be easily ruled out (e.g., [21, 22, 23, 24] keV). That said, there are situations when it is helpful to evaluate several different threshold combinations. For example: when the exact chemical make-up of a material under study is not precisely known. Or, when the k-edges of the elements making up the material are known but optimizing for noise versus threshold becomes important.

T. O'Donnell and A. Halaweish are with Siemens Healthineers.
Z. Fayad and V. Mani are with Translational Molecular Imaging Inistitute, Icahn School of Medicine at Mount Sinai.

Unfortunately, evaluating different combinations of thresholds can be time consuming as compared to conventional CT, not least because multiple image volumes are reconstructed as opposed to one. In this paper we ask the question: Is it possible to create a synthetic threshold combination from the components of native scanned volumes for the task of material decomposition? That is, if we scan thresholds [20, 40, 60, 80] keV and [20, 30, 70, 80] keV, do we need to actually scan [20 40 70 80] keV to determine how well it will decompose the material under study?

In theory, a threshold volume is independent of the other threshold volumes simultaneously acquired. A 20keV image volume should be the same whether it comes from a native [20, 40, 60, 80] keV combination or a synthetic [20, 45, 65, 85] keV combination. And we found this to be true. Therefore the accuracy of synthetic threshold combinations in the material decomposition task should comparable to that of native scans.

Here, we compare the ranking, form most accurate to least accurate, of a set of native acquisitions with their synthetic counterparts. In general, we find the rankings to correlate well. However, depending on the combination of materials, the solution to the image-based decomposition equation may be less conditioned when the characteristics of a material *mixture* are dominated by *one* of the materials. In these cases, the correlation is less reliable.

Related to our work, several authors have made contributions to the topic of material decomposition for photon counting CT [4,5,6,7]. However, none looked at whether synthetic energy combinations may be substituted for actually acquired energy combinations. O'Donnell et al., [8] examined the topic of ordering the accuracy of synthetic energy combinations. However, it was as applied to material discrimination as opposed to material decomposition.

In the following sections we describe materials and methods (Section II), results (Section III), followed by discussion (Section IV) and finally conclusions (Section V).

II. Materials and Methods

A. Image Acquisition

We examined 4 materials: bismuth (Pepto-Bismol®), calcium (calcium carbonate in suspension 1250mg/5ml), iodine (Isovue 370®), and gadolinium (Gadavist®). Two ml vials containing each material alone and equal parts (50/50) mixtures of all possible pairs were created. These were placed in a water bath and scanned on a prototype photon counting scanner (CounT®, Siemens Healthineers) at 140kV at 59mAs ("ClinicalDose") and 448mAs ("MaxDose"). The following threshold combinations were acquired:

[20, 28, 53, 85]keV
[20, 28, 53, 90]keV
[20, 28, 60, 90]keV
[20, 28, 75, 85]keV
[20, 35, 53, 85]keV
[20, 35, 53, 90]keV
[20, 35, 75, 85]keV
[20, 35, 75, 90]keV

Images were reconstructed with slice thickness 0.75mm, D30f kernel.

To create synthetic energy combinations, we grouped all energy images of a single threshold and averaged them (e.g., the eight 20keV images, the four 28keV images, etc.). Synthetic versions of the actually acquired energy combinations above were then formed by collecting these averaged images into a 4-dimensional vector.

Note that the average CT-numbers of ROIs drawn in the samples were used in calculations. Thus, noise correlation of actually acquired images (not present in synthetic examples) were avoided [8].

B. Decomposition

For the decomposition task we use the linear image space approach [9]. Determining the contributions of the archetypical basis materials to a mixture takes the form of solving a system of linear equations:

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \\ a_{41} & a_{42} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix} \qquad (1)$$

where the 4x2 matrix (made up of elements a) collects each of the 2 pure basis materials' archetypical normalized CT values acquired at the 4 threshold values in columns. The x vector (to be solved for) represents the contributions of the 2 materials for a given mixture. The p vector represents the normalized CT values of the mixture under study at the 4 different energies.

III. Results

A. Stability of CT Numbers at the Same Energy Threshold

To determine how stable the CT numbers are at a given energy threshold we computed the standard deviation of the HUs of a volume of interest, for each material, over all acquired images at that energy threshold. The results at "ClinicalDose" are in Table 1. Note that the threshold 60keV was acquired only once and therefore the standard deviation is 0. This table illustrates the stability of CT values of a particular threshold over different energy combinations. The MaxDose table yields even better, results (not shown due to space restrictions).

	20keV	28keV	35keV	53keV	60keV	75keV	85keV	90keV
Bi	**2.43**	**1.21**	**1.85**	**1.27**	**0.00**	**2.45**	**2.69**	**2.69**
Ca	**1.32**	**1.18**	**1.15**	**2.01**	**0.00**	**1.93**	**3.28**	**1.04**
I	**1.04**	**0.42**	**0.79**	**0.75**	**0.00**	**1.07**	**1.86**	**4.26**
Gd	**2.10**	**2.20**	**0.49**	**0.85**	**0.00**	**2.43**	**1.15**	**2.16**
BiI	**1.69**	**2.49**	**1.21**	**0.99**	**0.00**	**4.78**	**4.90**	**3.86**
CaI	**2.21**	**2.00**	**0.82**	**1.99**	**0.00**	**3.63**	**2.05**	**4.88**
GdI	**0.93**	**0.24**	**0.55**	**0.96**	**0.00**	**1.38**	**2.26**	**3.97**
CaBi	**2.65**	**3.06**	**1.51**	**4.42**	**0.00**	**2.83**	**1.42**	**0.58**
BiGd	**1.97**	**1.91**	**2.77**	**3.48**	**0.00**	**1.31**	**7.11**	**3.20**
CaGd	**1.42**	**1.25**	**0.85**	**1.71**	**0.00**	**1.11**	**0.92**	**4.37**

Table 1: Standard deviation of CT values of the same material at the same energy threshold over different acquisitions at "ClinicalDose."

B. Correlation of Accuracy of Decomposition Between Native and Synthetic Energy Combinations

We wish to determine if the ranking of energy combinations from best to worst is the same for a set of native scans as with synthetic scans using the same thresholds. To illustrate, we look at the mixture of bismuth and gadolinium.

All of the mixtures are, in fact, composed of 50% one material, 50% the other. The material decomposition reflects this ground truth more accurately for some energy threshold combinations than others. In Table 2, the 1st and 2nd columns, and 4th and 5th columns represent the recovered percentages of bismuth and gadolinium at different energy combinations for native and synthetic combinations respectively at MaxDose. The 3rd column is the difference between columns 1 and 2. For a perfect recovery, this difference would be zero (similarly for 6th column). Thus, the 3rd and 6th columns reflect the accuracies of the decompositions. Plotting the 3rd and 6th columns against each other (as in Figure 1) allows us to measure the correlation in ranking between native and synthetic scans.

	Acquired (native)			Synthetic		
	Bi	Gd	diff	Bi	Gd	diff
[20-28-53-85]keV	52.2	47.8	-4.3	51.4	48.6	-2.8
[20-28-53-90]keV	51.6	48.4	-3.3	51.9	48.1	-3.8
[20-28-60-90]keV	52.3	47.7	-4.7	51.7	48.3	-3.3
[20-28-75-85]keV	48.8	51.2	2.4	49.3	50.7	1.4
[20-35-53-85]keV	52.9	47.1	-5.8	53.2	46.8	-6.4
[20-35-53-90]keV	51.4	48.6	-2.8	51.2	48.8	-2.5
[20-35-75-85]keV	49.4	50.6	1.2	49.6	50.4	0.8
[20-35-75-90]keV	49.7	50.3	0.6	49.9	50.1	0.3

Table 2: Comparison of native and synthetic energy threshold combinations for decomposing a 50/50 mixture of bismuth and gadolinium at MaxDose. Columns 3 and 6 describe the difference from a perfect recovery of the ground truth. Figure 1 plots column 3 versus column 6.

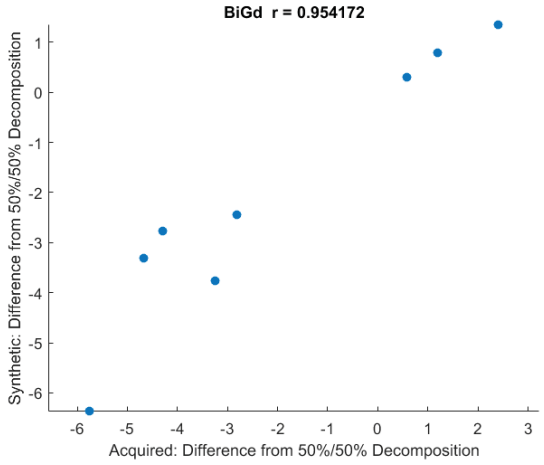


Figure 1: The correlation between the acquired (native) energy combinations and their corresponding synthetic ones for a bismuth-gadolinium mixture at MaxDose.

Table 3 collects the r-values for all of the mixtures.

	r-value	
	ClinicalDose	MaxDose
BiI	0.82	0.98
CaI	0.93	0.86
GdI	0.92	0.25
CaBi	0.79	0.58
BiGd	0.74	0.95
CaGd	0.88	0.59

Table 3: correlations for the different mixtures between native and synthetic scans at ClinicalDose and MaxDose.

C. *The Accuracy of Decomposition of the Native Scans*

From Table 3 we observe that the correlations between native and synthetic scans are poor for certain mixtures. This corresponds in part to the accuracy of the decompositions themselves. In Table 4 we show the average differences between the computed decomposition percentages of the two materials in the mixtures. Note that some (i.e, CaI, GdI, and CaBi) are far off.

	Clinical	MaxDose
BiI	17.6	19.8
CaI	46.6	47.7
GdI	-56.5	-49.8
CaBi	-58.3	-54.2
BiGd	-6.6	-2.1
CaGd	2.1	13.6

Table 4: The average accuracy over threshold combinations for native scans. The percentage of the second material is subtracted from the first, thus the appearance of negative numbers. An ideal value would be 0.

D. *An Example of Image Decomposition*

Using the MaxDose pure instances (Bi, Ca, I, Gd) as the archetypical CT-values for the materials, we decomposed ClinicalDose instances of Bi, Gd, and BiGd. Additionally we show a decomposition of MaxDose BiGd.

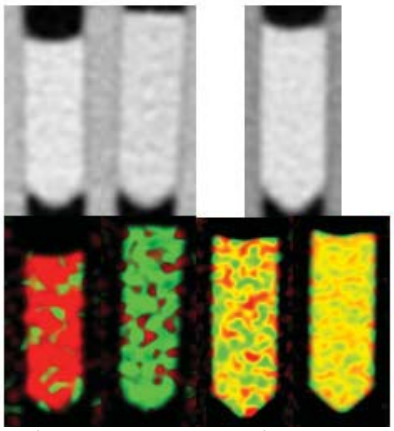


Figure 2: Top: Left to right, 20keV scans of Bi, Gd, and BiGd. Bottom: Material labelled reconstructions. Left to right [20, 35, 75, 85] keV Bi, Gd, BiGd (ClinicalDose), BiGd (MaxDose). Bismuth = red, gadolinium = green,

IV. Discussion

Given that the variance of the CT numbers is so low, as shown in Table 1, it is surprising that the correlation between the native and "identical" synthetic scans can be so different in Table 3 for certain mixtures. To examine this, we first look at the accuracy of the decompositions of the mixtures themselves (Table 4).

To see why the accuracy of the decompositions varies so widely in Table 4, we plot the *normalized centroids* of the two

component materials and their mixture for BiGd and CaBi in Figure 3. By way of explaining the figure itself: these are 4-dimensional plots, so a point (a centroid) is rendered as a series of connected line segments (see [8] for a more lengthy explanation of 4D plotting). And, by "normalized," we mean that the vector of HUs at [20, 35, 75, 90] keV are scaled to length one.

In the top panel of Figure 4, note that BiGd (yellow) lies midway between Bi (red) and Gd (green). The accuracy of this decomposition is 49.9% Bi, 50.1% Gd at MaxDose. Our object of interest, the correlation between synthetic and native material decompositions, is r=0.95 (Table 3). In the bottom panel, the mixture CaBi (magenta) practically lies on top of calcium (blue). It is almost indistinguishable from calcium. The decomposition is computed to be 78% Ca, 22% Bi at MaxDose. In this case, the correlation between synthetic and native material decompositions is r=0.58 (Table 3). The performance in Table 4 is related to the performance in Table 3 as explained below.

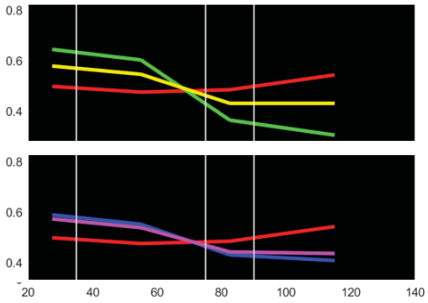


Figure 3: Top: 4 dimensional plot of the normalized centroids of Bi (red), Gd (green) and their mixture, BiGd, (yellow) for [20, 35, 75, 90] keV MaxDose. Bottom: Same for calcium (blue), and CaBi (magenta).

We make the observation that the decomposition equation (1) is less conditioned when the material to be decomposed is closer to the basis materials. To illustrate, we propose the following experiment: We have two materials A and B. Their archetypical normalized CT values are [1 0], and [0 1]. We plot the normalized vectors corresponding to mixtures (100% A, 0%B) to (0% A, 100%B) by 10 percent increments (see Figure 4).

Note that the distance between the (50% A, 50% B) and the (40% A, 60% B) is much larger than the distance between (90% A, 10% B). This indicates that noise in an ROI's CT values close to (50% A, 50% B) would result in less of a decomposition error than the same noise close to (100% A, 0% B).

Thus, we expect the correlation between native and synthetic ranking to be worse for CaI, GdI, and CaBi as compared to BiI, BiGd, and CaGd due to this effect.

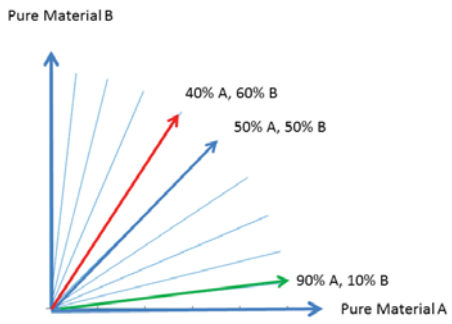


Figure 4: Ideal materials A and B and their mixtures at (100% A, 0% B) to (0%A, 100% B) by 10% increments. Note that (40% A, 60% B) is farther from (50% A, 50% B) than (90% A, 10% B) is from (0% A, 100% B). This reflects the reduced conditioning of equation (1) near the pure materials.

V. Conclusions

We have demonstrated that it is possible to substitute synthetic energy threshold combinations for native ones in order to perform material decomposition. The quality of this assessment is influenced by the how close the normalized mixture is to the archetypical basis materials. Closer mixtures are more susceptible to noise and thus the correlation between the native and "identical" synthetic combinations may be compromised.

For the thresholds we employed: 20, 28, 35, 53, 60, 75, 85, 90, there are an additional 31 combinations that may be created that were not natively acquired (e.g., [20 53 60 75] keV). The techniques described in this paper now allow these combinations to be accessed to determine if they merit further investigation.

References

[1] McCollough C, Leng S, Yu L and Fletcher J G 2015 Dual- and multi-energy CT: principles, technical approaches, and clinical applications Radiology 276(3).

[2] Pan D, Schirra C O, Senpan A, Schmieder A H, Stacy A J, Roessl E, Thran A, Wickline S A, Proska R and Lanza GM 2012 An early investigation of ytterbium nanocolloids for selective and quantitative 'multicolor' spectral CT imaging ACS Nano 6 3364–70.

[3] Anderson N G et al 2010 Spectroscopic (multi-energy) CT distinguishes iodine and barium contrast material in MICE Eur. Radiol. 20 2126–34.

[4] Nik S J 2013 Optimising the benefits of spectral x-ray imagingin material decomposition PhD Thesis University of Canterbury, Christchurch, New Zealand

[5] Wang A and Pelc N 2011 Sufficient statistics as a generalization of binning in spectral x-ray imaging IEEE Trans. Med. Imaging 30(1)

[6] Wang X, Meier D, Taguchi K, Wagenaar D J, Patt B E and Frey E C 2011 Material separation in x-ray CT with energy resolved photon-counting detectors Med. Phys. 38(3).

[7] Peng H B, Wei P F, Mianyi C and Deling M Material discrimination based on K-edge characteristics Comput. Math. Q7 Methods Med. 2013 308520, Vol. 2013.

[8] O'Donnell, T et al., "Optimal selection of thresholds for photon counting CT, SPIE Medical Imaging 2016.

[9] Rebuffel, V.; Dinten, J-M., "Dual-Energy X-Ray Imaging: Benefits and Limits," ECNDT 2006 - Th.1.3.1

Non-Convex Optimization-Based Reconstruction in Multispectral CT

Xiaochuan Pan, Buxin Chen, Emil Y. Sidky, Zheng Zhang, and Dan Xia

***Abstract*—In multi-spectral CT (MCT), the appropriate consideration of the spectral responses leads to a non-linear data model, and image reconstruction by solving the non-linear data model remains a topic of research interest. In the work, we report an optimization-based approach to reconstructing images in MCT through solving the non-linear data model. The approach entails three leading steps: identifying the non-linear component in the data model that results in the non-convexity of an optimization program, convexifying the program and tailoring an existing algorithm for convex programs to solve the convexified program, and modifying the tailored algorithm to compensate for the non-convexity of the original non-convex program. The approach may allow for the investigation of different designs of optimization programs and the associated iterative algorithms for solving the non-linear data model in MCT to yield image reconstructions of task-specific utility, and it can be exploited for enabling MCT with scanning configurations tailored to meet specific application constraints, including those special, non-standard scanning configurations of practical implication considered previously.**

I. Introduction

In multi-spectral CT (MCT), the appropriate consideration of the spectral responses leads to a non-linear data model. While methods have been investigated and developed to compensate for the non-linear spectral effect, image reconstruction remains a topic of research interest. In the work, we discuss an optimization-based approach to reconstructing images in MCT through solving the non-linear data model. The approach is a continuation of our recent works [1], [2] that reports a one-step method for image reconstruction directly from MCT data. In the approach, the image reconstruction in MCT is formulated as an optimization program with a data divergence formed by use of the non-linear data model. The specific non-linear form of the data model can lead to non-convex data divergence and consequently a non-convex optimization program. In an attempt to solve the non-convex program for image reconstruction in MCT, we first convexify the non-convex optimization program and then tailor an existing algorithm appropriate for solving the convexified program. The tailored algorithm is modified further compensate for the non-convexity of the original non-convex optimization program.

This summary is organized as follows. Following the introduction in Sec. I, we discuss in Sec. II the non-linear data model and its decomposition into linear and non-linear components. Using the data model, we form in Sec. III optimization programs, which are generally non-convex due to the non-linear component in the data model and discuss the convexification of the programs. In Sec. IV, we discuss how existing algorithms for convex programs can be tailored to solve the convexified programs and how the tailored algorithm can be modified further to compensate for the non-convexity of the original, non-convex optimization programs. Finally, the potential applications of the approach are discussed in Sec. V.

X. Pan is with the Departments of Radiology & Radiation and Cellular Oncology, The University of Chicago, Chicago, IL 60637, USA

B. Chen, E. Y. Sidky, Z. Zhang, and D. Xia are with the Department of Radiology, The University of Chicago, Chicago, IL 60637, USA

II. Non-Linear Data model

In MCT, the monochromatic linear-attenuation coefficient of interest is to be reconstructed from data collected. Let f_{mi} denote the linear-attenuation coefficient at energy m and voxel i, which can be expressed as

$$f_{mi} = \sum_{k}^{K} \mu_{km}\, b_{ki} \tag{1}$$

where b_{ki} denotes basis image k at voxel i, and μ_{km} the expansion coefficient, which is selected often as the mass-attenuation coefficient, at energy m for basis image k, $k = 1, 2, ..., K$, and K the number of basis images. For discussion convenience, we can form data and image vectors below. Let $\mathbf{b}_k$ denote basis-image vector k of size I with entries given by b_{ki}, where $i = 1, 2, ..., I$. Furthermore, we use $\mathbf{b}$ of size $K \times I$ to depict an aggregate basis-image vector formed by concatenating basis-image vectors $\mathbf{b}_k$, and simply refer to $\mathbf{b}$ as the basis image. In MCT, we assume that knowledge of μ_{km} is available. Therefore, the reconstruction of f_{mi} is tantamount to that of basis images b_{ki} (i.e., $\mathbf{b}$)

For an illumination along ray j with spectral s, model data $g_j^{[s]}$ can be written as

$$g_j^{[s]}(\mathbf{b}) = -\ln \sum_m q_{jm}^{[s]} \exp\left(-\sum_k \mu_{km} \sum_i a_{ji}^{[s]} b_{ki}\right), \tag{2}$$

where $q_{jm}^{[s]}$ denotes the normalized spectrum s at energy m, and $a_{ji}^{[s]}$ the discrete X-ray transform of ray j through voxel i. Clearly, the data model in Eq. (2) is non-linear in the sense that basis images b_{ki} are related non-linearly to model data $g_j^{[s]}$, and image reconstruction of b_{ki} is tantamount to solving the non-linear data model in Eq. (2). Basis image $\mathbf{b}$ is included in $g_j^{[s]}(\mathbf{b})$ as a variable to indicate that the reconstruction task is to determine $\mathbf{b}$ from knowledge of $g_j^{[s]}$ estimated from data measured.

III. Linear and Non-Linear Components in the Data Model

The non-linear component in the data model results in the non-convexity in the optimization program designed for image reconstruction in MCT [1], [2]. We have developed a decomposition of the data model into linear and non-linear components and exploit it to convexify the non-convex program [1], [2]. Clearly, the decomposition of the data model plays a key role in image reconstruction in MCT, as it can lead to considerably different reconstruction algorithms, as shown below. We present below a data-model decomposition that includes the previous result in Ref. [1], [2] as a special case.

We rewrite Eq. (2) as

$$g_j^{[s]}(\mathbf{b}) = -\ln \sum_m^M q_{jm}^{[s]} \exp\left(-\sum_k^K \mu_{km}\, A_{jk}^{[s]}(\mathbf{b})\right), \qquad (3)$$

where the discrete X-ray transform of basis image $\mathbf{b}$ is given by

$$A_{jk}^{[s]}(\mathbf{b}) = \sum_i^I a_{ji}^{[s]} b_{ki}. \qquad (4)$$

We perform a Taylor expansion of Eq. (3) with respect to $A_{jk}^{[s]}(\tilde{\mathbf{b}})$ for some obtainable basis image $\tilde{\mathbf{b}}$ of interest as

$$\begin{aligned} g_j^{[s]}(\mathbf{b}) &= g_j^{[s]}(\tilde{\mathbf{b}}) + \sum_k^K r_{jk}^{[s]}(\tilde{\mathbf{b}}) \sum_i^I a_{ji}^{[s]}(b_{ki} - \tilde{b}_{ki}) \\ &+ \Delta g_j^{[s]}(\mathbf{b}-\tilde{\mathbf{b}}), \end{aligned} \qquad (5)$$

where

$$r_{jk}^{[s]}(\tilde{\mathbf{b}}) = \frac{\sum_m^M q_{jm}^{[s]}\, \mu_{km} \exp\left(-\sum_k^K \mu_{km}\, A_{jk}^{[s]}(\tilde{\mathbf{b}})\right)}{\sum_m^M q_{jm}^{[s]} \exp\left(-\sum_k^K \mu_{km}\, A_{jk}^{[s]}(\tilde{\mathbf{b}})\right)} \qquad (6)$$

and

$$\begin{aligned} \Delta g_j^{[s]}(\mathbf{b}-\tilde{\mathbf{b}}) &= g_j^{[s]}(\mathbf{b}) - g_j^{[s]}(\tilde{\mathbf{b}}) \\ &+ \sum_k^K r_{jk}^{[s]}(\tilde{\mathbf{b}}) \sum_i^I a_{ji}^{[s]}(b_{ki} - \tilde{b}_{ki}). \end{aligned} \qquad (7)$$

As discussed in Ref. [2], we can form, for spectrum s, vector $\mathbf{g}^{[s]}(\mathbf{b})$ of size $J^{[s]}$, with elements $g_j^{[s]}(\mathbf{b})$ and then model-data vector $\mathbf{g}(\mathbf{b})$ of size $J = \sum_s^S J^{[s]}$ by concatenating $\mathbf{g}^{[s]}(\mathbf{b})$ for a total of S spectra. Similarly, we can also create aggregate vectors $\mathbf{g}(\tilde{\mathbf{b}})$ and $\Delta\mathbf{g}(\mathbf{b}-\tilde{\mathbf{b}})$ of size $J = \sum_s^S J^{[s]}$ with entries $g_j^{[s]}(\tilde{\mathbf{b}})$ and $\Delta g_j^{[s]}(\mathbf{b}-\tilde{\mathbf{b}})$, respectively.

We can then re-express Eq. (5) as

$$\mathbf{g}(\mathbf{b}) - \mathbf{g}(\tilde{\mathbf{b}}) - \Delta\mathbf{g}(\mathbf{b}-\tilde{\mathbf{b}}) = \mathcal{H}(\tilde{\mathbf{b}})(\mathbf{b}-\tilde{\mathbf{b}}), \qquad (8)$$

where matrix $\mathcal{H}$ of size $J \times (I \times K)$ is given by

$$\mathcal{H}(\tilde{\mathbf{b}}) = \begin{pmatrix} \mathcal{R}_1^{[1]}(\tilde{\mathbf{b}})\,\mathcal{A}^{[1]} & \mathcal{R}_2^{[1]}(\tilde{\mathbf{b}})\,\mathcal{A}^{[1]} & \cdots & \mathcal{R}_K^{[1]}(\tilde{\mathbf{b}})\,\mathcal{A}^{[1]} \\ \mathcal{R}_1^{[2]}(\tilde{\mathbf{b}})\,\mathcal{A}^{[2]} & \mathcal{R}_2^{[2]}(\tilde{\mathbf{b}})\,\mathcal{A}^{[2]} & \cdots & \mathcal{R}_K^{[2]}(\tilde{\mathbf{b}})\,\mathcal{A}^{[2]} \\ \vdots & \vdots & \ddots & \\ \mathcal{R}_1^{[S]}(\tilde{\mathbf{b}})\,\mathcal{A}^{[S]} & \mathcal{R}_2^{[S]}(\tilde{\mathbf{b}})\,\mathcal{A}^{[S]} & \cdots & \mathcal{R}_K^{[S]}(\tilde{\mathbf{b}})\,\mathcal{A}^{[S]} \end{pmatrix}, \qquad (9)$$

$\mathcal{R}_k^{[s]}(\tilde{\mathbf{b}})$ is a diagonal matrix of size $J^{[s]} \times J^{[s]}$ with diagonal element $r_{jk}^{[s]}(\tilde{\mathbf{b}})$ given in Eq. (6), and matrix $\mathcal{A}^{[s]}$ of size $J^{[s]} \times I$ with element $a_{ij}^{[s]}$ is defined in Eq. (4) above.

The data model in Eq. (8) is identical to that in Eq. (2) but has a different form in which the linear and non-linear components are explicitly decomposed. In image reconstruction, $\mathbf{g}(\mathbf{b})$ is replaced with measured data, and $\mathbf{g}(\tilde{\mathbf{b}})$ can be calculated for given $\tilde{\mathbf{b}}$ by use of Eq. (2). Furthermore, the right-hand side of Eq. (8) is linear relative to $\mathbf{b}$, and only term $\Delta\mathbf{g}(\mathbf{b}-\tilde{\mathbf{b}})$ embodies the non-linearity of the data model. Clearly, non-linear component $\Delta\mathbf{g}(\mathbf{b}-\tilde{\mathbf{b}})$ depends upon basis image $\tilde{\mathbf{b}}$, and linear component $\mathcal{H}(\tilde{\mathbf{b}})(\mathbf{b}-\tilde{\mathbf{b}})$ can also be impacted by $\tilde{\mathbf{b}}$ through matrix $\mathcal{H}(\tilde{\mathbf{b}})$. Therefore, the selection of $\tilde{\mathbf{b}}$ can impact the decomposition and thus the reconstruction to be discussed below.

When $\tilde{\mathbf{b}} = \mathbf{0}$ is selected, the decomposition in Eq. (8) and matrix in Eq. (9) become identical to their counterparts in Ref. [1], [2] even though the decomposition there was obtained from a completely different physical and mathematical perspective. Other types of $\tilde{\mathbf{b}}$ can also be explored. For example, one could use some prior images of relevance as $\tilde{\mathbf{b}}$. Or one can even use $\mathbf{b}^{(n)}$ obtained in the previous iteration as $\tilde{\mathbf{b}}$ in an iterative reconstruction. In the case, system matrix $\mathcal{H}(\tilde{\mathbf{b}})$ will thus vary from the previous iteration to the next iteration.

IV. Non-Convex Optimization Programs

It is unclear whether Eq. (2) can directly be inverted adequately for achieving the reconstruction of basis image $\mathbf{b}$ especially considering the large sizes of data and images involved in practical MCT. Instead, an optimization program is formulated for reconstruction of basis image $\mathbf{b}$. Optimization programs of various unconstraint or constraint forms can be devised, and the specific non-linearity of the data model in Eq. (2) generally leads to non-convex optimization programs.

Without loss of generality, we consider a non-convex constraint optimization program below and use it to illustrate the approach to dealing with a non-convex program:

$$\mathbf{b}^* = \underset{\mathbf{b}}{\operatorname{argmin}}\, D(\mathbf{g}(\mathbf{b}); \mathbf{g}_\mathcal{M}) \qquad \text{s.t.} \quad \Psi(\mathbf{b}), \qquad (10)$$

where aggregate vector $\mathbf{g}_\mathcal{M}$ of size $J = \sum_s^S J^{[s]}$ with entry $g_{\mathcal{M}j}^{[s]}$ denoting data measured with ray j and spectrum s; $D(\mathbf{g}(\mathbf{b}); \mathbf{g}_\mathcal{M})$ the data divergence between model data $\mathbf{g}(\mathbf{b})$ and measured data $\mathbf{g}_\mathcal{M}$; and $\Psi(\mathbf{b})$ a constraint set on basis images to be reconstructed.

For measured data $\mathbf{g}_\mathcal{M}$, data divergences of different forms can be devised, while a design of data divergence is motivated often by the specific statistic property of measured data. Data divergences of form below have been considered for image reconstruction:

$$D_{l_2}(\mathbf{g}(\mathbf{b}), \mathbf{g}_\mathcal{M}) = ||\mathbf{g}(\mathbf{b}) - \mathbf{g}_\mathcal{M}||_2 \qquad (11)$$

$$D_{l_1}(\mathbf{g}(\mathbf{b}), \mathbf{g}_\mathcal{M}) = ||\mathbf{g}(\mathbf{b}) - \mathbf{g}_\mathcal{M}||_1 \qquad (12)$$

$$\begin{aligned} D_{\mathrm{KL}}(\mathbf{g}(\mathbf{b}), \mathbf{g}_\mathcal{M}) &= \sum_i \big[\mathbf{g}(\mathbf{b}) - \mathbf{g}_\mathcal{M} \\ &+ \mathbf{g}_\mathcal{M} \ln(\mathbf{g}_\mathcal{M}) - \mathbf{g}_\mathcal{M} \ln(\mathbf{g}(\mathbf{b}))\big]_i \end{aligned} \qquad (13)$$

where D_{l_2} and D_{l_1} denote the l_2-norm and l_1-norm of the differences between model data and measured data, whereas

D_{KL} the Kullback-Leibler (KL) divergence of model and measured data. Clearly, each of the data divergences above is non-convex due to the non-linearity of the data model in Eq. (2). Therefore, when one of them is used in Eq. (10), the optimization program is thus non-convex, even if the set of image constraints, i.e., $\Psi(\mathbf{b})$, is convex.

V. Convexification and Algorithms

It remains challenging to derive an algorithm that solves mathematically the non-convex optimization programs discussed above. We discuss below a strategy for convexifying a non-convex program, then develop an algorithm to solve the program convexified, and finally modify the algorithm to compensate for the non-convexity of the original non-convex program. Without loss of generality, we consider the program in Eq. (10) in which the data divergence is one of those in Eqs. (11)-(13) and the image-constraint set is convex.

Using Eq. (8), we can obtain an optimization program that is equivalent to that in Eq. (10):

$$\mathbf{b}^* = \underset{\mathbf{b}}{\operatorname{argmin}}\, D(\mathcal{H}(\tilde{\mathbf{b}})\mathbf{b};\, \mathbf{g}'_{\mathcal{M}} - \Delta\mathbf{g}(\mathbf{b}-\tilde{\mathbf{b}}))\ \text{s.t.}\ \Psi(\mathbf{b}), \quad (14)$$

where

$$\mathbf{g}'_{\mathcal{M}} = \mathbf{g}_{\mathcal{M}} - \mathbf{g}(\tilde{\mathbf{b}}) + \mathcal{H}(\tilde{\mathbf{b}})\tilde{\mathbf{b}}, \quad (15)$$

and can readily be obtained from $\mathbf{g}_{\mathcal{M}}$ and $\tilde{\mathbf{b}}$ with Eqs. (2) and (9). It can be observed that the non-convexity of the program in Eq. (14) stems from term $\Delta\mathbf{g}(\mathbf{b}-\tilde{\mathbf{b}})$, and that if the term is known, the program become convex, and existing algorithms for convex programs can thus be applied to solving such a convex program. Therefore, we convexify the program in Eq. (14) by replacing $\Delta\mathbf{g}(\mathbf{b}-\tilde{\mathbf{b}})$ with constant term $\overline{\Delta\mathbf{g}}$, and then tailor an existing algorithm to solve the convexified program. Without loss of generality, let the update procedure of the algorithm selected to solve the convexified program be written as

$$\mathbf{b}^{(n+1)} = O(\mathbf{b}^{(n)}, \mathbf{g}'_{\mathcal{M}}, \overline{\Delta\mathbf{g}}), \quad (16)$$

where n indicates the iteration number, and function $O(\mathbf{b}^{(n)}, \mathbf{g}'_{\mathcal{M}}, \overline{\Delta\mathbf{g}})$ can be calculated from knowledge of $\mathbf{b}^{(n)}$, $\mathbf{g}'_{\mathcal{M}}$, and $\overline{\Delta\mathbf{g}}$.

In an attempt to compensate for the non-convex effect that stems from $\Delta\mathbf{g}(\mathbf{b}-\tilde{\mathbf{b}})$, we propose to replace $\overline{\Delta\mathbf{g}}$ with $\Delta\mathbf{g}(\mathbf{b}^{(n)}-\tilde{\mathbf{b}})$ in Eq. (16), leading to an algorithm with an update form

$$\mathbf{b}^{(n+1)} = O(\mathbf{b}^{(n)}, \mathbf{g}'_{\mathcal{M}}, \Delta\mathbf{g}(\mathbf{b}^{(n)}-\tilde{\mathbf{b}})), \quad (17)$$

where entry specified by j and s of vector $\Delta\mathbf{g}(\mathbf{b}^{(n)}-\tilde{\mathbf{b}})$ is computed by use of Eq. (7) with $\mathbf{b}^{(n)}$ as

$$\begin{aligned}\Delta g_j^{[s]}(\mathbf{b}^{(n)}-\tilde{\mathbf{b}}) &= g_j^{[s]}(\mathbf{b}^{(n)}) - g_j^{[s]}(\tilde{\mathbf{b}}) \\ &\quad + \sum_k^K r_{jk}^{[s]}(\tilde{\mathbf{b}})[A_{jk}^{[s]}(\mathbf{b}^{(n)}) - A_{jk}^{[s]}(\tilde{\mathbf{b}})]. \quad (18)\end{aligned}$$

Therefore, we obtain an algorithm, specified by Eqs. (17) and (18), to solve the non-convex optimization program in Eq. (10) for image reconstruction in MCT.

For a non-convex program with $\tilde{\mathbf{b}} = \mathbf{0}$ and $D_{l_2}(\mathbf{g}(\mathbf{b}), \mathbf{g}_{\mathcal{M}})$ in Eq. (11), we have developed a non-convex (NC) projection-onto-convex-sets (POCS) algorithm for lowering $D_{l_2}(\mathbf{g}(\mathbf{b}), \mathbf{g}_{\mathcal{M}})$. Quantitative studies suggest that, in combination with a steepest descend (SD) algorithm for reducing the image TV, the NC-POCS can solve the program numerically accurately for imaging conditions considered [2]. It is worthy and straightforward to investigate a NC-POCS algorithm applicable to the case of non-zero $\tilde{\mathbf{b}}$. This can be accomplished simply by replacing matrix $\mathcal{U}_k^{[s]}$ in the previous NC-POCS algorithm with $\mathcal{R}_k^{[s]}(\tilde{\mathbf{b}})$ defined in Eq. (9) above. However, when data divergence such as $D_{l_1}(\mathbf{g}(\mathbf{b}), \mathbf{g}_{\mathcal{M}})$ or $D_{\text{KL}}(\mathbf{g}(\mathbf{b}), \mathbf{g}_{\mathcal{M}})$ is considered, and/or when the optimization program in Eq. (10) is non-smooth, algorithms other than the NC-POCS algorithm need to be developed. For example, for the convexified program (even if it is non-smooth,) algorithms such as the primal-dual algorithm can be exploited for developing NC-primal-dual algorithms. Specifically, for a non-convex optimization program containing one of the data divergences in Eqs. (11)-(13), an instance of the primal-dual algorithm can readily be derived for its convexified program with constant term $\overline{\Delta\mathbf{g}}$. We can then replace $\overline{\Delta\mathbf{g}}$ with $\Delta\mathbf{g}(\mathbf{b}^{(n)}-\tilde{\mathbf{b}})$ in the specific algorithm instance to obtain a NC primal-dual algorithm for solving the non-convex program.

It is necessary to devise convergence conditions for the algorithm with an update procedure in Eq. (17). Clearly, its convergence depends upon the explicit forms of data divergence, image constraints, the update procedure, and various parameters involved. Without loss of generality, we consider a specific image constraint $\Psi(\mathbf{b})$ on the image TV, i.e., $||\mathbf{b}||_{\text{TV}} \leq t$, where $t > 0$ is a constraint parameter, and we can design two necessary convergence conditions as

$$\begin{aligned}\bar{D}(\mathbf{b}^{(n)}) &= \frac{\left|D(\mathbf{g}(\mathbf{b}^{(n+1)}), \mathbf{g}_{\mathcal{M}}) - D(\mathbf{g}(\mathbf{b}^{(n)}, \mathbf{g}_{\mathcal{M}})\right|}{\left|D(\mathbf{g}(\mathbf{b}^{(n+1)}), \mathbf{g}_{\mathcal{M}}) + D(\mathbf{g}(\mathbf{b}^{(n)}, \mathbf{g}_{\mathcal{M}})\right|} \to 0 \\ \bar{\Delta}_{\text{TV}}(\mathbf{b}^{(n)}) &= \Big|\,||\mathbf{b}^{(n)}||_{\text{TV}} - t\Big|/t \to 0.\end{aligned} \quad (19)$$

Additional necessary conditions such as the Karush-Kuhn-Tucker (KKT) conditions can be devised [3]. While the convergent conditions in Eq. (19) and additional KKT conditions represent necessary conditions for local or global minima, it remains unclear whether the algorithm with an update procedure in Eq. (17) can mathematically reach truly local and global minima of the non-convex optimization program in Eq. (10). In the work, we also seek to use the convergence conditions to aid the selection of algorithm parameters, i.e., convergence paths, to avoid saddle points that can lead to significantly divergent solutions. Furthermore, it is important also to appreciate that the reconstruction quality of interest is not a metric that would show whether the reconstruction is a truly local or global minimum; instead, it is characterized by, in addition to the practical convergence conditions, visual inspection and task-specific metrics.

VI. Discussion

In this summary, we report an approach to solving the nonlinear data model for achieving basis-image reconstruction in

MCT. The approach entails three leading steps: identifying the non-linear component in the data model that results in the non-convexity of an optimization program, convexifying the program and tailoring an existing algorithm for convex programs to solve the convexified program, and modifying the tailored algorithm to compensate for the non-convexity of the original non-convex program. We illustrate the three steps in detail by using an optimization program in a general constraint form containing data divergences and image constraints that are used widely for image reconstruction. In particular, one of our previous works on image reconstruction in MCT is a special case of the approach investigated here. The approach allows for the investigation of different designs of optimization programs and the associated iterative algorithms for solving the non-linear data model in MCT to yield image reconstructions of task-specific utility [4]. More importantly, the approach can be exploited for enabling MCT with scanning configurations tailored to meet specific application constraints, including those special, non-standard scanning configurations of practical implication considered in our previous work [2]. We plan to report some of the application details of the approach at the conference.

ACKNOWLEDGMENT

This work was supported in part by NIH R01 Grants Nos. CA182264 and EB018102. The contents of this article are solely the responsibility of the authors and do not necessarily represent the official NIH views.

REFERENCES

[1] X. Pan, B. Chen, Z. Zhang, E. Pearson, E. Sidky, and X. Han, "Optimization-based Reconstruction Exploiting Spectral Information in CT," in *The Third International Conference on Image Formation in X-Ray Computed Tomography*, 2014, pp. 228–232.

[2] B. Chen, Z. Zhang, E. Y. Sidky, D. Xia, and X. Pan, "Image reconstruction and scan configurations enabled by optimization-based algorithms in multispectral CT," *Physics in Medicine & Biology*, vol. 62, no. 22, p. 8763, 2017. [Online]. Available: http://stacks.iop.org/0031-9155/62/i=22/a=8763

[3] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK ; New York: Cambridge University Press, 2004.

[4] B. Chen, Z. Zhang, D. Xia, E. Y. Sidky, and X. Pan, "Algorithm-enabled partial-angular-scan configurations for dual-energy ct," *Medical Physics*, (in review), 2018.

Non-Convex Chambolle-Pock Algorithm for Multispectral CT

Buxin Chen, Zheng Zhang, Dan Xia, Emil Y. Sidky, and Xiaochuan Pan

***Abstract*—Chambolle-Pock (CP) algorithm is a first-order primal-dual algorithm for convex optimization problems. It has been used in CT and PET image reconstruction based on linear data model. Non-convex optimization programs, such as those based on a non-linear data model in multispectral CT, can not be solved by applying the CP algorithm. In this work, we propose a non-convex CP (ncCP) algorithm, inspired by our previous work on an algorithm for a non-convex optimization program of a specific form in multispectral CT. The proposed ncCP algorithm can deal with different optimization program designs including non-smooth objectives and/or constraints and also involves relatively fewer parameters for convergence. An algorithm instance for an example of non-convex optimization program used in multispectral CT reconstruction has been derived, together with the convergence conditions. A numerical study using computer simulation data with two spectra has been carried out for verifying the numerical convergence of the proposed ncCP algorithm.**

I. Introduction

Chambolle-Pock (CP) algorithm is a first-order primal-dual algorithm for convex optimization problems, including non-smooth objectives, unconstrained and constrained formulations [1], [2]. It has been used for CT image reconstruction [2]–[6] and PET image reconstruction [7] based on linear data models, for its prototyping ability to bring different convex optimization program designs to convergence and the relatively fewer parameters involved in the algorithm. Non-convex optimization programs can not be solved by simply applying the CP algorithms. However, inspired by our previous work on the ASD-NC-POCS algorithm [8] where the ASD-POCS algorithm for a convex problem in a specific form is combined with a non-linear correction step, in this work we propose non-convex CP (ncCP) algorithms for non-convex optimization programs based on the non-linear data model in multispectral CT. The proposed ncCP algorithm can deal with different optimization program designs including non-smooth objectives and/or constraints, while an example of non-convex optimization program is used in this work to demonstrate the design and mechanics of the proposed algorithm. Computer simulation data are used to verify the numerical convergence of the ncCP algorithm derived.

B. Chen, Z. Zhang, D. Xia, and E. Y. Sidky are with the Department of Radiology, The University of Chicago, Chicago, IL 60637, USA

X. Pan is with the Departments of Radiology & Radiation and Cellular Oncology, The University of Chicago, Chicago, IL 60637, USA

II. Materials and methods

A. Discrete-to-discrete Data Model

Using a basis decomposition, the discrete-to-discrete data model in multispectral CT can be expressed as [8],

$$g_j^{[s]}(\mathbf{b}) = -\ln \sum_m q_{jm}^{[s]} \exp\left(-\sum_k \mu_{km} \sum_i a_{ji}^{[s]} b_{ki}\right), \quad (1)$$

where $q_{jm}^{[s]}$ denotes the normalized X-ray spectrum at energy sampling m for ray j in spectral set s, μ_{km} the mass attenuation coefficient at energy m for basis material k, $a_{ji}^{[s]}$ the intersection length of ray j in spectral set s with voxel i, and b_{ki} the basis image of material k at voxel i. The indices $m \in 1, \cdots, M$, $k \in 1, \cdots, K$, $i \in 0, \cdots, I-1$ index energy, basis material, and voxels. $\mathbf{b}$ denotes an aggregate basis-image vector formed by concatenating individual basis-image vectors $\mathbf{b}_k$, whose elements are b_{ki}, in the ascending order of k, and $g_j^{[s]}(\mathbf{b})$ the model data for ray j in spectral set s as a function of the basis image $\mathbf{b}$.

Considering all of the measurements within spectral set s, we form vector $\mathbf{g}^{[s]}(\mathbf{b})$ of size $J^{[s]}$, with elements $g_j^{[s]}(\mathbf{b})$, and further obtain an aggregate model data vector $\mathbf{g}(\mathbf{b})$ by concatenating $\mathbf{g}^{[s]}(\mathbf{b})$ in the ascending order of s. Similarly, we use $g_{\mathcal{M}j}^{[s]}$ to denote the measured data for ray j in spectral set s, and form vectors $\mathbf{g}_{\mathcal{M}}^{[s]}$ and an aggregate vector $\mathbf{g}_{\mathcal{M}}$.

B. Non-convex Optimization Program

Based on the non-linear data model in equation (1), we form the following non-convex optimization program by minimizing the ℓ_2 distance between the model and measured data subject to constraints on a monochromatic image, as

$$\begin{aligned} \min_{\mathbf{b}} \ & \frac{1}{2}||\mathbf{g}(\mathbf{b}) - \mathbf{g}_{\mathcal{M}}||_2^2 \\ \text{s.t.} \ & ||(|\nabla(\mu_1(E_0)\mathbf{b}_1 + \mu_2(E_0)\mathbf{b}_2)|)||_1 \leq \gamma, \\ & \mu_1(E_0)\mathbf{b}_1 + \mu_2(E_0)\mathbf{b}_2 \succeq 0, \end{aligned} \quad (2)$$

where, without loss of generality, a two-basis decomposition model is used with $\mathbf{b} = (\mathbf{b}_1^\top, \mathbf{b}_2^\top)^\top$ and E_0 is the energy level at which the monochromatic image is constrained with total-variation (TV) and non-negativity; $\mu_1(E_0)$ and $\mu_2(E_0)$ are mass attenuation coefficients of the two basis materials at energy E_0, which will be simply referred at as μ_1 and μ_2 hereafter. $|\nabla u|$ is the gradient-magnitude image of u, where ∇ is a linear transform of finite differencing [2].

C. Non-convex Chambolle-Pock (CP) Algorithm

As the optimization problem in equation (2) is non-convex, the primal-dual CP algorithm can not be applied to solving the problem. Instead, we derive the CP algorithm instance for a convex problem which is based on the linear term in the data model in equation (1), apply the non-linear term correction to the data at the end of each iteration, and yield a heuristic algorithm, referred to as non-convex CP algorithm, or simply ncCP, which intends to numerically solve the non-convex optimization program in equation (2).

1) CP algorithm based on the linear model: We first derive the CP algorithm instance of a convex optimization problem which is based on the linear term the data model in equation (1). As we split the mass-attenuation coefficient μ_{km} into [8]–[10]

$$\mu_{km} = \bar{\mu}_{jk}^{[s]} + \Delta\mu_{jkm}^{[s]}, \tag{3}$$

where

$$\bar{\mu}_{jk}^{[s]} = \sum_m q_{jm}^{[s]}\mu_{km}, \quad \text{and} \quad \Delta\mu_{jkm}^{[s]} = \mu_{km} - \bar{\mu}_{jk}^{[s]}, \tag{4}$$

notice that $\bar{\mu}_{jk}^{[s]}$ is independent of energy as it is a spectrum-weighted average of μ_{km} over energy while $\Delta\mu_{jkm}^{[s]}$ remains energy dependent. Plugging equation (3) into equation (1) splits the data model into two terms as well, as

$$g_j^{[s]}(\mathbf{b}) = \bar{g}_j^{[s]}(\mathbf{b}) + \Delta g_j^{[s]}(\mathbf{b}), \tag{5}$$

where

$$\bar{g}_j^{[s]}(\mathbf{b}) = \sum_k \bar{\mu}_{jk}^{[s]} \sum_i a_{ji}^{[s]} b_{ki}, \tag{6}$$

and

$$\Delta g_j^{[s]}(\mathbf{b}) = -\ln\sum_m q_{jm}^{[s]} \exp\left(-\sum_k \Delta\mu_{jkm}^{[s]} \sum_i a_{ji}^{[s]} b_{ki}\right), \tag{7}$$

are a linear function and a non-linear function of $\mathbf{b}$, respectively. In the same concatenating fashion as aforementioned, two aggregate vectors $\bar{\mathbf{g}}(\mathbf{b})$ and $\Delta\mathbf{g}(\mathbf{b})$ can be formed and are connected as

$$\mathbf{g}(\mathbf{b}) - \Delta\mathbf{g}(\mathbf{b}) = \bar{\mathbf{g}}(\mathbf{b}) = \mathcal{H}\mathbf{b}, \tag{8}$$

where the linear system matrix $\mathcal{H}$ is defined as

$$\mathcal{H} = \begin{pmatrix} \mathcal{U}_1^{[1]}\mathcal{A}^{[1]} & \mathcal{U}_2^{[1]}\mathcal{A}^{[1]} & \cdots \\ \mathcal{U}_1^{[2]}\mathcal{A}^{[2]} & \mathcal{U}_2^{[2]}\mathcal{A}^{[2]} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}, \tag{9}$$

where matrix $\mathcal{A}^{[s]}$, of size $J^{[s]} \times I$ and with element $a_{ji}^{[s]}$, denotes the discrete X-ray transform for all measurements made in spectral set s, and $\mathcal{U}_k^{[s]}$ a diagonal matrix of size $J^{[s]}$ with diagonal elements $\bar{\mu}_{jk}^{[s]}$.

Note that, if the non-linear term $\Delta\mathbf{g}(\mathbf{b})$ were known, we can apply the non-linear correction by subtracting it from the measured data and use a linear model for the reconstruction with the corrected data. In reality, we can estimate the non-linear term with the current iteration of basis image $\mathbf{b}^{(n)}$ and subtract $\Delta\mathbf{g}(\mathbf{b}^{(n)})$ from the measured data iteratively [8], which is used as a non-linear correction step in addition to a algorithm for the convex problem based on the linear model. As a result, the rest is to derive a CP algorithm for the following constrained convex optimization program,

$$\begin{aligned} &\min_{\mathbf{b}} \ \frac{1}{2}||\mathcal{H}\mathbf{b} - \mathbf{g}_{\mathcal{M}}||_2^2 \\ &\text{s.t. } ||(|\nabla(\mu_1\mathbf{b}_1 + \mu_2\mathbf{b}_2)|)||_1 \le \gamma, \\ &\quad \mu_1\mathbf{b}_1 + \mu_2\mathbf{b}_2 \succeq 0, \end{aligned} \tag{10}$$

or, equivalently,

$$\begin{aligned} \min_{\mathbf{b}} \ &\frac{\lambda}{2}||\mathcal{H}\mathbf{b} - \mathbf{g}_{\mathcal{M}}||_2^2 \\ &+ \delta_{\text{diamond}(\alpha\gamma)}(\alpha|\nabla(\mu_1\mathbf{b}_1 + \mu_2\mathbf{b}_2)|) \\ &+ \delta_P(\beta(\mu_1\mathbf{b}_1 + \mu_2\mathbf{b}_2)), \end{aligned} \tag{11}$$

where $\delta_S(x)$ is an indicator function defined to be zero if $x \in S$ or ∞ if $x \notin S$ [2], the set P is all vectors with non-negative components, and λ, α, and β are tunable algorithm parameters that do not change the solution to the optimization problem, but can impact the path and convergence rate of the algorithm. α and β are pre-defined in this work as

$$\alpha = \frac{||\mathcal{H}||_2}{||(\mu_1\nabla, \mu_2\nabla)||_2}, \quad \beta = \frac{||\mathcal{H}||_2}{||(\mu_1\mathcal{I}, \mu_2\mathcal{I})||_2}. \tag{12}$$

The optimization program in equation (11) can be fitted into the primal minimization problem in the Chambolle-Pock (CP) algorithm [1], [2] as

$$\min_{\mathbf{x}} \ \{F(\mathcal{K}\mathbf{x}) + G(\mathbf{x})\} \tag{13}$$

where

$$\begin{aligned} F(\mathcal{K}\mathbf{x}) &= F(\mathbf{u}, \mathbf{v}, \mathbf{w}) = F_1(\mathbf{u}) + F_2(\mathbf{v}) + F_3(\mathbf{w}), \\ F_1(\mathbf{u}) &= \frac{\lambda}{2}||\mathbf{u} - \mathbf{g}||_2^2, \\ F_2(\mathbf{v}) &= \delta_{\text{diamond}(\alpha\gamma)}(|\mathbf{v}|), \\ F_3(\mathbf{w}) &= \delta_P(\mathbf{w}), \\ G(\mathbf{x}) &= 0, \end{aligned} \tag{14}$$

and

$$\begin{aligned} \mathbf{x} &= \mathbf{b}, \mathbf{y} = \mathcal{K}\mathbf{b} = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \\ \mathbf{w} \end{pmatrix}, \\ \mathbf{u} &= \mathcal{H}\mathbf{b}, \\ \mathbf{v} &= (\alpha\mu_1\nabla, \alpha\mu_2\nabla)\mathbf{b}, \\ \mathbf{w} &= (\beta\mu_1\mathcal{I}, \beta\mu_2\mathcal{I})\mathbf{b}, \\ \mathcal{K} &= \begin{pmatrix} \mathcal{H} \\ \alpha\mu_1\nabla, \alpha\mu_2\nabla \\ \beta\mu_1\mathcal{I}, \beta\mu_2\mathcal{I} \end{pmatrix}. \end{aligned} \tag{15}$$

As a result, through deriving the convex conjugates of F and G and deriving the proximal mappings of F^* and G, we can derive the CP algorithm instance for solving the convex optimization program in equation (11). The pseudo-code is listed in Algorithm 1 below. The operation $\text{ProjectOnto}\ell_1\text{Ball}_{\gamma_0}(\mathbf{x})$ is simply in a form of ℓ_2 projection onto the ℓ_1-ball of size γ_0 [11]. The mathematical convergence conditions for the CP

Algorithm 1 Pseudo-code of the CP algorithm instance based on the linear model for solving equation (11)

1: $\alpha \leftarrow ||\mathcal{H}||_2/||(\mu_1\nabla, \mu_2\nabla)||_2$;
   $\beta \leftarrow ||\mathcal{H}||_2/||(\mu_1\mathcal{I}, \mu_2\mathcal{I})||_2$; Select λ
2: $L \leftarrow ||\mathcal{K}(\alpha, \beta)||_2$; $\sigma \leftarrow 1/L$; $\tau \leftarrow 1/L$; $\theta \leftarrow 1$; $n \leftarrow 0$
3: Initialize $\mathbf{b}_0, \mathbf{p}_0, \mathbf{q}_0$ and $\mathbf{r}_0$ to zero
4: $\bar{\mathbf{b}}_0 \leftarrow \mathbf{b}_0$
5: **repeat**
6: $\quad \mathbf{p}^{(n+1)} = \dfrac{\mathbf{p}^{(n)} + \sigma(\mathcal{H}\bar{\mathbf{b}}^{(n)} - \mathbf{g}_{\mathcal{M}})}{1 + \sigma/\lambda}$
7: $\quad \mathbf{q}'^{(n)} = \mathbf{q}^{(n)} + \sigma\alpha\nabla(\mu_1\bar{\mathbf{b}}_1^{(n)} + \mu_2\bar{\mathbf{b}}_2^{(n)})$
8: $\quad \mathbf{q}^{(n+1)} = \mathbf{q}'^{(n)} - \sigma\dfrac{\mathbf{q}'^{(n)}}{|\mathbf{q}'^{(n)}|}\mathrm{ProjectOnto}\ell_1\mathrm{Ball}_{\alpha\gamma}(\dfrac{|\mathbf{q}'^{(n)}|}{\sigma})$
9: $\quad \mathbf{r}^{(n+1)} = \mathrm{neg}(\mathbf{r}^{(n)} + \sigma\beta(\mu_1\bar{\mathbf{b}}_1^{(n)} + \mu_2\bar{\mathbf{b}}_2^{(n)}))$
10: $\quad \mathbf{b}^{(n+1)} = \mathbf{b}^{(n)} - \tau\mathcal{K}^\top \begin{pmatrix} \mathbf{p}^{(n+1)} \\ \mathbf{q}^{(n+1)} \\ \mathbf{r}^{(n+1)} \end{pmatrix}$
11: $\quad \bar{\mathbf{b}}^{(n+1)} = \mathbf{b}^{(n+1)} + \theta(\mathbf{b}^{(n+1)} - \mathbf{b}^{(n)})$
12: **until** practical convergence conditions are satisfied

Algorithm 2 Pseudo-code of the ncCP algorithm instance based on the non-linear model for solving equation (2)

1: $\alpha \leftarrow ||\mathcal{H}||_2/||(\mu_1\nabla, \mu_2\nabla)||_2$;
   $\beta \leftarrow ||\mathcal{H}||_2/||(\mu_1\mathcal{I}, \mu_2\mathcal{I})||_2$; Select λ
2: $L \leftarrow ||\mathcal{K}(\alpha, \beta)||_2$; $\sigma \leftarrow 1/L$; $\tau \leftarrow 1/L$; $\theta \leftarrow 1$; $n \leftarrow 0$
3: Initialize $\mathbf{b}_0, \mathbf{p}_0, \mathbf{q}_0$ and $\mathbf{r}_0$ to zero
4: $\bar{\mathbf{b}}_0 \leftarrow \mathbf{b}_0$, $\mathbf{g}_{\mathrm{LN}} \leftarrow \mathbf{g}_{\mathcal{M}}$
5: **repeat**
6: $\quad \mathbf{p}^{(n+1)} = \dfrac{\mathbf{p}^{(n)} + \sigma(\mathcal{H}\bar{\mathbf{b}}^{(n)} - \mathbf{g}_{\mathrm{LN}})}{1 + \sigma/\lambda}$
7: $\quad \mathbf{q}'^{(n)} = \mathbf{q}^{(n)} + \sigma\alpha\nabla(\mu_1\bar{\mathbf{b}}_1^{(n)} + \mu_2\bar{\mathbf{b}}_2^{(n)})$
8: $\quad \mathbf{q}^{(n+1)} = \mathbf{q}'^{(n)} - \sigma\dfrac{\mathbf{q}'^{(n)}}{|\mathbf{q}'^{(n)}|}\mathrm{ProjectOnto}\ell_1\mathrm{Ball}_{\alpha\gamma}(\dfrac{|\mathbf{q}'^{(n)}|}{\sigma})$
9: $\quad \mathbf{r}^{(n+1)} = \mathrm{neg}(\mathbf{r}^{(n)} + \sigma\beta(\mu_1\bar{\mathbf{b}}_1^{(n)} + \mu_2\bar{\mathbf{b}}_2^{(n)}))$
10: $\quad \mathbf{b}^{(n+1)} = \mathbf{b}^{(n)} - \tau\mathcal{K}^\top \begin{pmatrix} \mathbf{p}^{(n+1)} \\ \mathbf{q}^{(n+1)} \\ \mathbf{r}^{(n+1)} \end{pmatrix}$
11: $\quad \bar{\mathbf{b}}^{(n+1)} = \mathbf{b}^{(n+1)} + \theta(\mathbf{b}^{(n+1)} - \mathbf{b}^{(n)})$
12: $\quad \mathbf{g}_{\mathrm{LN}} = \mathbf{g}_{\mathcal{M}} - \Delta\mathbf{g}(\mathbf{b}^{(n+1)})$
13: **until** practical convergence conditions are satisfied

algorithm include the conditional primal-dual gap, defined as

$$\begin{aligned}\mathrm{cPD}(\mathbf{b}^*, \mathbf{p}^*, \mathbf{q}^*) =& \frac{\lambda}{2}||\mathcal{H}\mathbf{b}^* - \mathbf{g}_{\mathcal{M}}||_2^2 + \frac{1}{2\lambda}||\mathbf{p}^*||_2^2 \\ & - \mathbf{g}^\top\mathbf{p}^* + \alpha\gamma||(|\mathbf{q}^*|)||_\infty,\end{aligned} \tag{16}$$

tending to zero and the following conditions are satisfied

$$\begin{aligned}||\alpha(|(\mu_1\nabla, \mu_2\nabla)\mathbf{b}^*)|)||_1 &\leq \alpha\gamma, \\ (\beta\mu_1\mathcal{I}, \beta\mu_2\mathcal{I})\mathbf{b}^* &\succeq 0, \\ \mathbf{r}^* &\preceq 0, \\ \mathcal{K}^\top \begin{pmatrix} \mathbf{p}^* \\ \mathbf{q}^* \\ \mathbf{r}^* \end{pmatrix} &= \mathbf{0}.\end{aligned} \tag{17}$$

2) ncCP algorithm based on the non-linear model: Based on the CP algorithm instance derived for the convex optimization program with the linear data model, we add the non-linear correction step at the end by subtracting the estimated non-linear term $\Delta\mathbf{g}(\mathbf{b}^{(n+1)})$ from the measured data $\mathbf{g}_{\mathcal{M}}$. The resulting pseudo-code for the non-convex CP (ncCP) algorithm for solving the non-convex optimization program in equation (2) is listed in Algorithm 2 below.

Note that the difference between algorithms 1 and 2 are in line 4, where $\mathbf{g}_{\mathrm{LN}}$ is initialized as $\mathbf{g}_{\mathcal{M}}$, in line 6, where $\mathbf{g}_{\mathrm{LN}}$ is used to updated the dual variable $\mathbf{p}$, and in line 12, where the non-linear correction step is added. For the ncCP algorithm, where mathematical convergence is not guaranteed, we monitor the same conditions as in equation (17) and the conditional primal-dual gap, which now becomes

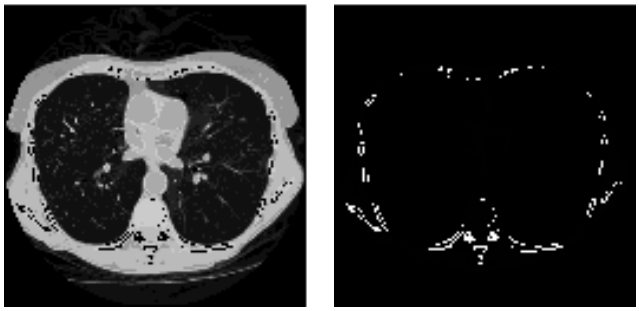
$$\begin{aligned}\mathrm{cPD}_{\mathrm{nc}}(\mathbf{b}^*, \mathbf{p}^*, \mathbf{q}^*) =& \frac{\lambda}{2}||\mathbf{g}(\mathbf{b}^*) - \mathbf{g}_{\mathcal{M}}||_2^2 + \frac{1}{2\lambda}||\mathbf{p}^*||_2^2 \\ & - \mathbf{g}^\top\mathbf{p}^* + \alpha\gamma||(|\mathbf{q}^*|)||_\infty.\end{aligned} \tag{18}$$

D. Study Design

We perform a computer simulation study to check the numerical convergence of the ncCP algorithm. The simulated data are collected from a fan-beam configuration with circular trajectory, with physical dimensions similar to those in a standard cone-beam CT geometry. The source-to-detector and source-to-center-of-rotation distances are 1500 mm and 1000 mm, respectively. The linear detector is 400 mm in length and consists of 256 1.56-mm bins, forming a field-of-view (FOV) of 265 mm in diameter. Two ($S = 2$) typical X-ray spectra of 80 and 140 kVp are generated using the TASMICS worksheet (v1.0) [12], assuming a tungsten anode and a 5-mm Al filter [8], and used for generating at 160 views evenly distributed over 2π for each of the 80-kVp and 140-kVp spectral data set. We consider a decomposition model with two ($K = 2$) basis materials, water and bone. The two digital truth basis images are on an 128×128 image array with 1.95-mm square pixels, as shown in Fig. 1.

Using equation (1) or (5), we generated data $\mathbf{g}_{\mathcal{M}}$, which contains the non-linear spectral response, and applied the ncCP algorithm in Algorithm 2 to recover the basis images from $\mathbf{g}_{\mathcal{M}}$. Meanwhile, for a comparison, we also generated consistent data, denoted by $\bar{\mathbf{g}}_{\mathcal{M}}$, using the linear model in equation (6), and applied the CP algorithm in Algorithm 1 to recover the basis image from $\bar{\mathbf{g}}_{\mathcal{M}}$.

For both algorithms, same parameters, such as spectra, pixel size, and system matrices, used in the data generation were used in the reconstruction as well. In the simulations study, the TV constraint parameter for both algorithms was selected to be the TV value of the truth monochromatic image as $\gamma = ||(|\nabla(\mu_1\mathbf{b}_1^{\mathrm{truth}} + \mu_2\mathbf{b}_2^{\mathrm{truth}})|)||_1$. Parameter λ was selected to be 100, after testing with different choices of 1, 10, 100 and 1000 and choosing the fastest convergence rate. Further, due to practical considerations such as limited computer precision, the practical convergence conditions for both algorithms are



(a) truth water (b) truth bone

Fig. 1: Digital truth basis images used to generate the data.

defined as

$$\begin{aligned} \mathrm{cPD}(\mathbf{b}^{(n)}, \mathbf{p}^{(n)}, \mathbf{q}^{(n)}) &\leq 10^{-8}, \\ \mathrm{TV}(\mathbf{b}^{(n)}) = |||(|(\mu_1\nabla, \mu_2\nabla)\mathbf{b}^{(n)})|)||_1/\gamma - 1| &\leq 10^{-8}, \\ \mathrm{monoNegMag}(\mathbf{b}^{(n)}) = |||\lfloor(\mu_1\mathcal{I}, \mu_2\mathcal{I})\mathbf{b}^{(n)}\rfloor_-||_2 &\leq 10^{-8}, \\ \mathrm{KTy}(\mathbf{p}^{(n)}, \mathbf{q}^{(n)}, \mathbf{r}^{(n)}) = ||\mathcal{K}^\top \begin{pmatrix} \mathbf{p}^{(n)} \\ \mathbf{q}^{(n)} \\ \mathbf{r}^{(n)} \end{pmatrix} ||_2 &\leq 10^{-8}, \end{aligned} \tag{19}$$

where the definitions of the cPD gap for Algorithms 1 and 2 are in equations (16) and (18), respectively, and operator $\lfloor \mathbf{x} \rfloor_-$ truncates the positive elements of $\mathbf{x}$ to zero. Note that one of the conditions in equation (17), $\mathbf{r}^* \preceq 0$, is automatically enforced by the $\mathrm{neg}()$ operation in line 9 of both Algorithms 1 and 2. In addition, we calculate the normalized image RMSE with respect to the truth basis images as

$$\bar{\Delta}_{\mathbf{b}}(\mathbf{b}^{(n)}) = ||\mathbf{b}^{(n)} - \mathbf{b}^{\mathrm{true}}||_2 / ||\mathbf{b}^{\mathrm{true}}||_2. \tag{20}$$

III. Results

We investigate the ncCP algorithm by plotting in Fig. 2 the convergence metrics. It can be observed that the practical convergence conditions in equation (15) are met for both ncCP and CP algorithms with the non-linear data $\mathbf{g}_{\mathcal{M}}$ and linear data $\bar{\mathbf{g}}_{\mathcal{M}}$, respectively. In addition, we also show in Fig. 2 the normalized image RMSE, which have reached smaller than 10^{-12} for both algorithms. It shows that, under the data condition studied in this work, the ncCP algorithm is able to recover the basis images. Because the reconstructed basis images are numerically identical to the truth basis ones, they are not shown here. Such observations suggest the numerical convergence of the ncCP algorithm instance in Algorithm 2 for the non-convex optimization program in equation (2), as well as the CP algorithm instance in Algorithm 1 for the convex optimization program in equation (10). Furthermore, it is shown that with the same parameter λ the convergence rates of the ncCP algorithm and its associated CP algorithm are almost the same for the convergence metrics considered in the study.

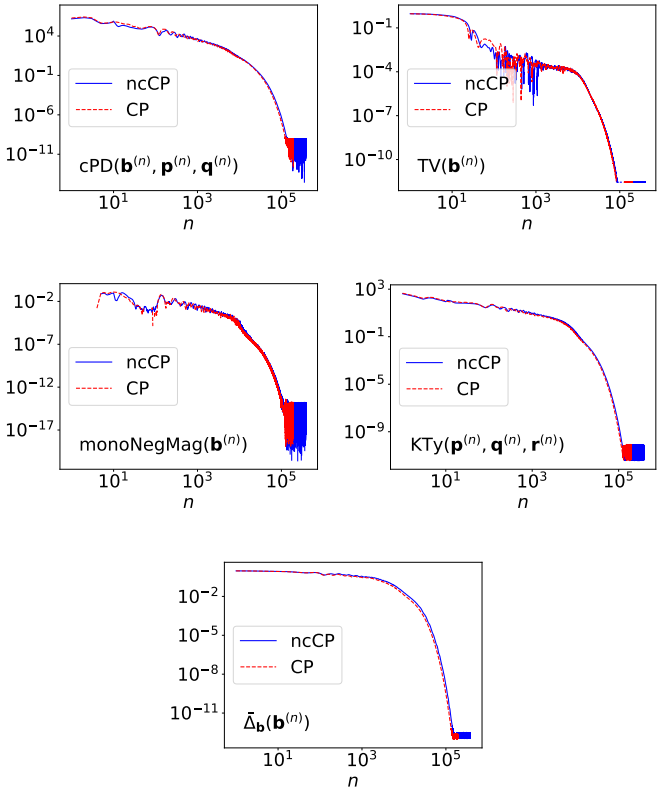


Fig. 2: Practical convergence metrics as defined in equation (19) (row 1 & 2) and the image RMSE (row 3), as functions of iteration n.

IV. Conclusion

In this work, we have proposed a non-convex optimization program for the reconstruction of multispectral CT data and derived a non-convex Chambolle-Pock (ncCP) algorithm for numerically solving the designed non-convex optimization program. The ncCP algorithm is based on the generic CP algorithm instance for a similar convex optimization program, in combination with a non-linear correction step for the measure data. We have also derived the convergence conditions, which can be used to check numerical convergence of the ncCP algorithm and to guide the selection of parameters to avoid potential saddle points of non-convex optimization programs. Using computer simulation data with two spectra, we have carried out a numerical study to verify the numerical convergence of the derived ncCP algorithm, by observing that the practical convergence conditions are all met and that the reconstructed basis images are numerically identical to the truth basis images.

Future work will focus on the application of the derived ncCP algorithm to real multispectral CT data with noise, decomposition error, and/or scattering. Also, it is of our interest to investigate and expand the ncCP algorithm instance for other optimization program designs, such as using KL divergence as the objective function, by incorporating the non-linear correction step with the CP algorithm derived for the convex optimization program based on the linear data model.

Acknowledgment

This work was supported in part by NIH R01 Grants Nos. CA182264 and EB018102. The contents of this article are

solely the responsibility of the authors and do not necessarily represent the official NIH views.

REFERENCES

[1] A. Chambolle and T. Pock, "A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 120–145, May 2011. [Online]. Available: http://link.springer.com/article/10.1007/s10851-010-0251-1

[2] E. Y. Sidky, J. H. Jorgensen, and X. Pan, "Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle-Pock algorithm," *Physics in Medicine and Biology*, vol. 57, no. 10, p. 3065, 2012. [Online]. Available: http://stacks.iop.org/0031-9155/57/i=10/a=3065

[3] E. Y. Sidky, J. S. Jrgensen, and X. Pan, "First-order convex feasibility algorithms for x-ray CT," *Medical Physics*, vol. 40, no. 3, pp. n/a–n/a, Mar. 2013. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.1118/1.4790698/abstract

[4] E. Y. Sidky, D. N. Kraemer, E. G. Roth, C. Ullberg, I. S. Reiser, and X. Pan, "Analysis of iterative region-of-interest image reconstruction for x-ray computed tomography," *Journal of Medical Imaging*, vol. 1, no. 3, pp. 031 007–031 007, 2014. [Online]. Available: http://dx.doi.org/10.1117/1.JMI.1.3.031007

[5] Z. Zhang, X. Han, E. Pearson, C. Pelizzari, E. Y. Sidky, and X. Pan, "Artifact reduction in short-scan CBCT by use of optimization-based reconstruction," *Physics in Medicine and Biology*, vol. 61, no. 9, p. 3387, 2016. [Online]. Available: http://stacks.iop.org/0031-9155/61/i=9/a=3387

[6] D. Xia, D. A. Langan, S. B. Solomon, Z. Zhang, B. Chen, H. Lai, E. Y. Sidky, and X. Pan, "Optimization-based image reconstruction with artifact reduction in C-arm CBCT," *Physics in Medicine and Biology*, vol. 61, no. 20, p. 7300, 2016. [Online]. Available: http://stacks.iop.org/0031-9155/61/i=20/a=7300

[7] Z. Zhang, J. Ye, B. Chen, A. E. Perkins, S. Rose, E. Y. Sidky, C.-M. Kao, D. Xia, Chi-Hua Tung, and X. Pan, "Investigation of optimization-based reconstruction with an image-total-variation constraint in PET," *Physics in Medicine and Biology*, vol. 61, no. 16, p. 6055, 2016. [Online]. Available: http://stacks.iop.org/0031-9155/61/i=16/a=6055

[8] B. Chen, Z. Zhang, E. Y. Sidky, D. Xia, and X. Pan, "Image reconstruction and scan configurations enabled by optimization-based algorithms in multispectral CT," *Physics in Medicine & Biology*, vol. 62, no. 22, p. 8763, 2017. [Online]. Available: http://stacks.iop.org/0031-9155/62/i=22/a=8763

[9] Y. Zou and M. D. Silver, "Analysis of fast kV-switching in dual energy CT using a pre-reconstruction decomposition technique," in *Proc. SPIE, Medical Imaging: Physics of Medical Imaging*, vol. 6913, 2008, pp. 691 313–691 313–12. [Online]. Available: http://dx.doi.org/10.1117/12.772826

[10] X. Pan, B. Chen, Z. Zhang, E. Pearson, E. Sidky, and X. Han, "Optimization-based Reconstruction Exploiting Spectral Information in CT," in *The Third International Conference on Image Formation in X-Ray Computed Tomography*, 2014, pp. 228–232.

[11] J. Duchi, S. Shalev-Shwartz, Y. Singer, and T. Chandra, "Efficient Projections Onto the L1-ball for Learning in High Dimensions," in *Proceedings of the 25th International Conference on Machine Learning*, ser. ICML '08. New York, NY, USA: ACM, 2008, pp. 272–279. [Online]. Available: http://doi.acm.org/10.1145/1390156.1390191

[12] A. M. Hernandez and J. M. Boone, "Tungsten anode spectral model using interpolating cubic splines: Unfiltered xray spectra from 20 kV to 640 kV," *Medical Physics*, vol. 41, no. 4, pp. 042 101–1–042 101–15, Apr. 2014. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.1118/1.4866216/abstract

Coronary artery calcium scoring using dual energy tomosynthesis

Scott S. Hsieh and Matthew J. Budoff

***Abstract*—Coronary artery calcium scoring is a controversial technique that can assess an individual's risk of cardiovascular disease, but which requires a CT scan that delivers a nontrivial amount of radiation. The goal of this work is to explore an alternative pathway to measuring coronary calcium using a combination of dual energy and tomosynthesis. We find that a tomosynthesis-based system is less sensitive but is simpler and can function at reduced dose, indicating its potential use for population-based screening of cardiovascular risk.**

***Index Terms*—calcium scoring, tomosynthesis, dual energy**

I. Introduction

CARDIOVASCULAR disease is the leading cause of death in the United States, and has been so for more than 90 years. Several risk factors can be used to predict cardiovascular disease, including, but not limited to: elevated blood pressure, smoking, diabetes, elevated cholesterol, male sex, and old age. From traditional risk factors it is possible to build a risk score (1), and this risk score can be used as a basis for prescribing treatment. For example, some guidelines call for the prescription of cholesterol-lowering statin therapy to be initiated or titrated on the patient's cholesterol levels and their risk score (2). Studies that have examined the efficacy of these statins have found that they deliver similar magnitude of response amongst those patients with intermediate and low levels of background LDL cholesterol (3), suggesting that patients that would ordinarily not receive statin therapy because of low cholesterol levels may still benefit from them.

Current risk scores are capable of predicting heart disease with moderate effectiveness, but there remains room for significant improvement. Accuracy of prediction is often assessed with the Area under the Receiver Operating Curve (AUC) metric. In this metric, random chance is associated with AUC = 0.5, and perfect prediction is AUC = 1. For the Women's Health Study cohort, it has been observed that current risk scores are able to predict cardiovascular disease with AUC = .78. More accurate risk scores that improve the AUC metric could lead to better prescription of medications, and would also allow the patient to choose from a wide variety of lifestyle choices that would reduce his or her risk to cardiovascular disease.

Amongst emerging risk factors, the most powerful predictor for cardiovascular disease is coronary artery calcium scoring (CACS) (5). Calcium scoring measures calcium deposits that arise from the atherosclerotic disease progression. A weakness of calcium scoring is that it requires a CT scan, exposing the patient to a nontrivial amount of radiation of about 1 mSv, equal to approximately 3 months of background radiation. In addition, incidental findings and high cost have kept calcium scoring from widespread utilization (6, 7).

For these reasons, alternative methods to obtain information similar to calcium scoring have been highly desired. The measurement of carotid media-intima thickness using ultrasound is another attempt to quantify atherosclerosis but fails to reach the level of risk score improvement as CACS (5). Projection-based x-ray techniques, including dual-energy fluoroscopy, have been proposed to assess calcium scoring (8). However, their efficacy in patients has not been evaluated. In this project, we use dual-energy tomosynthesis to improve performance. The goal of this work is to assess the feasibility of this approach.

II. METHODS

Conventional tomosynthesis would be difficult to implement for scanning the coronary arteries. The coronary arteries are subject to significant motion from both the beating heart and the breathing lung. Although it would be possible to stitch together data from multiple adjacent heartbeats, registration may be imperfect in the context of an improperly held breath. Misregistration on the order of 1 millimeter may make it difficult to perform the necessary dual energy subtraction.

Moreover, calcifications in the coronary arteries are typically of modest contrast and size. The central feasibility question was whether or not a fast tomosynthesis-based approach would have sufficient contrast to detect clinically significant calcium deposits.

A. System design

The real-time tomosynthesis system design has been described elsewhere (9). This design uses a single frame to read out multiple x-ray images; hence, it is a Single-Frame X-ray Tomosynthesis (SFXT) design. In this application, two acquisitions are necessary at two different energies. However, within each acquisition we use only 4 different projections to provide the necessary temporal resolution.

Figure 1 shows an example of SFXT when applied to a different task, the detection of a lung tumor. The raw detector

S. Hsieh and M. Budoff are with the University of California, Los Angeles.

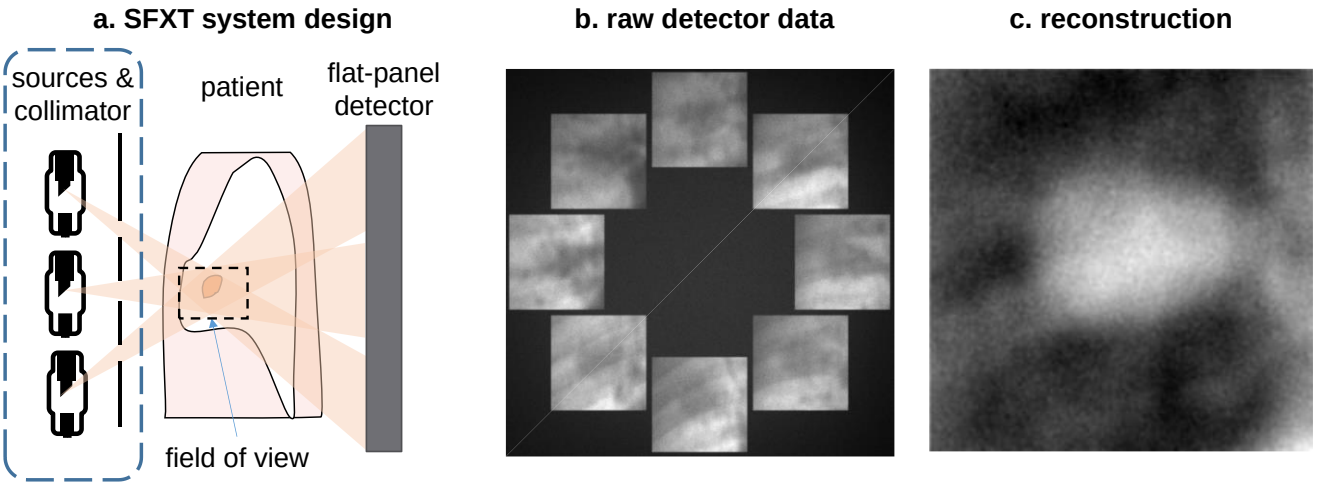


Figure 1. System design for SFXT. For clarity of presentation, this simulation shows a higher-contrast lung tumor rather than a more subtle calcification. (a) Multiple sources illuminate a small field of view. (b) The detector sees the multiple projections arriving at different sectors. (c) Shift-and-add reconstruction yields a higher fidelity image of the lung tumor compared to a single projection.

data in this case has sufficient contrast so that the tumor can be discerned in each source image, although the image quality is higher when the data is combined in the reconstruction process. The primary differences between the lung tumor tracking application and the calcium scoring application, from a system design perspective, is that the calcium scoring application requires a larger field of view and hence fewer projections can be used. We use four projections rather than 8 as shown in Figure 1. Furthermore, dual energy acquisition is necessary to detect calcium deposits, so two separate acquisitions must be recorded at different tube voltages.

The temporal resolution of such a system would be very high. The four stationary x-ray sources are energized at the same time to pass radiation through the field of view. The sources are toggled between two energies and a switching filter is used in tandem to improve dual energy contrast. For a detector operating at 45 fps, and assuming instant switching time, the acquisition time can be as short as 44 ms. A more conservative estimate, assuming a 30 fps detector that requires two frames at each energy level and one frame of switching time, is an acquisition time of 133 ms.

The field of view can be estimated as follows. Typical large-area fluoroscopic x-ray detectors today have an area of 43x43 cm. Assuming a magnification of 1.5, and using the fact that four quadrants halves the linear dimension of the field of view, we calculate that the FOV would be 14 cm.

B. Source data

Five clinical cases were examined from the Multi-Ethnic Study of Atherosclerosis (MESA). The MESA trial enrolled more than 6,800 men and women and included a calcium scoring CT scan for each individual.

Each scan was decomposed into water and calcium on a voxel-by-voxel basis. We assumed that each voxel less than 100 HU was composed entirely of water of the appropriate density, and voxels greater than 100 HU were composed of water at a density of 1.1 g/cc, with the remaining attenuation being provided by calcium.

The resolution of the simulated tomosynthesis scans was severely limited by the underlying MESA data, which was stored with a characteristic slice thickness of 3 mm.

The coronary arteries were sometimes located near the end of a dataset. The data was padded using zero-order hold (constant extrapolation) when necessary.

C. Simulations

Previous experimental data on dual-energy fluoroscopy suggested that low dose reconstruction may be possible (8). Hence, in the first set of experiments we sought to verify this in the MESA dataset. Best-case reconstructions were performed by forward projection without noise addition. We examined five cases from across a range of calcification levels.

Reconstruction was performed by shift-and-add. To enhance the detection of small calcifications with background, we used a difference of Gaussians (DoG) filter. In the noise-free images, we simply projected the calcium images as we assumed that the water content would be subtracted out.

Finally, to assess the impact of quantum noise, noise was injected assuming the low energy image was acquired at 60 kVp and the high energy image was acquired at 140 kVp, including 0.5 mm Sn. The spectrum was discretized and each energy was projected separately. The high and low energy images were weighted and subtracted to isolate calcium contrast. The mAs for the images was adjusted to visualize the effect on detectability.

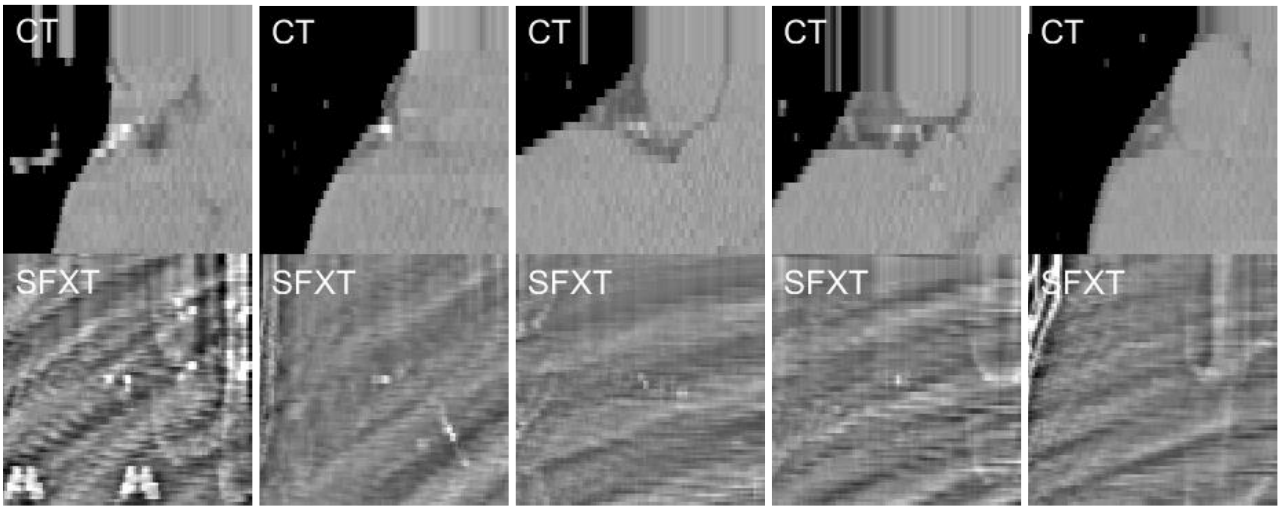


Figure 2. Five examples of coronary calcification with different severities. Each column corresponds to a distinct cases with CT on top and dual-energy SFXT on bottom. The left cases have more calcium than the right cases. From left to right, the calcium scores corresponds to calcium scores of 925, 734, 366, 124, and 10. The ACC/AHA guidelines describe abnormal CAC levels to be in excess of 300 points (2), at which point statin therapy could be initiated for intermediate-risk patients. [WL, WW] = [0, 600] HU for CT images.

III. Results

Figure 2 shows the detectability of coronary calcium for five distinct cases. The majority of these cases fall below the clinically actionable level of a coronary artery calcium score threshold of 300 (2), although sensitivity to lower scores may still be desired for other reasons. The ability to detect fairly subtle calcifications between 100 to 300 is limited. However, the intrinsic resolution of the data is poor, and higher resolution may aid detection. Cases with calcium above 300 can generally be seen. Because of the four projections used, calcifications that are out-of-focus are seen in a square pattern.

Figure 3 shows the impact of quantum noise on a separate case. In general, it seems that adequate noise statistics can be obtained with dose roughly comparable to a chest x-ray. To a large extent, detectability is limited by background anatomy rather than quantum statistics.

IV. Discussion

Cardiovascular disease is a leading cause of mortality and morbidity worldwide, and improved risk stratification using calcium scoring has the potential to improve patient outcomes. To date, calcium scoring has only been performed using CT. We have presented preliminary data showing that it may be possible to detect calcium using a combination of tomosynthesis and dual energy.

In the absence of overlap with bony structures, dual energy fluoroscopy would be sufficient to detect calcium and there would be no reason to resort to CT (8). In practice, overlap with the ribs is difficult to avoid in the thorax. Using four views in tomosynthesis, as is done here, provides some capability to detect calcium against background structure. However, it is clear that the results of dual energy SFXT are not as good as CT. In particular, it should be noted that dual energy tomosynthesis performs better at detecting single, large

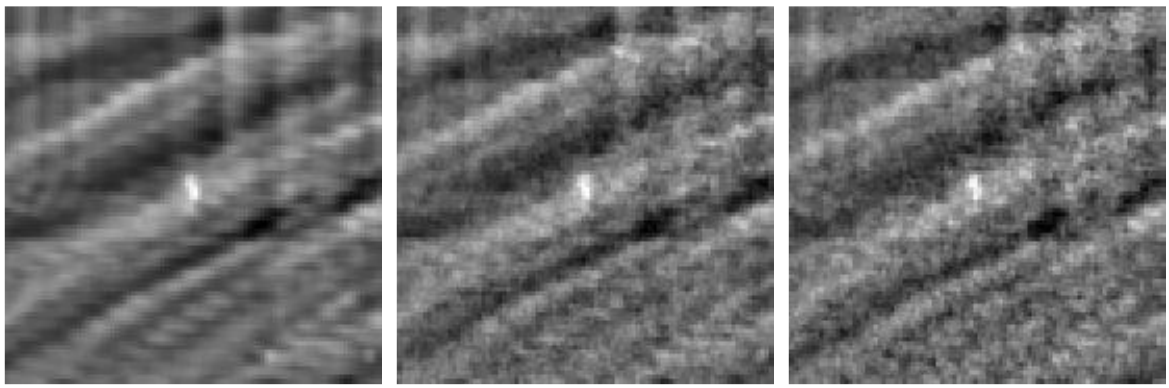


Figure 3. Effect of quantum noise on detectability. (Left) Noiseless case. (Middle) Noise assuming 60 kVp @ 6 mAs, and 140 kVp + Sn @ 1 mAs. This dose is roughly equivalent to a standard chest x-ray. (Right) Noise assuming 60 kVp @ 3 mAs, 140 kVp + Sn @ 0.5 mAs, or half dose.

calcifications than several independent calcium deposits. This results are intuitive if one assumes that there is an intrinsic threshold of detectability. A large deposit above the threshold of detectability may be lost if it is distributed into several small deposits, underneath the threshold of detectability. Assuming dose levels comparable with a chest x-ray, this detectability threshold is primarily limited by background structure and not from quantum statistics.

While our results do not show perfect agreement with CT, it should be noted that in most cases it was possible differentiate individuals with significant calcium from those that had minimal calcium. It has been noted that for low-risk women, advanced coronary calcium scores (above 300) were associated with a 22x increase in cardiovascular risk compared to those without calcium (10). A simple, low-dose, high-throughput device that can screen a large group of individuals may help to identify those that would benefit from more intensive interventions. This tomosynthesis-type device may fit these requirements.

This preliminary study has several limitations. Only a small number of datasets were examined. A future study will use higher-resolution data, examine a set of more than 30 cases, incorporate more sophisticated modeling of x-ray physics using Monte Carlo simulations, and use a blinded reader to better evaluate the effectiveness of SFXT for determining cardiovascular risk.

REFERNCES

1. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation.* 1998 May 12; 97(18): 1837-1847.

2. Stone NJ, Robinson JG, Lichtenstein AH, Bairey Merz CN, Blum CB, Eckel RH, Goldberg AC, Gordon D, Levy D, Lloyd-Jones DM, McBride P, Schwartz JS, Shero ST, Smith SC,Jr, Watson K, Wilson PW, Eddleman KM, Jarrett NM, LaBresh K, Nevo L, Wnek J, Anderson JL, Halperin JL, Albert NM, Bozkurt B, Brindis RG, Curtis LH, DeMets D, Hochman JS, Kovacs RJ, Ohman EM, Pressler SJ, Sellke FW, Shen WK, Smith SC,Jr, Tomaselli GF, American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: A report of the american college of cardiology/american heart association task force on practice guidelines. *Circulation.* 2014 Jun 24; 129(25 Suppl 2): S1-45.

3. Yusuf S, Bosch J, Dagenais G, Zhu J, Xavier D, Liu L, Pais P, López-Jaramillo P, Leiter LA, Dans A. Cholesterol lowering in intermediate-risk persons without cardiovascular disease. *N Engl J Med.* 2016; 374(21): 2021-2031.

4. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation.* 2007 Feb 20; 115(7): 928-935.

5. Yeboah J, McClelland RL, Polonsky TS, Burke GL, Sibley CT, O'leary D, Carr JJ, Goff DC, Greenland P, Herrington DM. Comparison of novel risk markers for improvement in cardiovascular risk assessment in intermediate-risk individuals. *JAMA.* 2012; 308(8): 788-795.

6. Blaha MJ, Silverman MG, Budoff MJ. Is there a role for coronary artery calcium scoring for management of asymptomatic patients at risk for coronary artery disease?: Clinical risk scores are not sufficient to define primary prevention treatment strategies among asymptomatic patients. *Circ Cardiovasc Imaging.* 2014 Mar; 7(2): 398-408; discussion 408.

7. Andersson C, Vasan RS. Is there a role for coronary artery calcium scoring for management of asymptomatic patients at risk for coronary artery disease?: Clinical risk scores are sufficient to define primary prevention treatment strategies among asymptomatic patients. *Circ Cardiovasc Imaging.* 2014 Mar; 7(2): 390-7; discussion 397. PMCID: PMC4154564.

8. Xu T, Ducote JL, Wong JT, Molloi S. Feasibility of real time dual-energy imaging based on a flat panel detector for coronary artery calcium quantification. *Med Phys.* 2006; 33(6): 1612-1622.

9. Hsieh SS, Ng LW. Real-time tomosynthesis for radiation therapy guidance. *Med Phys.* 2017.

10. Lakoski SG, Greenland P, Wong ND, Schreiner PJ, Herrington DM, Kronmal RA, Liu K, Blumenthal RS. Coronary artery calcium scores and risk for cardiovascular events in women classified as "low risk" based on framingham risk score: The multi-ethnic study of atherosclerosis (MESA). *Arch Intern Med.* 2007; 167(22): 2437-2442.

Precision Learning: Reconstruction Filter Kernel Discretization

C. Syben, B. Stimpel, K. Breininger, T. Würfl, R. Fahrig, A. Dörfler and A. Maier

***Abstract*—In this paper, we present substantial evidence that a deep neural network will intrinsically learn the appropriate way to discretize the ideal continuous reconstruction filter. Currently, the Ram-Lak filter or heuristic filters which impose different noise assumptions are used for filtered back-projection. All of these, however, inhibit a fully data-driven reconstruction neural network learning approach. In addition, the heuristic filters are not chosen in an optimal sense. To tackle this issue, we propose a formulation to directly learn the reconstruction filter. The filter is initialized with a Ramp filter as a strong pre-training and learned in frequency domain. We compare the learned filter with the Ram-Lak and the Ramp filter on a numerical phantom as well as on a real CT dataset. The results show that the network properly discretizes the continuous Ramp filter and converges towards the Ram-Lak solution. In our view these observations are interesting to gain a better understanding of deep learning techniques and traditional analytic techniques such as Wiener filtering and discretization theory. Furthermore, this will allow fully trainable data-driven reconstruction deep learning approaches.**

***Index Terms*—Computed Tomography, Deep Learning, Filtered Back-Projection, Filter Discretization, Filter Learning**

I. Introduction

Recently deep learning has shown promising results in the field of Computed Tomography (CT) reconstruction. In his perspective article, Wang [1] states that a reconstruction pipeline implemented as a deep neural network allows to access the capability of learning-based reconstruction. Wang identifies the data-driven knowledge-enhancing abilities as the strength of deep learning-based reconstruction. Würfl et al. [2] have proposed an implementation of the filtered back-projection algorithm (FBP) as a neural network. Similar to iterative reconstruction algorithms, their proposed implementation avoids explicitly storing the system matrix, which would render the algorithm infeasible. The approach utilized the data-driven capability by learning the compensation weights in case of limited-angle tomography. Hammernik et al. [3] proposed a two-level deep learning architecture to compensate for additional streak artifacts in the limited-angle tomography case. They showed that their approach allows for a joint optimization without any heuristic parameter tuning. In both approaches the necessary filtering to perform a FBP is done with a fixed layer using an analytical discretization.

This work has been supported by the project P3-Stroke, an EIT Health innovation project. EIT Health is supported by EIT, a body of the European Union.

C. Syben, B. Stimpel, K.Breininger, T. Würfl, R. Fahrig and A. Maier are with Friedrich-Alexander-Universität Erlangen-Nürnberg, Pattern Recognition Lab, Erlangen, Germany.

C. Syben, B. Stimpel, and A. Dörfler are also with Friedrich-Alexander-Universität Erlangen-Nürnberg, Department of Neuroradiology, Erlangen, Germany.

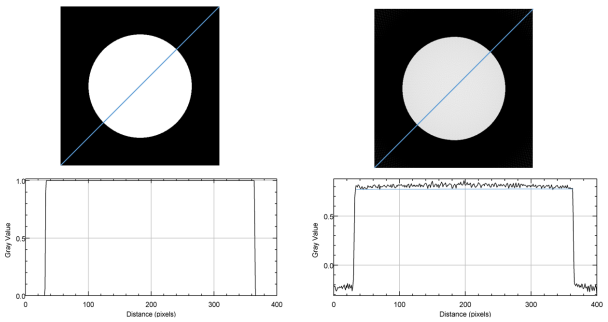


Fig. 1: Line profile through a 2D circle phantom. Ground truth (left) and FBP result based on Ramp filter (right).

The ideal filter for FBP can be derived using analytic reconstruction theory. Assuming an infinite number of projections and infinitely small detector pixels, it takes the form of the absolute value function in Fourier domain, commonly referred to as Ramp filter. In practice, however, Radon inversion has to be performed using a finite number of projections. This introduces discretization errors which need to be handled by replacing the Ramp filter with an appropriate discrete version. The occurring artifacts are commonly called cupping and dc shift artifacts [4], and an example of both is shown in Figure 1. The analytically derived discrete version of the Ramp filter is the well known Ram-Lak filter introduced by Ramachandran and Lakshminarayan [5]. Both, the ideal Ramp as well as the Ram-Lak filter are sensitive to noise. Thus, many different filters have been proposed which impose different noise assumptions. The most well known example is the Shepp-Logan filter [6], which incorporates a smoothing filter. However, while the choice of filter clearly matters, none of these heuristic filters are chosen in an optimal sense. CT vendors typically have specialized departments that aim at designing optimal filters for their users. The actual filter configurations are often deemed as company secrets. This shows the demand for specialized filters and raises the question whether data-driven methods can be used to deduce optimal filters.

Pelt et al. [7] proposed a method to learn such a data-dependent filter for the FPB algorithm. Their approach learns a filter approximation to increase reconstruction accuracy in specific cases. They optimize the filter with respect to the minimal error between the input data and the projection of the

FBP reconstructed input. Pelt et al. use an efficient formulation and exponential binning to handle the size of the system matrix. In order to do this, handling the discretization of the Ramp filter with respect to the details around the zero frequency as well as the boundaries is necessary.

We propose a formulation to learn a discrete optimal reconstruction filter directly in a deep learning context. We provide substantial evidence that a deep neural network will intrinsically learn the appropriate way to discretize the continuous filter. Furthermore, our proposed formulation leads to a straightforward implementation of the optimization using deep learning frameworks that automatically compute the gradient of the cost function using back-propagation.

II. Method & Materials

First, we describe the filtered back-projection in continuous and discrete form for the parallel-beam geometry and introduce the theoretical filter kernel. Afterwards, the optimization problem and the related gradient to learn the filter kernel are derived. In the last Sections we describe our experiments and discuss them.

A. Filtered Back-Projection

The filtered back-projection (FBP) algorithm is an efficient solution to the reconstruction problem. First the projection data $p(s,\theta)$ are filtered by a convolution with the filter kernel

$$q(s,\theta) = h(s) * p(s,\theta) \ , \tag{1}$$

$$h(s) = \int |\omega| e^{2\pi\omega s} d\omega \ , \tag{2}$$

subsequently, the filtered projection data $q(s,\theta)$ are back-projected to obtain the reconstruction result $f(x,y)$ with

$$f(x,y) = \int_0^\pi q(s,\theta)|_{s=x\cdot\cos\theta+y\cdot\sin\theta} d\theta \ . \tag{3}$$

In practice, we need a discrete description of the FBP presented in Eq. (1-3). The discrete reconstruction problem can be expressed by

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{p} \ , \tag{4}$$

where $\boldsymbol{A} \in \mathbb{R}^{N\times M\cdot P}$ is the system matrix, $\boldsymbol{x} \in \mathbb{R}^N$ is the volume and $\boldsymbol{p} \in \mathbb{R}^{M\cdot P}$ are the projections. Since the discrete representation of the Radon transform $\boldsymbol{A}$ is a tall matrix, it has no inverse. This means every reconstruction formula is, therefore, a unique pseudo inverse of $\boldsymbol{A}$,

$$\boldsymbol{x} = \underbrace{\boldsymbol{A}^\top}_{\text{Back-projection (Eq. 3)}} \underbrace{(\boldsymbol{A}\boldsymbol{A}^\top)^{-1}}_{\text{Filter (Eq.1)}} \boldsymbol{p} \ . \tag{5}$$

B. Learning the Filter Kernel

From the convolution theorem it is clear that a convolution in spatial domain is equal to a multiplication in the frequency space. Therefore, we can reformulate the filter $(\boldsymbol{A}\boldsymbol{A}^\top)^{-1}$ in form of a diagonal matrix in Fourier domain

$$\boldsymbol{x} = \boldsymbol{A}^\top \boldsymbol{F}^{\boldsymbol{H}} \boldsymbol{K} \boldsymbol{F} \boldsymbol{p} \ , \tag{6}$$

where $\boldsymbol{F} \in \mathbb{R}^{N\times N}$ and $\boldsymbol{F}^{\boldsymbol{H}} \in \mathbb{R}^{N\times N}$ represents the Fourier and inverse Fourier transform, respectively. $\boldsymbol{K} \in \mathbb{R}^{N\times N}$ is a diagonal matrix representing the filter in frequency domain. To learn the filter matrix $\boldsymbol{K}$, which is a discrete approximation of the Ramp filter, we use Eq. 6 to formulate a objective function $f(\boldsymbol{K})$ as a least-square minimization problem:

$$f(\boldsymbol{K}) = \frac{1}{2} \| \boldsymbol{A}^\top \boldsymbol{F}^{\boldsymbol{H}} \boldsymbol{K} \boldsymbol{F} \boldsymbol{p} - \boldsymbol{x} \|_2^2 \ . \tag{7}$$

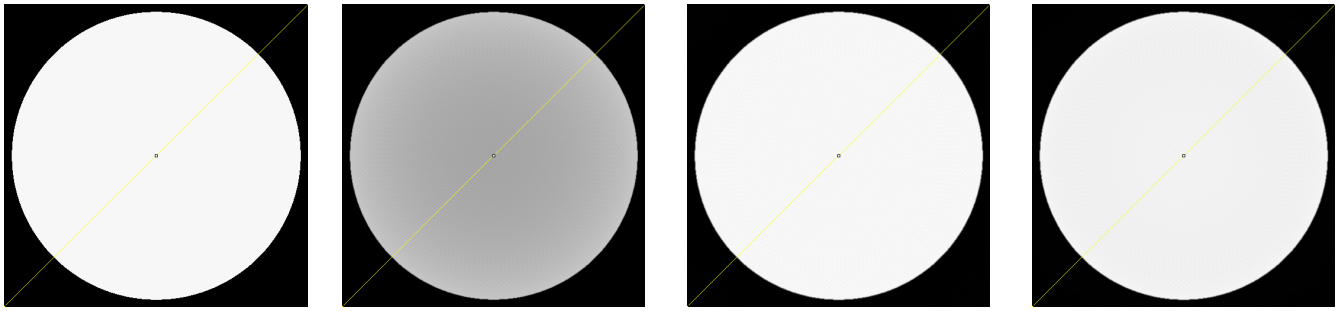
The gradient to our objective function $f(\boldsymbol{K})$ in Eq. 7 with respect to $\boldsymbol{K}$ is

$$\frac{\partial f(\boldsymbol{K})}{\partial \boldsymbol{K}} = \underbrace{\boldsymbol{F}\boldsymbol{A}\underbrace{(\boldsymbol{A}^T \boldsymbol{F}^{\boldsymbol{H}} \boldsymbol{K} \boldsymbol{F} \boldsymbol{p} - \boldsymbol{x})}_{\text{Error}} \underbrace{(\boldsymbol{F}\boldsymbol{p})}_{l-1}^\top}_{\text{Back-propagation}} \ . \tag{8}$$

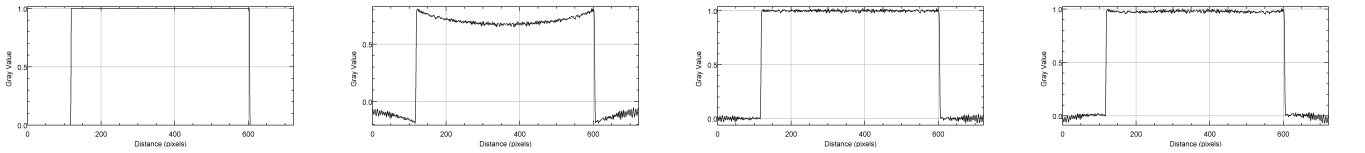
Note that this analytical gradient also has an interpretation in a neural network learning context. To describe Eq. 8 using the terms of back-propagation: Eq. 6 can be regarded as a network with input $\boldsymbol{p}$ and layers $\boldsymbol{F}$, $\boldsymbol{K}$, $\boldsymbol{F}^H$ and $\boldsymbol{A}^T$ with the identity as activation function between layers. $\boldsymbol{K}$ is the only layer containing trainable weights. Then, $\hat{\boldsymbol{x}} = \boldsymbol{A}^T \boldsymbol{F}^{\boldsymbol{H}} \boldsymbol{K} \boldsymbol{F} \boldsymbol{p}$ is the forward pass through the network following the considerations by Würfl et al. [2]. The gradient of the error function (Eq. 7) with respect to $\boldsymbol{K}$ is computed by multiplying two factors: 1) the partial derivative of the error function with respect to the output of the layer and 2) the transpose of the output of the previous layer $(l-1)$. The output of $l-1$ is readily described by $\boldsymbol{F}\boldsymbol{p}$. Using the recursive formulation of back-propagation, we yield $\boldsymbol{F}\boldsymbol{A}(\hat{\boldsymbol{x}} - \boldsymbol{x})$, with $(\hat{\boldsymbol{x}} - \boldsymbol{x})$ being the derivative of the error function. We consider these observations as interesting, as this gradient would be computed automatically in a deep learning framework such as TensorFlow.

C. Experiments

We implemented the cost function and the analytically derived gradient in CONRAD [8] and used stochastic gradient descent to learn the filter. Unmatched projectors were used for the reconstruction (pixel driven) and the forward projections (ray driven). 0 The filter matrix $\boldsymbol{K}$ is initialized with a slightly modified frequency domain representation of the ideal Ramp filter. By doubling the width of the zero-valued part of the ramp we emphasize the cupping artifacts. This serves exclusively to show the learning capability. For the training we use 10 numerical disc phantoms with increasing radii rendered on a 512×512 pixel grid. The learned filter is evaluated on a 512×512 slice of a real CT dataset showing the head of a pig acquired at Stanford University. For evaluation we compare the filtered back-projection using the modified Ramp Filter, the Ram-Lak and the learned filter, in the following referred to as Ramp-reco, Ram-Lak-reco and Learned-reco, respectively. The quantitative evaluation is done on the absolute difference between the ground truth (GT) and the respective reconstruction result. These difference images are evaluated using the mean, minimum and maximum difference as well as the standard deviation (std. dev.). Note that we use the original reconstruction as the ground truth for the pig experiment.



(a) GT: Phantom. (c) Ramp-reco. (e) Ram-Lak-reco. (g) Learned-reco.



(b) Line profile through GT. (d) Line profile through Ramp-reco. (f) Line profile through Ram-Lak-reco. (h) Line profile through Learned-reco.

Fig. 2: Reconstructions and respective line profile plots of numeric circle phantom using different filters. All images have the window/level set to: [1.03/0.50].

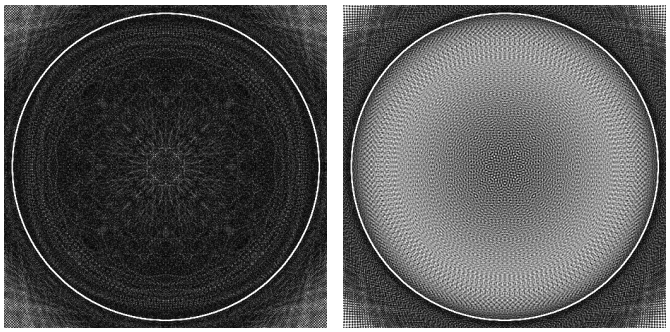
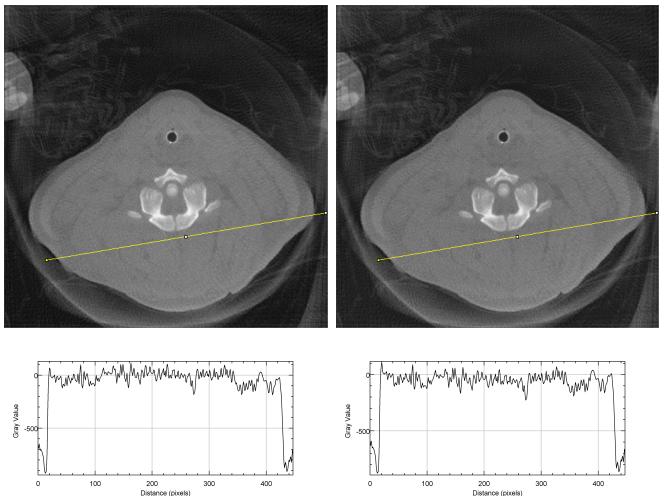


Fig. 3: Difference of the Ram-Lak-reco (left) and Learned-reco (right) to the GT. To enhance the difference between both images, the window/level was set to [0.02/0.04].



(a) Ram-Lak-reco. (b) Learned-reco.

Fig. 4: Reconstruction of a pig dataset filtered with Ram-Lak (left) and the learned filter kernel (right) and related line profiles along the yellow line below. For both images the window/level has set to: [2620/60] HU.

III. Results

A. Qualitative Results

Fig. 2 shows the results of the FBP on a numerical disc phantom (see ground truth in Fig. 2a), which was used for the training. Additionally, the results of Ramp-reco (see Fig. 2c), the Ram-Lak-reco (see Fig. 2e) and the Learned-reco (see Fig. 2g) and their respective line profiles are presented. The Ramp-reco leads to cupping artifacts in the homogeneous area of the disc that can be observed in Fig. 2c as well as in the respective line profile (see Fig. 2d). In Fig. 2f the Ram-Lak-reco is shown. The line profile in Fig. 2f illustrates that the homogeneous area of the disc is reconstructed properly. As displayed in Fig. 2g and 2h, the Learned-reco reconstructs the homogeneous area properly as well. Comparing the line profiles of the Ram-Lak-reco with the line profile from the Learned-reco, the Ram-Lak-reco shows nearly a straight line while the line profile of the Learned-reco still shows small deviations from the ideal line. Fig. 3a denotes the absolute difference between Ram-Lak-reco and the GT, while Fig. 3b denotes the absolute difference between the Learned-reco and the GT. Both absolute difference images show errors in the homogeneous area as well as at the edge of the disc phantom. Fig. 4 presents the Ram-Lak-reco and the Learned-reco of the pig dataset with their respective line profiles. Both line profiles, aligned over the same homogeneous area of the pig, show a similar behavior and no cupping.

B. Quantitative Results

In Tab. I the measurements of the absolute difference between the GT and the respective reconstructions are shown. The absolute mean error of the Ramp-reco is at 23.5%, while the absolute mean error of the Learned-reco is at 2.3%. The lowest absolute mean error can be observed by the Ram-Lak-reco. The standard deviation as well as the maximum

TABLE I: Measurements of the absolute difference between GT and respective circle reconstruction (in percent).

	mean	std. dev.	min	max
Ramp-reco	0.235	0.07	0.001	0.596
Ram-Lak-reco	0.01	0.031	0	0.41
Learned-reco	0.023	0.03	6.76E-09	0.409

TABLE II: Measurements of the absolute difference between GT and respective pig reconstruction in Hounsfield units (HU).

	mean	std. dev.	min	max
Ram-Lak-reco	66.99	61.401	6.10E-5	1634.82
Learned-reco	83.53	68.06	8.39E-5	1685.70

value of the Ramp-reco are higher compared to the other two reconstructions. The Ram-Lak-reco and the Learned-reco show a similar standard deviation of the absolute difference as well as the absolute maximum error. In Tab. II the absolute error measurements between the GT and the respective reconstructions are presented. Both the Ram-Lak-reco as well as the Learned-reco exhibit a similar absolute mean error. Also the standard deviation of the absolute error as well as the absolute maximum error are closely together.

IV. Discussion

The evaluation of the numerical disc phantom shows that homogeneous areas can be reconstructed properly without cupping artifacts when using the learned filter. However, the line profile reveals that the result is not as good as with the Ram-Lak filter. A possible explanation for this difference is our way of implementing the optimization directly and configuring the stochastic gradient parameters heuristically. Also note that our current training is performed with only 20 epochs and our data set consists only of 10 cylinders of different diameter. This is only a coarse approximation of the ideal training set. We assume that using a basis that spans the entire domain of $\boldsymbol{x}$ will do much better for this job. This will be subject of future experiments. Still, we consider 10 training samples as a good start for estimating such a complex relation that generalizes to other much more complex objects such as the pig data set. The minor difference between the Learned-reco and the Ram-Lak-reco in the real CT data experiment prove that the learned filter is not object dependent and not over-fitted to our training data. Utilizing discs with varying radii introduces two properties to the training process. First, narrow discs model the Dirac-impulse. Secondly, discs with larger radii exhibit large homogeneous regions. Occurring cupping and dc-shift artifacts in the homogeneous area will lead to strong gradients, which appear due to the wrong discretization. As a consequence, the weights converge towards the Ram-Lak solution, which is nothing else than learning the proper discretization of the continuous Ramp. Using the ideal Ramp less cupping is observed. This renders the dc-shift as the dominant artifact to compensate for. We expect that augmentation approaches of this method will lead to filters that are invariant / less prone to noise characteristics imposed by CT physics intrinsically. Thus, we believe that noise augmentation will lead to filters similar to the Shepp-Logan filter. Augmentation in this context will have a very similar result as a Wiener filter that is optimal given certain noise properties. Different approaches to learn the reconstruction filter were published in the past. Floyd [9] successfully learned the discrete version of the ramp filter for SPECT image reconstruction. However, compared to our presented method, their approach learns the filter in spatial domain using a neural network with fully connected layers. As a consequence of the system design, Floyd reported practical issues implementing the huge amount of trainable weights at the time of the publication. Furthermore, no detailed analysis of the discretization properties of the learned filter was performed. Even though similar approaches to learn the reconstruction filter were published, e.g. [7], [9], none of them explicitly uses domain knowledge to design the network topology. In contrast, our derivation of the network topology is based on the continuous analytical problem description. Furthermore, the transition to the discrete filter is intrinsically solved due to the discrete nature of the neural networks. In our view these observations are interesting to gain a better understanding of deep learning techniques and traditional analytic techniques such as Wiener filtering and discretization theory. To the best of our knowledge, we did not observe such links between analytical signal processing theory and deep learning so far.

V. Conclusion

We presented an approach to learn the discrete optimal reconstruction filter directly from the continuous Ramp filter. We have shown that the learning approach will automatically compensate for the errors inflicted by the discretization in an L2-sense optimal way with respect to our given training data. This is achieved by formulating a cost function to learn the filter in the frequency domain. This enables us to initialize the filter with the ideal Ramp, which can be seen as a very strong pre-training. Furthermore, the formulation can be straightforward transfered to a neural network architecture. Combining the proposed solution with the deep neural network suggested by Würfl et al. enables us to provide a fully trainable data-driven reconstruction deep learning approach. In future work, we want to apply noise models to the training data to learn an optimal discrete filter which is less sensitive to noise.

References

[1] Ge Wang, "A perspective on deep imaging," *IEEE Access*, vol. 4, pp. 8914–8924, 2016.

[2] Tobias Würfl, Florin Cristian Ghesu, Vincent Christlein, and Andreas Maier, "Deep Learning Computed Tomography," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016: 19th International Conference, Proceedings, Part III*, 2016, vol. 3, pp. 432–440.

[3] Kerstin Hammernik, Tobias Würfl, Thomas Pock, and Andreas Maier, "A deep learning architecture for limited-angle computed tomography reconstruction," in *Bildverarbeitung für die Medizin 2017: Algorithmen - Systeme - Anwendungen. Proceedings des Workshops vom 12. bis 14. März 2017 in Heidelberg*, Berlin, Heidelberg, 2017, pp. 92–97, Springer Berlin Heidelberg.

[4] Avinash C. Kak and Malcolm Slaney, *Principles of Computerized Tomographic Imaging*, IEEE Press, 1988.

[5] GN Ramachandran and AV Lakshminarayanan, "Three-dimensional reconstruction from radiographs and electron micrographs: application of convolutions instead of fourier transforms," *Proceedings of the National Academy of Sciences*, vol. 68, no. 9, pp. 2236–2240, 1971.

[6] Lawrence A Shepp and Benjamin F Logan, "The fourier reconstruction of a head section," *IEEE Transactions on Nuclear Science*, vol. 21, no. 3, pp. 21–43, 1974.
[7] Daniël M Pelt and Kees Joost Batenburg, "Improving filtered backprojection reconstruction by data-dependent filtering," *IEEE Transactions on Image Processing*, vol. 23, no. 11, pp. 4750–4762, 2014.
[8] Andreas Maier, Hannes Hofmann, Martin Berger, Peter Fischer, Chris Schwemmer, Haibo Wu, Kerstin Müller, Joachim Hornegger, Jang-Hwan Choi, Christian Riess, Andreas Keil, and Rebecca Fahrig, "CONRAD - A software framework for cone-beam imaging in radiology," *Medical Physics*, vol. 40, no. 11, 2013.
[9] C. E. Floyd, "An artificial neural network for spect image reconstruction," *IEEE Transactions on Medical Imaging*, vol. 10, no. 3, pp. 485–487, Sep 1991.

Deep Learning Interior Tomography for Region-of-Interest Reconstruction

Yoseob Han, Jawook Gu and Jong Chul Ye

***Abstract*—Interior tomography for the region-of-interest imaging has advantages of using a small detector and reducing X-ray radiation dose. However, standard analytic reconstruction suffers from severe cupping artifacts due to existence of null space in the truncated Radon transform. Existing penalized reconstruction methods may adress this problem but they require extensive computations due to the iterative reconstruction. Inspired by the recent deep learning approaches to low-dose and sparse view CT, here we propose a deep learning architecture that removes null space signals from the FBP reconstruction. Experimental results have shown that the proposed method provides near-perfect reconstruction with about 7-10 dB improvement in PSNR over existing methods in spite of significantly reduced run-time complexity.**

I. Introduction

X-ray Computed Tomography (CT) is one of the most powerful clinical imaging tools, delivering high-quality images in a fast and cost effective manner. However, the X-ray is harmful to the human body, so many studies has been conducted to develop methods that reduce the X-ray dose. Specifically, X-ray doses can be reduced by reducing the number of photons, projection views or the size of the field-of-view of X-rays. Among these, the CT technique for reducing the field-of-view of X-ray is called interior tomography. Interior tomography is useful when the region-of-interest (ROI) within a patient's body is small (such as heart), because interior tomography aims to obtain an ROI image by irradiating only the ROI with X-rays. Interior tomography not only can dramatically reduce the X-ray dose, but also has cost benefits by using a small-sized detector. However, the use of an analytic CT reconstruction algorithm generally produces images with severe artifacts due to the transverse directional projection truncation.

Sinogram extrapolation is a simple approximation method to reducte the artifacts. However, sinogram extrapolation method still generates sbiased CT number in the reconstructed image. Recently, Katsevich et al. [1] proved the general uniqueness results for the interior problem and provided stability estimates. Using the total variation (TV) penalty, the authos in [2] showed that a unique reconstruction is possible if the images are piecewise smooth. In a series of papers [3], [4], our group has shown that a generalized L-spline along a collection of chord lines passing through the ROI can be uniquely recovered [3]; and we further substantiated that the high frequency signal can be recovered analytically thanks to the Bedrosian identify, whereas the computationally expensive iterative reconstruction need only be performed to reconstruct the low frequency part of the signal after downsampling [4]. While this approach significantly reduces the computational complexity of the interior reconstruction, the computational complexity of existing iterative reconstruction algorithms prohibits their routine clinical use.

In recent years, deep learning algorithms using convolutional neural network (CNN) have been successfully used for low-dose CT [5], [6], sparse view CT [7], [8], etc. However, the more we have observed impressive empirical results in CT problems, the more unanswered questions we encounter. In particular, one of the most critical questions for biomedical applications is whether a deep learning-based CT does create any artificial structures that may mislead radiologists in their clinical decision. Fortunately, in a recent theory *deep convolutional framelets* [9], we showed that the success of deep learning is not from a magical power of a black-box, but rather comes from the power of a novel signal representation using non-local basis combined with data-driven local basis. Thus, the deep network is indeed a natural extension of classical signal representation theory such as wavelets, frames, etc; so rather than creating new informations, it attempts to extract the most information out of the input data using the optimal signal representation.

Inspired these findings, here we propose a deep learning framework for interior tomography problem. Specifically, we demonstrate that the interior tomography problem can be formulated as a reconstruction problem in an end-to-end manner under the constraints that remove the null space signal components of the truncated Radon transform. Numerical results confirmed the proposed deep learning architecture outperforms the existing interior tomography methods in image quality and reconstruction time.

II. Theory

A. Problem Formulation

Here, we consider 2-D interior tomography problem and follow the notation in [3]. The variable $\boldsymbol{\theta}$ denotes a vector on the unit sphere $\mathbb{S} \in \mathbb{R}^2$. The collection of vectors that are orthogonal to $\boldsymbol{\theta}$ is denoted as

$$\boldsymbol{\theta}^{\perp} = \{\mathbf{y} \in \mathbb{R}^2 \ : \ \mathbf{y} \cdot \boldsymbol{\theta} = 0\}.$$

We refer to real-valued functions in the spatial domain as images and denote them as $f(\mathbf{x})$ for $\mathbf{x} \in \mathbb{R}^2$. We denote the Radon transform of an image f as

$$\mathcal{R}f(\boldsymbol{\theta}, s) := \int_{\boldsymbol{\theta}^{\perp}} f(s\boldsymbol{\theta} + \mathbf{y}) d\mathbf{y} \tag{1}$$

Y. Han, J. Gu, and J. Ye are with the Department of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea (e-mail: {hanyoseob, jwisdom9299, jong.ye}@kaist.ac.kr).

where $s \in \mathbb{R}$ and $\boldsymbol{\theta} \in \mathbb{S}$. The local Radon transform for the truncated field-of-view is the restriction of $\mathcal{R}f$ to the region $\{(\boldsymbol{\theta}, s) \ : \ |s| < \mu\}$, which is denoted as $\mathcal{T}_\mu \mathcal{R} f$. Then, the interior reconstruction is to find the unknown $f(\mathbf{x})$ within the ROI from $\mathcal{T}_\mu \mathcal{R} f$.

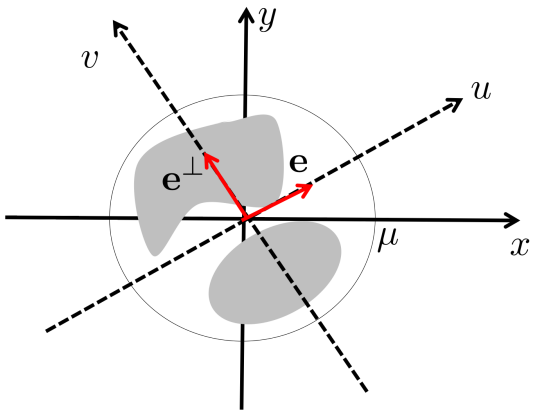


Fig. 1: The coordinate system for interior tomography.

B. Null Space of Truncated Radon Transform

The main technical difficulty of the interior reconstruction is the existence of the null space [3], [10]. To analyze the null space, we follow the mathematicla analysis in [3]. Specifically, the analytic inversion of $\mathcal{T}_\mu \mathcal{R} f$ can be equivalently represented using the differentiated backprojection followed by the truncated Hilbert transform along the chord lines, se we analyze the interior reconstruction problem to take advantages of this. More specifically, if the unit vector $\mathbf{e} \in \mathbb{R}^2$ along the chord line is a set as a coordinate axis, then we can find the unit vector $\mathbf{e}^\perp \in \mathbb{R}^2$ such that $V = [\mathbf{e}, \mathbf{e}^\perp]$ consists of the basis for the local coordinate system and $(u, v) \in \mathbb{R}^2$ denotes its coordinate value (see Fig. 1). We further define 1-D index set parameterized by the v:

$$I_\mu(v) := \{u' \in \mathbb{R} \mid \sqrt{(u')^2 + v^2} \leq \mu\}$$

Then, the null space of the $\mathcal{T}_\mu \mathcal{R} f$ is given by [3], [4],

$$\mathcal{N}_\mu := \left\{ g \mid g(u,v) = -\int_{u' \notin I_\mu(v)} \frac{du'}{\pi(u-u')} \psi(u', v) \right\}$$

for some functions $\psi(u, v)$. A typical example of the null space image g is illustrated in Fig. 2. This is often called as the cupping artifact. The cupping artifacts reduce contrast and interfere with clinical diagnosis.

Note that the null space signal $g \in \mathcal{N}_\mu$ is differentiable in any order due to the removal of the origin in the integrand. Accordingly, an interior reconstruction algorithm needs an appropriate regularization term that suppresses $g \in \mathcal{N}_\mu$ by exploiting this. Specifically, one could find an analysis transform L such that its null space $\mathcal{N}_\mathrm{L}$ is composed of entire function, and use it for an analysis-based regularization term. For example, the regularization using TV [2] and L-spline model [3], [4] correspond to this. The main result on

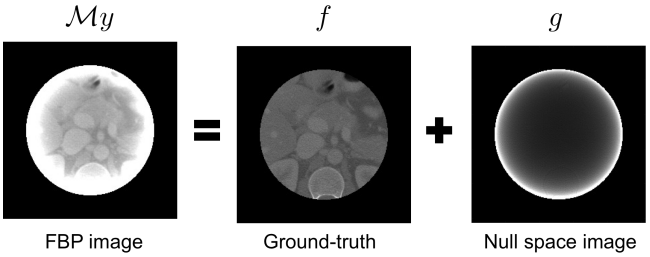


Fig. 2: Decomposition of the analytic reconstruction into null space component and the true image.

the perfect reconstruction in [3] is then stated as follows. If the null space component $g \in \mathcal{N}_\mu$ is equivalent to a signal $h \in \mathcal{N}_\mathrm{L}$ within the ROI, then g is identically zero due to the characterization of Hilbert transform pairs as boundary values of analytic functions on the upper half of the complex plane [3]; so TV or L-spline regularization provides the unique solution.

C. CNN-based Null Space Removal

Instead of designing a linear operator L such that the common null space of $\mathcal{N}_\mu$ and $\mathcal{N}_\mathrm{L}$ to be zero, we can design a frame $\mathcal{W}$ and its dual $\tilde{\mathcal{W}}$ such that $\tilde{\mathcal{W}}^\top \mathcal{W} = I$ and $\tilde{\mathcal{W}}^\top S_\lambda \mathcal{W}(f^* + g) = f^*$ for all $g \in \mathcal{N}_\mu$ and the ground-truth image f^*. This frame-based regularization is also an active field of research for image denoising, inpainting, etc [11].

One of the most important contributions of the deep convolutional framelet theory [9] is sthat $\mathcal{W}$ and $\tilde{\mathcal{W}}^\top$ correspond to the encoder and decoder structure of a CNN, respectively, and the shirinkage operator S_λ emerges by controlling the number of filter channels and nonlinearities. Accordingly, a convolutional neural network represented by $\mathcal{Q} = \tilde{\mathcal{W}}^\top S_\lambda \mathcal{W}$ can be designed such that

$$\mathcal{Q}(f^* + g) = f^*, \quad \forall g \in \mathcal{N}_\mu \tag{2}$$

Then, our interior tomography algorithm is formulated to find the solution f for the following problem:

$$y = \mathcal{T}_\mu \mathcal{R} f, \quad \mathcal{Q} f = f^* \tag{3}$$

where f^* denotes the ground-truth data available for training data, and $\mathcal{Q}$ denotes the CNN satisfying (2). Now, by defining $\mathcal{M}$ as a right-inverse of $\mathcal{T}_\mu \mathcal{R}$, i.e. $(\mathcal{T}_\mu \mathcal{R})\mathcal{M} y = y, \forall y$, we have

$$\mathcal{M} y = f^* + g$$

for some $g \in \mathcal{N}_\mu$, since the right-inverse is not unique due to the existence of the null space. See Fig. 2 for the decomposition of $\mathcal{M}y$. Thus, $\mathcal{M}y$ is a feasible solution for (3), since

$$\mathcal{Q}\mathcal{M} y = \mathcal{Q}(f^* + g) = f^* \tag{4}$$

and the data fidelity constraint is automatically satisfied due to the definition of the right-inverse. Therefore, the neural network training problem to satisfy (4) can be equivalently represented by

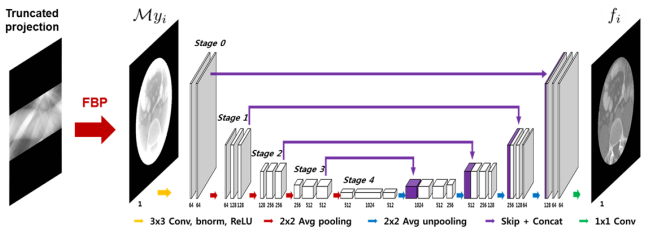


Fig. 3: The proposed deep learning architecture for inteior tomography.

$$\min_{\mathcal{Q}} \sum_{i=1}^{N} \|f_i^* - \mathcal{Q}\mathcal{M}y_i\|^2 \tag{5}$$

where $\{(f_i^*, y_i)\}_{i=1}^N$ denotes the training data set composed of ground-truth image an its truncated projection. A typical example of the right-inverse for the truncated Radon transform is the inverse Radon transform, which can be implemented by the filtered backprojection (FBP) algorithm. Thus, $\mathcal{M}y_i$ in (5) can be implemented using the FBP.

After the neural network $\mathcal{Q}$ is learned, the inference can be done simply by processing FBP reconstruction image from a truncated radon data y_t using the neural network $\mathcal{Q}$, i.e. $\hat{f} = \mathcal{Q}\mathcal{M}y_t$. The details of the network $\mathcal{Q}$ and the training procedure will be discussed in the following section.

III. Method

A. Data Set

Ten subject data sets from AAPM Low-Dose CT Grand Challenge were used in this paper. Out of ten sets, eight sets were used for network training. The other two sets were used for validation and test, respectively. The provided data sets were originally acquired in helical CT, and were rebinned from the helical CT to 360° angular scan fan-beam CT. The 512×512 size artifact free CT images are reconstructed from the rebinned fan-beam CT data using FBP algorithm. From the CT image, sinogram is numerically obtained using a forward projection operator. The number of detector in numerical experiment is 736. Only 350 detectors in the middle of 736 detectors are used to simulate the truncated projection data. Using this, we reconstruct 256×256 ROI images.

B. Network Architecture

The proposed network is shown in Fig. 3. The first layer is the FBP layer that reconstructs the cupping-artifact corrupted images from the truncated projection data, which is followed by a modified architecture of U-Net [12]. A yellow arrow in Fig. 3 is the basic operator and consists of 3×3 convolutions followed by a rectified linear unit and batch normalization. The yellow arrows between the seperate blocks at every stage are omitted. A red arrow is a 2×2 average pooling operator and is located between the stages. The average pooling operator doubles the number of channels and reduces the size of the layers by four. Conversely, a blue arrow is 2×2 average unpooling operator, reducing the number of channels by half and increasing the size of the layer by four. A violet arrow is the skip and concatenation operator. A green arrow is teh simple 1×1 convolution operator generating the final reconstruction image.

C. Network Training

The proposed network was implement MatConvNet toolbox in MATLAB R2015a environment. Processing units used in this research are Intel Core i7-7700 central processing unit and GTX 1080-Ti graphics processing unit. Stochastic gradient reduction was used to train the network. As shown in Fig. 3, the inputs of the network are the truncated projection data, i.e. y_i. The target data f_i corresponds to the 256×256 size center ROI image cropped from the ground-truth data. The number of epochs was 300. The initial learning rate was 10^{-3}, which gradually dropped to 10^{-5}. The regularization parameter was 10^{-4}. Training time lasted about 24 hours.

IV. Results

We compared the proposed method with existing iterative methods such as the TV penalizaed reconstruction [2] and the L-spline based multi-scale regularization method by Lee et al. [4]. Fig. 4 shows the ground-truth images and reconstruction results by FBP, TV, Lee method [4] and the proposed method. The graphs in the bottom for in Fig. 4 are the crosss-section view along the white lines on the each images. Fig. 5 shows the magnitude of difference images between the ground-truth image and reconstruction results of each method. The reconstructed images and the cut-view graphs in Fig.4 show that the proposed method results have more fine detalis than the other methods. The error images in Fig. 5 confirm that the high frequency components such as edges and textures are better restored in the proposed method than other method.

We also calculated the average values of the peak signal-to-noise ratio (PSNR) and the normalized mean square error (NMSE) in Table I. The proposed method achieved the highest value in PSNR and the lowest value in NMSE with about 7-10 dB improvement. The computational times for TV, Lee method [4] and the proposed method were 1.8272s, 0.3438s, and 0.0532s, respectively, for each slice reconstruction. The processing speed of the proposed method is about 34 times fater than the TV method and 6 times faster than Lee method [4].

TABLE I: Quantitative comparison of various methods.

	FBP	TV	Lee method [4]	Proposed
PSNR [dB]	9.4099	30.2004	27.0344	37.4600
NMSE	8.2941e-1	6.9137e-3	1.4332e-2	1.2994e-3

V. Conclusion

In this paper, we proposed a deep learning network for interior tomography problem. The reconstruction problem was formulated as a constraint optimization problem under data fidelity and null space constraints. Based on the theory of deep convolutional framelet, the null space constraint was implemented using the convolutional neural network with

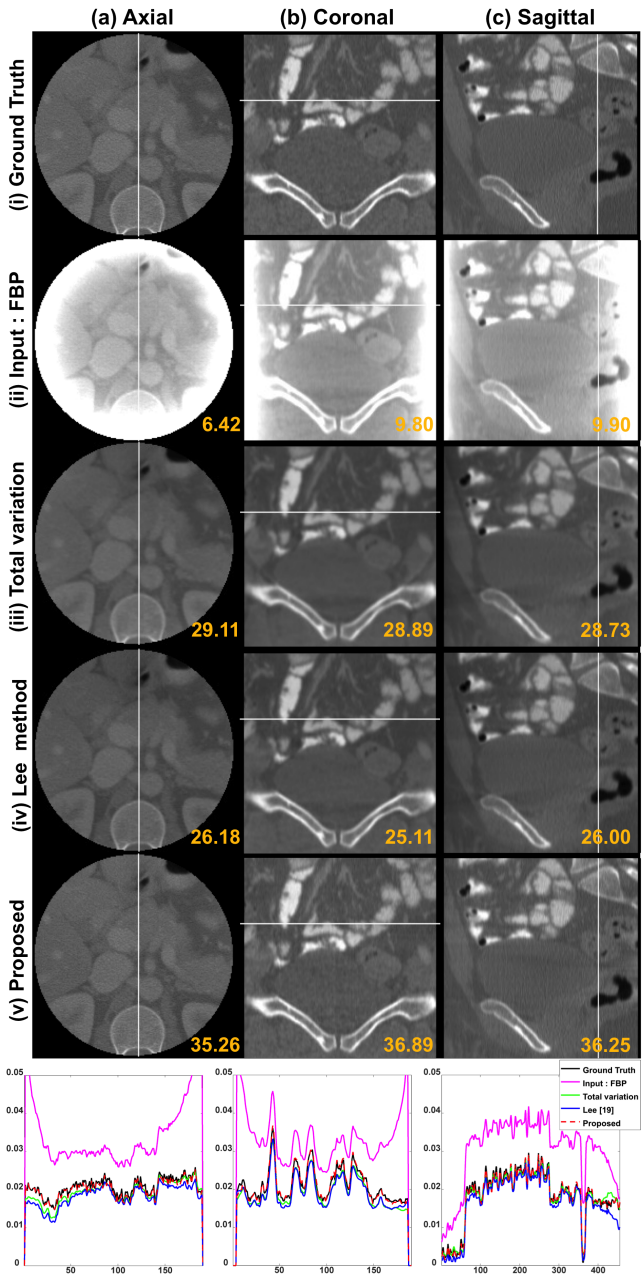


Fig. 4: Reconstruction images by the cone-beam simulation. The last row shows the cut-view plots of the white lines on the images. The number written in the images is the PSNR value in dB.

encoder and decoder architecture. Numerical results showed that the proposed method has the highest value in PSNR and the lowest value in NMSE and the fastest computational time.

VI. Acknowledgement

The authors would like to thanks Dr. Cynthia McCollough, the Mayo Clinic, the American Association of Physicists in Medicine (AAPM), and grant EB01705 and EB01785 from the National Institute of Biomedical Imaging and Bioengineering for providing the Low-Dose CT Grand Challenge data set. This work is supported by national Research Foundation of Korea, Grant number NRF-2016R1A2B3008104. Yoseob Han and Jawook Gu contributed equally to this work.

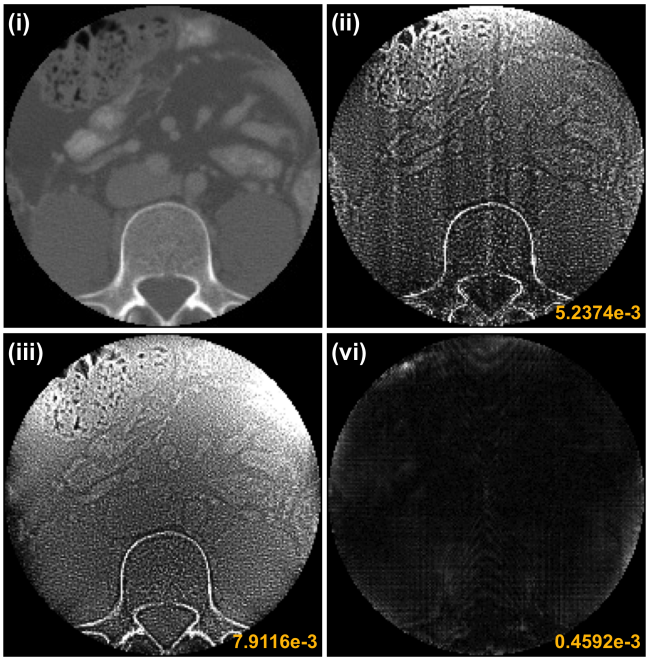


Fig. 5: Error images from (ii) TV, (iii) Lee method [4], and (iv) the proposed method. (i) is ground truth image. The number written in the images is the NMSE value.

References

[1] E. Katsevich, A. Katsevich, and G. Wang, "Stability of the interior problem with polynomial attenuation in the region of interest," *Inverse Problems*, vol. 28, no. 6, p. 065022, 2012.

[2] H. Yu and G. Wang, "Compressed sensing based interior tomography," *Physics in Medicine and Biology*, vol. 54, no. 9, p. 2791, 2009.

[3] J. P. Ward, M. Lee, J. C. Ye, and M. Unser, "Interior tomography using 1D generalized total variation. Part I: Mathematical foundation," *SIAM Journal on Imaging Sciences*, vol. 8, no. 1, pp. 226–247, 2015.

[4] M. Lee, Y. Han, J. P. Ward, M. Unser, and J. C. Ye, "Interior tomography using 1D generalized total variation. Part II: Multiscale implementation," *SIAM Journal on Imaging Sciences*, vol. 8, no. 4, pp. 2452–2486, 2015.

[5] E. Kang, J. Min, and J. C. Ye, "A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction," *Medical Physics*, vol. 44, no. 10, 2017.

[6] H. Chen, Y. Zhang, M. K. Kalra, F. Lin, Y. Chen, P. Liao, J. Zhou, and G. Wang, "Low-dose CT with a residual encoder-decoder convolutional neural network," *IEEE Transactions on Medical Imaging*, vol. 36, no. 12, pp. 2524–2535, 2017.

[7] Y. Han, J. Yoo, and J. C. Ye, "Deep residual learning for compressed sensing CT reconstruction via persistent homology analysis," *arXiv preprint arXiv:1611.06391*, 2016.

[8] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4509–4522, 2017.

[9] J. C. Ye, Y. S. Han, and E. Cha, "Deep convolutional framelets: A general deep learning framework for inverse problems," *arXiv preprint arXiv:1707.00372*, 2017.

[10] A. Katsevich and A. Tovbis, "Finite Hilbert transform with incomplete data: Null-space and singular values," *Inverse Problems*, vol. 28, no. 10, p. 105006, 2012.

[11] J.-F. Cai, R. H. Chan, and Z. Shen, "A framelet-based image inpainting algorithm," *Applied and Computational Harmonic Analysis*, vol. 24, no. 2, pp. 131–149, 2008.

[12] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2015, pp. 234–241.

Towards automatic Abdominal Multi-Organ Segmentation in Dual Energy CT using cascaded 3D Fully Convolutional Network

Shuqing Chen, Holger Roth, Sabrina Dorn, Matthias May, Alexander Cavallaro, Michael Lell, Marc Kachelrieß, Hirohisa Oda, Kensaku Mori and Andreas Maier

***Abstract*—Automatic multi-organ segmentation of the dual energy computed tomography (DECT) data can be beneficial for biomedical research and clinical applications. Recent advances in deep learning showed the feasibility to use 3-D fully convolutional networks (FCN) for voxel-wise dense predictions in single energy computed tomography (SECT). In this paper, we proposed a 3D FCN based method for automatic multi-organ segmentation in DECT. The work was based on a cascaded FCN for the major organs trained on a large set of SECT data. We preprocessed the DECT data by using linear weighting and fine-tuned the FCN for the DECT data. The method was evaluated using 42 torso DECT data acquired with a clinical dual-source CT system. Four abdominal organs (liver, spleen, left and right kidneys) were evaluated with cross-validation strategy. Effect of the weight on the accuracy was researched. In all the tests, we achieved an average Dice coefficient of 93% for the liver, 92% for the spleen, 91% for the right kidney and 89% for the left kidney, respectively. The results show that our method is feasible and promising.**

I. Introduction

The Hounsfield unit (HU) scale value depends on the inherent tissue properties, the x-ray spectrum for scanning and the administered contrast media [1]. In a SECT image, materials having different elemental compositions can be represented by identical HU values [2]. Therefore, SECT has challenges such as limited material-specific information and beam hardening as well as tissue characterization [1]. DECT has been investigated to solve the challenges of SECT. In DECT, two energy-specific image data sets are acquired at two different X-ray spectra, which are produced by different energies, simultaneously. The multi-organ segmentation in DECT can be beneficial for biomedical research and clinical applications, such as material decomposition [3], organ-specific context-sensitive enhanced reconstruction and display [4], [5], and computation of bone mineral density [6]. We are aiming at exploiting the prior anatomical information that is gained through the multi-organ segmentation to provide an improved context-sensitive DECT imaging [4], [5]. The novel technique offers the possibility to present evermore complex information to the radiologists simultaneously and bears the potential to improve the clinical routine in CT diagnosis.

Automatic multi-organ segmentation on DECT images is a challenging task due to the inter-subject variance of human abdomen, the complex 3-D intra-subject variance among organs, soft anatomy deformation, as well as different HU values for the same organ by different spectra. Recent researches show the power of deep learning in medical image segmentation [7]. To solve the DECT segmentation problem, we use the successful experience from multi-organ segmentation in volumetric SECT images using deep learning [8], [9]. The proposed method is based on a cascaded 3D FCN, a two-stage, coarse-to-fine approach [8]. The first stage is used to predict the region of the interest (ROI) of the target organs, while the second stage is learned to predict the final segmentation. No organ-specific or energy-specific prior knowledge is required in the proposed method. The cross-validation results showed that the proposed method is promising to solve multi-organ segmentation problem for DECT. To the best of our knowledge, this is the first study about multi-organ segmentation in DECT images based on 3D FCNs.

II. Materials and Methods

A. Network Architecture for DECT Prediction

As described by Krauss et al. [10], a mixed image display is employed in clinical practice for the diagnose using DECT. The mixed image is calculated by linear weighting of the images values of the two spectra:

$$I_{\text{mix}} = \alpha \cdot I_{\text{low}} + (1 - \alpha) \cdot I_{\text{high}} \tag{1}$$

where α is the weight of the dual energy composition, I_{mix} denotes the mixed image. I_{low} and I_{high} are the images at low and high kV, respectively.

We preprocessed the DECT images following Eq. 1 straightforwardly. Figure 1 illustrates the network architecture of the proposed method for the DECT multi-organ segmentation. To prepare network training, labeled segmentation is generated manually by experts for each training data. In the training phase, first of all, mixed image is calculated by combining the images at the low energy level and the high energy level using Eq. 1. Then, a binary mask is generated by thresholding the skin contour of the mixed image. Subsequently, the mixed

S. Chen is with Pattern Recognition Lab, Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen, Germany (e-mail: shuqing.chen@fau.de).

H. Roth, H. Oda, and K. Mori are with Nagoya University, Nagoya , Japan.

S. Dorn, and M. Kachelrießare with German Cancer Research Center (DKFZ), Heidelberg, Germany, and also with Ruprecht-Karls-University Heidelberg, Heidelberg, Germany.

M. May, and A. Cavallaro are with Department of Radiology, University Hospital Erlangen, Erlangen, Germany.

M. Lell is with University Hospital Nürnberg, Paracelsus Medical University, Nürnberg, Germany.

A. Maier is with Pattern Recognition Lab, Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen, Germany.

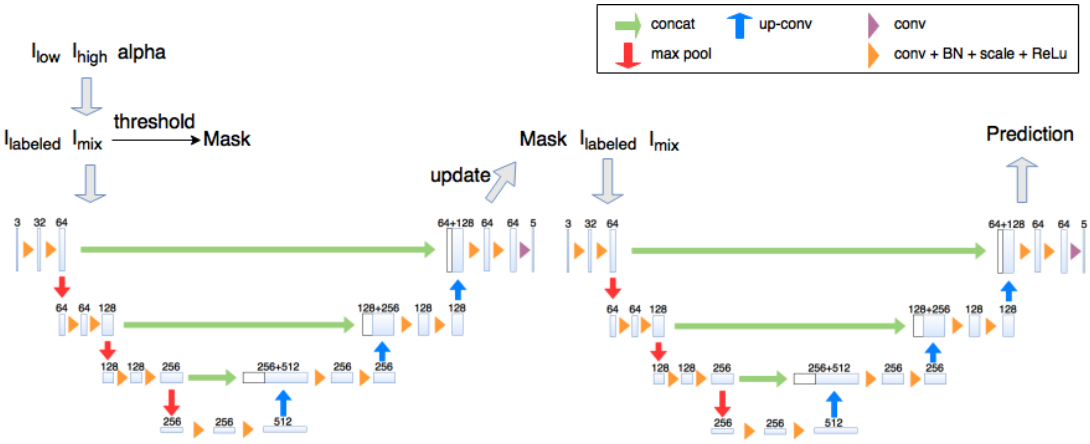


Fig. 1: Cascaded network architecture for DECT multi-organ segmentation

image, the binary mask and the labeled image are given into the network as multi-channel inputs. The network consists of two stages. The first stage is applied to generate the region of the interest (ROI) in order to reduce the search space for the second stage. The prediction result of the first stage is taken as the mask for the second stage. Each stage is based on a standard 3D U-Net [11], which is a fully convolutional network including an analysis and a synthesis path. We used the open-source implementation of two stages cascaded network [8] developed by Roth et al. based on the 3D U-Net [11] and the Caffe deep learning library [12]. The cascaded network was trained by Roth et al. [8] on a large set of SECT images including some of the major organ labels. Our model was trained by fine-tuning the pre-trained network with the mixed DECT images using the pre-trained weights as initialization. The difference between the network output and the ground truth labels are compared using softmax with weight voxel-wise cross-entropy loss [8], [11].

B. Experimental Setup

The proposed method was evaluated with 42 clinical torso DECT images scanned by the department of radiology, university hospital Erlangen. All of the images were taken from male and female adult patients who had different clinically oriented indication justified by the radiologist. Ultravist 370 was given as contrast agent with body weight adapted volumes. The images were acquired at different X-ray tube voltage setting of 70 kV (560 mAs) and Sn 150 kV (140 mAs, with Sn filter) using a Siemens SOMATOM Force CT system with Stellar, an energy integrating detector. The traning volumes contains 992-1290 slices with slice size 512x512 pixels. The voxel dimensions are [0.6895-0.959, 0.6895-0.959, 0.6] mm. Four abdominal organs were tested, including liver, spleen, right and left kidneys. Ground truth was generated by experts manually.

To avoid the bias of the data selection and to keep the dataset distribution similar, a malnifold learning-based technique [13] was applied to split the data into training dataset, validation dataset, and test dataset. First, the images were resized to the same image spacing (e.g.[3mm 3mm 5mm]). Then, the distribution of the images was calculated and plotted by using locally linear embedding (LLE) [14]. Subsequently, the images were clustered into 3 classes using k-means. Finally, training data, validation data, and test data were selected randomly from these classes with the ratio 5:1:1, i.e. in each test we used 2 images from each class (6 in total) for validation, 2 images from each class for test (6 in total), and the remaining 30 images for training.

		Liver	Spleen	r.Kidney	l.Kidney
DECT	Avg.	0.92	0.84	0.88	0.87
	SD	0.02	0.08	0.03	0.03
	Min.	0.84	0.62	0.80	0.78
	Max.	0.94	0.95	0.94	0.93

TABLE I: Dice coefficients of cross-validation with α_{training}=0.6 and α_{test} =0.6. SD is abbreviated for standard deviation.

III. RESULTS

A. Performance Estimation with Cross-Validation

NVIDIA GeForce GTX 1080 Ti with 11 GB memory was used for all of the experiments. The similarity between the segmentation result and the ground truth was measured with Dice metric by using the tool provided by VISCERAL [15]. First, the performance of the proposed method was estimated by 8-folds cross-validation, using 0.6 as α_{training} as well as α_{test}. Fig. 2 shows one segmentation results in 3-D. TableI summarizes the Dice coefficients of the segmentation results and compares DECT results with the SECT results. The proposed method under the above weight condition yielded an average Dice coefficient of 92% for the liver, 84% for the spleen, 88% for the right kidney and 87% for the left kidney, respectively. Fig. 3 plots the distributions of the Dice coefficients for different test scenarios and showed the high robustness of the proposed method. Though the Dice coefficients under above mentioned weight condition are less than SECT results in [9], we performed a second test which is focused on the weight alpha both for training and for test phase.

B. Study on the Weight α

We are aiming at exploiting the spectral information in the DECT data. Since the α mixing results basically in pseudo monochromatic images comparable to single energy scans, the influence of the weight α on the accuracy was further

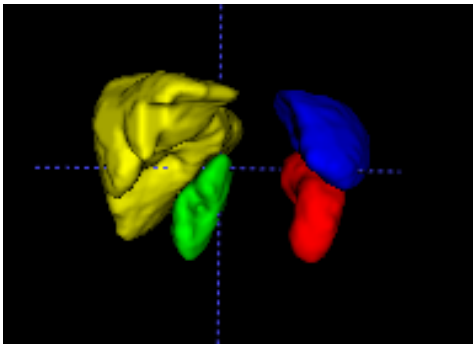


Fig. 2: 3D rendering of one DECT segmentation with yellow for liver, blue for spleen, green for right kidney and red for

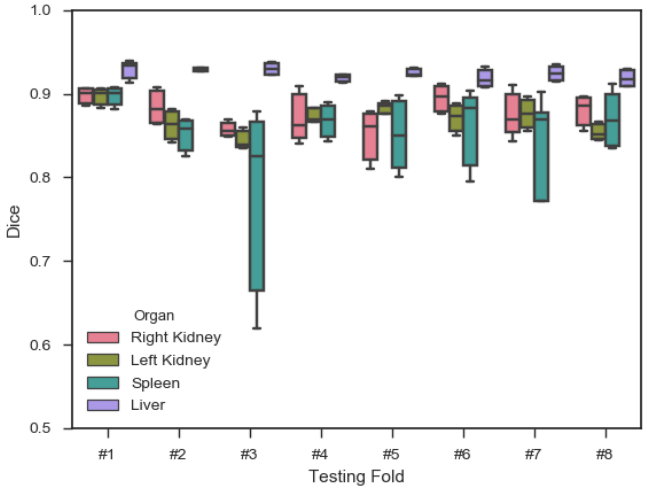


Fig. 3: Dice coefficients of the target organs with $\alpha_{\text{training}} = 0.6$ and $\alpha_{\text{test}} = 0.6$ for 8 different testing folds

researched. 0, 0.3, 0.6, 0.9 and 1 were chosen as α_{training} and α_{test} in this study. Table II lists the average Dice coefficient. For all of the cases, the liver had the highest accuracy (92%-93%). The segmentation of the right kidney was usually more accurate than the left kidney. The best Dice values per organ per training set are highlighted in Table II. The test with α_{training}=0.9 and α_{test}=0.9 obtained the highest accuracy for liver and right kidney. The test with weight combination 0.9-1 showed the best segmentation for spleen, the combination with 0.9-0.3 had the finest result for left kidney. High α_{test} generated better segmentation for liver and spleen. For most organs, the best Dice values of DECT are higher than the SECT results given in [9].

IV. DISCUSSION AND CONCLUSION

We proposed a deep learning based method for automatic abdominal multi-organ segmentation in DECT. The evaluation results show the feasibility of the proposed method. Compared to the results of the SECT images reported by Roth et al. [9], our method is promising and robust (see Table II). For most organs, the segmentation of our method is more accurate than the SECT [9] when an optimal fusion weight is selected. The results illustrate that the image fusion affects the segmentation of DECT. In the cross validation, the third testing fold had a large deviation. The reason could be that our image data were

α_{training}-α_{test}	Liver	Spleen	r.Kidney	l.Kidney
0-0	0.908	*0.878*	0.840	*0.852*
0-0.3	0.915	0.876	0.860	0.841
0-0.6	0.919	0.875	*0.865*	0.839
0-0.9	0.922	0.876	0.864	0.837
0-1	*0.923*	0.876	0.861	0.835
0.3-0	0.876	0.885	0.845	0.835
0.3-0.3	0.924	0.899	*0.900*	*0.891*
0.3-0.6	0.925	*0.902*	0.891	0.881
0.3-0.9	*0.926*	0.901	0.877	0.859
0.3-1	0.921	0.900	0.877	0.854
0.6-0	0.865	0.857	0.786	0.796
0.6-0.3	0.909	0.897	0.844	0.885
0.6-0.6	*0.922*	0.904	*0.895*	*0.887*
0.6-0.9	0.912	0.906	*0.895*	0.873
0.6-1	0.919	*0.908*	0.843	0.866
0.9-0	0.881	0.848	0.745	0.764
0.9-0.3	0.930	0.901	0.898	**0.892**
0.9-0.6	0.932	0.908	0.904	0.873
0.9-0.9	**0.933**	0.915	**0.906**	0.862
0.9-1	0.930	**0.917**	0.905	0.872
1-0	0.907	0.822	0.784	0.812
1-0.3	0.912	0.886	0.869	0.889
1-0.6	0.915	0.895	*0.879*	0.886
1-0.9	0.917	0.901	*0.879*	*0.891*
1-1	*0.918*	*0.902*	0.877	*0.891*
SECT [9]	0.95	0.90	0.90	0.88

TABLE II: Dice coefficients of different alpha for testing fold 1. Bold denotes the best organ results of DECT. Italic underline denotes the best results in the group with the same training weight. Notice that the DECT and SECT approaches used different data set.

taken from patients with different disease (liver tumor, spleen tumor, etc.). The disease type is not considered by the data selection. Training and test with inconsistent symptoms could have an impact on the accuracy.

The study on the weight can be divided into three groups with different α_{training}. α=0.9 is close to the low energy images which have on average the best soft-tissue contrast, α_{training}=0.9 worked thus better in general. The intra-group comparison showed that the cases with identical training and test conditions had a higher probability to get the best segmentation result. This is expected because the mixed images generated by the matched training and test conditions may have the highest similarity. Furthermore, the comparison of the case 0.3-0.9 (low-contrast model for high-contrast image) with the case 0.9-0.3 (high-contrast model for low-contrast image) showed that using a model trained on high-contrast images for segmenting low-contrast test images works better. In addition, liver is well segmented in middle to high α ranges. Spleen is segmented best at α=0.6. Kidneys work best in matched training and test conditions. This suggests that there is an optimal α for each organ for image segmentation.

The weight α for the mixed image calculation is currently a user-defined parameter in the preprocessing in our approach. The fact suggests that the alpha shall be regarded as organ specific parameter in the network and optimized in the training phase. It can be used to augment the data for the training in future. Also, the network could be modified with two image inputs. Furthermore, more organs and more scans from different patients could be used.

ACKNOWLEDGMENTS This work was supported by the

German Research Foundation (DFG) through research grant No. KA 1678/20, LE 2763/2-1 and MA 4898/5-1.

REFERENCES

[1] Dushyant Sahani, "Dual-energy CT: The technological approaches," Society of Computed Body Tomography and Magnetic Resonance (SCBT-MR), 2012.

[2] Cynthia H. McCollough, Shuai Leng, Lifeng Yu, and Joel G. Fletcher, "Dual- and multi-energy CT: Principles, technical approaches, and clinical applications," *Radiology*, vol. 276, no. 3, pp. 637–653, 2015.

[3] Stefan Kuchenbecker, Sebastian Faby, David Simons, Michael Knaup, Heinz-Peter Schlemmer, Michael M. Lell, and Marc Kachelriess, "Segmentation-assisted material decomposition in dual energy computed tomography (DECT)," in *Radiological Society of North America (RSNA)*, 2015.

[4] Sabrina Dorn, Shuqing Chen, Francesco Pisana, Joscha Maier, Michael Knaup, Stefan Sawall, Andreas Maier, Michael Lell, and Marc Kachelrieß, "Organ-specific context-sensitive single and dual energy CT (DECT) image reconstruction, display and analysis," in *103rd Scientific Assembly and Annual Meeting of the Radiological Society of North America (RSNA)*, 2017.

[5] Sabrina Dorn, Shuqing Chen, Stefan Sawall, David Simons, Matthias May, Joscha Maier, Michael Knaup, Heinz-Peter Schlemmer, Andreas Maier, Michael Lell, and Marc Kachelrieß, "Organ-specific context-sensitive CT image reconstruction and display," in *Medical Imaging Proc. SPIE*, 2018.

[6] S Wesarg, M Kirschner, M Becker, M Erdt, K Kafchitsas, MF Khan, et al., "Dual-energy CT-based assessment of the trabecular bone in vertebrae," *Methods of information in medicine*, vol. 51, no. 5, pp. 398, 2012.

[7] Marc Aubreville, Miguel Goncalves, Christian Knipfer, Nicolai Oetter, Tobias Würfl, Helmut Neumann, Florian Stelzle, Christopher Bohr, and Andreas K. Maier, "Patch-based carcinoma detection on confocal laser endomicroscopy images - A cross-site robustness assessment," *CoRR*, vol. abs/1707.08149, 2017.

[8] Holger R. Roth, Hirohisa Oda, Yuichiro Hayashi, Masahiro Oda, Natsuki Shimizu, Michitaka Fujiwara, Kazunari Misawa, and Kensaku Mori, "Hierarchical 3D fully convolutional networks for multi-organ segmentation," in *arXiv preprint arXiv:1704.06382*.

[9] Holger Roth, Ying Yang, Masahiro Oda, Hirohisa Oda, Yuichiro Hayashi, Natsuki Shimizu, Takayuki Kitasaka, Michitaka Fujiwara, Kazunari Misawa, and Kensaku Mori, "Torso organ segmentation in CT using fine-tuned 3D fully convolutional networks," in *36 (JAMIT)*, 2017.

[10] Bernhard Krauss, Bernhard Schmidt, and Thomas G. Flohr, *Dual Energy CT in Clinical Practice*, chapter Dual Source CT, Springer Berlin Heidelberg, 2011.

[11] Özgün Çiçek, Ahmed Abdulkadir, Soeren S. Lienkamp, Thomas Brox, and Olaf Ronneberger, "3D U-Net: Learning dense volumetric segmentation from sparse annotation," in *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, 2016.

[12] Yangqing Jia, Evan Shelhamer, Jeff Donahue, Sergey Karayev, Jonathan Long, Ross Girshick, Sergio Guadarrama, and Trevor Darrell, "Caffe: Convolutional architecture for fast feature embedding," *arXiv preprint arXiv:1408.5093*, 2014.

[13] Shuqing Chen, Sabrina Dorn, Michael Lell, Marc Kachelrieß, and Andreas Maier, "Manifold learning-based data sampling for model training," in *Informatik-Aktuell*, Bildverarbeitung für die Medizin: Algorithmen-Systeme-Anwendungen, Ed., Germany, 2018, pp. 269–274.

[14] Sam T. Roweis and Lawrence K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.

[15] Abdel Aziz Taha and Allan Hanbury, "Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool," *BMC Medical Imaging*, vol. 15, pp. 29, August 2015.

Noise Subtraction for Low-Dose CT Images Using a Deep Convolutional Neural Network

Andrew D. Missert, Shuai Leng, Lifeng Yu, and Cynthia H. McCollough,

***Abstract*—Minimizing the patient's exposure to ionizing radiation mandates that CT images be acquired using as few X-ray photons as reasonably possible, which increases the prominence of noise in the reconstructed image. This problem can be mitigated by applying denoising techniques to reduce image noise, while preserving the underlying anatomical information. We introduce a noise-subtraction algorithm which employs a deep convolutional neural network (CNN) that is trained to identify noise in low-dose CT images. The noise identified by the CNN is then subtracted to produce an output image with reduced noise. The CNN architecture is inspired by state-of-the-art networks for image classification, and employs sequential identical blocks of aggregated residual transformations to facilitate the training of many layers. The training step is performed by finding the network parameters which minimize the mean squared error difference between full dose and simulated quarter dose images, which are cropped from patient scans in the "*2016 NIH-AAPM-Mayo Clinic Low Dose CT Grand Challenge*" dataset. The trained CNN is able to significantly reduce the noise in the low-dose CT images, while also preserving sharp edges and fine anatomical details.**

I. Introduction

Computed tomography (CT) is a widely used imaging modality in the modern clinical setting. To acquire a CT image, the patient must be subjected to ionizing radiation in the form of X-ray photons. In order to keep patient risk from exposure to an absolute minimum, there is clinical pressure to acquire CT images using the smallest possible X-ray flux that doesn't compromise the diagnostic task. The reduced flux in low-dose CT (LDCT) scans increases the variance of the measured data, which results in increased noise in the reconstructed images. This CT image noise may impact diagnostic accuracy in a few key ways:

1) The noise may hide an anatomical detail or pathology.
2) The noise may simulate a false anatomical detail or pathology.
3) The noise may contribute to reader fatigue.

Denoising algorithms are employed to address the above points by reducing the image noise in LDCT images while preserving the relevant diagnostic information.

There exists a wide variety of approaches for denoising for LDCT scans. Projection-based methods [1] [2] [3] can access the full information of the acquired sinogram, which comes at the cost of folding the scanner-dependent and sometimes proprietary reconstruction method into the result. Image-based methods [4] [5] [6], on the other hand, are generally more portable, which comes at the cost of not having access to any information which lost in the reconstruction process.

In general, the quality of a particular noise reduction algorithm depends on how well it addresses each of the potential problems outlined above, with natural trade-offs existing when addressing these problems. To address problem 3, the noise textures should be "smoothed over" to make the anatomical features more obvious, a procedure which can make problem 1 and problem 2 worse. To address problem 1, fine structures in the noise should be amplified, which can exacerbate problems 2 and 3, and so on.

A current trend in LDCT noise reduction is to use deep learning techniques to train mathematical models for image-based denoising tasks. In particular, the application of convolutional neural networks (CNNs) trained on synthetic LDCT data has shown promising results [7]. In this work, we employ an image-based denoising algorithm that is based on a deep residual CNN that is trained to reduce LDCT image noise using image-based noise subtraction. The CNN architecture is fully convolutional and utilizes repeated blocks of aggregated residual convolutions inspired by state-of-the-art networks for image classification suck as ResNet [8] [9]. Training is performed using images clipped from both the full-dose (FD) and simulated quarter-dose (QD_{sim}) scans in the "*2016 NIH-AAPM-Mayo Clinic Low Dose CT Grand Challenge*" dataset. The resulting algorithm significantly reduces the LDCT noise, while also preserving fine structures and small contrast variations arising from the patient anatomy.

II. Methods

A. Denoising With A Convolutional Neural Network

The neural network used in this work is structured like a denoising autoencoder. An autoencoder is an approach to unsupervised learning where the network is trained to identically match its output to its input. The autoencoder network consists of two parts: an encoder $E : X \to Y$ that maps vectors from the input space X to feature vectors in some intermediate feature space Y, and a decoder $D : Y \to X$ that maps vectors from Y back into X. The autoencoder $F : X \to X$ is then defined as the composition:

$$F(x) = D(E(x)), \forall x \in X. \tag{1}$$

A. D. Missert is with the CT Clinical Innovation Center, Mayo Clinic, Rochester, MN, 55901. E-mail: missert.andrew@mayo.edu

S. Leng is with the CT Clinical Innovation Center and Department of Radiology, Mayo Clinic, Rochester, MN, 55901. E-mail: leng.shuai@mayo.edu

L. Yu is with the CT Clinical Innovation Center and Department of Radiology, Mayo Clinic, Rochester, MN, 55901. E-mail: yu.lifeng@mayo.edu

C. H. McCollough is with the CT Clinical Innovation Center and Department of Radiology, Mayo Clinic, Rochester, MN, 55901. E-mail: mccollough.cynthia@mayo.edu

Since the input and output of the autoencoder both belong to the same domain, a standard comparative loss metric, such as the mean squared error (MSE), can be used to train the autoencoder to reconstruct its input.

In our approach, we use an artificial neural network as an autoencoder that maps input images to similarly-shaped output images. Since the availability of both full-dose and simulated low-dose images allows for an ideal comparison of the same anatomical structures in each patient case, we use supervised learning when training the network. The use of supervised learning is a key difference between our work and the standard autoencoder approach. We treat the noise reduction problem as one of finding an autoencoder that can map LDCT images to full-dose CT images. Let x_{LD} be an 2D image from a LDCT scan, and let x_{FD} be the corresponding full-dose image. We seek a function $F(x_{LD})$ whose output minimizes the MSE loss L when evaluated using the full-dose image:

$$L(x_{FD}, F(x_{LD})) = \|x_{FD} - F(x_{LD})\|^2. \tag{2}$$

We employ a CNN as a flexible non-linear model for the mapping function F. CNNs have been successfully used for a wide variety of image processing tasks such as classification, segmentation, denoising, and super-resolution. Based loosely on the image processing procedure of the human brain, these networks have proven to be very powerful and general models for manipulating image data. Due to this great flexibility, we choose a CNN as a model for the mapping function that minimizes Eq. 2.

A CNN is essentially a parameterized nonlinear function of its inputs. Let $N(x_{LD}|\Theta)$ be the CNN output given a set of network parameters Θ and an input LDCT image x_{LD}. The problem then becomes finding the network parameters Θ^* that minimize the MSE loss function:

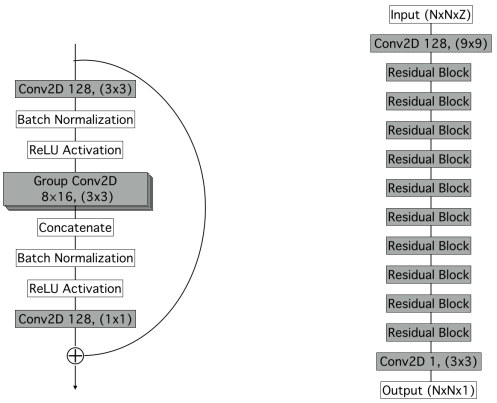
$$\Theta^* = \underset{\Theta}{\operatorname{argmin}}(\|x_{FD} - N(x_{LD}|\Theta)\|^2). \tag{3}$$

This minimization step can be performed using standard techniques such as stochastic gradient descent.

B. Network Architecture

The network architecture of the CNN determines the precise form of $N(x_{LD}|\Theta)$. Determining the best network architecture for a given problem is an area of active research in the field of computer vision, and few general principals exist to guide this choice.

The guiding principle in this work is that the denoising task is essentially one of feature extraction, and therefore CNN architectures that have proven to be powerful feature extractors for other tasks, such as image classification, should also work well here. For this reason, we model our network architecture after that of residual networks such as ResNet, which has achieved excellent performance for image classification tasks [8]. The key feature of residual networks is the "shortcut" connection which allows features to skip over convolution operations, thereby potentially allowing information from the initial layers to be passed directly to the deeper layers.



(a) (b)

Fig. 1. The structure of a single residual block is shown in (a). The numbers shown for the convolutional layers denote the number of filters and the filter shape, respectively. The core CNN architecture consists of 10 repeated residual blocks as shown in (b).

Following the structure of ResNet, our network architecture consists of repeated identical residual blocks as shown in Fig. 1(a). Each residual block performs three sets of 2D convolution operations, with a non-linearity imposed between them by application of a rectified linear unit (ReLU) activation function. Batch normalization is performed after each convolution layer to normalize the layer output. We also choose apply an aggregated group convolution in each residual block, which has been shown to improve the performance of residual networks [9]. In a grouped convolution, the input features are split into separate channels and convolution operations are applied to each channel, with the outputs then concatenated back together.

The overall structure of the network is shown in Fig. 1(b). An input image is fed into a neural network consisting of 10 residual blocks, which produces a set of 128 feature maps. These feature maps are then passed to a decoding layer, which simply projects the 128 feature maps back into the image domain by applying a single convolutional layer kernel size. The output of the CNN is then subtracted from the input to produce the final denoised image, which can then be compared to the full-dose image by evaluating the MSE.

Structuring the network to perform noise subtraction on the input has a couple of distinct advantages compared to directly reconstructing the de-noised image. By treating the CNN output as a perturbation to the original image, a CNN with randomly initialized parameters starts off with a much smaller loss, which saves time in the initial training phase. Furthermore, although the patient anatomy may vary substantially between scans, we expect the noise texture to be more consistent between scans for a given scanner and reconstruction method. Requiring the network to fit for the noise texture instead of the patient anatomy is therefore a potentially simpler and more general task.

C. Training Data

The data used for training this network is identical to the patient data made available for the "*2016 NIH-AAPM-Mayo Clinic Low Dose CT Grand Challenge*" [10]. This dataset consists of 10 full-dose clinical patient scans, which were obtained using similar scanner models (Somatom Definition AS+ and Somatom Definition Flash, operated in single-source mode, Siemens Healthcare, Forchheim, Germany). For the results shown here, the images were reconstructed by filtered backprojection using a medium-smooth kernel (B30). For each scan, a simulated quarter-dose (QD_{sim}) scan was produced using a validated projection-based noise insertion algorithm [11]. The use of simulated data taken from the full-dose scans allows an exact pixel-by-pixel comparison between the images. This is ideal for training the network using supervised learning with the QD_{sim} pixel values used as the network inputs, and the full-dose pixel values defined as the ground truth.

D. Training Methods

The CNN is trained using randomly-centered image crops of a various sizes which are generated from the full-dose and corresponding QD_{sim} data. Since the CNN effectively acts as a non-linear convolutional filter, it can accept an input image of any size, however, the use of zero-padding at the image edges during the convolution operations can produce small differences in the network output depending on the size of the input. We therefore apply three training steps: An initial training using 32×32 pixel images, followed by additional fine-tuning using 64×64 and 256×256 pixel images. The full training procedure exposes the network to 500000 distinct images. Testing is performed by evaluating the performance on full-sized 512×512 scan slices taken from a separate set of QD_{sim} data that is never used for training.

The network is implemented in TensorFlow [12] using Keras [13], and is trained using an NVIDIA GTX 1080 GPU. The optimization method used is the Adam optimizer [14] with default parameters.

III. Results

Results of the CNN noise subtraction algorithm can be qualitatively seen in the examples shown in Fig. 2. All of the results shown in this section use the exact same CNN, which is trained as described in the previous section. The input QD_{sim} images shown in these examples are taken from the validation data, which was not used during training and is therefore representative the network response to "new" input.

The CNN effectively identifies and subtracts noise textures from the QD_{sim} image, producing a much smoother output with reduced noise characteristics compared to both the QD_{sim} and FD images. The contrast-to-noise (CNR) ratio calculated in the liver region increases by 138% compared to the corresponding QD_{sim} image and by 55% compared to the FD image. Low-contrast objects, such as a small liver lesion found in Fig. 2 and shown profiled in Fig. 3, appear to be more readily detected in the denoised image compared to the low-dose image.

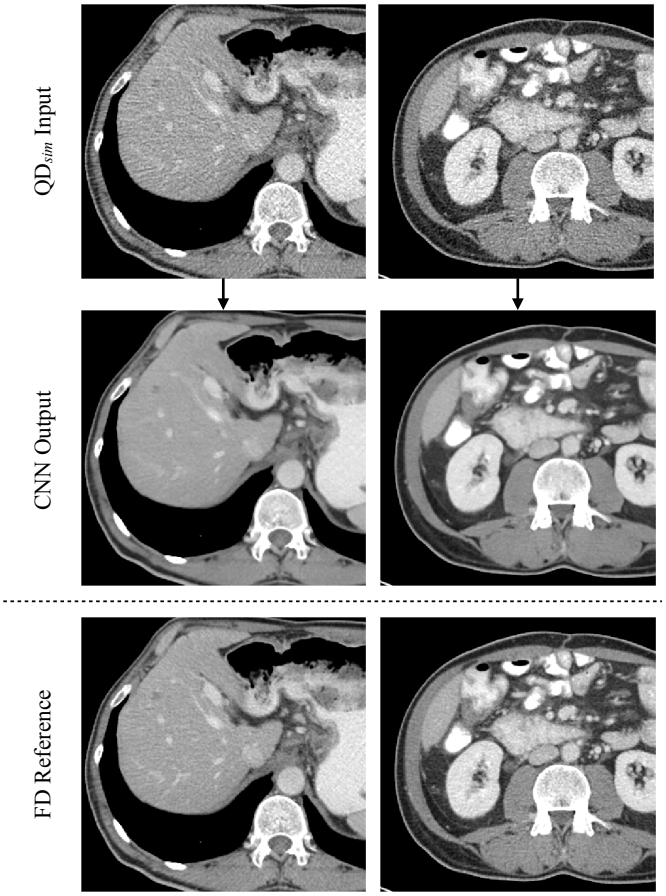


Fig. 2. Effect of CNN-based noise subtraction on abdominal CT images cropped from testing data that was not used for network training. The top row shows the simulated quarter-dose input images, the center row shows the output images after being passed through the CNN, and the bottom row shows the corresponding full-dose image for reference.

The noise reduction algorithm also preserves sharp edges and small-scale details found in the image. The latter can be seen in the noise-subtracted lung image shown in Fig. 4. The fine vascular structure of the lung is still clearly visible and undistorted, despite the noticeable degree of smoothing in the noise-subtracted image.

IV. Discussion

The images obtained using noise subtraction with a deep CNN exhibit many positive characteristics. Compared to the QD_{sim} images, the noise is sharply reduced. The texture of the noise-subtracted image is also perceptually smooth and absent of obvious artificial characteristic texture that distracts from the anatomical background.

It is often the case that image de-noising algorithms achieve smooth textures at the cost of reducing sharpness and/or distorting fine details. Remarkably, the CNN output is able to preserve these details while maintaining a high degree of smoothness, as can be seen in the noise-subtracted images of the lung.

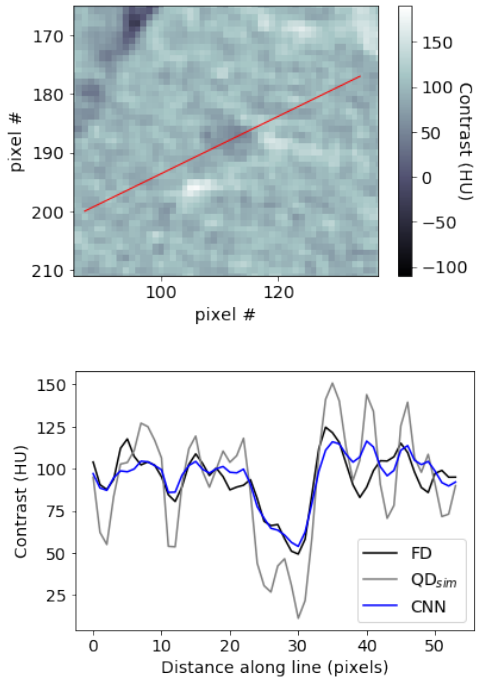


Fig. 3. Profile of the contrast across the small liver lesion that can be seen in Fig. 2. The FD image and the path along which the profile was taken are shown in the upper panel. The contrast profile along this path for both the FD (black), QD$_{sim}$ (grey), and CNN output (blue) images is shown below. The CNN output reduces the contrast fluctuations surrounding the lesion, while accurately matching the FD image across the lesion center.

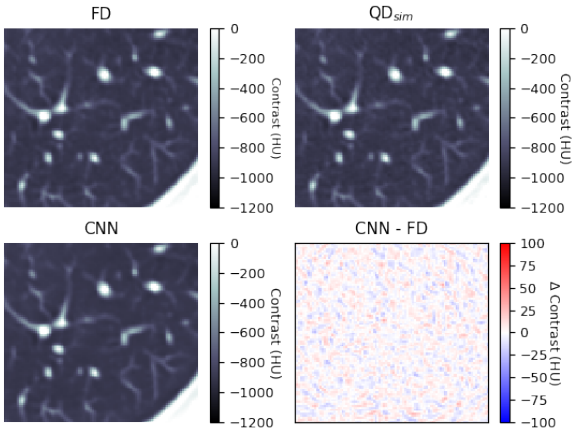


Fig. 4. Images of lung tissue cropped from Low-Dose CT Grand Challenge testing data. The upper right panel shows the simulated quarter-dose scan that was used as the network input, and the lower right panel shows the CNN output after noise subtraction. The corresponding full-dose image is shown for reference in the upper left panel. The difference between the noise-subtracted image and the full-dose image, which appears to be devoid of any anatomical features, is shown in the lower right panel.

Overall, the application of a deep residual CNN trained on the Low-dose CT Grand Challenge data has produced images with reduced noise that are perceived as being both smooth and sharp. The precise extent to which this noise reduction algorithm mitigates the problems with noise in LDCT images as outlined in Section I is currently being investigated.

While the current work is focused on subtracting noise textures for a particular type of scanner and a particular reconstruction method, the results from the testing data show that method generalizes well across different patient anatomies and different tissue types. Once the noise-subtraction network has been trained, the algorithm is able to process images very quickly, taking approximately 0.5 seconds to denoise a full scan slice using current GPU hardware (NVIDIA GTX 1080). The speed, quality, and generality of these results make noise subtraction via a deep neural network a promising candidate for future clinical use.

References

[1] A. Manduca *et al.*, "Projection space denoising with bilateral filtering and CT noise modeling for dose reduction in CT," *Med Phys*, vol. 36, no. 11, pp. 4911–4919, Nov 2009.

[2] J. Wang, T. Li, H. Lu, and Z. Liang, "Penalized weighted least-squares approach to sinogram noise reduction and image reconstruction for low-dose X-ray computed tomography," *IEEE Trans Med Imaging*, vol. 25, no. 10, pp. 1272–1283, Oct 2006.

[3] P. J. La Riviere, "Penalized-likelihood sinogram smoothing for low-dose CT," *Med Phys*, vol. 32, no. 6, pp. 1676–1683, Jun 2005.

[4] Z. Li *et al.*, "Adaptive nonlocal means filtering based on local noise level for CT denoising," *Med Phys*, vol. 41, no. 1, p. 011908, Jan 2014.

[5] H. Chen *et al.*, "Low-dose ct via convolutional neural network," *Biomed. Opt. Express*, vol. 8, no. 2, pp. 679–694, Feb 2017. [Online]. Available: http://www.osapublishing.org/boe/abstract.cfm?URI=boe-8-2-679

[6] Y. Zhang *et al.*, "Adaptive non-local means on local principle neighborhood for noise/artifacts reduction in low-dose ct images," *Medical Physics*, vol. 44, no. 9, pp. e230–e241, 2017. [Online]. Available: http://dx.doi.org/10.1002/mp.12388

[7] H. Chen *et al.*, "Low-dose ct with a residual encoder-decoder convolutional neural network (red-cnn)," vol. PP, 02 2017.

[8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *CoRR*, vol. abs/1512.03385, 2015. [Online]. Available: http://arxiv.org/abs/1512.03385

[9] S. Xie, R. B. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," *CoRR*, vol. abs/1611.05431, 2016. [Online]. Available: http://arxiv.org/abs/1611.05431

[10] C. H. McCollough *et al.*, "Low-dose CT for the detection and classification of metastatic liver lesions: Results of the 2016 Low Dose CT Grand Challenge," *Med Phys*, vol. 44, no. 10, pp. e339–e352, Oct 2017.

[11] L. Yu, M. Shiung, D. Jondal, and C. H. McCollough, "Development and validation of a practical lower-dose-simulation tool for optimizing computed tomography scan protocols," *J Comput Assist Tomogr*, vol. 36, no. 4, pp. 477–487, 2012.

[12] M. Abadi *et al.*, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015, software available from tensorflow.org. [Online]. Available: https://www.tensorflow.org/

[13] F. Chollet *et al.*, "Keras," 2015. [Online]. Available: https://github.com/fchollet/keras

[14] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *CoRR*, vol. abs/1412.6980, 2014. [Online]. Available: http://arxiv.org/abs/1412.6980

Determination of Algorithm Parameters by Using Input/Output Image Pairs

Larry Zeng

***Abstract* — Many image reconstruction algorithms contain user-adjusted parameters such as the number of iterations. The parameters are usually tedious to determine. The goal of the paper is to propose an automatic way to determine those parameters. First, an objective function is setup, which is the distance between the reconstructed image and the desired image. Second, an optimization algorithm is used to find the optimal parameters with the training pairs in the objective function. Third, the estimated parameters learned from the second step are ready to reconstruct other images. The methodology proposed in this paper is inspired by the rapid development of machine learning. Machine learning ideas can be extended into many areas, where the neural network structure is not necessary. Our methodology implies that any algorithm can learn to find its optimal parameters by using the training sets.**

I. Introduction

IN medical image reconstruction algorithms, especially in Bayesian algorithms, there are parameters for users to determine. A typical example of such parameters is the number of iterations in an iterative algorithm. Usually a trial-and-error approach is adopted in parameter selection. Inspired by the recent fast development in machine learning and its related fields, in this paper, we propose an automatic method to determine the parameters if pairs of the training images are available.

Neural network based machine learning is one form of artificial intelligence [1-7]. Nowadays, deep learning is a very active research field that uses multiple layers of neural network with non-linear activation functions [8, 9]. Some of the ideas from the field of machine learning can be borrowed and applied to image reconstruction algorithm design without using the neural network architecture at all.

The idea of using machine learning to design the parameters in an algorithm is not new. In 1960's, Professor B. Widrow of Stanford University was first to use machine learning to design linear filters; the learning was based on the least mean squares (LMS) adaptive algorithm [10-13]. The LMS algorithm led to the ADALINE and MADALINE artificial neural networks and to the backpropagation technique.

One advantage in machine learning is the availability of pairs of data: the inputs and the corresponding desired outputs. This paper borrows the idea of using the input-output pair and develops architecture to determine the algorithm parameters. The algorithm in concern can be treated as a black box. We only focus on the parameters to be determined.

In the case of image reconstruction, the training pairs contain the noisy projections and their corresponding desired reconstructions. In image de-noising applications, the training pairs contain the noisy images and their corresponding desired less-noisy images.

The black box is the algorithm between the input and the output. The algorithm can be iterative or analytic. Almost all algorithms contain some parameters to be adjusted according to the task. In the conventional algorithm development, a set of parameters is used to generate an image. This resultant image may be too blurry, too noisy, or containing too many artifacts. The human developer then replaces the old set of parameters with a new set of parameters. A new image is produced and evaluated. Another set of parameters is tested. This procedure is repeated many times until the human algorithm developer is satisfied with the results.

If training pairs are available, this recursive procedure can be performed by a computer, via machine learning. This method will be demonstrated by an example in the next section.

II. Methods

The purpose of this paper is NOT to provide a recipe for people to follow. Instead, this paper uses one example to show that it is possible to employ the machine learning strategy to determine some adjustable parameters in an algorithm for anybody's own project (see Fig. 1 for the setup)

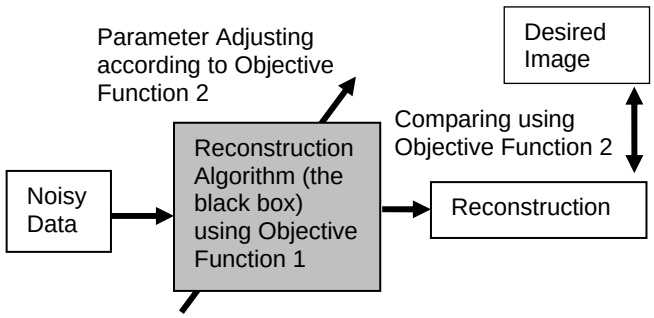


Figure 1. A general machine learning scheme for parameter determination.

Let us consider an x-ray CT de-noising problem as an illustrative example. In the training stage, we are given pairs of CT reconstructed images with low-dose and standard dose, respectively. The lower radiation dose is approximately 25% of the standard dose.

A. The image pair

In the setup, the input of the black box is the low-dose image; the output of the black box is the standard-dose image. The images were provided by Mayo Clinic in 2016 for the Low-Dose CT Grand Challenge competition. The low-dose images were generated in Mayo Clinic by adding noise to the

Gengsheng L. Zeng is with Department of Engineering, Weber State University, Ogden, UT 84408 USA (telephone: 801-626-6864, e-mail: larryzeng@weber.edu) and also with Department of Radiology and Imaging Sciences, University of Utah, Salt Lake City, UT 84108 USA (telephone: 801-581-3918, e-mail: larry.zeng@hsc.utah.edu).

regular-dose CT measurements. The "black box" is a de-noising algorithm described as follows.

B. The black box — de-noising algorithm

Many algorithms are available for the task of image de-noising, such as Gaussian smoothing, edge-preserving anisotropic filters, total variation minimization, Wiener filter, wavelet thresholding, bilateral filter, guided filter, non local means algorithms, and so on [14-16].

In this paper, we use a special way to de-noise by using an iterative image reconstruction method, similar to the idea used in [17]. One advantage of using this method is that we are able to model the projection noise which is ray-sum dependent.

The Mayo Clinic provided us pairs of low-dose and regular-dose x-ray CT patient torso images. The image slice was 512 × 512 and the slice thickness was 3 mm. The image volumes were first forward projected using the parallel-beam geometry, to generate projection sinograms. The parallel-beam sinograms were then used as the inputs for the iterative Landweber algorithm for reconstruction [18]. One form of the iterative Landweber algorithm can be expressed as [19]

$$x_i^{(k+1)} = x_i^{(k)} - \frac{\sum_j a_{ji} w_j (\sum_n a_{jn} x_n^{(k)} - p_j)}{\sum_j a_{ji} w_j \sum_n a_{jn}}, \qquad (1)$$

where $x_i^{(k)}$ is the i^{th} image pixel at the k^{th} iteration, p_j is the j^{th} line-integral (ray-sum) measurement value, a_{ji} is the contribution of the i^{th} image pixel to the j^{th} measurement, and w_j is weighting factor for the jth ray-sum p_j. For transmission tomography, the weighting factor is given as the exponential function of the negative of the ray-sum, that is, $w_j = \exp(-p_j)$. The purpose of the denominator $\sum_j a_{ji} \sum_n a_{jn}$ is to normalize the step size so that the step size is independent from the system matrix A. The summation over the index n is the projector and the summation over the index j is the backprojector.

It is well understood that the reprojected sinogram does not carry exactly the same information as the original sinogram. The spatial resolution may get degraded. The original projection data generally do not satisfy the data consistency conditions, while the reprojected data always satisfy the data consistency conditions. The noise in the original data can be assumed to be independent; however, the noise in the reprojected data is not independent.

There are two objective functions involved in our methodology. The objective function 1 is the discrepancy between the reporjected sinogram and the measurements

$$\text{Objective Function \#1} = \sum_j w_j (\sum_n a_{jn} x_n - p_j)^2. \qquad (2)$$

The iterative Lawdweber algorithm (1) is a gradient descend algorithm that aims to reduce the objective function 1 at each iteration. The second objective function is the discrepancy between the reconstructed image and the desired image

$$\text{Objective Function \#2} = \sum_{n \in ROI} (x_n - x_n^{Full-dose})^2. \qquad (3)$$

The desired image is the regular-dose image in our case. The desired iteration number k is determined when the objective function 2 reaches the minimum.

One has freedom to set up the second objective function according to the application in mind. The purpose of our CT images is lesion detection. A large lesion was located in the liver of a patient. We set up a 40×50×10 rectangular region-of-interest (ROI) that contained the lesion. The second objective function was the sum of the square of the distances between the reconstructed ROI voxels and the desired ROI voxels. Both ROI images were normalized to [0, 1] before the discrepancy was evaluated. Image values outside the ROI were not included in calculation the 2nd objective function.

C. Determination of the parameters

In most machine learning projects, large data sets are used. We only used one pair of images in selection of the optimal iteration number k, by minimizing the 2nd objective function. Since the iteration number k is an integer, the easiest way to find the optimal k is to run the iterative Landweber algorithm for, say, 800 iterations, evaluate the 2nd objective function at each iteration, and the find the minimum among the 800 function values.

If one wants to determine many parameters, it is more efficient to use a gradient decent method to minimize the 2nd objective function.

D. Applying to other images

After the parameters are determined, the black box is considered trained. The de-noising or image reconstruction algorithm is ready to process new inputs in the situations where the desired outputs are not available. Of course, the adjustable parameters are not universal; they can only be used as a guideline for the cases with similar patient size and similar noise levels.

III. RESULTS

Fig. 2 shows a typical a slice of input/output pair. Fig. 3 shows a converged image (same slice) with the given training input. The machine-selected optimal iteration number in our particular example was k = 342. The images are displayed using the window [-200, 400] in Hounsfield units.

The algorithm with k = 342 was then applied to a different patient data set which was not in the training set. Two images using other iterations numbers are shown in Fig. 4. Small iteration numbers make the image blurry and large iteration numbers produce noisier images. The resultant image using k = 342 and the low-dose input image for a different patient are shown in Fig. 5.

Fig. 6 shows how the objective function #1's value changes as the number of iteration increases from k = 1 to k =800. The natural logarithm of the function values are ploted in Fig. 6. This plot implies that the iterative Landweber algorithm effectively minimizes the discrepancy between the measurements and the reprojected data.

Fig. 7 shows how the objective function #2's value changes as the number of iteration increases from k = 1 to k =800. The natural logarithm of the function values are ploted in Fig. 7. This plot implies that the iterative Landweber algorithm does not know where the optimal solution according to the objective function #2.

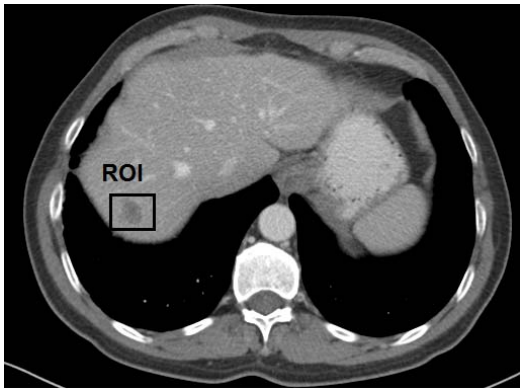
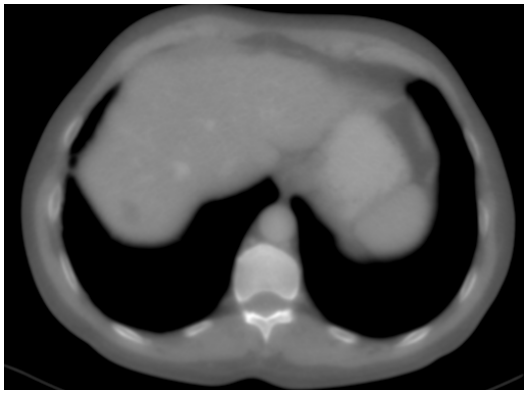


Figure 2. Training pair: (Upper) Low dose; (Lower) Standard dose

Figure 3. Reconstructed image using machine-selected iteration number (k = 342).



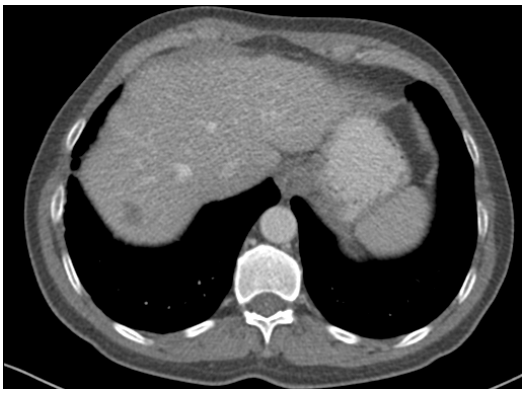


Figure 4. Reconstructed image using other iteration numbers. (Upper) k = 50; (Lower) k = 500.

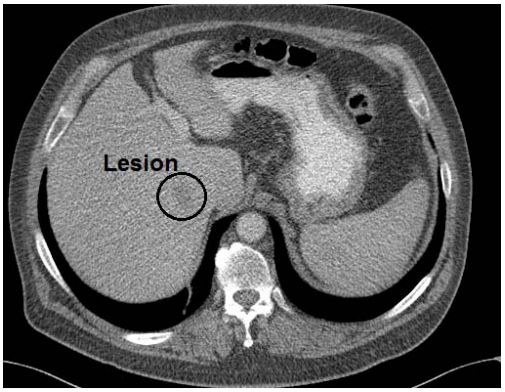


Figure 5. Application of the learned algorithm to a different patient: (Upper) Low dose image; (Lower) Processed with k = 342.

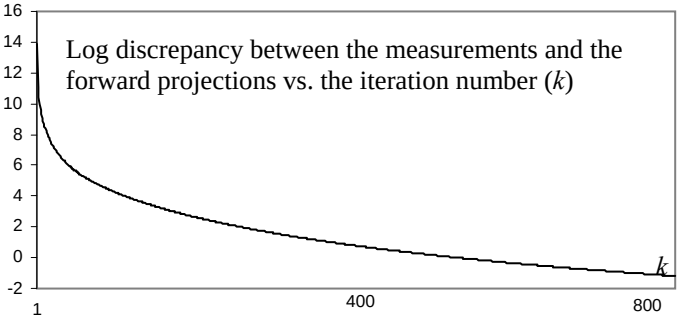


Figure 6. As the iteration number k increases the objective function #1 is minimized by the iterative Landweber algorithm (1).

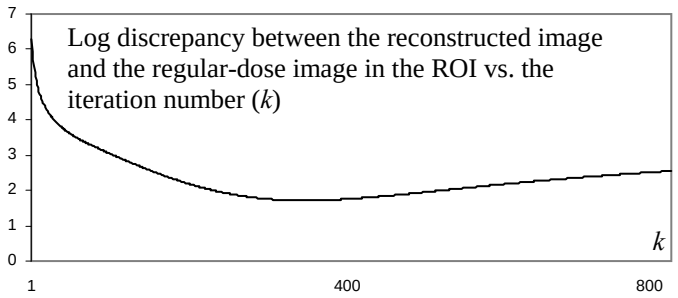


Figure 7. The value in the objective function #2 changes as the iteration number k increases. The minimum is reached at k = 342.

IV. Conclusion

In the 1960's, Widrow proposed a machine-learning method to design a finite impulse response (FIR) filter. An FIR filter is a linear filter using convolution to produce the output. The convolution kernel is trained by pairs of the input/output signals. This paper expends Widrow's machine learning method to more general applications where the neural network architecture may not be used. The key concept in this paper is that a given input/output pair or pairs can help to determine the adjustable parameters automatically, saving lots of trial-and-error time. The objective function does not have to use the L_2 measure. The users should be creative in designing the objective function to fit their application tasks.

The parameters of a "black box" are determined by the input/output pairs. The "black box" does not need to be linear or any specific structure. There may not be any closed-form formula to calculate the gradient with respect to the parameters to be determined. If the gradient cannot be evaluated, the users are encouraged to use some optimization algorithms that do not require the calculation of the gradient.

In this sense, machine learning can be applied to any algorithm that contains some adjustable parameters. Once we have pairs of training sets, any algorithm can learn. We gave a simple example in this paper to illustrate this extension. In our example, the image de-nosing task is achieved by an iterative image reconstruction algorithm, which can model the projection noise. The parameter to be determined is the stopping rule. The machine-learned stopping rule can to used as a guideline for similar patient studies.

Finally, an import observation is that there is no requirement of using a large number of training pairs. This observation is quite different from the conventional neural network that is used for labeling or classifying the inputs. For labeling or classifying tasks, a large number of data is needed for the training stage. In our simple illustrative example, only one image volume pair was used for training. The image volume contained 10 slices in our example. In parameter determination tasks, the size of the training data set is not a critical issue. The fundamental requirement is the availability of the input/output pair.

V. ACKNOWLEDGMENTS

The author thanks Mayo Clinic for providing the patient data.

References

[1] K. Gurney, *An Introduction to Neural Networks*, London: Routledge. ISBN 1-85728-673-1 (hardback) or ISBN 1-85728-503-4 (paperback), 1997.

[2] J. Lawrence, *Introduction to Neural Networks*, California Scientific Software Press, ISBN 1-883157-00-5, 1994.

[3] C. M. Bishop, *Neural Networks for Pattern Recognition*, Oxford: Oxford University Press. ISBN 0-19-853849-9 (hardback) or ISBN 0-19-853864-2 (paperback), 1995.

[4] B. Scholkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge, MA: MIT Press, p. 626. 2001.

[5] M. Egmont-Petersen, D. de Ridder, and H. Handels, "Image processing with neural networks – a review," *Pattern Recognition*, vol. 35 (10), pp. 2279–2301. doi:10.1016/S0031-3203(01)00178-9, 2002.

[6] T. Masters, *Signal and Image Processing with Neural Networks*, John Wiley & Sons, Inc. ISBN 0-471-04963-8, 1994.

[7] B. D. Ripley, *Pattern Recognition and Neural Networks*, Cambridge, 1996.

[8] L. Deng, and D. Yu, "Deep Learning: Methods and Applications," *Foundations and Trends in Signal Processing*, vol. 7 (3–4), pp. 1–192, 2014, doi:10.1561/2000000039.

[9] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press. Online, 2016.

[10] K. Steinbuch and B. Widrow, "A critical comparison of two kinds of adaptive classification networks," *IEEE Transactions on Electronic Computers*, pp. 737–740, 1965

[11] B. Widrow and S. D. Stearns. *Adaptive Signal Processing*. New Jersey: Prentice-Hall, Inc., 1985.

[12] B. Widrow, "An adaptive 'ADALINE' neuron using chemical 'memistors'," *Technical Report Nl. 1553-2 for Office of Naval Research*, 1960, http://www-isl.stanford.edu/ ~widrow/papers /t1960anadaptive.pdf

[13] B. Widrow and M. A. Lehr, "30 years of adaptive neural networks: perceptron, madaline, and backpropagation," *Proceedings of the IEEE*, vol. 78 (9), pp. 1415–1442, 1990. doi:10.1109/5.58323

[14] A. Buades, B. Coll, and J.-M. Morel, "A review of image denoising algorithms, with a new one," *SIAM Journal on Multiscale Modeling and Simulation: A SIAM Interdisciplinary Journal*, vol. 4 (2), pp.490-530, 2005.

[15] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," In: *Proceedings of the 1998 IEEE International Conference on Computer Vision*. Bombay, India: IEEE; 1998.

[16] K. He, J. Sun, and X. Tang, "Guided Image Filtering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, Issue 6, pp. 1397-1409, 2013.

[17] G. L. Zeng, "A fast method to emulate an iterative POCS image reconstruction algorithm," *Med. Phys*., vol. 44, pp. e353-e359, 2017.

[18] L. Landweber, "An iteration formula for Fredholm integral equations of the first kind," *Amer. J. Math*., vol. 73, no. 3, pp. 615–624, 1951.

[19] G. L. Zeng and W. Wang, "Noise weighting with an exponent for transmission CT," *Biomedical Physics & Engineering Express*, vol. 2, no. 045004, 2016.

Deep Learning Reconstruction for 9-View Dual Energy CT Baggage Scanner

Yoseob Han[1], Jingu Kang[2], and Jong Chul Ye[1]

***Abstract*—For homeland and transportation security applications, 2D X-ray explosive detection system (EDS) have been widely used, but they have limitations in recognizing 3D shape of the hidden objects. Among various types of 3D computed tomography (CT) systems to address this issue, this paper is interested in a stationary CT using fixed X-ray sources and detectors. However, due to the limited number of projection views, analytic reconstruction algorithms produce severe streaking artifacts. Inspired by recent success of deep learning approach for sparse view CT reconstruction, here we propose a novel image and sinogram domain deep learning architecture for 3D reconstruction from very sparse view measurement. The algorithm has been tested with the real data from a prototype 9-view dual energy stationary CT EDS carry-on baggage scanner developed by GEMSS Medical Systems, Korea, which confirms the superior reconstruction performance over the existing approaches.**

***Index Terms*—Explosive detection system (EDS), sparse-view X-ray CT, convolutional neural network (CNN)**

I. Introduction

In homeland and aviation security applications, there has been increasing demand for X-ray CT EDS system for carry-on baggage screening. A CT-EDS can produce an accurate 3D object structure for segmentation and threat detection, which is often not possible when a 2D-EDS system captures projection views in only one or two angular directions. There are currently two types of CT EDS systems: gantry-based CT and stationary CT. While gantry-based CT EDS is largely the same as medical CT, baggage screening should be carried out continuously, so it is often difficult to continuously screen carry-on bags because of the possible mechanical overloading of the gantry system. On the other hand, a stationary CT EDS system uses fixed X-ray sources and detectors, making the system suitable for routine carry-on baggage inspection.

For example, Fig. 1 shows source and detector geometry of the prototype stationary CT-EDS system developed by GEMSS Medical Systems, Korea. As shown in Fig. 1(a), nine pairs of X-ray source and dual energy detector in the opposite direction are distributed at the same angular interval. For seamless screening without stopping convey belt, each pair of source and detectors are arranged along the z-direction as shown in Fig. 1(b) so that different projection view data can be collected while the carry-on baggages moves continuously on the conveyor belt. Then, 9-view fan beam projection data is obtained for each z-slice by rebinning the measurement data. This type of stationary CT system is suitable for EDS applications because it does not require a rotating gantry, but with only 9 projection views it is difficult to use a conventional filtered backprojection (FBP) algorithm due to severe streaking artifacts. Therefore, advanced reconstruction algorithms with fast reconstruction time are required.

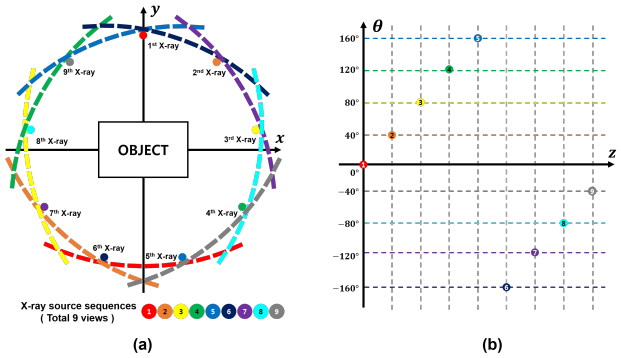


Fig. 1. X-ray source positions in our prototype 9 view dual energy CT EDS: (a) $x-y$ direction and (b) $\theta-z$ direction, respectively.

For sparse-view CT EDS, model-based iterative reconstruction (MBIR) with the total variation (TV) penalty have been extensively investigated [1], [2]. Inspired by the recent success of deep learning approach for sparse view and limited angle CT [3], [4], [5], [6] that outperform the classical MBIR approach, this paper aims at developing a deep learning approach for real-world sparse view CT EDS. However, neural network training using the retrospective angular subsampling as in the existing works [3], [4], [5], [6] is not possible for our prototype system, since there are no ground-truth data for the real world sparse view CT EDS. We therefore propose a novel deep learning approach composed of image domain and sinogram domain learning that compensate for the imperfect label data.

II. Theory

A. Problem Formulation

Recall that the forward model for sparse view CT EDS system can be represented by

$$g_\Theta = \mathcal{P}_\Theta \mathcal{R} f \tag{1}$$

where $\mathcal{R}$ denotes the 3D projection operator from an $x-y-z$ volume image to a $s-\theta-z$ domain sinogram data with s, θ and z denoting the detector, projection angle, and the direction of the conveyor belt travel, respectively. See Fig. 2 for the coordinate systems. In (1), $\mathcal{P}_\Theta$ denotes the view sampling

[1]Y. Han and J.C. Ye are with the Department of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon 34141, Republic of Korea (e-mail: {hanyoseob,jong.ye}@kaist.ac.kr).

[2]J. Kang is with GEMSS Medical Co. Seongnam, Republic of Korea (e-mail: jingu.kang@gemss-medical.com).

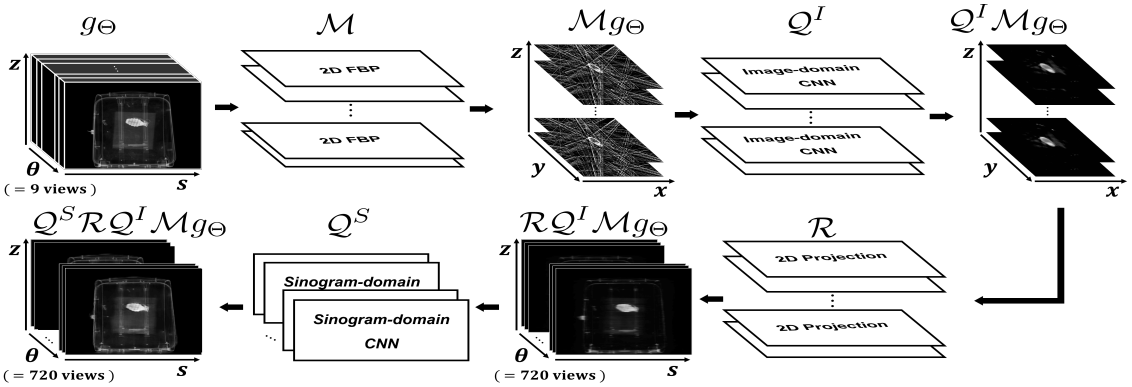


Fig. 2. Sinogram interpolation flow for the proposed method. The final reconstruction is obtained by applying the FBP for the interpolated sinogram data.

operator for the measured angle set Θ, and g_Θ refers to the measured sinogram data. For each projection view data, we use the notation g_θ and $\mathcal{P}_\theta$, where θ denotes the specific view.

The main technical issue of the sparse view CT reconstruction is the non-uniqueness of the solution for (1). More specifically, there exists a null space $\mathcal{N}_\Theta$ such that

$$\mathcal{P}_\Theta\mathcal{R}h = 0, \quad \forall h \in \mathcal{N}_\Theta,$$

which leads to infinite number of feasible solutions. To avoid the non-uniqueness of the solution, constrained form of the penalized MBIR can be formulated as :

$$\min_{f\in\mathbb{R}^3} \|\mathrm{L}f\|_1, \quad \text{subject to} \quad g_\Theta = \mathcal{P}_\Theta\mathcal{R} \ , \tag{2}$$

where L refers to a linear operator and $\|\cdot\|_1$ denotes the l_1 norm. For the case of the TV penalty, L corresponds to the derivative. Then, the uniqueness of (2) is guaranteed that if the $\mathcal{N}_\mathrm{L} \cap \mathcal{N}_\Theta = \{0\}$, where $\mathcal{N}_\mathrm{L}$ denotes the null space of the operator L.

Instead of designing a linear operator L such that the common null space of $\mathcal{N}_\Theta$ and $\mathcal{N}_\mathrm{L}$ to be zero, we can design a frame $\mathcal{W}$, its dual $\tilde{\mathcal{W}}$, and shrinkage operator S_λ such that $\tilde{\mathcal{W}}^\top\mathcal{W} = I$ and

$$\tilde{\mathcal{W}}^\top S_\lambda \mathcal{W}(f^* + g) = f^* \quad \forall g \in \mathcal{N}_\Theta$$

for the ground-truth image f^*. This frame-based regularization is also an active field of research for image denoising, inpainting, etc [7]. One of the most important contributions of the deep convolutional framelet theory [8] is that $\mathcal{W}$ and $\tilde{\mathcal{W}}^\top$ correspond to the encoder and decoder structure of a convolutional neural network (CNN), respectively, and the shrinkage operator S_λ emerges by controlling the number of filter channels and nonlinearities. More specifically, a convolutional neural network can be designed such that $\mathcal{Q} = \tilde{\mathcal{W}}^\top S_\lambda \mathcal{W}$ and

$$\mathcal{Q}(f^* + h) = f^*, \quad \forall h \in \mathcal{N}_\Theta \quad . \tag{3}$$

In other word, (3) directly removes the null space component. Eq. (3) is the constraint we use for training our neural network.

B. Derivation of Image and Projection Domain CNNs

More specifically, our sparse view reconstruction algorithm finds the unknown $f \in \mathbb{R}^3$ that satisfy both data fidelity and the so-called frame constraints [8]:

$$g_\Theta = \mathcal{P}_\Theta\mathcal{R}f, \quad \mathcal{Q}^I(f) = f^* \quad , \tag{4}$$

where $\mathcal{Q}^I$ is the image domain CNN that satisfies (3) and f^* denotes the ground-truth images that are available for training data. Now, by defining $\mathcal{M}$ as a right-inverse of $\mathcal{P}_\Theta\mathcal{R}$, i.e. $(\mathcal{P}_\Theta\mathcal{R})\mathcal{M}g_\Theta = g_\Theta, \forall g_\Theta$, we have

$$\mathcal{M}g_\Theta = f^* + h$$

for some $h \in \mathcal{N}_\Theta$, since the right inverse is not unique due to the existence of the null space. Thus, we can show that $\mathcal{M}g_\Theta$ is the feasible solution for (4), since we have

$$\mathcal{Q}^I\mathcal{M}g_\Theta = \mathcal{Q}^I\left(f^* + h\right) = f^* \quad , \tag{5}$$

for the training data, and

$$\mathcal{P}_\Theta\mathcal{R}\mathcal{M}g_\Theta = \mathcal{P}_\Theta\mathcal{R}(f^* + h) = g_\Theta \quad . \tag{6}$$

Therefore, the neural network training problem to satisfy (4) can be equivalently represented by

$$\min_{\mathcal{Q}^I} \sum_{i=1}^{N} \|f^{*(i)} - \mathcal{Q}^I\mathcal{M}g_\Theta^{(i)}\|^2 \tag{7}$$

where $\{(f^{*(i)}, g_\Theta^{(i)})\}_{i=1}^N$ denotes the training data set composed of ground-truth image an its sparse view projection. Since a representative right inverse for the sparse view projection is the inverse Radon transform after zero padding to the missing view, $\mathcal{M}g_\Theta^{(i)}$ in (7) can be implemented using the standard FBP algorithm. In fact, this is the main theoretical ground for the success of image domain CNN when the ground-truth data is available [3], [4], [5], [6]. Moreover, the fan-beam rebinning makes the problem separable for each z slices, so we use the 2D FBP for each slice as shown in Fig. 2.

However, the main technical difficulties in our 9-view CT EDS system is that we do not have ground-truth image $\{f^{*(i)}\}_{i=1}^N$. One could use physical phantoms and atomic number to form a set of ground-truth images, but those data set may be different from the real carry-on bags, so we need a

new method to account for the lack of ground-truth for neural network training. Thus, to overcome the lack of the ground-truth data, the approximate label images are generated using an MBIR with TV penalty. Then, using MBIR reconstruction as label data $\{f^{*(i)}\}_{i=1}^{N}$, an 2D image domain network $\mathcal{Q}^I$ is trained to learn the mapping between the artifact-corrupted 2D image and MBIR reconstruction in $x-y$ domain.

One downside of this approach is that the network training by (7) is no more optimal, since the label data is not the ground-truth image. Thus, the generated sinogram data from the denoised 3D volume may be biased. Thus, we impose additional frame constraint to the sinogram data in addition to (4):

$$g_\theta^* = \mathcal{Q}^S\left(g_\theta\right), \tag{8}$$

for the measured angle θ, where $\mathcal{Q}^S$ is the $s-z$ sinogram domain CNN and g_θ^* denotes the ground-truth sinogram data measured at θ. Then, Eq. (8) leads to the following network training:

$$\min_{\mathcal{Q}^S} \sum_{\theta\in\Theta}\sum_{i=1}^{N} \| g_\theta^{*(i)} - \mathcal{Q}^S\left(\mathcal{P}_\theta \mathcal{R} \mathcal{Q}^I \mathcal{M} g_\Theta^{(i)}\right) \|^2 \tag{9}$$

More specifically, as shown in Fig. 2, 3D sinogram data is generated in the $s-\theta-z$ domain by applying the forward projection operator along 720-projection views after stacking the image domain network output over multiple slices to form 3D reconstruction volume in the $x-y-z$ domain. Next, a 2D sinogram domain network $\mathcal{Q}^S$ is trained so that it can learn the mapping between the synthetic $s-z$ sinogram data and the real projection data in the $s-z$ domain. Since the real projection data is available only in 9 views, this sinogram network training is performed using synthetic and real projection data in the measured projection views. The optimization problems (7) and (9) can be solved sequentially or simultaneously, and in this paper we adopt the sequential optimization approach for simplicity.

After the neural networks $\mathcal{Q}^I$ and $\mathcal{Q}^S$ are trained, the inference can be done simply by obtaining $x-y-z$ volume images from the 9 view projection data by slice-by-slice FBP algorithm, which are then fed into $\mathcal{Q}^I$ to obtain the denoised 3D volume data. Then, by applying projection operator, we generate 720 projection view data in $s-\theta-z$ domain, which are fed into the $\mathcal{Q}^S$ to obtain denoised sinogram data for each θ angle. Then, the final reconstruction is obtained by applying FBP algorithms. One could use post-processing using additional TV-based denosing. This algorithmic flow is illustrated in Fig. 2.

III. Methods

A. Real CT EDS data Acquisition

We collected CT EDS data using the prototype stationary 9 view dual energy CT-EDS system developed by GEMSS Medical Systems, Korea as shown in Fig. 1. the distance from source to detector (DSD) and the distance from source to object (DSO) are 1202.6mm and 648.2mm, respectively. The number of detector is 384 with a pitch of 1.5mm. The region of interest (ROI) is 256×256 and the pixel size is 2mm^2.

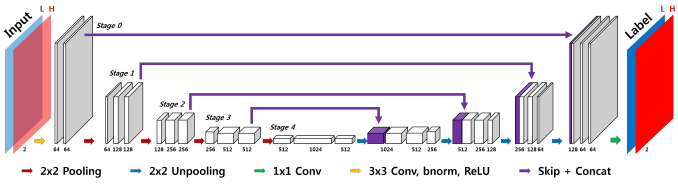


Fig. 3. CNN architecture for our image and singoram domain networks.

The detectors collect low and high energy X-ray at 80 KVp and 120 KVp, respectively.

We collect 47 sets of projection data from the prototype CT EDS baggage scanner. Among the 47 sets, 32 dataset are simple-objects and the other set are realistic carry-on bags. The 47 set of 28 simple- and 13 baggage-objects was used during the training phase, and the validation was performed by two simple- and one baggage-object. The other set was used for test.

B. Network Architecture and Training

Fig. 3 illustrates modified the U-Net structure [9] for the image domain and the sinogram domain networks. To account for the multi-energy image and sinogram data, the input for the network is two channel multi-energy image and sinogram data. The proposed network consists of convolution layer, batch normalization, rectified linear unit (ReLU) [10], and contracting path connection with concatenation [9]. A detail parameters are illustrated as shown in Fig. 3.

The proposed networks were trained by stochastic gradient descent (SGD). The regularization parameter was $\lambda = 10^{-4}$. The learning rate has been set from 10^{-3} to 10^{-5}, which has been reduced step by step in each epoch. The number of epoch was 200. The batch size was 12 and the patch size for image and projection data are $256\times256\times2$ and $768\times384\times2$, respectively. The network was implemented using MatConvNet toolbox (ver.24) [11] in the MATLAB 2015a environment (MathWorks, Natick). Central processing unit (CPU) and graphic processing unit (GPU) specification are i7-7700 CPU (3.60 GHz) and GTX 1080 Ti GPU, respectively.

IV. Experimental Results

To evaluate the performance of the proposed method, we perform image reconstruction from real 9-view CT EDS prototype system. Fig. 4 illustrates image reconstruction results of bag using various methods such as FBP, MBIR with TV penalty, image domain CNN [3], [5], and the proposed method. The FBP reconstruction results suffered from severe streaking artifacts, so it was difficult to see the threats in the tomographic reconstruction and 3D rendering. The MBIR and image domain CNN were slight better in their reconstruction quality, but the detailed 3D structures were not fully recovered and several objects were not detected as indicated by the red arrow in Fig. 4. Moreover, the 3D rendering results in Fig. 4 correctly identify the shape of grenade and knife as well

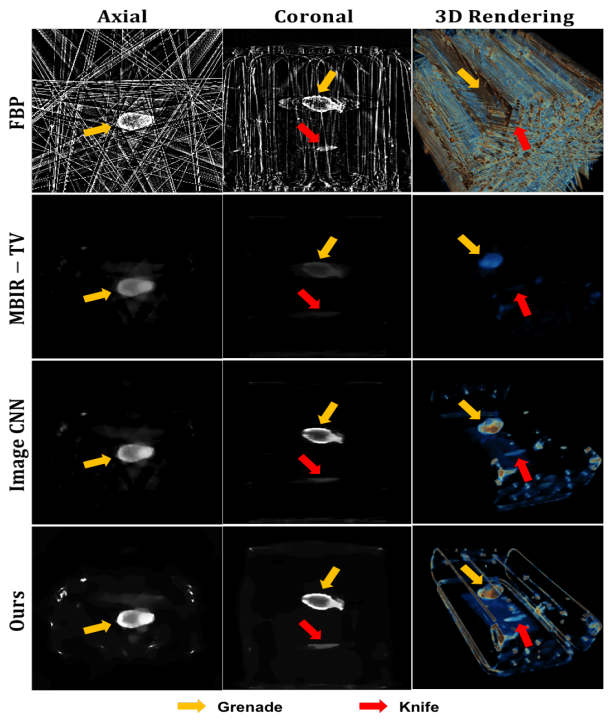


Fig. 4. Reconstruction results by various methods from 9-views CT-EDS.

as the frame of the bag, which was not possible using other methods.

Because we do not have the ground-truth in the image domain, we perform quantitative evaluation using normalized mean squares error (NMSE) in the sinogram domain. More specifically, after obtaining the final reconstruction, we perform the forward projection to generate the sinogram data in the measured projection view and calculated the normalized mean square errors. Table I showed that the proposed method provides the most accurate sinogram data compared to the other methods. Moreover, the $s-z$ projection data in Fig. 5 showed that the projection data from the proposed method is much closer to the ground-truth measurement data.

TABLE I
NMSE VALUE COMPARISON OF VARIOUS METHODS.

Energy level	FBP	MBIR-TV	Image CNN	Ours
80 KVp	1.6647e+1	5.8247e-1	3.3207e-1	0.6845e-1
120 KVp	1.0536e+1	6.0440e-1	3.2249e-1	0.5450e-1

V. Conclusion

In this paper, we proposed a novel deep learning reconstruction algorithm for a prototype 9-view dual energy CT EDS for carry-on baggage scanner. Even though the number of projection view was not sufficient for high equality 3D reconstruction, our method learns the relationships between the 2D tomographic slices in $x-y$ domain as well as the 2D projections in $s-z$ domain such that the artifact-corrupted

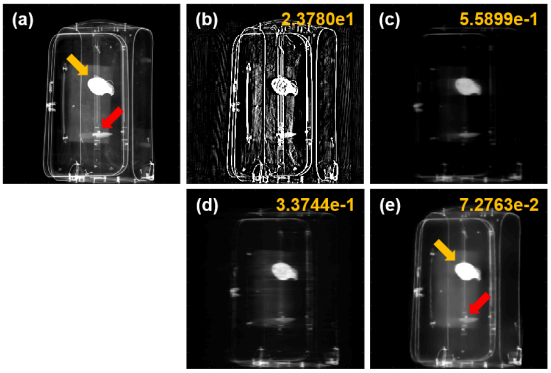


Fig. 5. A $s-z$ domain sinogram data from (a) measurement, (b) 9-views FBP (c) MBIR, (d) image CNN, and (e) the proposed method. The number written in the images is the NMSE value. Yellow and red arrows indicate grenade and knife, respectively.

image and sinogram data can be successively refined to obtain high quality images. Using real data from our prototype 9-view CT EDS system, we demonstrated that the proposed method outperforms the existing algorithms, delivering high quality three reconstruction for threat detection.

Acknowledgment

This work is supported by Korea Agency for Infrastructure Technology Advancement, Grant number 17ATRP-C071164-05-000000.

References

[1] S. Mandava, D. Coccarelli, J. A. Greenberg, M. E. Gehm, A. Ashok, and A. Bilgin, "Image reconstruction for view-limited x-ray ct in baggage scanning," in *Anomaly Detection and Imaging with X-Rays (ADIX) II*, vol. 10187. International Society for Optics and Photonics, 2017, p. 101870F.

[2] S. J. Kisner, E. Haneda, C. A. Bouman, S. Skatter, M. Kourinny, and S. Bedford, "Limited view angle iterative ct reconstruction," in *Computational Imaging X*, vol. 8296, 2012.

[3] Y. Han, J. Yoo, and J. C. Ye, "Deep residual learning for compressed sensing ct reconstruction via persistent homology analysis," *arXiv preprint arXiv:1611.06391*, 2016.

[4] Y. Han and J. C. Ye, "Framing U-net via deep convolutional framelets: Application to sparse-view CT," *arXiv preprint arXiv:1708.08333*, 2017.

[5] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4509–4522, 2017.

[6] J. Gu and J. C. Ye, "Multi-scale wavelet domain residual learning for limited-angle CT reconstruction," *arXiv preprint arXiv:1703.01382*, 2017.

[7] J.-F. Cai, R. H. Chan, and Z. Shen, "A framelet-based image inpainting algorithm," *Applied and Computational Harmonic Analysis*, vol. 24, no. 2, pp. 131–149, 2008.

[8] J. C. Ye, Y. S. Han, and E. Cha, "Deep convolutional framelets: A general deep learning framework for inverse problems," *arXiv preprint arXiv:1707.00372*, 2017.

[9] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2015, pp. 234–241.

[10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[11] A. Vedaldi and K. Lenc, "Matconvnet: Convolutional neural networks for matlab," in *Proceedings of the 23rd ACM international conference on Multimedia*. ACM, 2015, pp. 689–692.

Convolutional Neural Network Based CT Image Post-processing From FBP to ADMIRE

Yanbo Zhang, *Senior Member, IEEE*, Robert D. MacDougall, and Hengyong Yu*, *Senior Member, IEEE*

***Abstract*—Iterative reconstruction methods achieve better image quality than conventional filtered backprojection (FBP) for low-dose CT data. However, due to multiple forward and backward projections, the computational cost of iterative reconstruction methods is much higher. In this work, we apply the residual learning of a convolutional neural network (CNN) to reduce noise in the FBP images. We apply images reconstructed by the FBP and an iterative method to train the network with the goal of producing images that match the quality of the iterative images. The developed method is validated on clinical patient data acquired from a Siemens SOMATOM Force CT system. The experimental results demonstrate that the noise can be reduced remarkably and the obtained images are very close to those produce with commercial iterative reconstruction.**

***Index Terms*—Computed tomography, iterative reconstruction, residual learning, convolutional neural network**

I. Introduction

CT imaging at the lowest possible dose while maintaining the diagnostic accuracy is the central tenet in practice, referring as the As Low as Reasonably Achievable (ALARA) principle. While the carcinogenic risks of ionizing radiation at the diagnostic imaging levels remains controversial, it remains a priority to minimize the cumulative exposure over a lifetime especially for pediatric imaging and lung-cancer screening.

Advances in scanning technology (*e.g.*, low kV imaging [1-3], automatic tube current modulation (ATCM) [4], automatic kV-adaption [5], improved detector sensitivity [6]) have contributed to reduce dose for the state-of-the-art CT scanners. Meanwhile, iterative reconstruction (IR) algorithms, including image-based, model-based and hybrid algorithms, have had a measurable effect on reducing dose that is needed to acquire a diagnostic scan [7, 8]. The latest reconstruction technology is typically restricted to premium systems with major upgrades happening every 3-4 years. However, there exists a large install base of legacy (> 7 year old) scanners, as it is cost prohibitive for a hospital to replace a CT fleet every few years. In addition to the financial restrictions, it also has workflow restrictions. For example, some model-based iterative reconstruction algorithms can take 30-60 mins for reconstruction, far too long in an emergency setting.

Very recently, deep learning has been applied to low-dose CT reconstruction and CT image denoising [9-13]. In these works, high-quality images obtained from high-dose data or iterative reconstruction are adopted as the reference to train the designed network. Particularly, Wang summarized the applications of deep learning in radiology applications and gave an example of CNN based image post-processing [13]. In his work, FBP images and model based iterative reconstruction (MBIR) images were respectively used as the input and target for neural network training. However, this perspective article did not provide implementation information or results. Here, we are concerned to use a neural network to improve CT reconstruction for all scanners It is independent of scanner software and proprietary IR algorithms. To this end, we investigate a method of reconstructing images from a filtered back-projection (FBP) DICOM data set to simulate the appearance of an arbitrary IR algorithm. By training a convolutional neural network on a patient-matched FBP and IR image dataset, a vendor-neutral reconstruction algorithm can be created. This approach allows for exams performed on legacy CT scanner to provide the image quality of state-of-the-art IR algorithms.

II. Method

Inspired by the residual learning of deep CNN [12], we propose a residual network model for the low-dose CT image processing. Fig. 1 shows the architecture of the proposed network. The developed method consists of two phases: network training and post-processing. In the training phase, high-quality images are obtained with the advanced modelled iterative reconstruction (ADMIRE, Siemens Healthineers, Forchheim Germany) method [14]. The FBP reconstructed image is the input of the network, and residual between FBP and ADMIRE images is applied as the target. Hence, in the post-processing phase, the well-trained network is able to predict the difference between the input FBP image and the expected ADMIRE image. Then the high-quality image is obtained by subtracting the generated residual from the FBP image.

There are five convolutional layers in this network. Each

This work was supported in part by the NIH/NIBIB U01 grant EB017140 and R21 grant EB019074.

Y. B. Zhang is with the Department of Electrical and Computer Engineering, University of Massachusetts Lowell, Lowell, MA 01854, USA (e-mail: yanbo_zhang@ uml.edu).

R. D. MacDougall, is with the Boston Children's Hospital, Boston, MA 02115, USA (e-mail: Robert.D.MacDougall@childrens.harvard.edu).

H. Y. Yu is with the Department of Electrical and Computer Engineering, University of Massachusetts Lowell, Lowell, MA 01854, USA (e-mail: hengyong-yu@ieee.org).

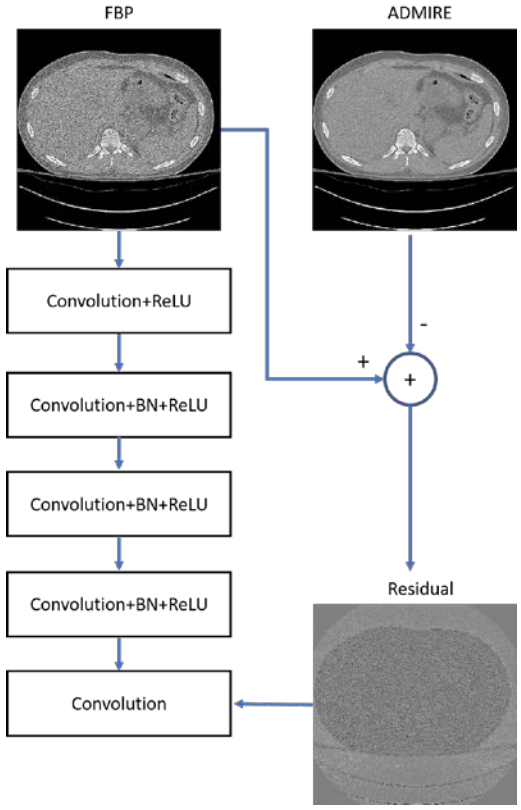


Fig. 1. Architecture of the proposed residual network.

layer consists of 32 convolutional kernels and each of them is with the size of 3×3. Padding is performed to ensure that the size of the output image is the same to the input. In the second to the fourth layers, batch normalization (BN) is used to solve the internal covariate shift after the convolution. Batch normalization brings several advantages to the method, such as fast training and better performance. Finally, ReLU is applied to the first four layers after the convolution or convolution+BN.

III. Experiments

A. Data

Two patient data were acquired on a third-generation dual-source CT system (SOMATOM Force, Siemens). One patient data was used for training, and the other patient data was applied to test the proposed algorithm. The focal spot size is 0.8 mm. The source-to-patient distance is 595 mm, and the source-to-detector distance is 1085.6 mm. The CT protocol used CARE kV and CARE Dose4D for kV selection and mA modulation, respectively. The tube voltage ranged from 100-110 kVp, the Quality Reference mAS was 150 and the exposure time was 250 ms. All data were reconstructed using both the FBP and ADMIRE. A total of 150 slices were reconstructed for each patient, and the slice thickness was 2mm. The reconstructed images were 512×512 with 0.7324 × 0.7324 mm^2 for each pixel. The convolutional kernel "Bl64d" was adopted for reconstruction.

B. Implementation

The code is implemented in Matlab (R2016a) environment running on a PC with AMD FX(tm)-6300 CPU and an Nvidia GeForce GTX 970 GPU. In addition, MatConvNet toolbox [15] is used for the network training.

To train the network, the FBP and ADMIRE reconstructed images are divided into 64×64 image patches. Then 10,000 FBP-ADMIRE-pair patches are randomly selected as the training data. We train 400 epochs for the proposed network.

C. Evaluation

We apply the relative root mean square error (rRMSE) and structural similarity (SSIM) [16] to quantitatively evaluate the results. The rRMSE is defined as

$$rRMSE(\mathbf{f}, \mathbf{f}_{\text{ref}}) = \frac{\|\mathbf{f}-\mathbf{f}_{\text{ref}}\|_2}{\|\mathbf{f}_{\text{ref}}\|_2}, \quad (1)$$

where $\mathbf{f}$ is the reconstructed image and $\mathbf{f}_{\text{ref}}$ is the reference image. The SSIM is calculated as

$$\text{SSIM}(\mathbf{x}, \mathbf{y}) = \frac{(2\mu_x\mu_y+C_1)(2\sigma_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)}, \quad (2)$$

where $\mathbf{x}$ and $\mathbf{y}$ are two images, respectively. μ_x and μ_y are local means, σ_x and σ_y are local standard deviations and σ_{xy} is cross-correlation.

IV. Results

Fig. 2 shows the volumetric images reconstructed by the FBP, ADMIRE and the proposed method. Representative axial, sagittal and coronal images are present. It can clearly be observed from the first two columns that the FBP images are contaminated by severe noise and the ADMIRE images have higher signal-to-noise ratio. By comparison, the images processed by the network are very close to the ADMIRE images. Fig. 3 plots image profiles indicated by the lines in Fig. 2. The profile in Fig. 3(a) passes through two flat regions, and we can see that the noise level is reduced after the post-processing. The line in Fig. 3(b) passes through different image slices, and fat, muscle, bone and lung are on this profile. It can be seen that the plot of our result is very close to the ADMIRE plot.

To quantitatively evaluate the capability of the proposed algorithm, we assume the ADMIRE images as the references and calculate the rRMSE and SSIM at each image slice as shown in Fig. 4. Overall, the rRMSEs of FBP images are 0.1236 ± 0.0109, which is reduced as low as 0.0435 ± 0.0041. Additionally, the SSIM scores increase from 0.7649 ± 0.0149 to 0.8922 ± 0.0016 after the post-processing.

V. Discussion and Conclusion

The network training is time-consuming. It takes 6.18 hours to complete the training in this work. On the contrary, the post-processing phase is very efficient, taking only 4.48 minutes to process 150 slices. Hence, we can train and store the

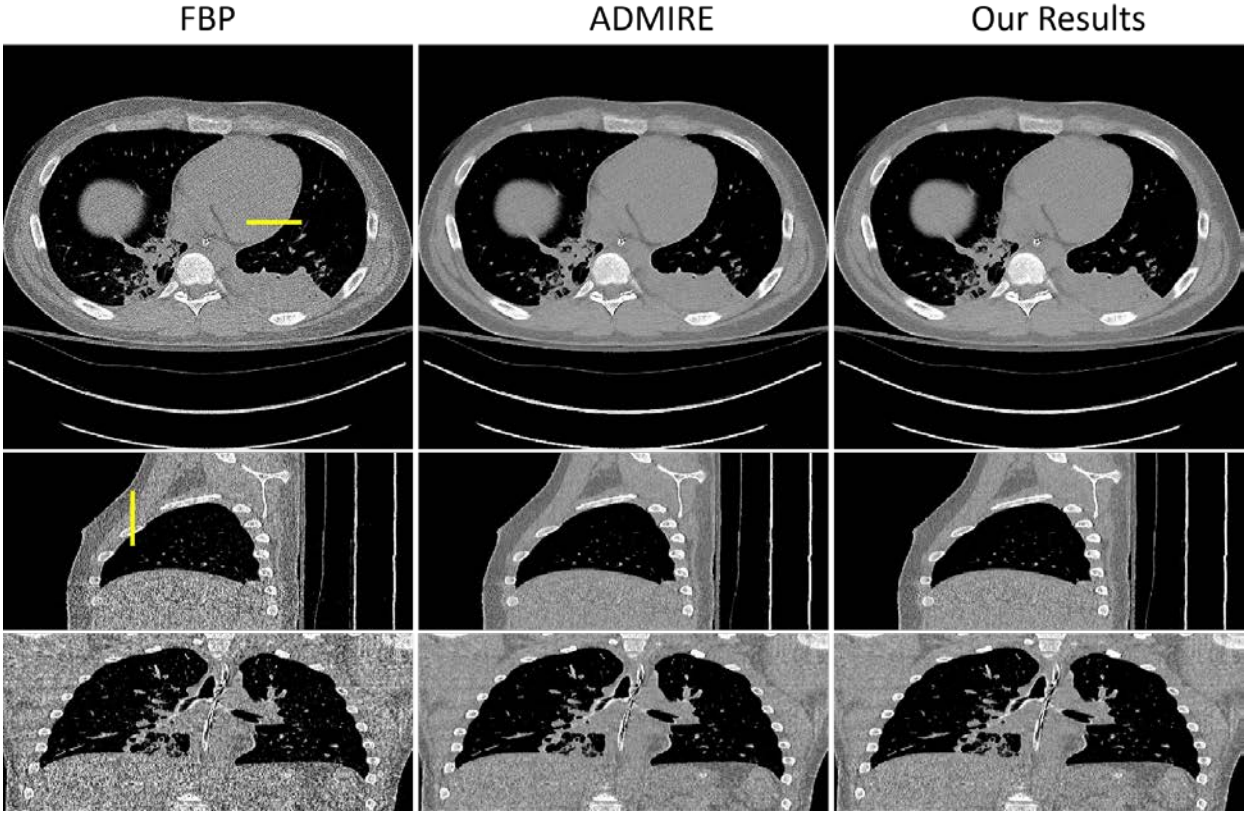


Fig. 2 Representative images of FBP (left), ADMIRE (middle) and post-processed (right). From top to bottom, images are axial, sagittal and coronal views, respectively. The display window is (-500, 500) HU.

network in advance for fast processing. Because the computational cost of the post-processing is dramatically less than that of the iterative reconstruction methods, the proposed method has a great potential in clinical applications.

In the network training phase, the training patches are collected from only one patient data. Hence, the trained network only works for the data with the same scanning and reconstruction parameters, such as noise level, filter strength, *etc*. Thus, we need to train different networks for specific imaging parameters to deal with various imaging cases. Alternatively, we can also adopt training data with variety of noise levels and filter strengths to extend its application range. However, its performance might be compromised to some extent and a deeper network is expected to achieve the same capability.

This preliminary work is evaluated on one patient data. In the near future, we will carry out a comprehensive assessment with more clinical data. Furthermore, medical physicists and radiologists will perform a subjective evaluation.

In conclusion, we have proposed a residual learning of convolutional neural network based post-processing method. Experimental results have confirmed that the developed network has the capability to significantly suppress image noise and to make the obtained images close to the iteratively reconstructed images. In the near future, extensive patient data will be applied to validate the effectiveness of the proposed method and radiologists will be involved for the subjective clinical evaluation.

REFERENCES

[1] R. D. MacDougall, P. L. Kleinman, L. Yu, and E. Y. Lee, "Pediatric thoracic CT angiography at 70 kV: a phantom study to investigate the effects on image quality and radiation dose," *Pediatric radiology*, vol. 46, pp. 1114-1119, 2016.

[2] M. Weis, T. Henzler, J. W. Nance Jr, H. Haubenreisser, M. Meyer, S. Sudarski, *et al.*, "Radiation Dose Comparison Between 70 kvp and 100 kvp With Spectral Beam Shaping for Non–contrast-enhanced Pediatric Chest Computed Tomography: A Prospective Randomized Controlled Study," *Investigative radiology*, vol. 52, pp. 155-162, 2017.

[3] T. Niemann, S. Henry, A. Duhamel, J.-B. Faivre, A. Deschildre, L. Colas, *et al.*, "Pediatric chest CT at 70 kVp: a feasibility study in 129 children," *Pediatric radiology*, vol. 44, pp. 1347-1357, 2014.

[4] C. H. McCollough, M. R. Bruesewitz, and J. M. Kofler Jr, "CT dose reduction and dose management tools: overview of available options," *Radiographics*, vol. 26, pp. 503-512, 2006.

[5] L. Yu, H. Li, J. G. Fletcher, and C. H. McCollough, "Automatic selection of tube potential for radiation dose reduction in CT: a general strategy," *Medical physics*, vol. 37, pp. 234-243, 2010.

[6] F. Mafalanka, C. Etard, J. Rehel, D. Pesenti-Rossi, F. Amrar-Vennier, N. Baron, *et al.*, "Establishment of diagnostic reference levels in cardiac CT in France: a need for patient dose optimisation," *Radiation protection dosimetry*, vol. 164, pp. 116-119, 2014.

[7] S. Singh, M. K. Kalra, M. D. Gilman, J. Hsieh, H. H. Pien, S. R. Digumarthy, *et al.*, "Adaptive statistical iterative reconstruction technique for radiation dose reduction in chest CT: a pilot study," *Radiology*, vol. 259, pp. 565-573, 2011.

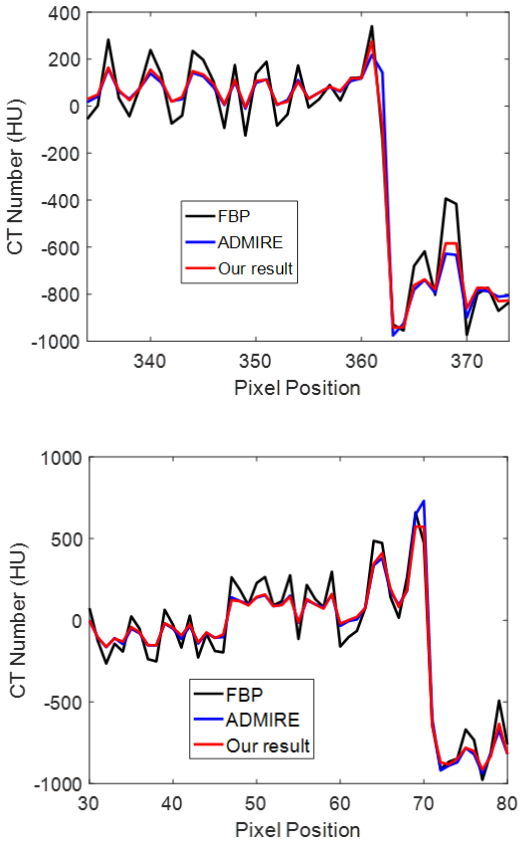


Fig. 3. Plots at the lines indicated by the horizontal and vertical lines in Fig. 2, respectively.

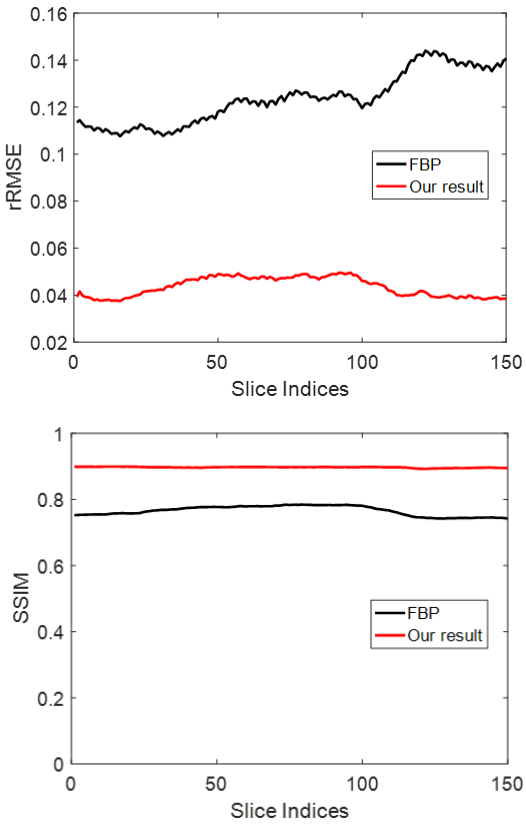


Fig. 4. RRMASE and SSIM before and after the post-processing for each slice.

[8] A. Padole, S. Singh, J. B. Ackman, C. Wu, S. Do, S. Pourjabbar, *et al.*, "Submillisievert chest CT with filtered back projection and iterative reconstruction techniques," *American Journal of Roentgenology*, vol. 203, pp. 772-781, 2014.

[9] E. Kang and J. C. Ye, "Wavelet Domain Residual Network (WavResNet) for Low-Dose X-ray CT Reconstruction," *arXiv preprint arXiv:1703.01383*, 2017.

[10] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, *et al.*, "Low dose CT image denoising using a generative adversarial network with wasserstein distance and perceptual loss," *arXiv preprint arXiv:1708.00961*, 2017.

[11] W. Yang, H. Zhang, J. Yang, J. Wu, X. Yin, Y. Chen, *et al.*, "Improving Low-Dose CT Image Using Residual Convolutional Network," *IEEE Access*, vol. 5, pp. 24698-24705, 2017.

[12] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising," *IEEE Transactions on Image Processing*, 2017.

[13] G. Wang, "A Perspective on Deep Imaging," *IEEE Access*, vol. 4, pp. 8914-8924, 2016.

[14] S. Gordic, L. Desbiolles, P. Stolzmann, L. Gantner, S. Leschka, D. Husarik, *et al.*, "Advanced modelled iterative reconstruction for abdominal CT: qualitative and quantitative evaluation," *Clinical radiology*, vol. 69, pp. e497-e504, 2014.

[15] A. Vedaldi and K. Lenc, "Matconvnet: Convolutional neural networks for matlab," in *Proceedings of the 23rd ACM international conference on Multimedia*, 2015, pp. 689-692.

[16] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, pp. 600-612, 2004.

Ring Reduction for Micro CT using Deep Residual Learning

Matthew Holbrook, Darin P. Clark, Cristian T. Badea

Abstract— **Computed tomography (CT) ring correction is an active area of research. Many strategies exist for reducing ring artifacts including experimental and image processing procedures. Deep learning (DL) has achieved success in many applications for image corrections, including correction for CT-specific noise and reconstruction artifacts. We propose a DL approach to ring artifact correction using a class of convolutional neural network known as a residual U-net. The U-net contains collapsing encoder layers and expanding decoder layers. Pooling and unpooling was performed more aggressively in the angle dimension of the sinogram to aid in the detection and correction of continuous ring artifacts. The U-net was coded in Python using Tensorflow. Functions found in the Keras API were employed for convenience. Our network was trained on experimental data which had been altered to include a variety of random ring artifacts. The results confirm successful removal or attenuation of ring artifacts in experimentally acquired micro-CT sets. The performance of U-net ring correction on experimental data can be further improved by increasing the quality of simulated training data and the variety of rings it contains.**

I. Introduction

Ring artifacts are concentric patterns which often appear in images acquired by third generation computed tomography (CT) systems. For micro-CT acquisitions in which image resolution is pushed to the extreme, rings artifacts are especially prone to occur and to degrade image quality. Ring artifacts can stem from a variety of causes, including defective detector pixels, non-linear pixel responses, inconsistent detector gain, detector and x-ray source structural patterns, and changes in system performance over time and temperature conditions.

Many experimental methods for reducing ring artifacts have been explored. Minimization of ring artifacts is possible by using flat-field correction [1], moveable detector arrays [2], and adequate scanning protocols (e.g., dual gain calibration technique [3]). However, it is difficult to totally eliminate such artifacts and attain high-quality reconstructions through purely experimental measures. Image processing techniques for correcting ring artifacts have been implemented in the sinogram domain [4], the image domain [5], [6], and in combinations of the two [7]. The performance of these methods is dependent on the quantity and frequency of the ring artifacts. These methods generally achieve their best results with few, high frequency rings. Low-frequency rings are more difficult to detect and correct. More robust approaches to ring artifact correction are needed which can adequately minimize the effect of low and high frequency artifacts in reconstructions.

Recently, deep learning (DL) has generated enthusiasm in various imaging applications, including image denoising, deblurring, super-resolution, segmentation, detection, and recognition [8]. Convolutional neural networks (CNNs) are powerful DL models recently tested in CT applications. For example, Chen et al. utilized a CNN-based framework for artifact reduction in low-dose CT imaging [9]. An algorithm using a deep CNN was also applied to the wavelet transform coefficients of low-dose CT images to suppress CT-specific noise [10]. A broad class of CNN architectures, referred to as U-Nets, has been effectively used to reduce reconstruction artifacts in low-dose clinical CT acquisitions [11]. These studies show that DL techniques can be effectively used to detect and correct structural patterns present in CT images.

To our knowledge, no study has applied DL techniques for ring artifact correction in CT. There are several variants of CNNs for image-to-image tasks that could be considered for ring artifact corrections including encoder-decoder configurations. The information consolidated by the encoder encrypts the global structure of the input image, discarding non-essential details. The decoder section of the network then interprets the encoded data to produce an image with the desired features. In this study we use an approach based on a U-Net CNN. This approach takes advantage of the structure of ring artifacts in the sinogram domain; ring artifacts appear as stripes spanning the angle dimension. We focus here on developing and testing a CNN-based ring artifact correction method for our in-house developed

M. Holbrook, D. P. Clark, and C. T. Badea are with the Center for In Vivo Microscopy, Department of Radiology, Duke University, Durham, NC 27710.

All work was performed at the Duke Center for In Vivo Microscopy, an NIH/NIBIB national Biomedical Technology Resource Center (P41 EB015897). This work was also supported by the NIH National Cancer Institute (R01 CA196667, U24 CA220245) and NIH Grant #T32-EB001040.

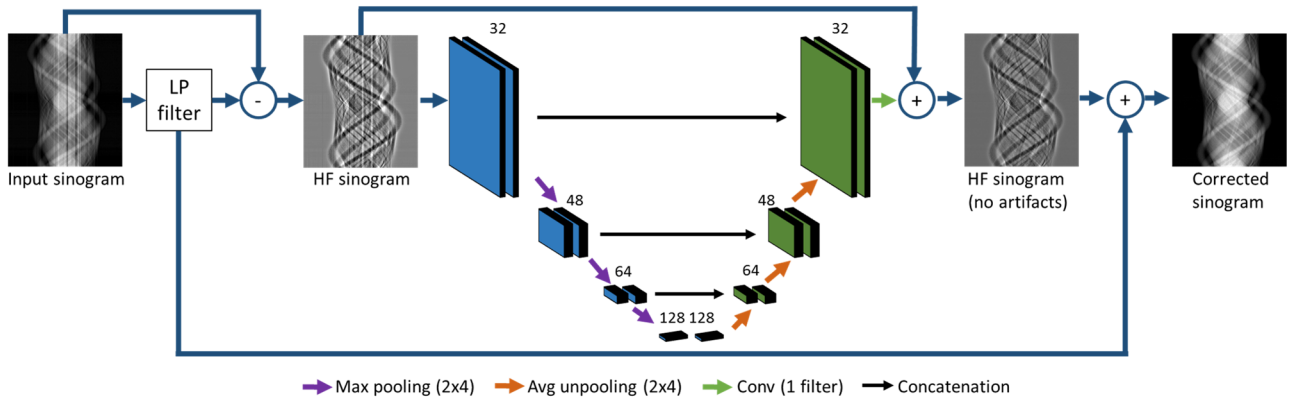


Fig. 1. Map of the network architecture used to remove rings from normalized sinograms. Prior to the trainable layers of the network each input image is separated into high and low frequency components via a low-pass filter. The high frequency image is normalized by its own standard deviation and sent through the network. At the end of the network the sum of the output and high frequency sinogram are scaled by the old standard deviation. The corrected sinogram is obtained by adding back the low frequency image. The network consists of encoder (blue) and decoder (green) halves. There are two convolution and ReLU operations performed for each set of maps shown. The number of feature maps after each pair of operations is given above the stacks. Concatenation operations between the encoder and decoder allow recovery of high frequency information.

micro-CT scanner [12], but a similar approach could be applied, with appropriate training, to any other CT scanner.

II. Methods

A. CNN U-Net architecture

The network used in this work is shown in Fig 1. The first operation performed by the network is to separate input sinograms into high and low frequency components using a low-pass filter. Care was taken in selecting a cutoff frequency for the filter such that all ring artifacts ended up in the high frequency map, which is passed through the network. This operation effectively zero-centers the data. After this operation the high frequency image is normalized by its standard deviation. This scaling is reversed at the end of the network before adding high and low frequency components together to form the corrected sinogram.

The U-net is divided in encoder and decoder halves. The encoder half encodes features relevant to ring detection. The decoder portion of the network organizes these features into an output image. Each stage in the U-net consists of a pair of convolutions and rectified linear unit (ReLU) activations. The network contains four distinct levels, separated by max pooling and average unpooling operations. In the encoding portion of the network, feature maps undergo max pooling, reducing the size of the maps and allowing the network to find coarser features. In the decoding half of the network, feature maps are average un-pooled, bringing the feature map dimensions back to match those of the input image. The number of feature maps increases as the size of the feature maps decreases, allowing for features to be preserved at lower resolutions.

Ring artifacts manifest as continuous linear stripes in the sinogram domain, making them a good candidate for detection by a CNN. To encourage the network to find ring artifacts as continuous lines, the pooling and unpooling operations in this network are anisotropic. Max pooling and average unpooling are performed in a 2x4 pixel window; the larger window dimension corresponds to the sinogram angle dimension. Ring artifacts found in the encoding half of the network are stretched and combined along the angle dimension to create continuous stripes in the output image. A ring artifact detected in enough places will be corrected across the whole sinogram.

The network presented here is a residual network, meaning that the network is trained to find the residual image which connects the corrected and uncorrected images. The sum of the input high frequency sinogram and the network output will produce a high frequency sinogram free of ring artifacts. Residual networks have been situationally shown to train faster than networks which attempt to find a direct solution [11], [13]. These networks converge faster because most kernels and weights are initialized to be close to zero, resulting in an initial near-zero output from the network. In cases in which the label has small values or is sparse relative to the input image, this near-zero initial guess will start much closer to the correct solution. In this work, a residual network was employed to find image based rings because the desired residual output will be both small and sparse.

The U-net was coded in Python using Tensorflow [14]. Functions found in the Keras API [15] were employed for convenience in coding portions of the network.

B. Training data

Data for training was comprised of a 2:1 mix of experimental and simulated sinograms, both altered to contain ring artifacts. Realistic artifacts were generated through the use of flat field images. A set of 600 flat fields were captured over a period of 2 hours with an imaging system whose performance varies with time and temperature. The flat fields were normalized with one another, generating sets of low and high frequency artifact maps where the flat fields disagreed, similar to those artifacts that would appear in an imaging application after flat field normalization. These artifact maps were applied to clean, normalized sinograms via multiplication prior to log transforming projection data. In addition, bad pixels and patches of bad pixels were randomly simulated into other sinograms. In this way we

created pairs of artifacted sinograms and their equivalent clean labels i.e. sinograms without artifacts.

C. *Training the U-net*

The network was trained on 3600 pairs of sinograms. The training set dimensions were 2000 by 360 (detector row by projection angles) pixels. The network was trained in batches of 48. The Adam optimization algorithm was used to train the network by minimizing the mean square error (MSE) of the output. The network trained for 24 hours, or 375 epochs. Training was ended after 24 hours to prevent overfitting. The network was trained on a stand-alone workstation. Computations were performed in parallel on four NVIDIA Titan Xp graphics cards.

III. Results

A. *Training performance*

The MSE of the training set during training is given in Fig 2. By stopping training before it reached its minimum training loss, we prevented overfitting.

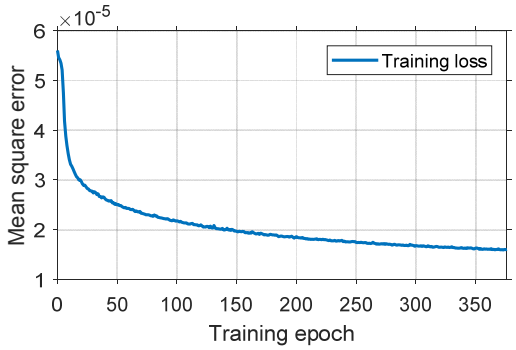


Fig. 2. Training error as a function of epoch. The network was trained for 24 hours on a mix of experimental data with inserted ring artifacts and simulated data. Training was ended after one day to prevent overfitting the training data.

B. *Network performance on experimental data*

The ability of the U-net to detect and correct organic ring artifacts was tested on experimental datasets afflicted with ring artifacts. The artifacted data was captured on the same micro CT system used to acquire the flat field images for the generation of training data. The test dataset contained 1000 equiangular projections over a single 360 degrees of rotation. The data was reconstructed with 44 µm isotropic voxels and contained both high and low frequency rings.

Prior to processing the data with the CNN, each sinogram was subsampled by a factor of 3 along the angular dimension by selecting every third projection. This was done to give the sinograms of the test data a similar size and angular step as contained in the training data, which contained 360 angles. The subsampled sinograms were then individually processed by the trained network. The sinograms were restored to their original size by interleaving the angles of the subsampled sinograms. Filtered backprojection reconstructions (FBP) using the original and corrected projections are shown in Fig 3. Sample images from the reconstructions using uncorrected projections illustrate the amplitudes and frequencies of ring artifacts that exist in the dataset. The overall amplitude of the rings is visibly reduced in the corrected reconstruction. The effectiveness of the correction can be better assessed by the residual images showing the difference between the original and corrected images. These images demonstrate that the network has only altered the ring artifacts as image features are not present. The line profiles, plotted through the center of the reconstructions, show how the artifacts have been altered by the network. The largest amplitude rings are most affected by the network. Variation in the line profile due to vessels, lung, and bone remain visible in the corrected reconstruction. However, both high and low frequency rings, though diminished, are still present in the corrected image.

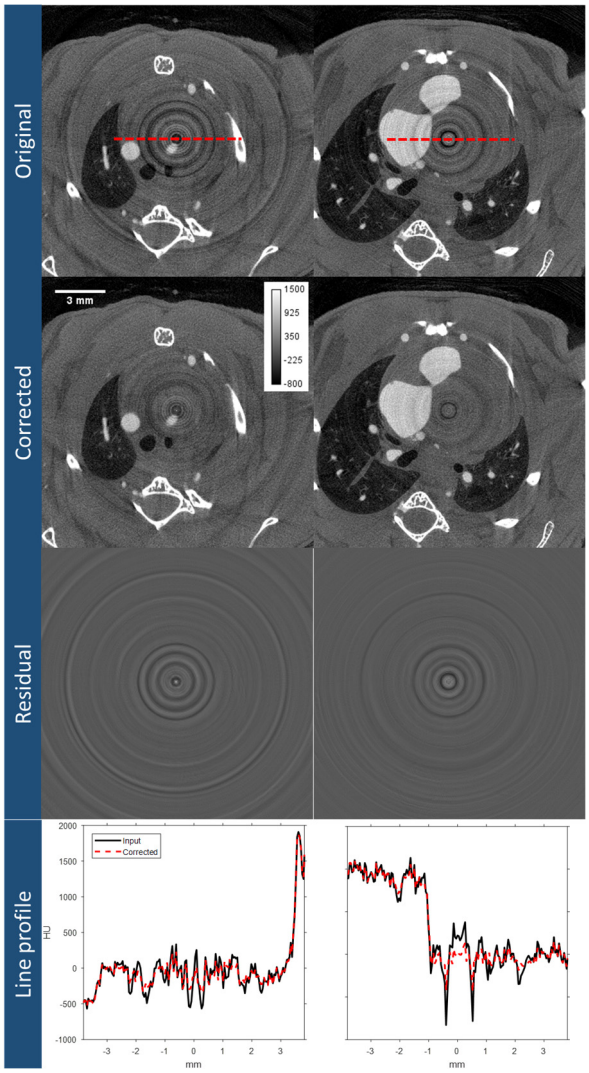


Fig. 3. Reconstructions of experimental data with ring artifacts. The images were reconstructed to 44 µm voxels and contain a mix of high and low frequency ring artifacts. Residual images between the original and corrected reconstructions demonstrate that image features are left undisturbed by the network, only rings are affected. The line profiles are taken from the dashed red lines and demonstrate that variation due to structures in the images are preserved in the original and corrected reconstructions.

IV. Discussion

In this work, we have examined a class of CNN known as a residual U-net for the purpose of ring artifact reduction in micro-CT data. This type of network has been successfully used to improve CT reconstruction and shows promise for the reduction of ring artifacts.

The network was trained on a mix of experimental and simulated data altered to include a variety of random ring artifacts. As shown from the MSE of the training (Fig. 2), our network was able to learn and correct the pattern of ring artifacts in the training data. To help with the task of reducing ring artifacts, the simulated rings in the training dataset were created with the same artifacts and frequencies captured in the test set. This allowed the network to learn the range of ring amplitudes and frequencies present in the test data. The success of this method is demonstrated in the ring corrected reconstruction by a decrease in amplitude of low frequency rings, which are generally difficult to identify and correct.

The network was designed to pool and unpool more aggressively in the angle dimension of the sinogram. This was done to facilitate identification and correction of artifacts represented as continuous stripes. The residual images of the test reconstruction (Fig. 3), which are shown as continuous circles, demonstrate the success of the network in creating near-continuous corrections. The lack of image features beyond the rings in the residual images show the ability of the network to reduce rings while preserving image structure.

Convergence of the U-net was difficult to declare because there was no good validation set on which to monitor training progress. Pairs of simulated features and labels do not guarantee that the network will generalize well to data outside the simulations. Thus the performance of the network was determined qualitatively by its ability to treat sinograms with naturally occurring ring artifacts. By checking the performance of the network on experimental ring data and monitoring the effectiveness of the ring artifact reduction it was determined to stop training after 24 hours. We note this method for determining convergence as a limitation of this work, as it is inefficient and subject to human judgement.

The network performed well on experimental data (Fig. 3, top), but there is room for improvement. The test sets were selected to contain a variety of rings, both high and low frequency. The highest amplitude rings, both low and high frequency, were more successfully addressed by the network and show reduced amplitude. However, the highest frequency rings were left largely uncorrected as shown in the line profiles, particularly if those rings were low amplitude. Low amplitude, high frequency rings would affect the MSE cost function of the network less than large amplitude or low frequency rings. The presence of these rings in the corrected images could indicate that the network was not fully trained at the time of testing.

The shortcomings of the network can be largely attributed to the training data and the lack of good validation data. The ring artifacts in the experimental data which most affect the MSE loss of the network saw the most correction. For example, the high amplitude rings saw the most correction by the network. It is possible that with more training time, better results could be attained. The availability of good training data is one of the biggest difficulties in machine learning. In many cases such data may not exist and must be simulated. This is the case for CT ring correction; there is no perfect correction for ring artifacts which can produce clean label data on which to train, and such a technique would invalidate the need for DL ring correction. Training a network with data corrected using some other technique would only serve to make a network approximate that technique. However, using simulations we can create adequate training sets provided that we can model the whole range of ring artifacts that can appear experimentally.

The network presented here has been shown to have the ability to correct certain types of ring artifacts in the sinogram domain. The performance of the network on experimental data can be improved by increasing the quality of the training data and validation data.

References

[1] L. R. Benedetti *et al.*, "A comparison of 'flat fielding' techniques for x-ray framing cameras," *Rev. Sci. Instrum.*, vol. 87, no. 11, p. 11D622, Nov. 2016.

[2] Y. Zhu, M. Zhao, H. Li, and P. Zhang, "Micro-CT artifacts reduction based on detector random shifting and fast data inpainting," *Med. Phys.*, vol. 40, no. 3, p. n/a-n/a, Mar. 2013.

[3] P. G. Roos *et al.*, "Multiple-gain-ranging readout method to extend the dynamic range of amorphous silicon flat-panel imagers," presented at the Medical Imaging 2004: Physics of Medical Imaging, 2004, vol. 5368, pp. 139–149.

[4] B. Münch, P. Trtik, F. Marone, and M. Stampanoni, "Stripe and ring artifact removal with combined wavelet — Fourier filtering," *Opt. Express*, vol. 17, no. 10, pp. 8567–8591, May 2009.

[5] L. Yan, T. Wu, S. Zhong, and Q. Zhang, "A variation-based ring artifact correction method with sparse constraint for flat-detector CT," *Phys. Med. Biol.*, vol. 61, no. 3, p. 1278, 2016.

[6] Y. Kyriakou, D. Prell, and W. A. Kalender, "Ring artifact correction for high-resolution micro CT," *Phys. Med. Biol.*, vol. 54, no. 17, p. N385, 2009.

[7] M. E. Eldib, M. Hegazy, Y. J. Mun, M. H. Cho, M. H. Cho, and S. Y. Lee, "A Ring Artifact Correction Method: Validation by Micro-CT Imaging with Flat-Panel Detectors and a 2D Photon-Counting Detector," *Sensors*, vol. 17, no. 2, p. 269, Jan. 2017.

[8] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.

[9] H. Chen *et al.*, "Low-dose CT via convolutional neural network," *Biomed. Opt. Express*, vol. 8, no. 2, pp. 679–694, Jan. 2017.

[10] E. Kang, J. Min, and J. C. Ye, "A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction," *ArXiv161009736 Cs*, Oct. 2016.

[11] Y. S. Han, J. Yoo, and J. C. Ye, "Deep Residual Learning for Compressed Sensing CT Reconstruction via Persistent Homology Analysis," *ArXiv161106391 Cs*, Nov. 2016.

[12] C. T. Badea, S. Johnston, B. Johnson, M. Lin, L. W. Hedlund, and G. A. Johnson, "A dual micro-CT system for small animal imaging," *Proc. SPIE*, vol. 6913, p. 691342, Apr. 2008.

[13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *ArXiv151203385 Cs*, Dec. 2015.

[14] M. Abadi *et al.*, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," *ArXiv160304467 Cs*, Mar. 2016.

[15] F. Chollet and others, *Keras*. GitHub, 2015.

CT Sinogram Analysis Using Deep Learning

Eri Haneda, Bernhard Claus, Paul FitzGerald, Ge Wang, Bruno De Man

***Abstract*— This paper presents an approach to extract clinically useful information directly from sinograms using a neural network approach. In recent years, deep neural networks have been applied in a wide variety of image processing fields for denoising, detection, recognition, and understanding. Recent studies in CT imaging show that neural networks can improve image quality, feature extraction, and lesion detectability. Most CT applications of neural networks have been in the image domain. A few groups have presented so-called end-to-end neural networks, which operate directly in sinogram space and perform an initial 'learnt' reconstruction step, possibly followed by a 'learnt' analysis step. To our knowledge, the application of neural networks to extract diagnostic information directly from sinogram data, without an explicit reconstruction step, has not yet been presented.**

In this paper, we demonstrate sinogram analysis with neural networks in the context of vessel lumen cross-sectional area size estimation. This is a proof-of-concept study of neural networks toward feature extraction from projection data. We simulated blood vessels with different diameters filled with iodine contrast agent and applied a fully connected network to the resulting sinograms to estimate the vessel cross-sectional area. We also applied a similar network to sinograms from a CT measurement of a vessel phantom we created using a thorax phantom. Our results suggest that the neural network can estimate vessel cross-sectional area accurately with reasonable error ranges in both cases.

***Index Terms*—Computed tomography, Neural networks, Sinogram, Raw data, Estimation**

I. Introduction

Automatic extraction of parameters that characterize structures and lesions within CT images is becoming increasingly important – with the goal of improving workflow and providing the radiologist with information that was previously not readily available. The associated processing approaches are partly designed to mimic and augment the role of the radiologist when identifying and extracting characteristics of structures contained within the image (e.g. computer-aided detection (CAD)), while a more recent goal is to extract image information that is not visually apparent for characterization by the radiologist (e.g. radiomics).

When extracting quantitative information from CT data with high accuracy, the focus shifts away from the traditional image quality considerations. Indeed, for this task, the reconstruction step represents only an intermediate step, and a more direct path consists of estimating parameters directly from the projection data / sinogram. Thereby one hopes to avoid loss of information due to various factors (e.g. resolution loss due to interpolation in the reconstruction, statistical information, degradation due to motion, etc.) but the price is that we must deal with data in the projection domain (superimposition of structures, etc.).

Recent advances in deep learning networks show that this type of processing is well suited to dealing with highly complex data that are difficult to capture with traditional algorithmic approaches. As a first step in this direction we present initial results in the development of a solution for estimating vessel lumen cross-sectional area directly from sinogram data. This is of practical interest, for example, in the imaging and evaluation of coronary atherosclerosis – for example for characterization of coronary stenosis or for creating a model of the coronary vessels for use in computational fractional flow reserve (cFFR).

For static object analysis, some studies have already demonstrated the theoretical advantage of sinogram-domain analysis over image-domain analysis in some applications. Chesler et al. [1] analytically proved that sinogram-domain SNR is greater than image-domain SNR under simple conditions where background and target are exactly known. There are also simulation studies that support this claim in the context of attenuation-coefficient estimation and lesion detectability [2], [3]. For dynamic analysis, the data-driven approach is sometimes used for cardiac and respiratory motion estimation and correction [4]. A direct-sinogram approach was also used for automatic cardiac phase selection in retrospective ECG gating [5].

II. Method

A. Neural network model

In this study, we used a generic neural network to show the potential of estimating vessel lumen cross-sectional area directly from sinogram data. We used fully connected networks with 3 hidden layers, where each layer consists of 48 neurons for all studies. According to the universal approximation theorem, any smooth function can be described by one hidden-layer feedforward network [6] - [8]. The theorem allows a general expressibility but guarantees neither efficiency nor learnability. Extremely large numbers of neurons and iterations may be required to fit an arbitrary function. Since multiple-layer networks can approximate polynomials more efficiently than single-layer networks [6], [9], deeper networks are typically used. However, extremely deep networks are also difficult to train. We used ReLU

Eri Haneda (haneda@ge.com), Bernhard Claus, Paul FitzGerald, and Bruno De Man are with GE Global Research, Niskayuna, NY. Ge Wang is with Rensselaer Polytechnic Institute, Troy, NY.

activation functions in all neurons and used SGD (stochastic gradient descent) for finding the weights in training. The network was implemented using Keras with a Tensorflow backend.

B. Simulations

1) Simple phantom simulation

We initially performed simple phantom simulations using ellipses (representing vessel cross-sections) with different sizes, positions, and densities. The major and minor axes (diameters) were randomly chosen from a uniform distribution from 0 to 5.0 mm, and the center position was also randomly selected from a uniform distribution from -2.0 to 2.0 mm in X and Y coordinates. The ellipse density was randomly selected from a uniform distribution from 1.0 to 1.6 density/cm^3. No background material was used. The angular orientation of the elliptical cross-section in the axial plane was varied by circularly shifting each sinogram by a random offset along the view direction. We used CatSim [10] to simulate a 2D parallel-beam geometry with 300 mA, 120 kVp, 1.0 second per gantry rotation, and 984 views per rotation. The parallel beam spacing was set to 1.0 mm and only the 10 central detector pixels were used to capture the projection of the ellipses.

We generated 30,000 training samples and 1,000 test samples. Ten percent of the training samples were used for validation. The ground-truth ellipse area (mm^2) was mathematically computed from the major/minor axes' lengths.

2) Advanced phantom simulation

To simulate more realistic scenarios, we set up a slightly more complicated phantom and CT scan settings. We created 2D elliptical objects (representing vessel cross-sections) with major and minor axes (diameters) randomly selected from a beta distribution with a range from 0.5 to 5.0 mm such that the mode of the size distribution was at about 1.5 mm. The ellipse orientations were randomly selected; the center positions (X and Y coordinates) were randomly selected from a uniform distribution from -50.0 to 50.0 mm. The ellipse density was uniformly selected from 1.2 to 1.6 g/cm^3. A 200-mm-diameter uniform water background was positioned at the center to create more realistic quantum noise in the simulated sinogram. For each simulation, four random vessel-representing ellipses were embedded in the background. We generated 5,000 training scans and 500 test scans, thereby providing 20,000 vessel training samples and 2,000 vessel test samples.

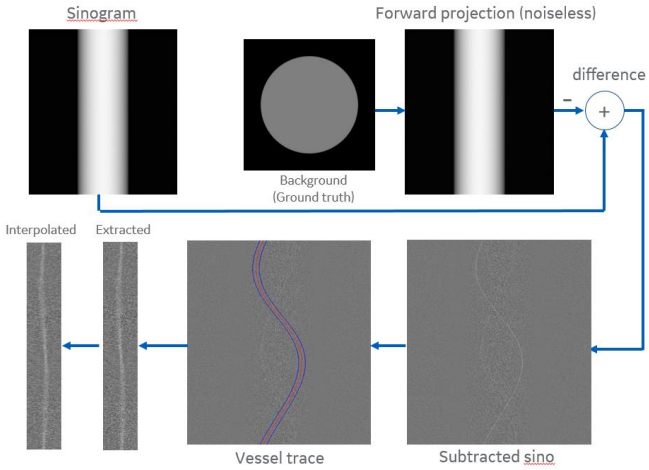


Fig. 1. Flowchart of vessel trace extraction from sinogram in a simulation. The ground truth background sinogram was first subtracted from the original sinogram. Then, the vessel trace location was found by projecting target vessel position (red line). The trace was extracted with blue line boundaries, then interpolated to align to true position in each view. Then the results were supplied to the neural network.

We used CatSim to simulate a 2D axial fan-beam geometry with large bowtie filter, 300 mA, 120 kVp, 0.28 seconds per gantry rotation, and 984 views per rotation. The detector was 888 pixels wide with 1.0 mm pitch. To simplify the experiment, we assumed that ground truth background and target vessel position were known. With this information, each target vessel "trace" (i.e., a small range of detector pixels centered around the nominal position of the vessel projection on the detector) can be extracted from each sinogram and supplied to a neural network. The flowchart for creating the network input is shown in Fig. 1. First, the ground truth background sinogram was subtracted from the sinogram that included the target vessel. Then, the vessel trace position was found in the sinogram by projecting the target vessel (indicated by the red line). The trace between two boundaries (indicated by blue lines) was extracted with an interval of 33 pixels around the red line. The 33-pixel interval was wide enough to cover all vessel projections at any position with any size. The extracted sinogram was then interpolated in each view for the alignment to the true trace position in mm.

C. Phantom measurements

A custom-made vessel phantom was constructed consisting of vessel inserts in a homogeneous cylindrical background, which was then inserted in a QRM pseudo-anthropomorphic thorax phantom, and scanned by a 256-slice GE Revolution CT (Fig. 2). Twelve cylindrical holes with diameters of 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.7, 2.0, 2.5, 3.0, 4.0, 5.0 mm were filled with 30-mg/mL iodine-water solution. (The hole in the center with an 8.0-mm diameter was not analyzed.) The phantom was placed at 9 different vertical and horizontal positions. The outer and inner six vessel inserts were also swapped for each position. Three axial scans were performed in each position using 120 kVp, 350 mA, and 1.0 sec/rotation. Sinograms were created from the data from the two central detector rows.

The procedure of target vessel trace extraction is shown in Fig. 3. Starting from the acquired sinogram data (b), a reconstruction was performed and a target vessel was

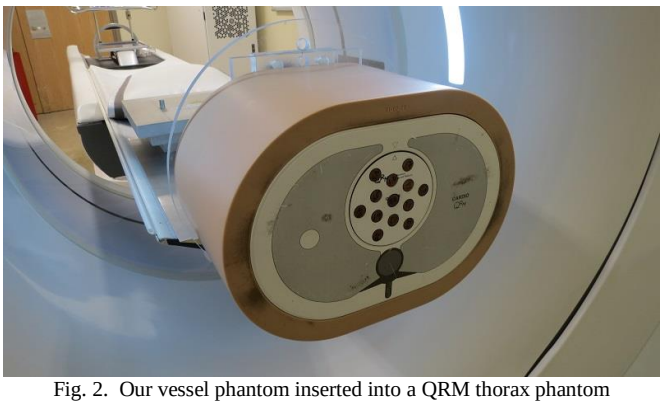


Fig. 2. Our vessel phantom inserted into a QRM thorax phantom

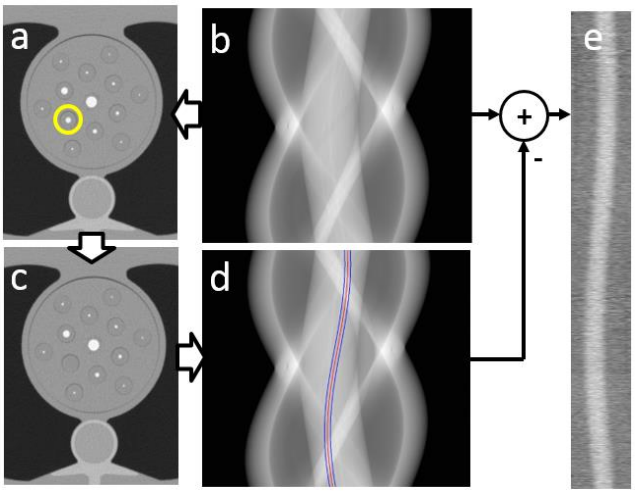


Fig. 3. Flowchart of vessel trace extraction in an experimental scan. The target vessel was erased from the reconstructed image (a) by neighboring pixel averaging (c). The image was then forward projected (d) and the sinogram was subtracted from the original sinogram (b). From the resulting sinogram, the vessel trace between the two boundaries (blue line) was extracted to obtain (e).

identified in the resulting image (a). The location of the target vessel determined the corresponding vessel trace in sinogram (d, red line). Confounding background information was removed from the acquired sinogram by first erasing the vessel in the reconstruction using neighboring pixel averaging (c), forward projecting the image and subtracting the background sinogram (d, without the colored lines) from the original sinogram (b). From the resulting sinogram we extracted the vessel trace between the two boundaries with an interval of 19 pixels (d, blue lines) to obtain (e). Note that the first-pass reconstruction may not be strictly necessary, but it greatly facilitates identifying the region of interest and subtracting the background sinogram.

After the 12 vessel traces were extracted from each sinogram, the vessel sinograms were circularly shifted in increments of 22.5 degrees to emulate 16 different start-scan angles. A total of 20,736 vessel sinograms were analyzed (12 vessel phantoms × 18 phantom positions × 3 scans × 2 detector rows × 16 angles). One half of the datasets were used for training, and the other half were used for testing.

III. EXPERIMENT RESULTS

A. *Simulations*

1) *Simple phantom simulation*

Fig. 4 shows a scatter plot of the estimate versus ground truth for the vessel cross-sectional area (mm^2). The red line indicates the ideal. The variance of estimate was almost uniform for all sizes. The accuracy was slightly improved as the vessel became larger. Fig. 4 also shows the trained weights in the first hidden layer. For improved display, these weights correspond to a 41-view case instead of a 984-view case. It can be seen that trained network typically contains edge detection filters in the first layer.

2) *Advanced phantom simulation*

Fig. 5 shows four scatter plots of estimate versus ground truth for the advanced phantom simulation. The top left shows the noiseless case (infinite mA) while the top right shows results with noise (finite mA) but without a 200-mm water background phantom. The bottom left is with noise and with the background phantom. We can observe that the noise level impacts the estimation accuracy substantially. The variance of estimation becomes larger as quantum noise increases. However, we found that the error could be dramatically reduced by partially using the weights obtained from pre-training with noiseless data. The bottom right scatter plot is the results from the modified network with the same data. More specifically, we first trained a full network with noiseless data and L2 regularization. Then all other layer weights were trained with noisy data while the first layer was frozen. With this hybrid network, the error was reduced and became comparable to the experiment without a background phantom. This result implies that quantum noise makes network training difficult (in general, the first few layers are more difficult to train due to the vanishing gradient problem). However, the results can be improved by utilizing a hybrid training with noiseless data, which can be substituted by high mA data.

B. *Phantom measurements*

Note that there were substantial differences between the actual scan and simulation experiments. First, the number of vessel diameters was limited, so "ground truth" data in the scatter plot became discrete. Second, the minimum diameter of our circular vessels was 1.0 mm due to the mechanical challenges of phantom construction. We also tested only a single density of contrast agent (30 mg/mL). As for the vessel trace extraction, we developed a new scheme since target and background were not exactly known for the actual scan.

Fig. 5 shows a scatter plot of the estimate versus ground truth for the actual scans. The variance of estimation is uniform for all sizes and reasonably small compared to the errors in the simulation results. The background subtraction and vessel trace extraction did not seem to degrade the estimate accuracy.

IV. CONCLUSION

In this paper, we demonstrated the feasibility of blood vessel lumen cross-sectional area estimation directly from CT sinogram data using a trained neural network. In simulations, we tested ellipses with different sizes, orientations, positions, and densities under realistic conditions. Although training becomes more challenging as quantum noise increases, the estimation error could be dramatically reduced by partially using weights obtained from pre-training with noiseless data, which can be substituted by high mA data. We also performed a similar evaluation using a custom-made vessel phantom scanned by a GE Revolution CT scanner. Although the vessel characteristics of the phantom were limited and simplified compared to the simulated scenarios, the estimation error was quite reasonable. The capabilities of sinogram-domain analysis versus image-domain analysis still need to be carefully and extensively investigated; however, our study

clearly demonstrated that it may be possible to use CT sinograms and neural networks to estimate blood vessel cross-sectional area.

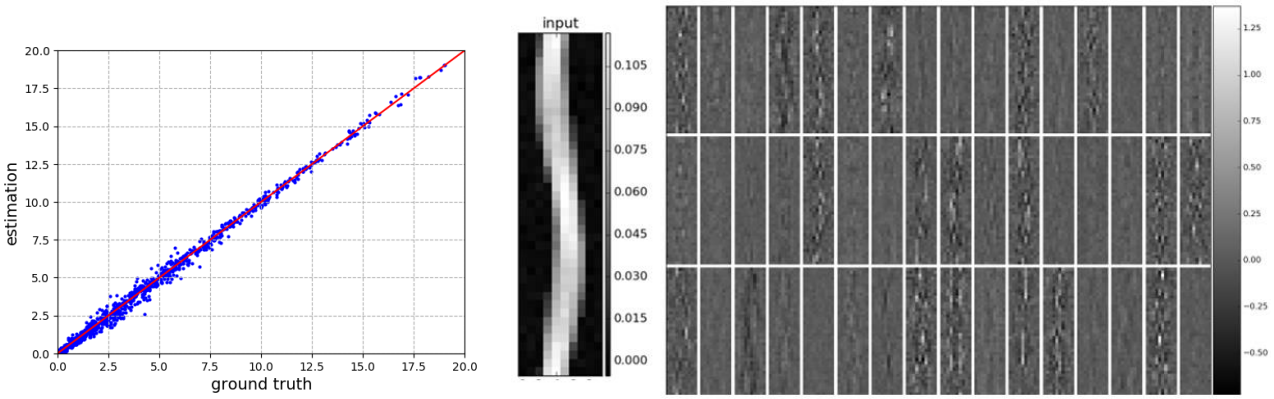


Fig. 4. Left: Scatter plot of estimation vs. ground truth for vessel cross section area (mm^2) in a simple phantom simulation. Middle: An example of a simulated sinogram. Right: The weights in the first layer of the trained network.

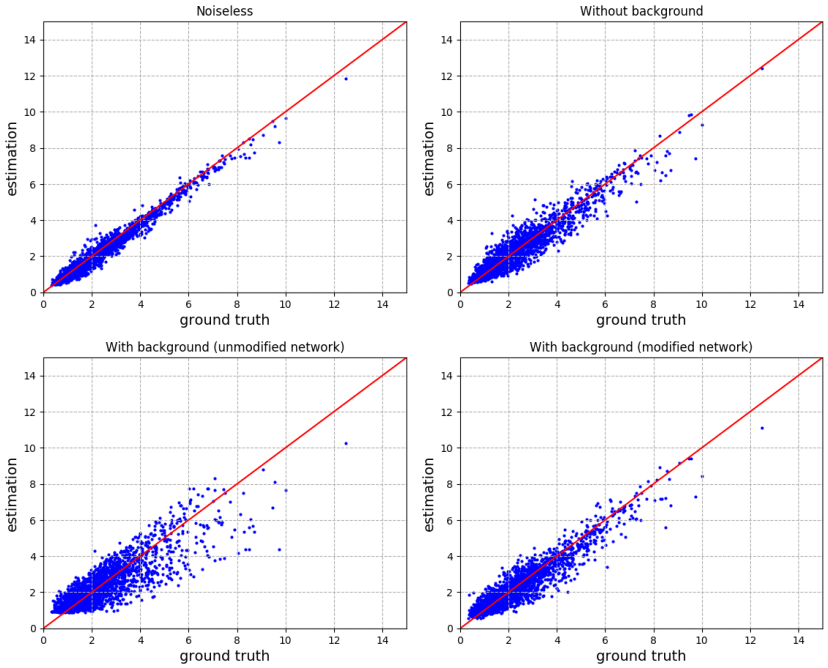


Fig. 5. Scatter plots of estimation vs. ground truth for ellipse area (mm^2) in an advanced phantom simulation. Top left is the noiseless case; top right is with noise but without 200-mm water background phantom; bottom left is with noise and with background phantom. Bottom right is the results with the same data, but the network partially reuses weights obtained from training with noiseless data.

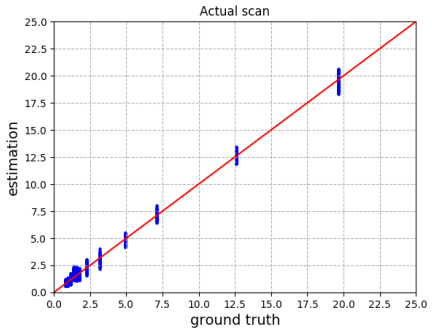


Fig. 6. Estimation vs. ground truth for circular cross section area (mm^2) using a custom-made vessel phantom. Twelve cylindrical holes with different diameters were filled with 30-mg/mL iodine-water solution, installed in a QRM anthropomorphic thorax phantom, and scanned by a GE 256-slice Revolution CT.

REFERENCES

[1] D.A. Chesler, S.J Riederer, and N.J. Pelc, "Noise due to photon counting statistics in computed X-ray tomography," in *J. Computer. Assisted. Tomography*, vol. 1, no.1, pp. 64-74, 1978

[2] S. Hsieh, D.A. Chesler, D. Fleischmann, and N. Pelc, "A limit on dose reduction possible with CT reconstruction algorithms without prior knowledge of the scan subject," in *Medical Physics*, vol. 43, no. 3, pp. 1361-1368, 2016

[3] R. De Man, G. Wang, M. K. Kalra, A. Otrakji, S. Hsieh and N. Pelc, "Upper-Bound on Dose Reduction in CT Reconstruction for Nodule Detection," in *IEEE Access*, vol. 4, pp. 4247-4253, 2016

[4] R. Mooser, F. Forsberg, E. Hack, G. Szekely, U. Sennhauser, "Estimation of affine transformations directly from tomographic projections in two and three dimensions", in *Machine vision and applications*, vol. 24, no. 2, pp. 419-434, 2013

[5] D. Ertel, T. Pflederer, M. Kachelrieß, S. Achenbach, P. Steffen, and W.A. Kalender, "Validation of a Rawdata-based Synchronization Signal (Kymogram) for a Phase-Correlated Cardiac Image Reconstruction," in *European Radiology*, vol 18, no. 2, pp.253-262, 2008

[6] G. Cybenko, "Approximation by superpositions of sigmoidal functions," in *Mathematics of Control, Signals, and Systems*, vol. 2 (4), 303-214, 1989

[7] Ba, J. and Caruana, "Do Deep Nets Really Need to be Deep?," in *Neural Information Processing Systems (NIPS)*, pp. 2654–2662. 2014

[8] A. Romero, N. Ballas, S. E. Kahou, A. Chassang, C. Gatta, and Y. Bengio, "FitNets: Hints for Thin Deep Nets," ICLR, 2015

[9] H. Lin, M. Tegmark, and D. Rolnick," Why does deep and cheap learning work so well ?", in *Journal of Statistical Physics*, vol. 168 (6), pp 1223-1247, 2017

[10] B. De Man, S. Basu, N. Chandra, B. Dunham, P. Edic, M. Iatrou, et al., "CATSIM: a new computer assisted tomography SIMulation environment," Medical Imaging 2007: Physics of Medical Imaging, Pts 1-3, vol. 6510, pp. U1191-U1198, 2007

Optimize interpolation-based MAR for practical dental CT with deep learning

Kaichao Liang, Li Zhang, Hongkai Yang, Yirong Yang, Yuxiang Xing

***Abstract*-Metal artefacts is common in diagnostic dental CT images. Due to the high attenuation of heavy matters such as metals, severe global artefacts occur in reconstructions. Typical metal artefact reduction (MAR) techniques segment out metal regions and estimate the corrupted projection data by various interpolation methods. To improve the inaccuracy of interpolation, we proposed a three-step MAR which is easy to implement: firstly, metal-related projection data are recognized and a preliminary image reconstruction is done using linearly interpolated data for metal-related regions; Secondly, use deep-learning network to remove artefacts from linear interpolation and recover non-metal information; Thirdly, add ROI reconstruction of metals. The advantage of this method is that the error in reconstructions from linear interpolation can be easily simulated and can be well generalized to real interpolation-based MAR results. The proposed method was validated with practical clinical dental CT data. The image quality was significantly improved with little metal artefacts left.**

I. Introduction

Metal artefacts reduction (MAR) is a persistent problem in modern clinical computed tomography (CT) imaging especially for dental CT. The high attenuation of metal components and beam-hardening effects introduce severe global streaking and shading artefacts so that significantly influence diagnosis. In last decades, various techniques have been proposed for MAR. Dual-energy CT can restrain beam-hardening artefacts [1]. One commonly used type of methods is to recognize metal projection and estimate it with interpolation or from forward projections [2, 3, 4]. Removing CT image artefacts by deep neural networks have be studied recently [5- 8]. For supervised learning, reference images are needed to form loss functions (usually L2-norm) in network training. In simulation, metal components are added to metal-free images to form labels. Then forward projection and reconstruction are implemented to acquire images with metal artefacts used as network inputs. It has been demonstrated that an image-domain U-Net structure can easily remove simulated metal artefacts and significantly improve image quality [7]. However, real metal artefacts are complicated by beam-hardening, scattering, high level of noise and other factors. It is difficult to generalize simulation results to real situation. Hence, we would like to find a new method to bridging simulation and real situation.

In this work, we proposed a new MAR method based on optimizing coarse reconstructions from interpolated non-metal projection data. A U-net structure was trained in image domain to remove interpolation artefacts. It took interpolation-reconstructed images as input. For training, corresponding metal-free images were used as labels. The trained U-Net significantly improved real metal interpolation-reconstructed image quality.

Kaichao Liang, Li Zhang, Kongkai Yang, Yuxinag Xing are members of Key Laboratory of Particle & Radiation Imaging and Department of Engineering Physics, Tsinghua University. Yirong Yang is member of Department of Engineering Physics, Tsinghua University. Yuxiang Xing is the corresponding writer (Xingyx@mail.tsinghua.edu.cn)..

II. Method

A. Interpolation-based MAR

In an X-ray CT, we use $\mathbf{p}=\mathbf{H}\boldsymbol{\mu}+\mathbf{n}$ to denote the discretized projection process with where $\boldsymbol{\mu} \in \mathbb{R}^{N\times 1}$ being a vector denoting a $\sqrt{N}\times\sqrt{N}$ image of effective attenuation map within the field of view (FOV). The matrix $\mathbf{H} \in \mathbb{R}^{M\times N}$ is a system matrix with its elements Hij descripting the contribution from pixel j to ith ray path, and n is a zero-mean noise vector. $\mathbf{p} \in \mathbb{R}^{M\times 1}$ is the acquired projection. When imaging an object with metals, we can separate the attenuation map into metal and non-metal components: $\boldsymbol{\mu}=\breve{\boldsymbol{\mu}}+\boldsymbol{\mu}^{\text{metal}}$. Correspondingly, we define $\breve{\mathbf{p}} \equiv \mathbf{H}\breve{\boldsymbol{\mu}}$ and $\mathbf{p}^{\text{metal}} \equiv \mathbf{H}\boldsymbol{\mu}^{\text{metal}}$. For interpolation-based MAR, a preliminary reconstruction image $\hat{\boldsymbol{\mu}}^{0}$ is firstly acquired by a simple analytic reconstruction. Next, metal components are segmented out by a threshold: $\hat{\boldsymbol{\mu}}^{\text{metal}}=\{\hat{\boldsymbol{\mu}}:\hat{\boldsymbol{\mu}}\geq \text{threshold}\}$. Forward projection of metal components $\hat{\mathbf{p}}^{\text{metal}}=\mathbf{H}\hat{\boldsymbol{\mu}}^{\text{metal}}$ is used to locate metal traces. Then, the metal projections are replaced by interpolation in projection domain:

$$\left[\hat{\breve{\mathbf{p}}}\right]_i=\begin{cases}\left[\mathbf{p}\right]_i, & \left[\hat{\mathbf{p}}^{\text{metal}}\right]_i=0\\ \varphi^{\text{interp}}(\mathbf{p}), & \left[\hat{\mathbf{p}}^{\text{metal}}\right]_i>0\end{cases}$$

Here, $\left[\mathrm{g}\right]_i$ denotes the ith element in a vector, and φ^{interp} the interpolation function. The preliminary estimation $\hat{\breve{\boldsymbol{\mu}}}^{0}$ are reconstructed from $\hat{\breve{\mathbf{p}}}$. Fig.1 illustrates the interpolation-based MAR process for a 2D fan-beam CT. Obviously, the $\hat{\breve{\boldsymbol{\mu}}}^{0}$ has severe artefact. Our method use deep learning method to get $\breve{\boldsymbol{\mu}}=\Phi^{\text{DL}}\left(\hat{\breve{\boldsymbol{\mu}}}^{0}\right)$. Here we use Φ^{DL} to denote the operation of deep learning defined by our network.

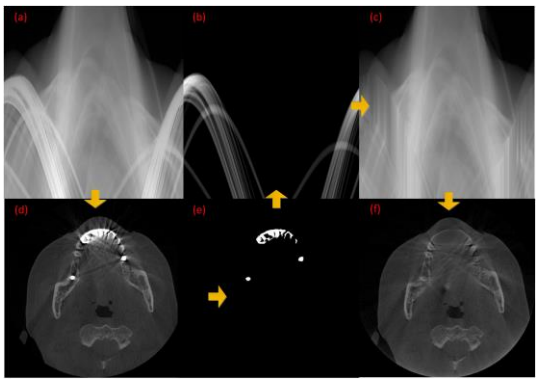


Fig.1 Interpolation-based MAR process with the orange arrows indicate the steps. (a) Acquired projection. (b) Forward projection of metal components from (e). (c) Interpolated projection. (d) FBP reconstruction from (a). (e) Segmented metal components. (f) FBP reconstruction from (c).

B. *U-Net structure*

Convolutional neural networks (CNN) is a commonly used type of deep neural networks in imaging field. It reduces numbers of parameters by weights-sharing [9]. U-Net is a net structure based on CNN. It was firstly proposed for segmentation task [10]. U-Net achieves bigger receptive field with down-sampling process, and recovers spatial resolution by concatenating un-pooled feature maps with up-sampled feature maps. Considering that error from interpolation in projection is non-local, we pick U-Net to be our network structure. Our U-Net is as in Fig. 2. We use four down sampling and four up sampling in the network structure with 18 convolutional layers in total. Adjacent convolutional layers are connected with active functions [11] and batch normalization layers [12].

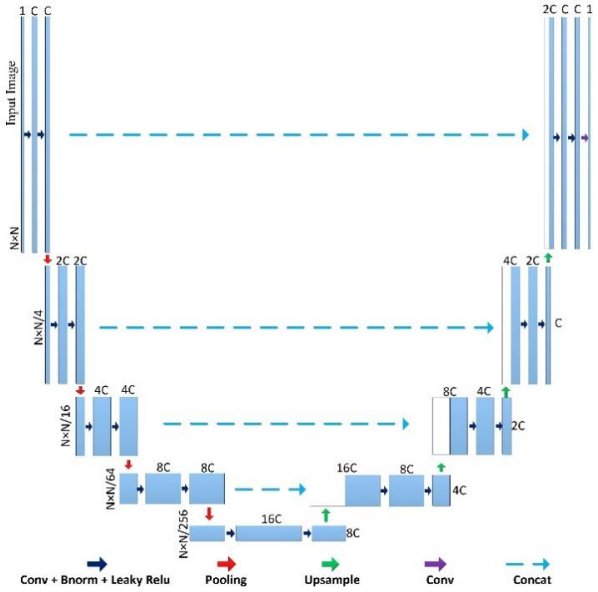


Fig. 2 U-Net structure.

C. *Preparation of training dataset and training*

For real CT images containing metal components, there are no ground-truth available. We simulated training dataset containing various artefacts from projection-domain interpolation. Firstly, metal-free images of high quality were prepared. Then, we randomly added metal on teeth and determine locations of metal projections. After that, linear interpolation was used to obtain metal projections. FBP reconstructions of these interpolated projections formed the simulated dataset of $\hat{\boldsymbol{\mu}}^0$. The high quality images before metal-insertion were used as labels. Please notice that the manually planted metal components were only used to locate metal traces so that only the planted metal shapes influence results. Fig. 3 shows some examples of metal insertions.

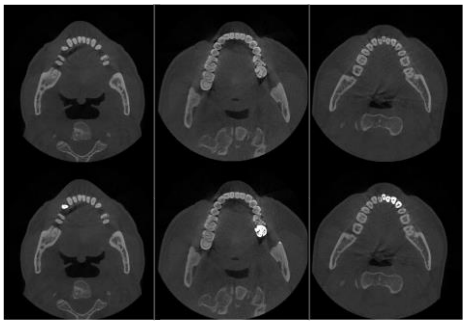


Fig. 3 Examples of high quality images without and with metal insertions. Top row: original meatal-free images. Bottom row: metal-inserted images.

We trained our network by minimizing L2-norm loss function between the outputs of U-Net and labels. Instead of minimizing the L2-norm loss over all pixels within the images, we ignored the differences in metal regions so that the network can focuses more on global artefacts from projection-domain interpolation rather than local metal shapes.

III. EXPERIMENT RESULTS

We acquired realistic projection data from a commercial Dental CT system. For our 2D experiment, we took metal-free CT images of 1383 patients. Each images were planted with metal components of different shapes to train our network. Among them, 1000 were used as training set, and the left were used as validation set. We also collected 1000 real data with metal components to test the trained networks. It took about 80 epochs for the network to converge. The error curves are showed in Fig. 4. After training, we test our network on real patient data with metal plants. Results are shown in Fig. 5. We demonstrated four cases with complicate components of metal. In all four cases, the streaking and shading artefacts from metal around metal area are significantly reduced. We can see that our simulation-data trained network can be well generalized to real cases.

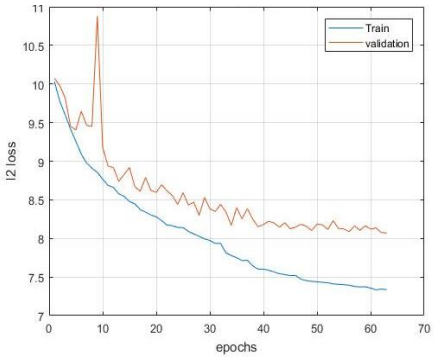


Fig. 4 Convergence of the loss function for the U-Net

IV. CONCLUSION

In this work, we proposed a strategy to build a transferable learning from simulation to practical system for metal artefact reduction. A supervised deep-learning method can thus be applied directly for practical usage. Our strategy is to transforms MAR problem to interpolation artefacts reduction problem so that simulation training data with ground-truth labels can easily model the same features with real data. The network significantly optimized both simulation data and real data. Our experimental results demonstrated that projection-domain interpolation + FBP + U-Net frame is an effective MAR method that can be used to practical dental CT. While we carried out our experiments in 2D, the method can be extended to 3D easily.

REFERENCES

[1] Bamberg, F., Dierks, A., Nikolaou, K., Reiser, M. F., Becker, C. R., & Johnson, T. R. (2011). Metal artifact reduction by dual energy computed tomography using monoenergetic extrapolation. European Radiology, 21(7), 1424-1429.
[2] Gjesteby, L., Man, B. D., Jin, Y., Paganetti, H., Verburg, J., & Giantsoudi, D., et al. (2016). Metal artifact reduction in CT: where are we after four decades?. IEEE Access, 4(99), 5826-5849.
[3] Kratz, B., Oehler, M., & Buzug, T. M. (2009). On limitations of 1d interpolation-based metal artefact reduction approaches – a comparison of fbp versus mlem. Ifmbe Proceedings, 25(2), 398-401.
[4] Mahnken, A. H., Raupach, R., Wildberger, J. E., Jung, B., Heussen, N., & Flohr, T. G., et al. (2003). A new algorithm for metal artifact reduction in computed tomography: in vitro and in vivo evaluation after total hip replacement. Investigative Radiology, 38(12), 769-775.
[5] Chen, H., Zhang, Y., Zhang, W., Liao, P., Li, K., Zhou, J., & Wang, G. (2016). Low-Dose CT via Deep Neural Network.
[6] Zhang, H., Li, L., Qiao, K., Wang, L., Yan, B., & Li, L., et al. (2016). Image prediction for limited-angle tomography via deep learning with convolutional neural network.
[7] Zhang, C., & Xing, Y. (2018). CT artifact reduction via U-net CNN. Medical Imaging 2018: Image Processing (Vol.10574). International Society for Optics and Photonics.
[8] Han, Y., & Ye, J. C. (2017). Framing u-net via deep convolutional framelets: application to sparse-view CT.
[9] Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324.
[10] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: convolutional networks for biomedical image segmentation. , 9351, 234-241.
[11] Ye, J. C., & Han, Y. S. (2017). Deep Convolutional Framelets: A General Deep Learning for Inverse Problems.
[12] Loffe, S., & Szegedy, C. (2015). Batch normalization: accelerating deep network training by reducing internal covariate shift. International Conference on International Conference on Machine Learnin, 448-456.

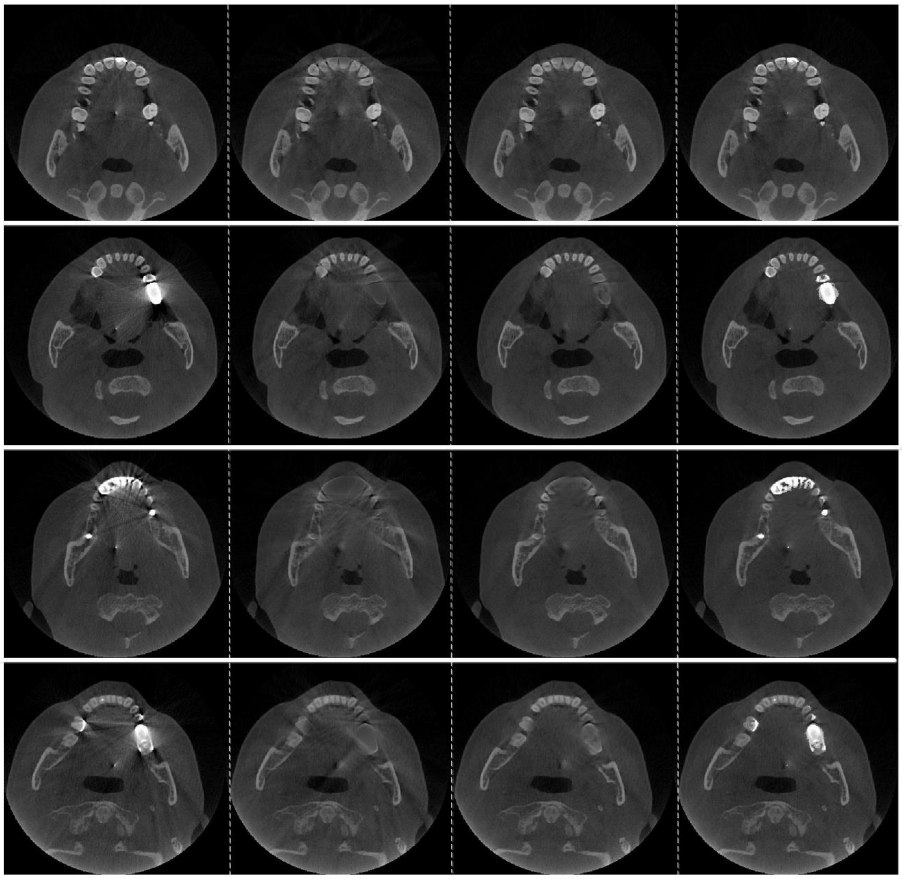


Fig. 5: Four cases of MAR with trained U-net. From left to right: Reconstructions with no MAR, MAR with interpolation, proposed Interpolation + U-Net，proposed Interpolation + U-Net together with metals.

Performance comparison of deep learning based denoising techniques in low-dose CT images

Byeongjoon Kim, Hyunjung Shim, Jongduk Baek

***Abstract*—Deep learning has become a promising tool for reducing noise in low-dose CT images. However, it has not been explored the dependency of the network modifications on denoising performance. In this work, we compare the denoising performance of various network structures and loss functions (i.e., L1, L2, and perceptual loss function). The network training was conducted using images from XCAT phantom data with different quantum noise levels, and tested using independently generated XCAT images. The results show that UNet with a perceptual loss function reduces the noise effectively while preserving details in ultra-low dose CT images.**

Keywords—low dose, computed tomography, deeper network, loss functions

I. INTRODUCTION

DURING the last decade, reducing a patient dose in CT scans has become an important issue, and thus there have been much efforts in developing techniques to improve image quality of low dose CT images. While employing iterative techniques [1-3], or adapting filtering based denoising techniques into sinogram [4,5] and reconstruction space [6,7] were effective to reduce CT noise, these techniques often introduced image blurs or required high computational costs.

Recently, convolutional neural networks (CNNs) have been shown the impressive performance for color image denoising. Upon this success, the initial result in low dose CT denoising using a 3-layer CNN was presented by Hu Chen et. al. [8]. While the 3-layer CNN shows promising results with improved computational efficiency and sharpness of the object structures, when the noise level is higher or the object structure is more complex, more powerful model using a deeper network or different loss functions would be required for effective denoising in low dose CT images.

The naïve approach to constructing a deeper network is to include additional layers on top of the 3-layer network, but this entails difficulty in network optimization, such as gradient vanishing. To implement the deeper network architecture without gradient vanishing problem, residual blocks have been introduced in object classification problem [9]. Residual blocks force the networks to learn residual mapping functions, effectively increasing the gradient during training such that the gradient vanishing can be alleviated. Alternative approach is to utilize the autoencoder architecture such as denoising autoencoder or UNet, which combines the autoencoder with skip connections and is widely used in bio-medical applications [10].

Choosing an appropriate loss function is also important to improve the quality of results. The effect of loss functions has been investigated in the problem of color image denoising [11]. However, this issue has not been discussed in the domain of medical imaging. In this work, we investigate the denoising performance in ultra-low dose CT images using a variant of deeper networks with various loss functions, and compare them with the 3-layer CNN.

The authors are with the School of Integrated Technology and Yonsei Institute of Convergence Technology, Yonsei University, Incheon 406-840 South Korea (e-mail: jongdukbaek@yonsei.ac.kr).

II. METHODS

A. Network Architecture

A-1. Deep Residual Network

The deep residual network consists of a single 3x3 convolution layer followed by six residual blocks and a single 1x1 convolution layer. The residual blocks are composed of batch normalization, ReLU and 3x3 convolution layers. Each residual block has a skip connection that adds output of a previous residual block to output of a present residual block. For each 3x3 convolution layer, 32 filters are learned. We denote this network as a DResNet as shown in Fig. 1(a).

A-2. UNet

The UNet in [10] was modified by replacing up-convolution with nearest neighbor interpolation in order to reduce checkerboard artifacts induced by deconvolution. Thus, the modified UNet is composed of 23 convolutional layers. As illustrated in Fig. 1(b), UNet includes 9 sub-unit, called a module, and each module with downsampling (or upsampling) would be the input to the next module. In addition, UNet also introduces a skip connection, which concatenates the output of encoder module to the input of the decoder module with the same scale. In this way, details from encoder layers can be directly transferred to the decoder, preventing information loss caused by down-sampling.

A-3. Loss functions

DResNet and UNet were trained with different loss functions: L1 loss, L2 loss, and the perceptual loss. Perceptual loss measures the quality of image using differences between feature representations instead of pixel

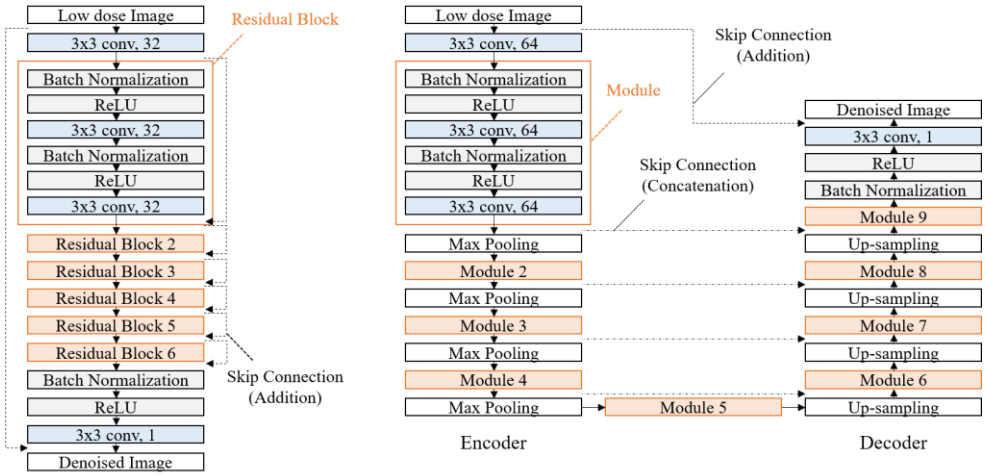


Fig. 1. Network architecture of (a) DResNet (left) and (b) UNet (right)

differences. This strategy has been successfully applied to image generation tasks, such as texture synthesis [12] and style transfer [13]. In this paper, the perceptual loss was calculated at the block3_1 layer of the VGG-16, denoted as VGG loss. The intermediate layer, block3_1, was chosen because calculating loss at the higher feature domain produces distortions in texture and object shape in denoising results [14], while the lower feature domain does not capture abstract information that helps considering semantically meaningful information. Because of pooling layers used in the VGGNet, our network with perceptual loss alone produces a checkerboard artifact in denoising results. To reduce this artifact, our networks were fine-tuned with L1 or L2 loss, denoted as VGG-L1 and VGG-L2, respectively.

A-4. Trained networks

We trained DResNet with L1 and L2 loss functions, each denoted by DResNet-L1, DResNet-L2. The UNet was trained with L1 and L2 loss functions, and perceptual loss: UNet-L1, UNet-L2, UNet-VGG, UNet-VGG-L1, UNet-VGG-L2. For comparison with shallow networks, the 3-layer CNN was also trained using L2, denoted as Hu-L2.

TABLE I. Simulation parameters

Parameters		Values
Phantom	Size	51.2 × 51.2 cm^2 (512 × 512)
	Pixel width	0.1 cm
Reconstructed Images	Size	43.52 × 43.52 cm^2 (256 × 256)
	Pixel width	0.17 cm
Geometry	Source-isocenter distance	50 cm
	Source-detector distance	50 cm
	Data acquisition angle	360°
	Number of views	1024
	Detector array length	102.4 cm (512 × 1)
	Detector cell width	0.2 cm
Photon	Energy	70 keV
	Blank scan flux	100, 300, 500, 1000

B. Dataset

The abdomen part of an XCAT phantom was used for data generation. 12 different patients were simulated by varying patient size and amounts of fat and muscle. For each patient, 50 abdomen slices were extracted, and projection data were acquired using Siddon's method [15]. To generate different noise levels, 100, 300, 500, and 1000 detected photons per detector cell (i.e., N_0) were used by assuming the ideal bowtie filter. The images were reconstructed by direct fan beam reconstruction with an equally spaced detector array [16]. The simulation parameters are summarized in Table I. The noise generation process was conducted twice independently for each projection data, which generates 4800 images in total. For the network training, 3600 images from 9 patients were used, and the rest 1800 images were used for test. A sample image is presented in Fig. 2.

C. Training Configurations

Patches with a size of 32x32 pixels were extracted from each image by sliding window at every 16 pixels. Also, the DResNet was trained with those of 32x32 patch while UNet was trained with a full-scale image. A batch size of DResNet was 256 while that of UNet was 1. UNet-VGG-L1

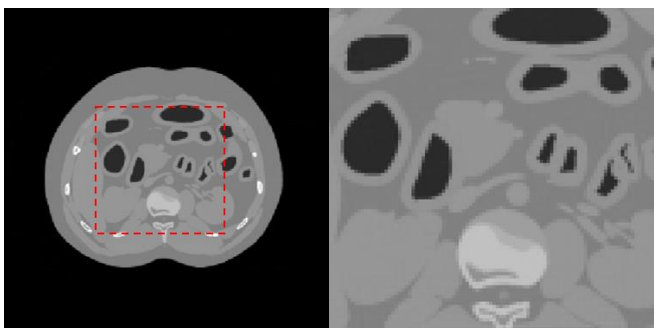


Fig. 2. (a) Sample abdomen CT image (left) and (b) region of interest extracted from the red dotted square (right).

and UNet-VGG-L2 were trained for 50 epochs and all other networks were trained for 40 epochs. The number of iterations was determined empirically by observing the convergence of the loss function. For the solver, we chose an Adam optimizer [17], and the learning rate initially to be 0.005, and gradually reduced to 0.0001. For each network, training time took 2 to 4 hours.

III. RESULTS

Fig. 3 and Fig. 4 compare the original FBP image, noiseless image, and denoised images from the various network structures with different noise levels (i.e. 100 and 1000 detected photons per detector cell cases). In Fig. 3, Hu-L2 loses details of the original images as shown by two red circles, but all other networks preserve the original details with reduced noise. When the noise level is high, denoising performance varies for different networks as shown in Fig. 4. It is observed that the denoising performance of Hu-L2 is worse than other networks. While the DResNet-L1 and DResNet-L2 eliminate the noise effectively, small features are also blurred out. UNet-L1 and UNet-L2 preserve details of the original structures, but the effect of using different loss function is reflected as preserving the sharpness of the original structures, including small features. UNet-VGG shows the checkboard artifacts as addressed earlier, which is removed effectively by UNet-VGG-L1 and UNet-VGG-L2 networks. Overall, UNet-VGG-L1 and UNet-VGG-L2 networks provide good denoising performance without losing small features of the

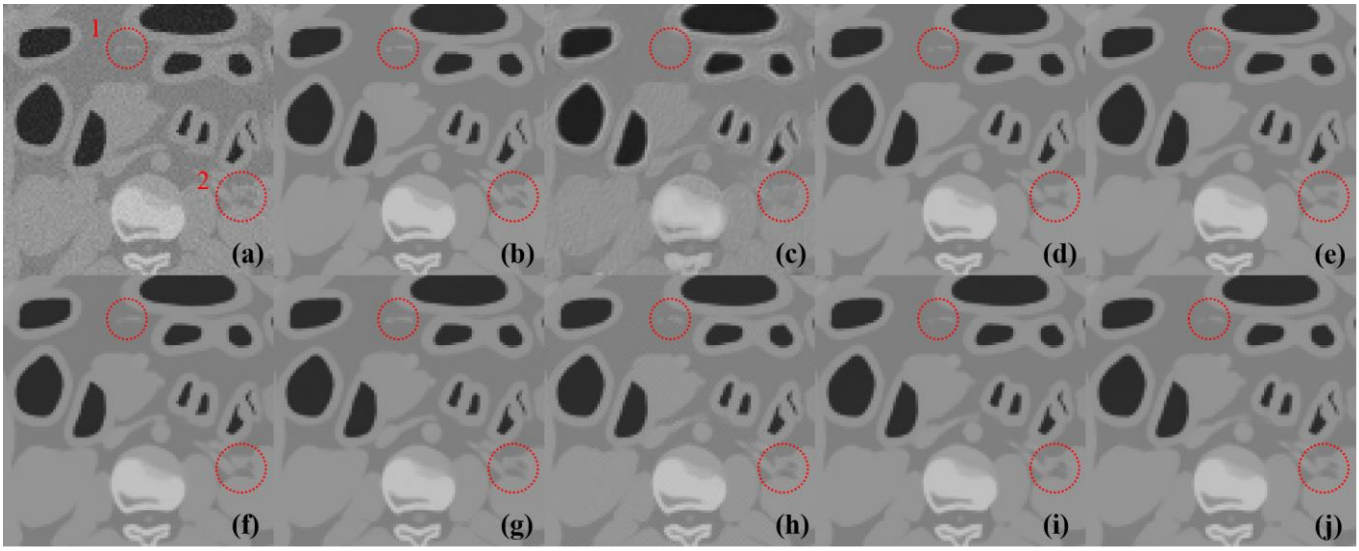


Fig. 3. The region of interest in (a) low dose reconstructed image (i.e., 1000 detected photons per detector cell), (b) noiseless reconstructed image, output of (c) Hu-L2, (d) DResNet-L1, (e) DResNet-L2, (f) UNet-L1, (g) UNet-L2, (h) UNet-VGG, (i) UNet-VGG-L1, and (j) UNet-VGG-L2. The red dotted circles indicate details in interest.

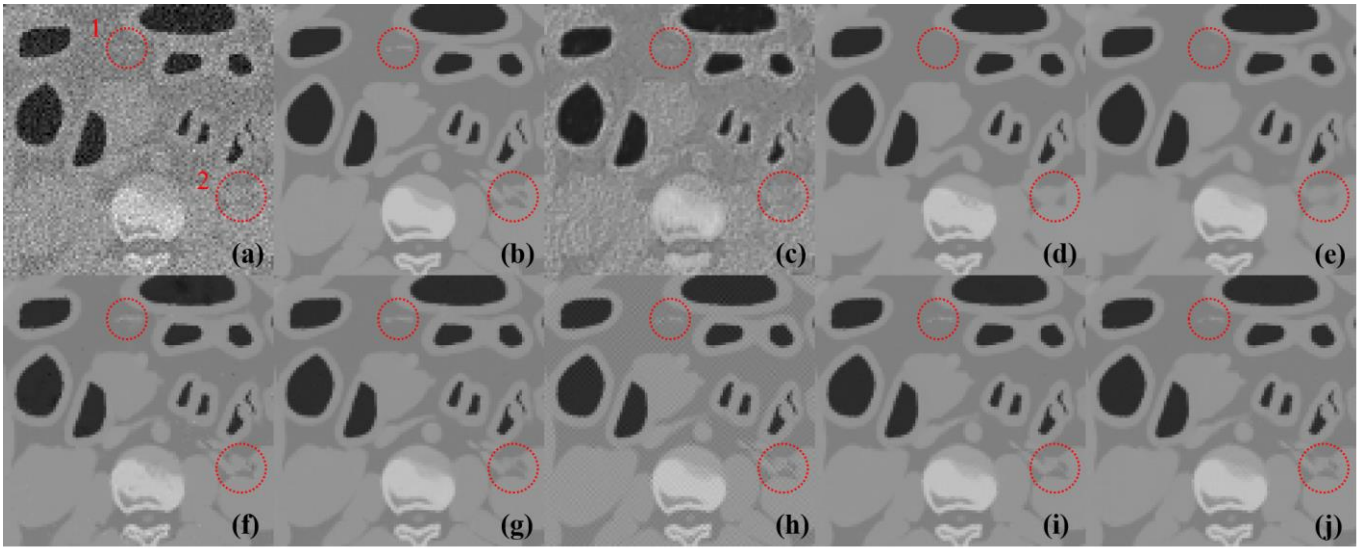


Fig. 4. The region of interest in (a) ultra-low dose reconstructed image (i.e., 100 detected photons per detector cell), (b) noiseless reconstructed image, output of (c) Hu-L2, (d) DResNet-L1, (e) DResNet-L2, (f) UNet-L1, (g) UNet-L2, (h) UNet-VGG, (i) UNet-VGG-L1, and (j) UNet-VGG-L2. The red dotted circles indicate details in interest.

TABLE II. Results of quantitative evaluation for test images with different noise levels (N_0).

	Input		Hu-L2		DResNet-L1		DResNet-L2		UNet-L1		UNet-L2		UNet-VGG-L1		UNet-VGG-L2	
N_0	1000	100	1000	100	1000	100	1000	100	1000	100	1000	100	1000	100	1000	100
RMSE (1e-3)	6.37 ±0.05	20.27 ±0.16	8.64 ±0.69	13.27 ±0.70	2.21 ±0.11	4.80 ±0.26	2.11 ±0.09	4.50 ±0.24	3.54 ±0.27	5.40 ±0.24	3.88 ±0.53	4.74 ±0.56	3.86 ±0.31	4.57 ±0.36	3.64 ±0.31	**4.45 ±0.36**
PSNR	43.92 ±0.06	33.86 ±0.07	41.30 ±0.69	37.56 ±0.46	53.12 ±0.43	46.39 ±0.47	53.51 ±0.37	46.96 ±0.47	49.04 ±0.65	45.36 ±0.39	48.30 ±1.09	46.54 ±0.96	48.31 ±0.70	46.84 ±0.67	48.80 ±0.74	**47.06 ±0.70**
SSIM	0.9692 ± 0.0009	0.7670 ± 0.0054	0.9755 ± 0.0027	0.9015 ± 0.0085	0.9974 ± 0.0003	0.9891 ± 0.0013	0.9977 ± 0.0002	0.9904 ± 0.0011	0.9947 ± 0.0015	0.9876 ± 0.0015	0.9940 ± 0.0023	0.9913 ± 0.0027	0.9940 ± 0.0012	0.9919 ± 0.0014	0.9945 ± 0.0011	**0.9921 ± 0.0013**

original image.

Table II summarizes quantitative results (i.e., root mean square error (RMSE), peak signal to noise ratio (PSNR), and structural similarity (SSIM) index) of denoising performance with different network structures. It is observed that UNet-VGG-L1 and UNet-VGG-L2 networks provide the highest denoising performance for both low dose and ultra-low dose CT images. Note that UNet-VGG was not considered in the quantitative evaluation due to the checkerboard artifacts. Our results show that UNet shows better denoising performance than DResNet, which would be attributed to its larger receptive field and various relative kernel sizes by pooling layers. Since the CT noise is highly correlated, using a larger receptive field in UNet can be more effective in denoising of CT images.

IV. DISCUSSION AND CONCLUSION

In this work, we investigate the denoising performance in low dose and ultra-low dose CT images using various network structures and loss functions of deep neural network. Our results showed the strength of the deeper network with a perceptual loss function for denoising of ultra-low dose CT images. While we used simple image quality metrics (i.e., RMSE, PSNR, SSIM) to evaluate the network performance, it would be more appropriate to use mathematical observer models due to the non-linear nature of the deep neural network. In the future, we plan to conduct a task-based assessment using mathematical observer models for each network.

ACKNOWLEDGMENT

This research was supported by Ministry of Science, ICT and Future Planning (IITP-2017-2017-0-01015) and National Research Foundation of Korea (2017M2A2A4A01070302, 2017M2A2A6A01019663). The authors would like to thank Dr. Cynthia McCollough, the Mayo Clinic, the American Association of Physicists in Medicine, and grants EB017095 and EB017185 from the National Institute of Biomedical Imaging and Bioengineering.

REFERENCES

[1] Thibault, Jean-Baptiste, et al. "A three-dimensional statistical approach to improved image quality for multislice helical CT." Medical physics 34.11 (2007): 4526-4544.

[2] Sidky, Emil Y., and Xiaochuan Pan. "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization." Physics in medicine and biology 53.17 (2008): 4777.

[3] Xu, Qiong, et al. "Low-dose X-ray CT reconstruction via dictionary learning." IEEE Transactions on Medical Imaging 31.9 (2012): 1682-1697.

[4] Manduca, Armando, et al. "Projection space denoising with bilateral filtering and CT noise modeling for dose reduction in CT." Medical physics 36.11 (2009): 4911-4919.

[5] Balda, Michael, Joachim Hornegger, and Bjoern Heismann. "Ray contribution masks for structure adaptive sinogram filtering." IEEE transactions on medical imaging 31.6 (2012): 1228-1239.

[6] Li, Zhoubo, et al. "Adaptive nonlocal means filtering based on local noise level for CT denoising." Medical physics 41.1 (2014).

[7] Chen, Yang, et al. "Improving abdomen tumor low-dose CT images using a fast dictionary learning based processing." Physics in medicine and biology 58.16 (2013): 5803.

[8] Chen, Hu, et al. "Low-dose CT via convolutional neural network." Biomedical optics express 8.2 (2017): 679-694.

[9] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[10] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "UNet: Convolutional networks for biomedical image segmentation." International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2015.

[11] Zhao, Hang, et al. "Loss functions for image restoration with neural networks." IEEE Transactions on Computational Imaging 3.1 (2017): 47-57

[12] Yosinski, Jason, et al. "Understanding neural networks through deep visualization." arXiv preprint arXiv:1506.06579 (2015).

[13] Gatys, Leon, Alexander S. Ecker, and Matthias Bethge. "Texture synthesis using convolutional neural networks." Advances in Neural Information Processing Systems. 2015.

[14] Johnson, Justin, Alexandre Alahi, and Li Fei-Fei. "Perceptual losses for real-time style transfer and super-resolution." European Conference on Computer Vision. Springer International Publishing, 2016.

[15] Siddon, Robert L. "Fast calculation of the exact radiological path for a three-dimensional CT array." Medical physics 12.2 (1985): 252-255.

[16] Hsieh, Jiang. "Computed tomography: principles, design, artifacts, and recent advances." Bellingham, WA: SPIE, 2009.

[17] Kingma, Diederik, and Jimmy Ba. "Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980 (2014)

Variational network learning for low-dose CT

Erich Kobler, Matthew J. Muckley, Baiyu Chen, Florian Knoll, Kerstin Hammernik,
Thomas Pock, Daniel K. Sodickson, and Ricardo Otazo

***Abstract*—The reconstruction of high quality images from low-dose X-ray CT scans data is a topic of significant technical and clinical relevance. In this paper, we develop learning-based variational networks (VNs) to reconstruct low-dose 3D helical CT data. We consider two dose reduction methods: (1) x-ray tube current reduction and (2) x-ray beam interruption also known as SparseCT. In the first case we train a VN to denoise the current-reduced reconstruction to account for the smaller signal-to-noise ratio, whereas, in the second case the VNs learn reconstruction schemes that suppress undersampling artifacts. We use 4 clinical abdominal 3D scans to train VNs for 4-fold dose reduction and compare against state-of-the-art model-based denoising and sparse reconstruction methods on a 5th clinical abdominal test scan. The proposed VNs improve performance over state-of-the-art iterative model-based denoising and sparse reconstruction techniques. VNs for SparseCT compare favorably to VNs for current reduction, particularly for reconstruction of small low-contrast features.**

***Index Terms*—CT image reconstruction, SparseCT, low-dose CT, compressed sensing, machine learning, variational networks**

I. Introduction

The increasing utilization of CT scanners in clinical imaging examinations, has triggered the need to reduce the radiation dose, particularly for recurrent studies. One of the most common approaches is to reduce the tube current, e.g., tube current modulation [1], or lower tube currents in conjunction with iterative model-based denoising methods [2]. These techniques have been successfully integrated in commercial scanners, but they only offer moderate radiation dose reductions of 30-40% in practice, due to compromises between denoising and smoothing.

The radiation dose can also be mitigated without reducing the tube current by decreasing the number of X-rays that penetrate a patient during a CT scan. The compressed sensing (CS) theory [3] supports this approach, since CT images are compressible in a transform domain and reducing the number of X-ray projections results in small additive incoherent streaking artifacts. A simple way to omit projections is to perform angular undersampling, i.e., just acquire projections for a fraction of the angular views, as proposed by [4]. The SparseCT method [5] extended this idea by blocking a subset of X-rays in an incoherent way across the angular and slice dimensions, which divides the overall undersampling along multiple dimensions and thus increases the perfromance of CS for reconstruction of the whole volume.

Recent low-dose CT reconstruction algorithms for low-current and/or undersampled data are typically model-based iterative methods that incorporate prior knowledge to increase image quality. These prior models are typically rather simple and model just a small subset of the CT image statistics, e.g., the popular total variation (TV) prior enforces sparsity in the image gradient domain. In addition, the balance between a regularizing prior term and a data fidelity term has to be empirically tuned to generate suitable reconstructions. In accelerated magnetic resonance imaging, deep learning was introduced to overcome this empirical tuning and to learn image models that are tailored towards medical imaging, demonstrating significant improvements over standard compressed sensing algorithms [6]. Likewise, recent work on deep learning for low-dose CT demonstrated improved performance compared to standard denoising and sparse reconstructions [7]–[9]. The U-net-like structures of [8] and [9] as well as the residual encoding network of [7] learn a mapping from low-dose filtered back-projection images to reference images that encodes and decodes the relevant information, in contrast to the step-wise refinement structure of [6].

In this work, we propose to learn variational networks for low-dose CT data acquired with tube current reduction and SparseCT. We train the VNs on four clinical abdominal data sets and evaluate the reconstruction quality of the proposed VNs on a test data set and compare it to state-of-the-art model-based reconstructions.

II. Model-based CT Reconstruction

The process of acquiring CT data of a volume $u \in \mathbb{R}^{M\times N\times D}$ can be formalized as

$$d = Au + n\,, \tag{1}$$

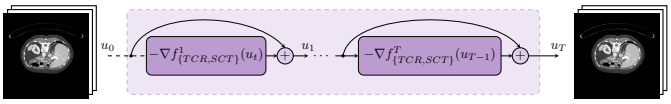
where $d \in \mathbb{R}^P$ is the post-log measured data of P X-ray projections. The random variable n models the effects of quantum and electronic noise and is assumed to be Gaussian due to preprocessing. The linear forward operator $A : \mathbb{R}^{M\times N\times D} \mapsto \mathbb{R}^P$ implements the mapping from the volume to the measurement data that is defined by the scanner geometry. For SparseCT A additionally implements the undersampling pattern.

*Erich Kobler is with the Institute of Computer Graphics and Vision, Graz University of Technology, Graz, Austria (e-mail: erich.kobler@icg.tugraz.at).

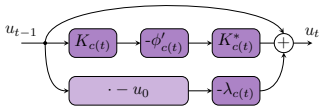
Matthew J. Muckley, Baiyu Chen, Florian Knoll, Daniel K. Sodickson and Ricardo Otazo are with the New York University School of Medicine, New York, NY, 10016 USA.

Kerstin Hammernik and Thomas Pock are with the Institute of Computer Graphics and Vision, Graz University of Technology, Graz, Austria.

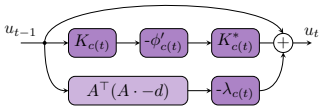
We acknowledge support from the Austrian Science Fund (FWF) under the START project BIVISION, No. Y729, the European Research Council under the Horizon 2020 program, ERC starting grant HOMOVIS, No. 640156. and the National Institutes of Health (NIH) grants U01-EB018760 and P41-EB017183.



(a) Variational Network (VN) structure for CT



(b) VU for CT denoising



(c) VU for CT reconstruction

Fig. 1. (a) Illustration of the VN for CT and the variational units (VU) for (b) CT denoising and (c) CT reconstruction.

For a given noisy and possibly undersampled CT scan data d, the inverse problem of recovering the volume u is usually defined by a variational minimization problem such as

$$\min_u F(u) := \beta\|\nabla u\|_1 + \frac{1}{2}\|Au - d\|_2^2 \,. \tag{2}$$

Here the scalar $\beta \geq 0$ is used to balance the solution between smoothness, which is enforced by the total variation (TV), i.e., ℓ_1-norm of the image gradients, and data fidelity. A suitable algorithm to solve (2) is the primal-dual approach with line search [11], since it requires just a view evaluations of the operator A that are computationally expensive.

III. Variational Networks for CT

Typical optimization schemes for variational imaging models, such as (2), can be implemented using convolutional networks. This observation, inspired [12] to train all parameters of a gradient descent scheme for variational image reconstruction models, i. e., analysis operators, potential functions, weighting and step sizes, from data. Variational networks (VNs) [13] connect this scheme, convolutional neural networks and variational minimization. To adapt VNs for CT, we apply fields-of-experts-type priors [14] of the form

$$R_c(u) = \langle 1, \phi_c(K_c u; W_c)\rangle \tag{3}$$

that are parameterized by a convolution operator $K_c : \mathbb{R}^{M\times N\times D} \mapsto \mathbb{R}^{M\times N\times D\times N_k}$, which stacks N_k 3D convolutions $K_c^i : \mathbb{R}^{M\times N\times D} \mapsto \mathbb{R}^{M\times N\times D}$, and corresponding potential functions $\phi_c^i(\cdot; w_c^i) : \mathbb{R} \mapsto \mathbb{R}$. These functions are point-wisely applied to the corresponding filter response and are parameterized by the weights $w_c^i \in \mathbb{R}^{N_w}$. For the sake of simplicity, we group all these functions into $\phi_c(\cdot, W_c)$ and their parameters $(w_c^i)_{i=1}^{N_k}$ into W_c.

We use this prior model to construct a variational energy that fits into the VN framework [13] and define it as

$$F_{\{TCR,SCT\}} := \sum_{c=1}^{C} f^c_{\{TCR,SCT\}}(u) \tag{4}$$

$$f^c_{\{TCR,SCT\}}(u) = R_c(u) + \frac{\lambda_c}{2} D_{\{TCR,SCT\}}(u) \,, \tag{5}$$

where the data term $D_{\{TCR,SCT\}}(u)$ is adapted according to the dose-reduction approach. In the case of tube current reduction (TCR) we learn to denoise initial low-dose reconstruction, hence we use a simple ℓ_2-norm denoising data term

$$D_{TCR}(u) = \|u - u_0\|_2^2 \,. \tag{6}$$

In the case of SparseCT (SCT) we use the forward operator A and the undersampled data d to enforce data consistency to the undersampled data and facilitate the reconstruction scheme

$$D_{SCT}(u) = \|Au - d\|_2^2 \,. \tag{7}$$

For both low-dose VNs we use a cyclic component selection function, i.e., $c(t) = \mod(t, C)$, and follow [13] to define a variational unit (VU) as

$$u_t = u_{t-1} - \nabla f^{c(t)}_{\{TCR,SCT\}}(u_{t-1}) \,, \tag{8}$$

where the gradients of the energy components are given by

$$\nabla f^c_{TCR}(u) = K_c^* \phi_c'(K_c u; W_c) + \lambda_c(u - u_0) \tag{9}$$

$$\nabla f^c_{SCT}(u) = K_c^* \phi_c'(K_c u; W_c) + \lambda_c A^\top(Au - d) \,. \tag{10}$$

The adjoint operator of K_c is denoted as K_c^* and it is defined as a convolution with all 180 rotated filter kernels followed by a point-wise summation. Figure 1 illustrates the computation outline of a VN for low-dose CT. The input u_0 is transformed into the output u_T by applying T steps of the form (8).

A. Training of VNs for CT

To train a VN for a set of training samples $(u_0^s, u_{tar}^s)_{s=1}^S$, we minimize the problem

$$\min_{\theta\in\mathcal{T}} \frac{1}{2}\sum_{s=1}^{S} \|b^s \odot (u_T^s - u_{tar}^s)\|_2^2 \,, \tag{11}$$

where $\theta = \{W_c, K_c, \lambda_c,\, c = 1 \dots C\}$ holds all the parameters of the VN. As [13], we constrain the parameters to an admissible set $\mathcal{T}$ that enforces $\lambda_c \geq 0$ and that each convolution filter has zero-mean and its ℓ_2-norm lies on the unit ball. We are only interested in reconstructing the central scan regions because of the missing ray density at border regions. Thus we apply a binary mask $b^s \in \{0,1\}^{M\times N\times D}$ that selects the 9 central slices where $u_{tar} \in [0,1]$ and $\odot$ indicates a point-wise multiplication. Note that we rescaled the images such that the HU interval $[-200, 280]$ is mapped to $[0,1]$ to ease training and account for the desired HU range. We solve the constrained training problem (11) by using the Adam optimizer [15] extended by an additional back projection step onto $\mathcal{T}$ after each gradient step. We perform 1000 gradient steps using the default moments of the Adam optimizer and a step size of 1×10^{-2}.

B. Experimental Setup

For the reconstruction of low-dose CT data we apply $T = C = 10$ variational units and use $N_k = 32$ convolution filters of size $11 \times 11 \times 3$ and their corresponding activation functions are parameterized by $N_w = 31$ Gaussian radial basis functions. We scaled the volumes for both training and test data such that the interesting Hounsfield unit interval $[-200, 280]$ is mapped onto $[0,1]$ to ease the training of the parameters. We use 8 filter-function-pairs that are defined

TABLE I
TRAINING AND TEST DATA SETS

	reference tube current mAs	tube voltage kV	radiation dose CTDIvol	gantry rotations -
train	240	120	21.19	16
	240	120	19.01	26
	240	120	22.26	19
	350	120	29.63	20
test	320	100	12.90	17

on the interval $[-4, 4]$ to regularized the entire HU range, whereas, the remaining 24 filter-function-pairs are defined on $[-1, 1]$ to account for the details in the desired tissue interval. In total 126,090 parameters were trained for each VN.

We used four clinical 3D in vivo abdominal CT scans of different patients of a Siemens Definition AS scanner. Table I shows acquisition properties of the train and test scans. In order to fit the CT data reconstruction onto a single GPU, we split the data of each CT scan after a full gantry rotation and ended up with 81 batches for training and 17 test samples. For every sample we reconstructed an imaged volume of size $384 \times 384 \times 30$. The target volumes u_{tar}^s were computed by solving (2) with $\beta = 1$ using [11] on the full-dose CT data. Likewise, the initial reconstructions u_0^s were generated with $\beta = 1 \times 10^{-9}$ using either simulated fully-sampled low-dose data [16] or binary subsampled full-dose data for SCT. We apply the same W1S4 undersampling pattern as in [17] for a 4-fold dose reduction.

IV. RESULTS

We used the test data set to evaluate the reconstruction quality of the learned VNs for both TCR and SCT for 4-fold radiation dose reduction. Table II depicts a quantitative evaluation of the root mean squared error (RSME) of the proposed VNs and state-of-the-art model-based denoising and reconstruction approaches. In Fig. 2, we qualitatively compare representative abdominal slices reconstructed by the proposed VNs to the full-dose reference, SAFIRE [2] and TV reconstruction.

In the case of tube current reduction, the proposed VN for TCR outperforms SAFIRE [2] in terms of RMSE and also in reconstruction quality. The VN presents a higher noise reduction of the imaged volume, while keeping the fine structure of the vessels in the liver. The resulting images are slightly smoothed though. Since SAFIRE applies an edge-enhancing kernel to highlight edges in the reconstructions, we removed the skin region from the binary mask b in the evaluation process to perform a fair comparison. Fig. 3 depicts the difference to a corresponding reference slice for the considered methods. Clearly, SAFIRE yields higher differences at edge regions but also the remaining regions are rather noisy.

In the case of SparseCT, the trained VN yields a lower RMSE than the TV model-based reconstruction using 4-fold undersampled test data. The VN for SCT removes the aliasing artifacts better than the TV reconstruction, while maintaining the fine vessels in the liver. Moreover, the reconstructions of the VN for SCT present more details than those of the VN for TCR and are also sharper, highlighting the advantages of SparseCT over tube current reduction for the same dose reduction factor. In addition, the reconstructions of a VN for SCT using 6-fold undersampling are shown on the right in Fig. 2. Despite the increased dose reduction, the VN for SCT is able to reconstruct the fine details and remove aliasing artifacts and yield reconstructions with a similar quality.

TABLE II
QUANTITATIVE COMPARISON OF THE DIFFERENT 1/4-DOSE CT METHODS BY MEANS OF RMSE TO THE TARGET u_{tar}, MEASURED IN HU.

SAFIRE [2]	TV	VN TCR	VN SCT
17.75 ± 2.11	8.84 ± 1.20	7.91 ± 0.90	7.72 ± 0.82

V. CONCLUSION

In this work, we extended variational networks to reconstruct CT volumes from low-dose data. We learned VNs for two popular radiation dose reduction methods, namely tube current reduction and SparseCT. The proposed VNs yield reconstructions that outperform state-of-the-art denoising and sparse reconstruction methods for low-dose CT. The VNs present a higher noise and artifact reduction, while fine details such as vessels are properly reconstructed. The learned reconstructions for undersampled data (SCT) show more details and are sharper than the learned denoising scheme for reduced-current data (TCR). Our experiments suggest that the proposed VNs increase the image quality for a given radiation dose and would enable higher radiation dose reductions. Future work includes the extension of the binary undersampling masks of SparseCT to more realistic undersamling masks as in [18]. Additionally, we work on speeding up the training and reconstruction process by means of ordered-subsets.

REFERENCES

[1] Cynthia H. McCollough, Michael R. Bruesewitz, and James JM. Kofler, "CT Dose Reduction and Dose Management Tools: Overview of Available Options," *RadioGraphics*, vol. 26, no. 2, pp. 503–512, 2006.

[2] Katharine Grant and Rainer Raupach, "SAFIRE: Sinogram Affirmed Iterative Reconstruction," *Siemens Medical Solutions Whitepaper*, 2012.

[3] Emmanuel J Candès, Justin Romberg, and Terence Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[4] Guang-Hong Chen, Jie Tang, and Shuai Leng, "Prior image constrained compressed sensing (PICCS): a method to accurately reconstruct dynamic CT images from highly undersampled projection data sets," *Medical Physics*, vol. 35, no. 2, pp. 660–663, 2008.

[5] Thomas Koesters, Florian Knoll, Aaron Sodickson, Daniel K. Sodickson, and Ricardo Otazo, "SparseCT: Interrupted-beam acquisition and sparse reconstruction for radiation dose reduction," *Proc.SPIE*, vol. 10132, pp. 10132 – 10132 – 7, 2017.

[6] Kerstin Hammernik, Teresa Klatzer, Erich Kobler, Michael P Recht, Daniel K Sodickson, Thomas Pock, and Florian Knoll, "Learning a Variational Network for Reconstruction of Accelerated MRI Data," *Magnetic Resonance in Medicine*, 2017, in press.

[7] Hu Chen, Yi Zhang, Jiliu Zhou, and Ge Wang, "Deep learning for low-dose ct," in *Developments in X-Ray Tomography XI*. International Society for Optics and Photonics, 2017, vol. 10391, p. 103910I.

[8] Kyong Hwan Jin, Michael T McCann, Emmanuel Froustey, and Michael Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Transactions on Image Processing*, vol. 26, no. 9, pp. 4509–4522, 2017.

[9] Han Yo Seob and Jong Chul Ye, "Deep Residual Learning Approach for Sparse-view CT Reconstruction," in *Proceedings of the International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, 2017.

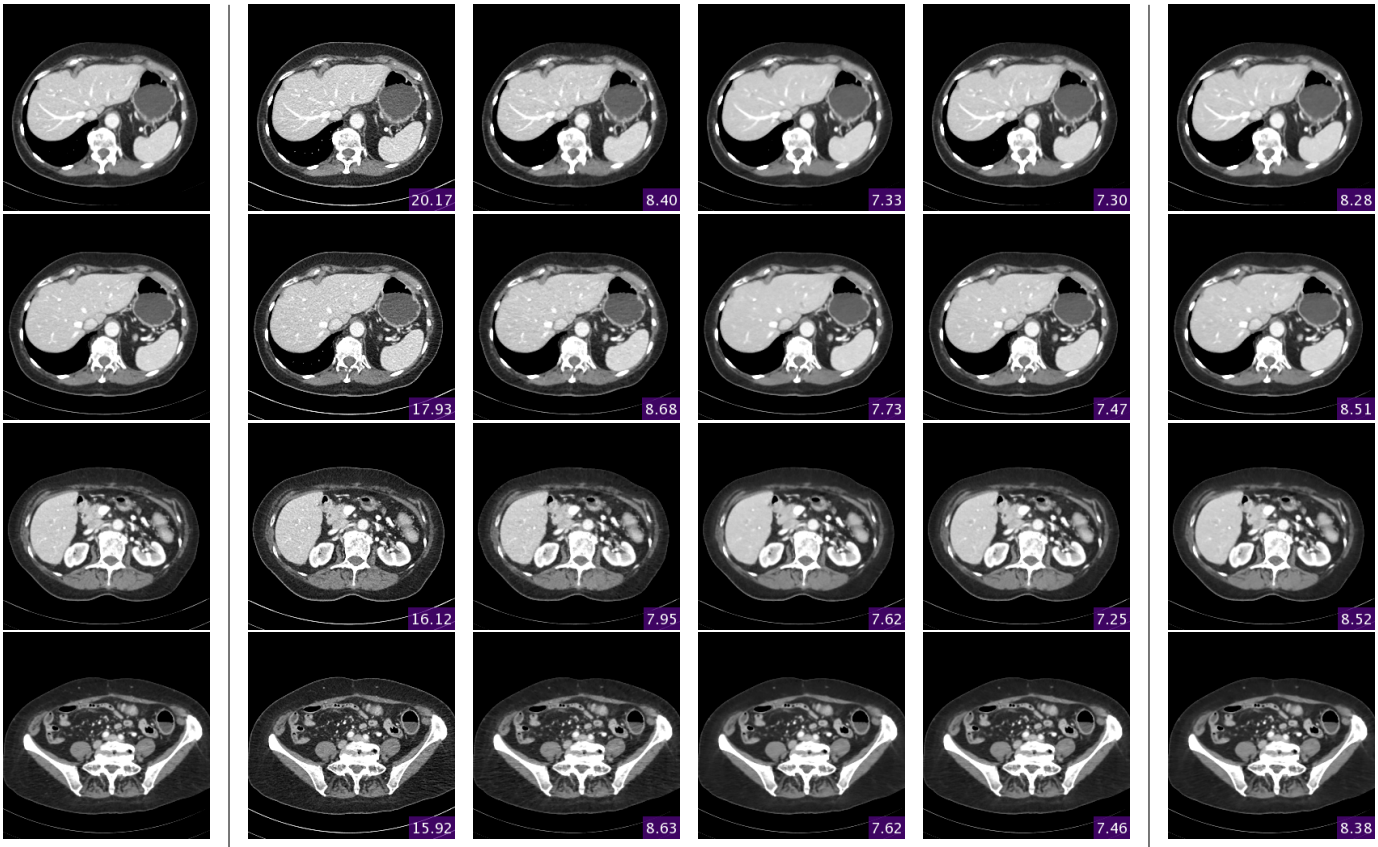


Fig. 2. Representative slices for reconstruction of in vivo abdominal test data for low-dose CT. The purple boxes report RMSE values. (a) Target: TV ($\beta = 1$) reconstruction of the fully-sampled high dose data, (b) SAFIRE [2] using $1/4$ dose, (c) TV ($\beta = 1.75$) reconstruction using 4-fold undersampling, (d) VN for TCR reconstruction using $T = 10$ steps and $1/4$ dose, (e) VN for SCT reconstruction using $T = 10$ steps and 4-fold undersampling, and (f) VN for SCT reconstruction using $T = 10$ steps and 6-fold undersampling.

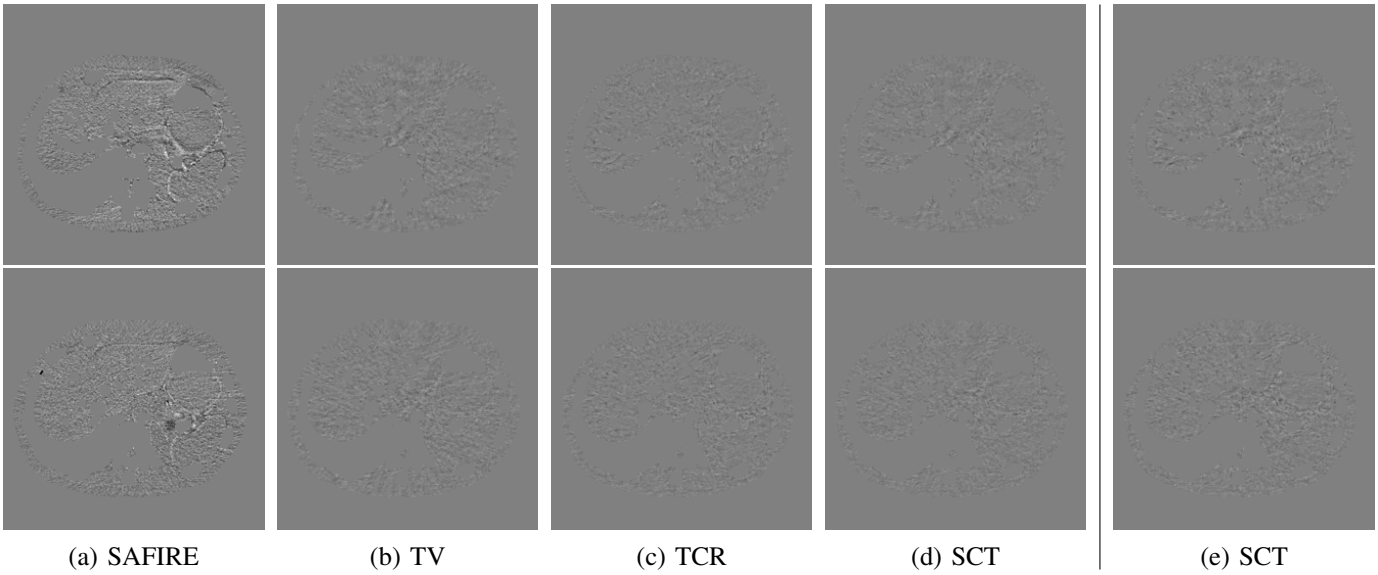


Fig. 3. Error to the reference reconstruction u_{tar} for the first two slices presented in Fig. 2. (a) SAFIRE [2] using $1/4$ dose, (b) TV ($\beta = 1.75$) reconstruction using 4-fold undersampling, (c) VN for TCR reconstruction using $T = 10$ steps and $1/4$ dose, (d) VN for SCT reconstruction using $T = 10$ steps and 4-fold undersampling, and (e) VN for SCT reconstruction using $T = 10$ steps and 6-fold undersampling. Note that we mapped the HU interval [-150, 150] to [0, 1] to ease visualization.

[10] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2015, pp. 234–241.
[11] Yura Malitsky and Thomas Pock, "A first-order primal-dual algorithm with linesearch," *arXiv preprint arXiv:1608.08883*, 2016.
[12] Yunjin Chen and Thomas Pock, "Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1256–1272, 2017.
[13] Erich Kobler, Teresa Klatzer, Kerstin Hammernik, and Thomas Pock, "Variational Networks: Connecting Variational Methods and Deep Learning," in *German Conference on Pattern Recognition*, 2017.
[14] Stefan Roth and Michael J Black, "Fields of Experts," *International Journal of Computer Vision*, vol. 82, no. 2, pp. 205–229, 2009.
[15] Diederik P. Kingma and Jimmy Lei Ba, "Adam: A method for stochastic optimization," in *International Conference on Learning Representations*, 2015.
[16] Lifeng Yu, Maria Shiung, Dayna Jondal, and Cynthia H McCollough, "Development and validation of a practical lower-dose-simulation tool for optimizing computed tomography scan protocols," *Journal of Computer assisted Tomography*, vol. 36, no. 4, pp. 477–487, 2012.
[17] Matthew Muckley, Baiyu Chen, Thomas Vahle, Aaron Sodickson, Florian Knoll, Danieln Sodickso, and Ricardo Otazo, "Regularizer performance for sparsect image reconstruction with practical subsampling," in *The 14th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, 2017, vol. 14.
[18] Baiyu Chen, M. Muckley, T. ODonnell, A. Sodickson, T. Flohr, K. Stierstorfer, B. Schmidt, F. Knoll, A. Primak, D. Faul, D. Sodickson, and R. Otazo, "Realistic undersampling model for compressed sensing using a multi-slit collimator," in *International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, 2017.

MR to X-ray Projection Image Synthesis

Bernhard Stimpel, Christopher Syben, Tobias Würfl, Katrin Mentl, Arnd Dörfler, and Andreas Maier

***Abstract*—Hybrid imaging promises large potential in medical imaging applications. To fully utilize the possibilities of corresponding information from different modalities, the information must be transferable between the domains. In radiation therapy planning, existing methods make use of reconstructed 3D magnetic resonance imaging data to synthesize corresponding X-ray attenuation maps. In contrast, for fluoroscopic procedures only line integral data, i.e., 2D projection images, are present. The question arises which approaches could potentially be used for this MR to X-ray projection image-to-image translation. We examine three network architectures and two loss-functions regarding their suitability as generator networks for this task. All generators proved to yield suitable results for this task. A cascaded refinement network paired with a perceptual-loss function achieved the best qualitative results in our evaluation. The perceptual-loss showed to be able to preserve most of the high-frequency details in the projection images and, thus, is recommended for the underlying task and similar problems. The abstract goes here.**

***Index Terms*—Medical image synthesis, multi-modality fusion, machine learning, Fluoroscopy**

I. Introduction

Promising concepts on how a combined magnetic resonance (MR) and computed tomography (CT) imaging device may look like were proposed in the past. Wang et al. [1] published a top-level design of an MR-CT scanner consisting of two superconducting electromagnets surrounding multiple, rotatable X-ray sources. The desired application for their model is combined image reconstruction for plaque characterization. In contrast, [2] focused on the interventional applicability of a hybrid MR-X-ray system and showed the great potential of this application. Assuming an imaging device that is capable of acquiring corresponding X-ray and MR projection images simultaneously, or at least consecutively in the same state of motion, the combined information would be highly useful for fluoroscopic procedures. On the one hand, overlay strategies of both modalities in their respective form could be used to simultaneously visualize soft- and dense-tissue or -material. On the other hand, the information of one modality could be transferred to the domain of its counterpart. This information could then be used for further processing and image enhancement. A possible application would be to exploit the high signal-to-noise ratio of MR imaging, especially in soft-tissue regions, to apply denoising methods on the corresponding X-ray images. Considering that the noise level in X-ray Fluoroscopy is directly related to the applied radiation dose, a higher tolerance for noise could lead to reduction of harmful patient radiation exposure. Furthermore, it allows for investigations in the field of super-resolution. Most of the mentioned applications would require corresponding images in the same domain. The acquisition of projection images that match the typical projective distortion directly from the MR is possible, as shown by [3], [4]. To allow for further downstream processing, a possibility to transfer the information between the projection images in the distinct domains would be useful. Similar methods are already used in radiation therapy planning, where attenuation maps are estimated from pseudo-CT scans that are synthesized from corresponding MR data [5], [6], [7]. However, all these methods are based on 3D tomographic image data. In contrast, for fluoroscopic procedures this transfer between the domains must be performed based on line integral data, i.e., 2D projection images, and not on reconstructed images. Motivated by its possible applications and inspired by existing methods from radiation therapy and natural image synthesis, we investigate different deep learning-based methods for X-ray projection image synthesis from MR projections.

This work has been supported by the project P3-Stroke, an EIT Health innovation project. EIT Health is supported by EIT, a body of the European Union.

B. Stimpel, C. Syben, T. Würfl, K. Mentl, and A. Maier are with Friedrich-Alexander-Universität Erlangen-Nürnberg, Pattern Recognition Lab, Erlangen, Germany.

B. Stimpel, C.Syben, and A. Dörfler are also with Friedrich-Alexander-Universität Erlangen-Nürnberg, Department of Neuroradiology, Erlangen, Germany.

II. Methods

Convolutional neural networks have shown great results in natural and medical image synthesis [6], [8]. Based on this, three different generator network architectures are used in the underlying work with the goal to generate X-ray projections $\boldsymbol{G}$ from input MR projection images $\boldsymbol{I}$. Training and evaluation are done using corresponding MR and label X-ray projections $\boldsymbol{L}$. All models have been adapted to our specific application. An overview of the investigated network architectures is given in Figure 1. Furthermore, we examined the impact of two different loss-functions on the generated results.

A. Model Architecture

Convolutional auto-encoders are a popular choice for generator networks in image synthesis. In general, an auto-encoder consists of an encoder and a decoder path. In the encoder path the image's resolution is decreased and the filter dimension is increased. The subsequent decoder path reverts this process to reach the initial resolution and dimension again. Enhancing the encoder-decoder structure with skip-connections between corresponding resolution levels has proven to be beneficial regarding the conservation of spatial information lost during down-sampling. Our first network model is close to the well-known "U-net" introduced by Ronneberger et al. [9]. Instead of maximum pooling layers we use strided convolution with stride two for up- and down-sampling. In addition to the

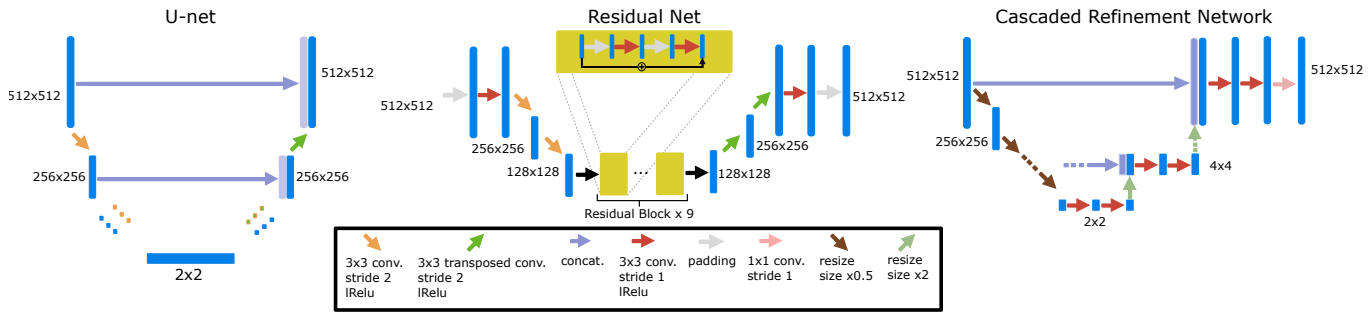


Fig. 1: Schematic architectures of the different generator networks.

architecture presented in Figure 1, the first three layers of the synthesis path use dropout with a keep probability of 50 percent.

The second generator network is a deep residual network (ResNet) [10] which was initially proposed for image recognition. The key component of this approach are residual connections that allow for more robust training of deeper networks than before. Besides the original application, this network architecture proved to yield good result in generative tasks. We use the model proposed by [11] for style transfer to generate our estimated X-ray projections. Deviating from their proposal, we add nine residual blocks instead of the originally proposed five.

Finally, a cascaded refinement network (CRN) is used as image generator. This model was recently proposed by Chen et al. [12] and yielded good results on natural image synthesis from a semantic layout. In contrast to many currently proposed approaches, their model does not use adversarial training but relies on a single feedforward network. The semantic layout as input is replaced by MR projection images in our case. The network consists of multiple refinement modules that work in a multi-scale strategy from coarse to fine as presented in Figure 1. The full model is built from 8 single refinement modules and the final 1 x 1 convolution layer maps the output to a single channel image. A major difference to the first two network architectures is that Chen et al. relinquished convolutional layers in the down-scaling path and, instead, only use resizing operations. Input information from higher resolution scales is solely incorporated using concatenation. By this, additional model capacity can be used for the subsequent up-scaling path.

B. Objective Functions

The choice of the objective function is a key aspect in every machine learning application. Multiple functions have been used for the task of image-to-image translation and image synthesis in the past. We picked two different loss-functions to compare them in our approach. Since a one-to-one correspondence is given by the matching image pairs, a simple but suitable loss function for image generation tasks is the ℓ_1-norm [13]. Pixel-wise comparison of the generated and label image intensities via the ℓ_1-loss function can be done by calculating

$$\mathrm{E}_{\ell_1}(\boldsymbol{L}, \boldsymbol{G}) = \sum_{i}^{N} |\boldsymbol{L}(i) - \boldsymbol{G}(i)| \ , \tag{1}$$

where i denotes one image pixel, $i \in N$, and N is the number of all pixel in one image.

A second loss function that was recently proposed for natural image synthesis without corresponding image pairs is the perceptual-loss [11]. The perceptual-loss does not calculate the error between the estimated and real intensity values. Instead, the generated and the label image are fed into a pre-trained image classification network that we will refer to as evaluation network in the following. While the resulting classification scores are not of interest, the raw feature activations between the different input images are compared. The underlying theory is that similarly looking images activate the same units in the image classification network, i.e. the higher the accordance between both feature activations the more similar the generated and label image are. The loss function can be written as

$$\mathrm{E}_{\mathrm{p}}(\boldsymbol{L}, \boldsymbol{G}) = \sum_{k}^{K} \left(\boldsymbol{V}_k(\boldsymbol{L}) - \boldsymbol{V}_k(\boldsymbol{G}) \right) \ , \tag{2}$$

where $\boldsymbol{V}_k(\boldsymbol{L})$ and $\boldsymbol{V}_k(\boldsymbol{G})$ is the feature activation map of the evaluation network for the label image $\boldsymbol{L}$ and the generated image $\boldsymbol{G}$ at the current layer k, $k \in K$. In this approach, the perceptual-loss is computed on the VGG-19 network [14] which was pre-trained on the ImageNet data set [15].

All generators are trained with an ADAM optimizer [16] and a learning rate of 0.004 for 100 epochs.

III. Experiments

Experiments were conducted using data of a realistic MR and X-ray sensitive phantom of the human head. Data was acquired on a 1.5 T Aera MR and a Axiom-Artis C-arm CT scanner (Siemens Healthcare GmbH, Forchheim, Germany). An ultra-short echo time sequence was used for the MRI scans. The reconstructed images' resolution is 320 x 320 x 250 with a spacing of 0.93 x 0.93 x 0.93 mm^3. The X-ray scans of the same phantom exhibit a voxel size of 0.48 x 0.48 x 0.48 mm^3 and a resolution of 512 x 512 x 399. Image registration of the corresponding scans was performed using elastix. The input (MR) and label (CT) images were generated by forward projecting the registered stack from various angulations using the CONRAD framework [17]. In this manner, 3200 different projection image pairs of both modalities were created and randomly divided into 3000 training and 200 testing images.

	MAE	SSIM	PSNR
U-net - p-loss	0.083	0.891	26.994
ResNet - p-loss	0.077	0.924	27.675
CRN - p-loss	0.071	0.931	28.353
U-net - l1-loss	0.068	0.917	28.506
ResNet - l1-loss	0.058	0.938	30.067
CRN - l1-loss	0.084	0.920	27.097

TABLE I: Quantitative results of the different network architectures and loss functions

The evaluation of the output can be done by calculating the deviation of the generated X-ray $\boldsymbol{G}$ from the real X-ray images $\boldsymbol{L}$. The mean squared error (MSE) can be used to this end. It is computed as

$$\text{MSE}(\boldsymbol{L}, \boldsymbol{G}) = \frac{1}{N} \sum_{i}^{N} \left\| \boldsymbol{L}(i) - \boldsymbol{G}(i) \right\|_2^2 . \tag{3}$$

Yet, not only the absolute difference of estimated values is of interest in projection image synthesis. The generated projection images must also correspond to each other from a visual point of view, which cannot be determined entirely be pixel-wise comparison of the image pairs. To this end, the structural similarity (SSIM) index [18], a perception-based metric, is computed. Assuming two patches $\boldsymbol{g}$ and $\boldsymbol{l}$ of the generated and label image. The SSIM is then computed as

$$\text{SSIM}(\boldsymbol{g}, \boldsymbol{l}) = \frac{(2\mu_{\boldsymbol{g}}\mu_{\boldsymbol{l}} + c_1)(2\sigma_{\boldsymbol{gl}} + c_2)}{(\mu_{\boldsymbol{g}}^2 + \mu_{\boldsymbol{l}}^2 + c_1)(\sigma_{\boldsymbol{g}}^2 + \sigma_{\boldsymbol{l}}^2 + c_2)} , \tag{4}$$

where μ is the mean, σ^2 the variance, and σ the covariance. To avoid instabilities, the constants c_1 and c_2 are introduced that are defined as $c_i = (K_i \mathcal{L})^2$, $i \in \{1, 2\}$, with $\mathcal{L}$ being the dynamic range of the intensity values and $K_1 = 0.01$ and $K_2 = 0.03$. Computing Equation 4 for all pairs of patches $\boldsymbol{g}$ and $\boldsymbol{l}$ yields the final SSIM measure for the whole image.

The third evaluation metric that is computed is the peak signal-to-noise ratio (PSNR). The PSNR measures the ratio between the highest intensity value and the occuring noise and is often applied to measure image quality, especially regarding reconstruction and compression loss. It is computed by

$$\text{PSNR}(\boldsymbol{L}, \boldsymbol{G}) = 20 \log_{10} \frac{\max(\boldsymbol{G})}{\text{MSE}(\boldsymbol{L}, \boldsymbol{G})} . \tag{5}$$

In the subsequent chapter results for all metrics will be presented. To present comparable absolute numbers, all images were scaled from -1 to 1 prior to the error metric calculations.

IV. Results and Discussion

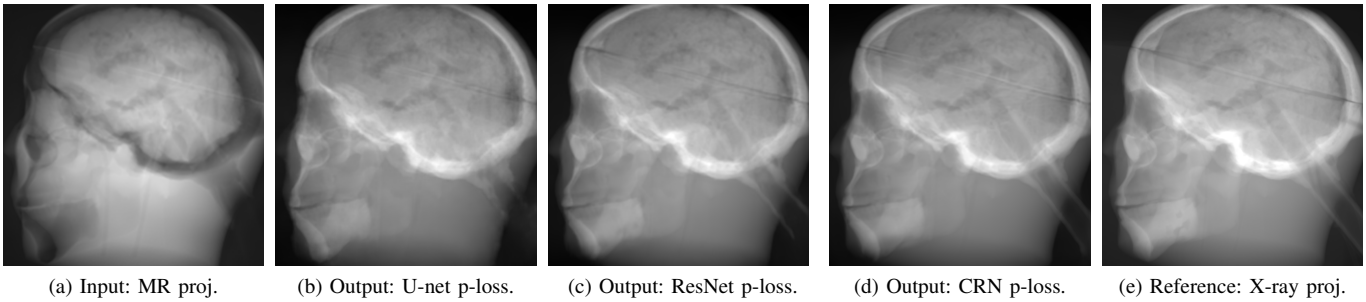
The quantitative and qualitative results of the proposed experiments are presented in Table I and Figure 2. By examining these it can be observed that the differences in the calculated MSE of all network architectures and incorporated loss functions are only small. The best results in terms of pixel-wise deviation could be achieved with the ResNet architecture combined with the ℓ_1-loss function. This network achieves a deviation from the reference of only 0.058, i.e., 2.4 percent. Also the results of the U-net and CRN networks are still good with deviations of 2.6 and 2.9 percent. Similarly small variation can be observed in the structured similarity measure. The ResNet and CRN exhibit approximately equal quality with SSIM measures of 0.938 and 0.920 for the ℓ_1-loss and 0.924 and 0.931 for the perceptual-loss, respectively. The results generated with the U-net are slightly worse. The highest peak signal-to-noise ratio is achieved by the ResNet (ℓ_1-loss), followed by the U-net (ℓ_1-loss) and CRN (p-loss). It is noteworthy that the ResNet and U-net both achieve the highest results in all error metrics using the ℓ_1-loss while the opposite is the case for the CRN which works best with the perceptual-loss function.

Overall, the perceptual-loss achieves competitive and in some cases even better results than the ℓ_1-loss when comparing the pixel-wise error metrics. For example, the cascaded refinement network's MSE is 0.013 smaller for the perceptual- than for the ℓ_1-loss. This might be suspicious at first sight, considering that the ℓ_1-loss purely optimizes for this pixel-wise error in the training process while the perceptual-loss compares the raw feature activations of the evaluation network. Contrarily, this behavior cannot be observed for the U-net and ResNet. The results produced with the ℓ_1-loss achieve higher values for all error measures for these networks. An explanation for this obervation is that the intensity values of the input image still cause an impact on the respective layers output in the evaluation network when computing the perceptual-loss. Consequently, these differences also transition to the computed loss value for all feature layers. Even though the perceptual-loss incorporates the raw intensity values, it is not guaranteed that the scaling of these is conserved in this process. By this, the relative changes can be similar, whereas the absolute range of values changes and, correspondingly, also the pixel-wise error metrics.

Another observation is that the perceptual-loss is able to conserve high-frequency details in the image. The fine line in the projection images that forms a circle around the cranium is visible in the input (Figures 2a & 2f), as well as in the label images (Figures 2e & 2j), and also in the images generated with the perceptual-loss function (Figures 2b, 2c, and 2d). In contrast, all generators "loose" this line when the ℓ1-loss is applied (Figures 2g, 2h, and 2i). This effect is also qualitatively observable in other parts of the images. Despite achieving equal or better results regarding the error metrics, the generally less sharp look of the results generated with the ℓ1-loss function is apparent. This behavior is in accordance with previous observations that concluded that an perceptual-loss leads to sharper images than a comparable ℓ1-loss [19]. Considering the common applications of X-ray Fluoroscopy, e.g., interventional guidance for stents and similar devices, high spatial resolution is a key requirement. Utilizing a loss function that is able to preserve high-frequency details in the images is desirable to this end. The perceptual-loss appears to be suited for this task as presented in our evaluation.

V. Conclusion

We showed the feasibility of image-to-image translation from MR projection images to corresponding X-ray projections. Three generator networks and two different loss



(a) Input: MR proj. (b) Output: U-net p-loss. (c) Output: ResNet p-loss. (d) Output: CRN p-loss. (e) Reference: X-ray proj.

(f) Input: MR proj. (g) Output: U-net ℓ_1-loss. (h) Output: ResNet ℓ_1-loss. (i) Output: CRN ℓ_1-loss. (j) Reference: X-ray proj.

Fig. 2: Results of the projection synthesis. Top row: Results generated with the perceptual-loss function. Bottom row: Results generated with the ℓ_1-loss function.

functions were implemented and evaluated to this end. All examined network architectures achieved good results on the proposed task. When comparing the generated projection images of all networks it became apparent that the loss function has a greater impact on the images' quality than the actual architectures of the network. The perceptual-loss proved to be able to conserve even small high-frequency details in the course of the image-to-image transfer. Because high-spatial resolution is desired in most fluoroscopic procedures, we recommend using this perceptual-loss function for the underlying task. The best quantitative and qualitative results with this loss function could be achieved by a cascaded refinement model in this work. The high-quality of the generated projection images unveils large potential regarding the applicability to multi-modal denoising, super-resolution, and more. As a next step, we plan to transfer this approach to real patient data. Additionally, the effect of combining multiple different MR acquisition protocols and weighting schemes will be investigated.

REFERENCES

[1] G. Wang, F. Liu, F. Liu, G. Cao, H. Gao, and M. W. Vannier, "Top-Level Design of the first CT-MRI scanner," in *Proc. 12th Fully 3D Meet.*, 2013, pp. 5–8.

[2] R. Fahrig, K. Butts, J. A. Rowlands, R. Saunders, J. Stanton, G. M. Stevens, B. L. Daniel, Z. Wen, D. L. Ergun, and N. J. Pelc, "A truly hybrid interventional MR/x-ray system: Feasibility demonstration," *J. Magn. Reson. Im.*, vol. 13, no. 2, pp. 294–300, 2001.

[3] S. Napel, S. Dunne, and B. K. Rutt, "Fast Fourier projection for MR angiography," *Magn. Reson. Med.*, vol. 19, no. 2, pp. 393–405, 1991.

[4] C. Syben, B. Stimpel, M. Leghissa, A. Dörfler, and A. Maier, "Fan-beam Projection Image Acquisition using MRI," in *3rd Conf. Image-Guided Interv. Fokus Neuroradiol.*, M. Skalej and C. Hoeschen, Eds., 2017, pp. 14–15.

[5] B. K. Navalpakkam, H. Braun, T. Kuwert, and H. H. Quick, "Magnetic ResonanceBased Attenuation Correction for PET/MR Hybrid Imaging Using Continuous Valued Attenuation Maps," *Invest. Radiol.*, vol. 48, no. 5, pp. 323–332, 2013.

[6] D. Nie, R. Trullo, J. Lian, C. Petitjean, S. Ruan, Q. Wang, and D. Shen, "Medical Image Synthesis with Context-Aware Generative Adversarial Networks," *Med. Image Comput. Comput. Interv. MICCAI 2017*, pp. 417–425, 2017.

[7] J. M. Wolterink, A. M. Dinkla, M. H. F. Savenije, P. R. Seevinck, C. A. T. van den Berg, and I. Išgum, "Deep MR to CT Synthesis Using Unpaired Data," in *Int. Work. Simul. Synth. Med. Imaging*, 2017, pp. 14–23.

[8] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks," *IEEE Conf. Comput. Vis. pattern Recognit.*, pp. 2414–2423, 2016.

[9] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *Med. Image Comput. Comput. Interv. MICCAI 2015*, pp. 234–241, 2015.

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016.

[11] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual Losses for Real-Time Style Transfer and Super-Resolution," *arXiv:1603.08155*, 2016.

[12] Q. Chen and V. Koltun, "Photographic Image Synthesis with Cascaded Refinement Networks," in *Int. Conf. Comput. Vis.*, 2017.

[13] H. Zhao, O. Gallo, I. Frosio, and J. Kautz, "Loss Functions for Image Restoration with Neural Networks," *IEEE Trans. Comput. IMAGING*, vol. 3, no. 1, 2017.

[14] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv:1409.1556*, 2015.

[15] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet Large Scale Visual Recognition Challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, 2014.

[16] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in *Int. Conf. Learn. Represent.*, 2015.

[17] A. Maier, H. G. Hofmann, M. Berger, P. Fischer, C. Schwemmer, H. Wu, K. Müller, J. Hornegger, J.-H. Choi, C. Riess, A. Keil, and R. Fahrig, "CONRAD - A software framework for cone-beam imaging in radiology," *Med. Phys.*, vol. 40, no. 11, pp. 111914, 2013.

[18] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004.

[19] A. Dosovitskiy and T. Brox, "Generating Images with Perceptual Similarity Metrics based on Deep Networks," *Adv. Neural Inf. Process. Syst. 29 (NIPS 2016)*, pp. 658–666, 2016.

Deep Neural Network for CT Metal Artifact Reduction with a Perceptual Loss Function

Lars Gjesteby, Hongming Shan, Qingsong Yang, Yan Xi, Bernhard Claus,
Yannan Jin, Bruno De Man, and Ge Wang

***Abstract*—Machine learning and deep learning are rapidly finding applications in the medical imaging field. In this paper, we address the long-standing problem of metal artifacts in computed tomography (CT) images by training a deep convolutional neural network for streak removal. While many metal artifact reduction methods exist, even state-of-the-art algorithms fall short in some clinical applications. Specifically, proton therapy planning requires high image quality with accurate tumor volumes to ensure treatment success. We explore a deep network structure with residual learning to correct metal streak artifacts after a first-pass by a state-of-the-art interpolation-based algorithm, NMAR. Our experiments compare a mean squared error loss function with a perceptual loss function to emphasize preservation of image features and texture. Both visual and quantitative metrics are used to assess the resulting image quality for metal implant cases. This study shows that image-domain deep learning can be highly effective for metal artifact reduction (MAR), and highlights the benefits and drawbacks of different loss functions for solving a major CT reconstruction challenge.**

***Index Terms*—Computed tomography (CT), convolutional neural network (CNN), metal artifact reduction (MAR), deep learning, proton therapy**

I. Introduction

Metal artifacts are a classical problem in CT that greatly hinder image formation accuracy. Methods to reduce these artifacts have had varying degrees of success over the past four decades (see [1] for a comprehensive overview). However, there remain important applications in which even state-of-the-art correction algorithms may fail to provide sufficient image quality. Specifically, proton therapy planning requires precise localization of tumors because energy lost by charged particles is inversely proportional to the square of their velocity, so most of the dose from the proton beam is delivered in a very small area. Tumor volume estimation is very sensitive to image reconstruction errors, and miscalculation due to metal artifacts may result in tumor recurrence or radiation toxicity [2], [3].

Projection completion algorithms are the primary MAR tool currently in use. They aim to correct/replace corrupted data inside the metal trace of the CT sinogram with data synthesized by an interpolation technique [4]–[7], reprojection from a prior image [8]–[11], or a combination of both that involves normalization [12]–[14]. Among these, NMAR is considered a state-of-the-art method that employs interpolation and normalization with a prior image to replace data in the metal trace [13].

Additional MAR techniques can be categorized as scan acquisition improvements, physics-based pre-processing, iterative reconstruction, and image post-processing. While image post-processing algorithms have had only limited success [15], [16], their merits are better seen when combined with projection-domain correction [17]. While these methods remain unsatisfactory for the most challenging applications, the addition of deep learning could change that.

The field of machine learning/deep learning has risen rapidly to solve increasingly complicated problems [18]. This technology presents a novel way to reduce CT metal artifacts, with our initial efforts reported within the past year [19]–[21]. Particularly, Convolutional neural networks (CNNs) are powerful in their ability to extract detailed features and learn patterns from large datasets, which has enabled great successes in image processing and analysis tasks [22]–[24]. The supervised learning process relies on labeled images so that the network learns how to map features between the input and the label/target. Once trained, the network uses forward prediction to estimate an output given an unlabeled input.

The main objective of this study is to use deep learning to reduce streak artifacts remaining after initial MAR correction in critical image regions around the metal object. Inspired by a rain streak removal method in natural images [25], [26], we developed a deep convolutional neural network architecture that utilizes a detail layer/image, which is a sparse representation of the artifacts to simplify the learning process. This approach differs from our most recent work in which the networks learned direct image-to-image mapping from metal-corrupted CT images to their corresponding artifact-free images [21]. The network is trained with clinical-like datasets generated from real CT scans with hip prostheses and spinal fixation rods added. We investigated both a mean squared error loss function and a perceptual loss function to preserve features. The resulting image quality and visual appearance from both networks is quantified and compared to that of the artifact-free ground truth, the NMAR algorithm, and our previous direct CNN method.

II. Methods

A. Data Synthesis

The Visible Human Project dataset was used as the basis for all training and validation data [27]. Two sets of voxelized phantoms were created from the volumetric data in the pelvic

L. Gjesteby, H. Shan, Q. Yang, Y. Xi, and G. Wang are with the Biomedical Imaging Center, Department of Biomedical Engineering, Rensselaer Polytechnic Institute, Troy, NY, USA (contact: wangg6@rpi.edu)

B. Claus, Y. Jin, and B. De Man are with Imaging, General Electric Global Research Center, Niskayuna, NY, USA (contact: deman@ge.com)

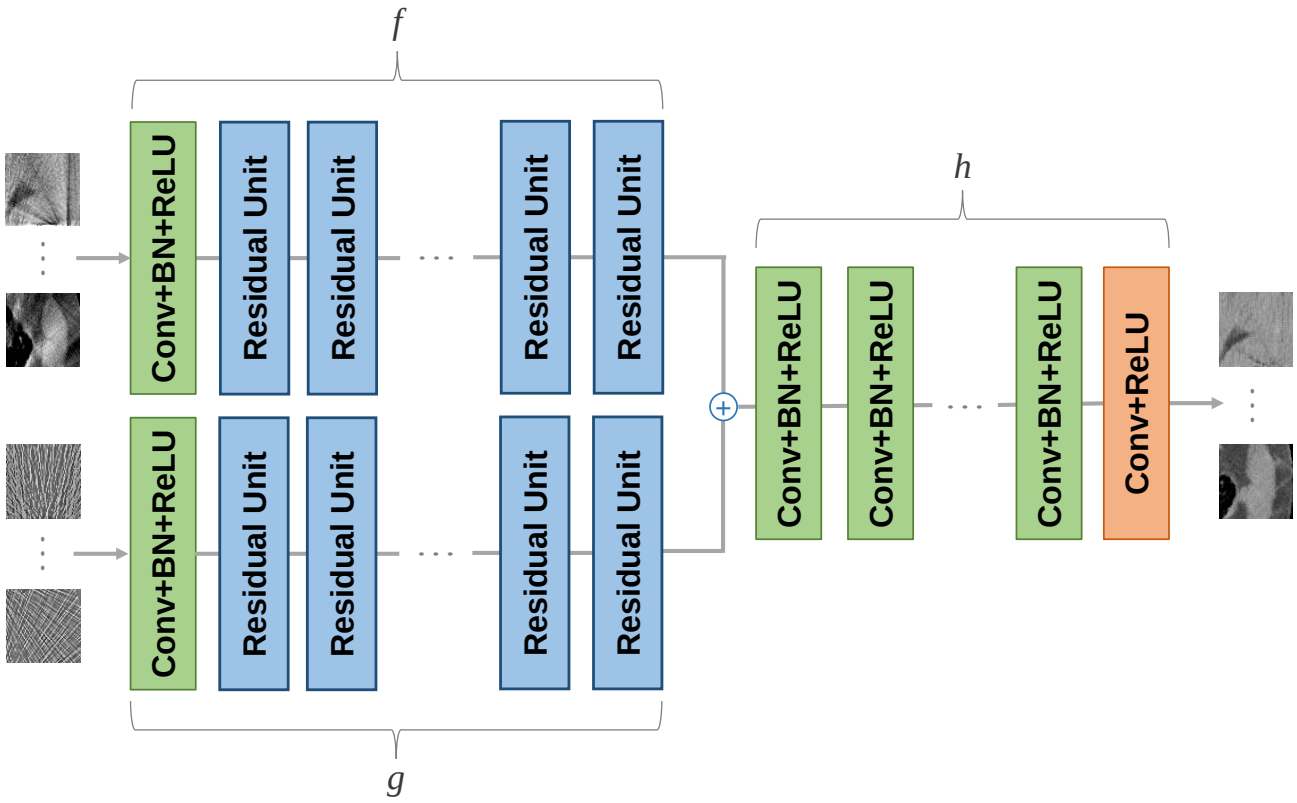


Fig. 1. **Network structure for *DestreakNet***. There are two parallel streams, f and g, each containing an initial layer followed by 20 residual units. Patches from the NMAR image and patches from the detail image are input to f and g, respectively. The outputs of these streams are merged in the feature space, and then passed through h, which contains eight parameter layers and a final layer, to yield the final output. All convolution layers have 32 filters (except for the final layer, which has only one filter), each of which has a 3×3 kernel and uses zero-padding. Batch normalization (BN) is used after each convolution layer (except for the final convolution layer), and is followed by a rectified linear unit (ReLU). The input and output patches are of size 56×56.

and spinal regions. Titanium was added in the femoral head region to represent a hip prosthesis (up to 20 mm radius) and next to vertebrae to represent spinal fixation rods (up to 10 mm radius). Then, industrial CT simulation software, *CatSim* [28], was used to scan both sets of phantoms to generate the metal-corrupted and artifact-free projections, and reconstruction was performed with filtered back-projection. The scan geometry simulated a *GE LightSpeed VCT* system architecture, with key parameters including a tube voltage of 120 kVp, a tube current of 300 mA, 10^8 photons, 888 detector columns, and 720 views at uniform projection angles between 0-360 degrees. The phantoms without metal were scanned with a 100 keV monoenergetic beam to minimize general noise and beam hardening artifacts.

The 512×512 reconstructed images of the metal-added phantoms contained severe artifacts. The NMAR algorithm was used to perform initial correction on the data via sinogram interpolation and provided baseline artifact reduction in the subsequent reconstructed images. This NMAR result served as the input to one stream of the network. Scans on the datasets without metal yielded images without artifacts to serve as the "ground truth" and target/label of the network. The phantom pairs were scanned for a total of 50 image slices at 1mm-thickness. Eight images were reserved for testing/validation and not used in the training process. From the full-size images, approximately 150,000 patches of size 56×56 were extracted from the images to form the dataset to train the CNNs.

B. Network Design

Our network structure is shown in Fig. 1, which we refer to as *DestreakNet*. Some parameter layers are grouped in pairs that we refer to as *residual units*. Within each residual unit, the network learns the residual error between the input and output of the unit. Training on the residual difference reduces the mapping range of pixel values used for learning [29]. Each residual unit (Fig. 2) contains two convolution layers (Conv), each of which is followed by batch normalization (BN) and a rectified linear unit (ReLU) activation function to maintain non-linearity. All convolution layers contain 32 filters with a 3×3 kernel and use zero padding. There are two parallel streams of the network, f and g, both with identical structures. Each stream begins with one parameter layer, followed by 20 residual units. The feature-maps of the streams f and g are then summed together to pass though h, which contains eight more parameter layers and a final layer without batch normalization.

In the work of [25], the image is modeled as follows:

$$\boldsymbol{I} = \boldsymbol{I}_{\text{detail}} + \boldsymbol{I}_{\text{base}}, \tag{1}$$

where $\boldsymbol{I}_{\text{detail}}$ denotes the detail image, and $\boldsymbol{I}_{\text{base}}$ denotes the base image. The base image can be obtained by low-pass

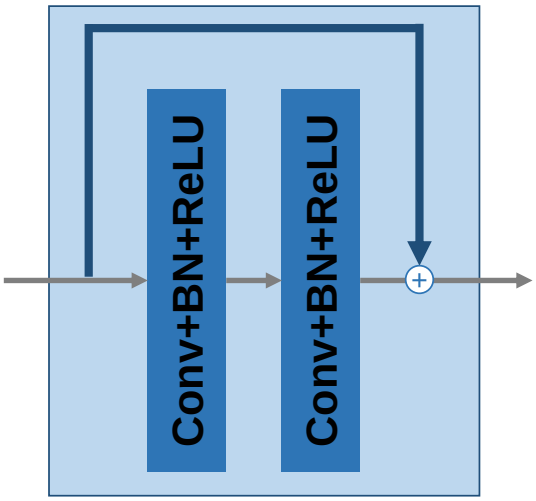


Fig. 2. **Residual Unit.** Each residual unit contains two convolution layers, each of which is followed by batch normalization (BN) and a rectified linear unit (ReLU).

filtering $\boldsymbol{I}$, after which the detail image is represented as $\boldsymbol{I}_{\text{detail}} = \boldsymbol{I} - \boldsymbol{I}_{\text{base}}$.

We adopt a similar approach, but make several key changes that better fit our specific task. The NMAR images have more content information but less streaks, while the uncorrected images have more metal streaks but less content. To harness the content information from NMAR images and the streak information from uncorrected images, two streams of our network, f and g, take data from the NMAR images and detail images, respectively. To create our detail image, $\boldsymbol{D}$, the original uncorrected image before NMAR correction, $\boldsymbol{U}$, is passed through a low-pass guided filter [30]. Then, the filtered image or base image, $\mathcal{F}(\boldsymbol{U})$, is subtracted from the original uncorrected image:

$$\boldsymbol{D} = \boldsymbol{U} - \mathcal{F}(\boldsymbol{U}). \tag{2}$$

After subtracting the base image from the uncorrected image, the background interference is removed and only metal streaks and object structures remain in the detail image. Given the NMAR image, $\boldsymbol{I}$, and the detail image, $\boldsymbol{D}$, the output of our network can be expressed as:

$$\widetilde{\boldsymbol{Y}} = h\Big(f\big(\boldsymbol{I}\big) + g\big(\boldsymbol{D}\big)\Big), \tag{3}$$

where h, f, and g are streams of layers in the network.

C. Loss Functions and Training

We explored two loss functions for optimizing the network. A mean squared error (MSE) loss function calculates the pixel-by-pixel error between the output, $\widetilde{\boldsymbol{Y}}_i$, and the target, $\boldsymbol{Y}_i$:

$$\mathcal{L}_{\text{MSE}} = \frac{1}{N}\sum_{i=1}^{N}\left\|\widetilde{\boldsymbol{Y}}_i - \boldsymbol{Y}_i\right\|_F^2, \tag{4}$$

where N is the number of training samples and subscript i denotes the i-th sample in training set.

We also employed a perceptual loss function to mitigate over-smoothing effects that can result from the MSE loss function. The perceptual loss function from the pre-trained VGG network [31] is used to optimize the network output in the feature space to preserve texture:

$$\mathcal{L}_{\text{P}} = \frac{1}{N}\sum_{i=1}^{N}\left\|\phi(\widetilde{\boldsymbol{Y}}_i) - \phi(\boldsymbol{Y}_i)\right\|_F^2, \tag{5}$$

where ϕ is a feature extractor to indicate loss calculation in the feature space. In our experiment, we use the output of the 16^{th} convolution layer in the VGG network as the feature extractor.

During training, the network adjusts its parameter weights in each layer to minimize the loss function. Training was run on an Nvidia GeForce GTX 1080 Ti graphics card using the TensorFlow framework in Python for 60 epochs, which took approximately seven hours. The network learning rate α was initialized to 10^{-3}. In the t-th training epoch, the learning rate α_t was decayed proportional to the epoch number: $\alpha_t = \alpha/\sqrt{t}$.

III. Results

To validate the network performances, hip and spine image slices withheld from training were used. We present a hip prostheses case in Fig. 3, which shows the metal-free truth, the initial uncorrected reconstruction image, the NMAR-corrected image, and the results of our networks with MSE and perceptual loss functions. We also compared with a contracting path convolutional encoder-decoder (CPCE-MSE) trained on a direct image-to-image mapping without incorporating a detail image [32]. The initial reconstruction from the raw projection data yielded a poor image with many streaks, a dark band along the lines of greatest attenuation, and regions of missing data. The pre-correction artifacts are severe in the hip case due to the large implant size. Our network significantly reduced streak artifacts that remained in the image after initial correction by the NMAR algorithm. The linear interpolation steps of NMAR appeared to introduce artifacts that could be due to incomplete segmentation of the metal.

The image quality metrics are displayed in Table I for all results in Fig. 3. We compared structural similarity index (SSIM) and peak signal-to-noise ratio (PSNR) in reference to the artifact-free truth. Both networks increase the SSIM and PSNR scores by a substantial amount. The *DestreakNet-MSE* output has the highest image quality scores, but an over-smoothing effect is evident. The *DestreakNet-Perceptual* output better preserves features and texture to give a more visually-pleasing result, although the image quality scores are slightly lower than those of both networks using MSE loss functions.

IV. Discussion and Conclusion

This MAR study explores a deep neural network structure for reducing metal artifacts in CT images. We provide the network with information on both the streak artifacts and the image feature content to strengthen the learning process. We show the benefits and drawbacks of different loss functions from a visual and quantitative standpoint, with an emphasis

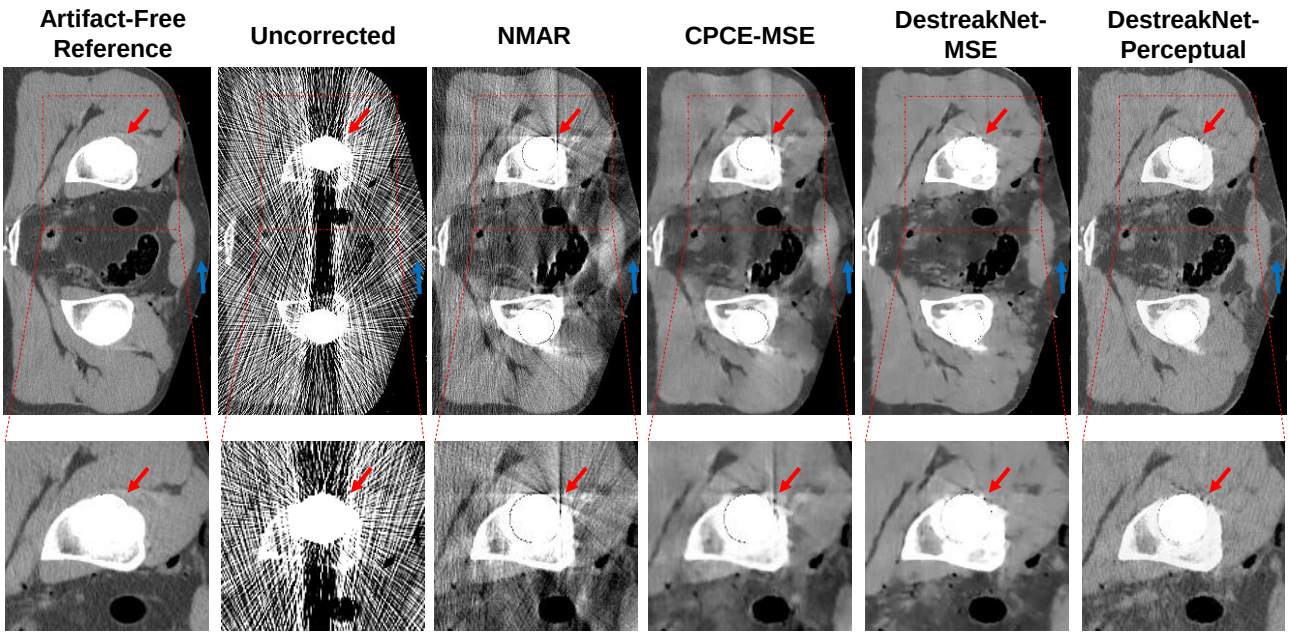


Fig. 3. **Deep learning results for a hip prostheses case.** The full images are displayed with zoomed ROIs on a critical image region. Left to right: Artifact-free truth; Initial uncorrected reconstruction; NMAR-corrected image; CPCE network using MSE loss [32]; Our deep neural network using MSE loss; Our deep neural network using perceptual loss.

TABLE I
STRUCTURAL SIMILARITY INDEX (SSIM) AND PEAK SIGNAL-TO-NOISE RATIO (PSNR) FOR RESULTS SHOWN IN FIG. 3 COMPARED TO ARTIFACT-FREE TRUTH.

	SSIM	PSNR
Uncorrected	0.2382	9.1830
NMAR	0.7014	18.8975
CPCE-MSE	0.8398	22.7963
DestreakNet-MSE	**0.8636**	**23.8582**
DestreakNet-Perceptual	0.8264	22.1685

on perceptual loss to preserve image features and texture. Our work is based on a superiority principle of data-driven machine learning, which seeks to improve upon any state-of-the-art traditional MAR result by using a well-trained network. In principle, learning-based MAR techniques will always outperform the best result, and could be repeatedly applied. This can be viewed as a monotonically-increasing sequence that possibly approaches the ground truth as the limiting point.

The key improvements to our image-domain deep learning process are the detail image and training on the residual error. With a detail image, the network can learn where the prominent streaks are in the uncorrected image. With residual error training, we reduce the solution range for the network to learn, which yields better results than our previous direct network methods. Our deep neural network pipeline for MAR can be further developed with an improved network architecture and refined parameters, and can include more diverse and enlarged implant types. The larger and the more challenging the dataset is, the better the MAR network can be trained. We are also exploring ways to better guide the network to distinguish between artifacts and features.

In conclusion, we have demonstrated how deep networks can help MAR in a significant way by providing a substantial gain on the results achieved with some of the best existing methods. The importance of the loss function is evident when comparing results among different networks. We continue to pursue advances for tomographic imaging in the machine learning framework, especially artifact reduction to help improve clinical applications such as radiation therapy planning.

REFERENCES

[1] L. Gjesteby, B. De Man, Y. Jin, H. Paganetti, J. Verburg, D. Giantsoudi, and G. Wang, "Metal Artifact Reduction in CT: Where Are We After Four Decades?," *IEEE Access*, vol. 4, 2016.

[2] G. X. Ding and W. Y. Christine, "A study on beams passing through hip prosthesis for pelvic radiation treatment," *International Journal of Radiation Oncology* Biology* Physics*, vol. 51, no. 4, pp. 1167–1175, 2001.

[3] C. Reft, R. Alecu, I. J. Das, B. J. Gerbi, P. Keall, E. Lief, B. J. Mijnheer, N. Papanikolaou, C. Sibata, and J. Van Dyk, "Dosimetric considerations for patients with HIP prostheses undergoing pelvic irradiation. Report of the AAPM Radiation Therapy Committee Task Group 63," *Medical physics*, vol. 30, no. 6, pp. 1162–1182, 2003.

[4] R. M. Lewitt and R. H. T. Bates, "Image-reconstruction from projections. III. Projection completion methods (theory)," *Optik*, vol. 50, no. 3, pp. 189–204, 1978.

[5] T. Hinderling, P. Rüegsegger, M. Anliker, and C. Dietschi, "Computed Tomography Reconstruction from Hollow Projections: An Application to In Vivo Evaluation of Artificial Hip Joints," *Journal of Computer Assisted Tomography*, vol. 3, no. 1, pp. 52–57, 1979.

[6] G. H. Glover and N. J. Pelc, "An algorithm for the reduction of metal clip artifacts in CT reconstructions.," *Medical physics*, vol. 8, no. 6, pp. 799–807, 1981.

[7] W. A. Kalender, R. Hebel, and J. Ebersberger, "Reduction of CT artifacts caused by metallic implants," *Radiology*, vol. 164, no. 2, pp. 576–577, 1987.

[8] C. R. Crawford, J. G. Colsher, N. J. Pelc, and A. H. R. Lonn, "High speed reprojection and its applications," in *SPIE Medical Imaging II*, pp. 311–318, International Society for Optics and Photonics, 1988.

[9] R. Naidu, I. Bechwati, S. Karimi, S. Simanovsky, and C. Crawford, "Method of and system for reducing metal artifacts in images generated by x-ray scanning devices," 2004.

[10] C. S. Olive, M. R. Klaus, V. Pekar, K. Eck, and L. Spies, "Segmentation aided adaptive filtering for metal artifact reduction in radio-therapeutic CT images," in *SPIE Medical Imaging 2004*, vol. 5370, pp. 1991–2002, 2004.

[11] C. Lemmens, D. Faul, and J. Nuyts, "Suppression of metal artifacts in CT using a reconstruction procedure that combines MAP and projection completion," *IEEE Transactions on Medical Imaging*, vol. 28, no. 2, pp. 250–60, 2009.

[12] J. Müller and T. M. Buzug, "Spurious structures created by interpolation-based CT metal artifact reduction," in *SPIE Medical Imaging*, pp. 72581Y–72581Y, International Society for Optics and Photonics, 2009.

[13] E. Meyer, F. Bergner, R. Raupach, T. Flohr, and M. Kachelrieß, "Normalized metal artifact reduction (NMAR) in computed tomography," *Medical Physics*, vol. 37, no. 10, pp. 5482–5493, 2010.

[14] E. Meyer, R. Raupach, B. Schmidt, A. H. Mahnken, and M. Kachelrieß, "Adaptive Normalized Metal Artifact Reduction (ANMAR) in Computed Tomography," in *2011 IEEE Nuclear Science Symposium Conference Record*, pp. 2560–2565, 2011.

[15] G. Henrich, "A simple computational method for reducing streak artifacts in CT images," *Computerized tomography*, vol. 4, no. 1, pp. 67–71, 1980.

[16] M. Bal, H. Celik, K. Subramanyan, K. Eck, and L. Spies, "A radial adaptive filter for metal artifact reduction," in *SPIE Medical Imaging 2005*, vol. 5747, pp. 2075–2082, 2005.

[17] O. Watzke and W. A. Kalender, "A pragmatic approach to metal artifact reduction in CT: Merging of metal artifact reduced images," *European Radiology*, vol. 14, no. 5, pp. 849–856, 2004.

[18] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, may 2015.

[19] L. Gjesteby, Q. Yang, Y. Xi, Y. Zhou, J. Zhang, and G. Wang, "Deep learning methods to guide CT image reconstruction and reduce metal artifacts," in *SPIE Medical Imaging*, pp. 101322W–101322W, International Society for Optics and Photonics, 2017.

[20] L. Gjesteby, Q. Yang, Y. Xi, B. Claus, Y. Jin, B. De Man, and G. Wang, "Reducing metal streak artifacts in CT images via deep learning: Pilot results," in *International Conference on Fully 3D Image Reconstruction*, (Xi'an, China), pp. 611–614, 2017.

[21] L. Gjesteby, Q. Yang, Y. Xi, H. Shan, B. Claus, Y. Jin, B. De Man, and G. Wang, "Deep learning methods for CT image-domain metal artifact reduction," in *SPIE Developments in X-Ray Tomography XI*, (San Diego, CA), 2017.

[22] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural computation*, vol. 1, no. 4, pp. 541–551, 1989.

[23] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, pp. 1097–1105, 2012.

[24] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9, 2015.

[25] X. Fu, J. Huang, D. Zeng, Y. Huang, X. Ding, and J. Paisley, "Removing rain from single images via a deep detail network," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[26] X. Fu, J. Huang, X. Ding, Y. Liao, and J. Paisley, "Clearing the skies: A deep network architecture for single-image rain removal," *IEEE Transactions on Image Processing*, vol. 26, no. 6, pp. 2944–2956, 2017.

[27] M. J. Ackerman, "The visible human project," *Proceedings of the IEEE*, vol. 86, no. 3, pp. 504–511, 1998.

[28] B. De Man, S. Basu, N. Chandra, B. Dunham, P. Edic, M. Iatrou, S. McOlash, P. Sainath, C. Shaughnessy, B. Tower, *et al.*, "Catsim: a new computer assisted tomography simulation environment," in *Medical Imaging 2007: Physics of Medical Imaging*, vol. 6510, p. 65102G, International Society for Optics and Photonics, 2007.

[29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, 2016.

[30] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 6, pp. 1397–1409, 2013.

[31] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[32] H. Shan, Y. Zhang, Q. Yang, U. Kruger, W. Cong, and G. Wang, "3D Convolutional Encoder-Decoder Network for Low-Dose CT via Transfer Learning from a 2D Trained Network," *arXiv preprint arXiv:1802.05656*, 2018.

Author Index

H

I

J

K

L

M

N

O

P

R

S

T

U

V

W

X

Y

Z