Medical Physics Letter

4D cone beam CT via spatiotemporal tensor framelet

Hao Gao[a)]
Departments of Mathematics and Computer Science, and Radiology and Imaging Sciences, Emory University, Atlanta, Georgia 30322

Ruijiang Li
Department of Radiation Oncology, Stanford University, Stanford, California 94305

Yuting Lin
Department of Radiological Sciences, University of California, Irvine, California 92697

Lei Xing
Department of Radiation Oncology, Stanford University, Stanford, California 94305

(Received 25 July 2012; revised 4 October 2012; accepted for publication 5 October 2012; published 25 October 2012)

Purpose: On-board 4D cone beam CT (4DCBCT) offers respiratory phase-resolved volumetric imaging, and improves the accuracy of target localization in image guided radiation therapy. However, the clinical utility of this technique has been greatly impeded by its degraded image quality, prolonged imaging time, and increased imaging dose. The purpose of this letter is to develop a novel iterative 4DCBCT reconstruction method for improved image quality, increased imaging speed, and reduced imaging dose.
Methods: The essence of this work is to introduce the spatiotemporal tensor framelet (STF), a high-dimensional tensor generalization of the 1D framelet for 4DCBCT, to effectively take into account of highly correlated and redundant features of the patient anatomy during respiration, in a multilevel fashion with multibasis sparsifying transform. The STF-based algorithm is implemented on a GPU platform for improved computational efficiency. To evaluate the method, 4DCBCT full-fan scans were acquired within 30 s, with a gantry rotation of 200°; STF is also compared with a state-of-art reconstruction method via spatiotemporal total variation regularization.
Results: Both the simulation and experimental results demonstrate that STF-based reconstruction achieved superior image quality. The reconstruction of 20 respiratory phases took less than 10 min on an NVIDIA Tesla C2070 GPU card. The STF codes are available at https://sites.google.com/site/spatiotemporaltensorframelet.
Conclusions: By effectively utilizing the spatiotemporal coherence of the patient anatomy among different respiratory phases in a multilevel fashion with multibasis sparsifying transform, the proposed STF method potentially enables fast and low-dose 4DCBCT with improved image quality.
 [http://dx.doi.org/10.1118/1.4762288]

Key words: low-dose, 4D cone beam CT, tensor framelet, GPU, reconstruction

I. INTRODUCTION

On-board 4D cone beam CT (4DCBCT) offers respiratory phase-resolved volumetric imaging,[1–4] and improves the accuracy of target localization in image guided radiation therapy.[5] However, the clinical utility of 4DCBCT has been greatly impeded by its degraded image quality, prolonged imaging time, and increased imaging dose. Conventional approaches for respiration-correlated or 4DCBCT imaging reconstruct each respiratory phase independently, and thus often have a very limited amount of projection data available for image reconstruction. Although some progress has been made to improve the 4DCBCT image quality,[6–8] there is clearly an unmet need for improving the current 4DCBCT imaging technique.

Considering the fact that the patient anatomy during respiration is highly correlated and redundant, the efficient 4DCBCT reconstruction should utilize the spatiotemporal coherence of the patient anatomy among different respiratory phases,[9–13] which is inherently superior to the conventional phase-independent reconstructions. In this letter, we present a novel method that effectively utilizes such spatiotemporal coherence in a multilevel fashion with multibasis sparsifying transform, with the attempt for fast and low-dose 4DCBCT with improved image quality.

II. METHODS

Specifically, a new sparse representation method–the spatiotemporal tensor framelet (STF)—is proposed to

characterize the spatiotemporal coherence of the 4D images. This new method goes as follows. Let us consider a 4D image as a tensor $X = \{x_{ijkm},\ i \le N_x,\ j \le N_y,\ k \le N_z,\ m \le N_t\}$ and define X^x, X^y, X^z, X^t as the unfolded matrices of X along x, y, z, t dimension, respectively. For example, X^y is a N_y by $N_x \times N_z \times N_t$ matrix with each column $X^y{}_{ikm} = \{x_{ikm}, j \le N_y\}$. Let W_1 be the standard 1D framelet transform (see the Appendix). Then the STF transform W and its adjoint W^T are defined by

$$Y = WX = \left\{W_1 X^x_{jkm}, W_1 X^y_{ikm}, W_1 X^z_{ijm}, W_1 X^t_{ijk}, i \le N_x, j \le N_y, k \le N_z, m \le N_t\right\}, \quad (1)$$

$$W^T Y = \left(\sum_{jkm} W_1^T Y^x_{jkm} + \sum_{ikm} W_1^T Y^y_{ikm} + \sum_{ijm} W_1^T Y^z_{ijm} + \sum_{ijk} W_1^T Y^t_{ijk}\right) \Big/ 4. \quad (2)$$

The STF via Eqs. (1) and (2) is a natural generalization of 1D framelet to the high-dimensional space through the tensor approach and it is easy to verify that many nice features of the framelet are kept, such as $W^T(WX) = X$. STF has mainly two advantages: STF takes into account the temporal similarity of the 4D images by the 1D framelet operator along the temporal dimension (e.g., $W_1 X^t$) in addition to the spatial sparsity (e.g., $W_1 X^x$); in terms of the required memory, STF (e.g., $\sim 4\,nN$) is more feasible for high-dimensional problems (e.g., 4DCBCT) than the framelet (e.g., $\sim n^4 N$). Here n is the number of framelet bases (refinement masks) and $N = N_x \times N_y \times N_z \times N_t$.

With STF, 4DCBCT is formulated as a L1 minimization problem

$$X = \arg\min_X \|AX - Y\|_2^2 + \lambda \|WX\|_1. \quad (3)$$

In Eq. (3), AX is the x-ray transform of X with the scan geometry specified later, Y represents the 4DCBCT data, and $\|WX\|_1$ is the L1 norm of the STF of X, i.e.,

$$\|WX\|_1 = \sum_{j,k,m} \left\|W_1 X^x_{jkm}\right\|_1 + \sum_{i,k,m} \left\|W_1 X^y_{ikm}\right\|_1 + \sum_{i,j,m} \left\|W_1 X^z_{ijm}\right\|_1 + \sum_{i,j,k} \left\|W_1 X^t_{ijk}\right\|_1. \quad (4)$$

The detail of $\|W_1 \cdot\|_1$ is in the Appendix, which clearly indicates that STF is a generalization of the first-order differencing operator, i.e., spatiotemporal total variation (STV),[12,13] which will be compared with STF next as a state-of-art reconstruction method.

This typical nondifferentiable L1 problem (3) can be solved through the split Bregman method.[14,15] Here, we implement a GPU-based parallel solver based on our prior algorithms[10,11] for both STF and STV, in which the recently proposed fast and highly parallelizable algorithms[16] are used to accelerate the x-ray transform AX and its adjoint $A^T X$ that are two most time-consuming components. Here l is scaled with the L1 norm of A and the data are scaled so that the image intensity is roughly between 0 and 1, for which $l = 0.1$ serves as a good choice for image reconstruction. Please note that STF is slightly faster than STV due to the left-inverse property $W^T W = 1$, since otherwise $W^T W$ needs to be computed repeatedly for STV.

III. MATERIALS

The 4DCBCT reconstruction was evaluated first using a digital NCAT phantom,[17] and then with experimental data. To generate the 30-s simulation data for the NCAT phantom, 40 respiratory phases were utilized, which corresponded to the maximal temporal resolution (100 ms) that was allowed by the simulation setting. For physical experiments, the CIRS dynamic thorax phantom was used, which provides known, accurate (<0.1 mm), and reproducible target motion inside a tissue equivalent phantom.[18] The phantom body represents an average human thorax in shape, proportion, and composition. A lung equivalent rod containing a spherical target was inserted into the phantom. The TrueBeam™ LINAC (Varian Medical Systems, Palo Alto, CA) equipped with an on-board kV imaging system was used to scan the phantom. The kV detector has a physical size of 39.7×29.8 cm^2. The kV imaging parameters were set to be 100 kVp, 80 mA, and 10 ms pulse width.

In both cases, the distance from source to the isocenter was 100 cm, the distance from the isocenter to the detector plane was 50 cm. The projection dataset was a 256×64 array (central slices of the measurement data) with the pixel size 0.16×0.16 cm^2. The respiratory signal during scanning was a sinusoidal curve with a period of 4 s. The full-fan scan was performed with a single 200° gantry rotation that takes ~30 s. The total number of projections was 400 and 372 for the digital and physical phantom, respectively.

For reconstruction, the field of view was $33 \times 33 \times 8$ cm^3. The reconstructed image had the spatial dimension $256 \times 256 \times 64$ with the spatial resolution $0.13 \times 0.13 \times 0.13$ cm^3. A total of 20 respiratory phases were reconstructed to provide a temporal resolution of ~200 ms. The stopping criterions for both STF and STV were based on the minimal data fidelity discrepancy. There were 18–20 projections of data corresponding to each reconstructed phase.

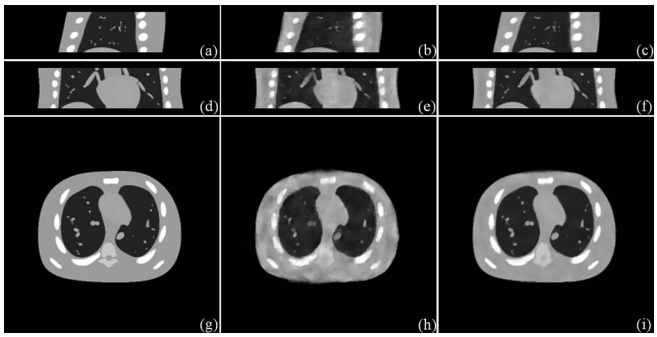

FIG. 1. Simulation results (the exhale phase). Sagittal view: (a) Phantom; (b) STV; (c) STF. Coronal view: (d) Phantom; (e) STV; (f) STF. Transverse view: (g) Phantom; (h) STV; (i) STF.

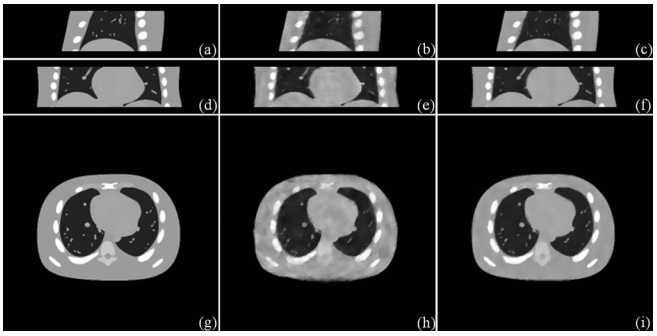

FIG. 2. Simulation results (the inhale phase). Sagittal view: (a) Phantom; (b) STV; (c) STF. Coronal view: (d) Phantom; (e) STV; (f) STF. Transverse view: (g) Phantom; (h) STV; (i) STF.

IV. RESULTS

With the simulation data, the phantom images, the STV results, and the STF results are displayed in Fig. 1 (the exhale phase) and Fig. 2 (the inhale phase). The relative differences between the reconstructed image X and the phantom image X_0 through $\sum_{ijkm}|X\text{-}X_0|/\sum_{ijkm}|X_0|$, were 8.0% for STV and 3.5% for STF. With the experimental data, the STV results and the STF results are displayed in Fig. 3 (Phase 8) and Fig. 4 (Phase 16). Both the simulation and experimental results (Fig. 1–4) suggest that STF enables the 30-s 200° full-fan 4DCBCT scan, and it offers the superior image quality than STV.

In terms of the temporal resolution, the reconstructed 4D images had 20 phases, while the maximally allowed phases for the setting were roughly 40. On the other hand, the conventional 4DCBCT reconstruction often has roughly ten phases. In this sense, STF is able to improve the temporal resolution of the reconstructed 4D image, or equivalently reduce the number of projections for the same temporal resolution. The further improvement of the 4D reconstruction with the maximal phases or the minimal number of projections will be explored.

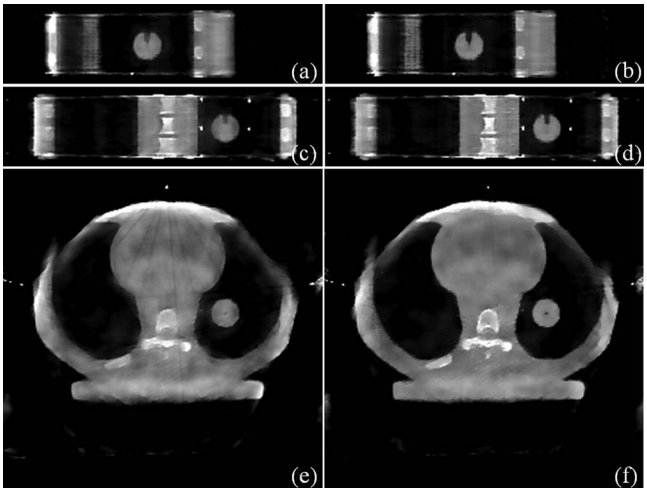

FIG. 3. Experimental results (Phase 8). Sagittal view: (a) STV; (b) STF. Coronal view: (c) STV; (d) STF. Transverse view: (e) STV; (f) STF.

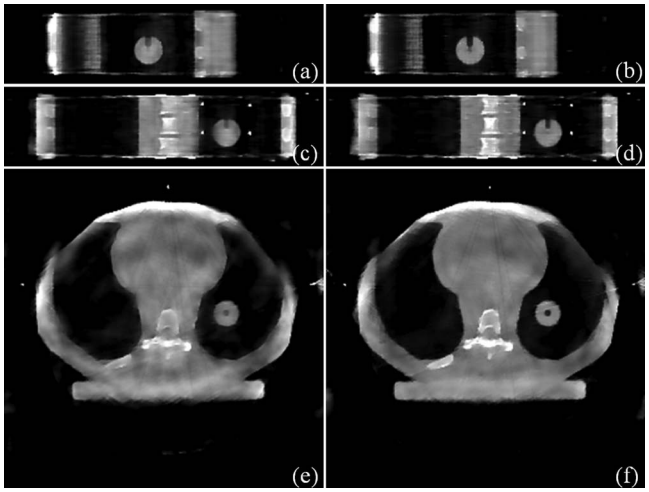

FIG. 4. Experimental results (Phase 16). Sagittal view: (a) STV; (b) STF. Coronal view: (c) STV; (d) STF. Transverse view: (e) STV; (f) STF.

The GPU-based reconstruction was implemented with a NVIDIA Tesla C2070 GPU card (448 cores and 5.25 GB device memory). The required device memory was less than 1 GB, while the total computer memory was roughly 5 GB (mainly due to STF). Two most time-consuming components are the x-ray transform (∼2 s each) and its adjoint (∼13 s each). Since the image reconstruction takes roughly 30 iterations and each iteration involves one pair of the x-ray transform and its adjoint, the total computational time for a 4D reconstruction via STF is currently less than 10 min. Further acceleration is possible with the GPU card with more cores or the multi-GPU parallelization, since the computational cost of the used x-ray transform and its adjoint per parallel thread is O(1).[16]

V. CONCLUSION

We have proposed a novel 4DCBCT image reconstruction technique based on STF, a high-dimensional tensor generalization of the 1D framelet that effectively utilizes the spatiotemporal coherence of 4DCBCT images in a multilevel and fashion with multibasis sparsifying transform. The proposed STF method potentially enables fast- and low-dose 4DCBCT with improved image quality. The STF codes are available online.[19]

ACKNOWLEDGMENT

This work is partially supported by National Institutes of Health (NIH) Grant No. EB013387 and NIH Grant No. CA166186.

APPENDIX: 1D FRAMELET TRANSFORM AND ITS ADJOINT

In this study, W_1 is the 1D piecewise-linear framelet or tight frame transform (TF) constructed by three following

refinement masks:

$$w_0 = \frac{1}{4}[1\ 2\ 1], \quad w_1 = \frac{\sqrt{2}}{4}[1\ 0\ -1], \quad w_2 = \frac{1}{4}[-1\ 2\ -1]. \tag{A1}$$

Please note that w_0, w_1, and w_2 correspond to the averaging operator, the first-order differencing operator, and the second-order differencing operator, respectively. Therefore, the TF transform contains the total variation (TV) transform on multilevels.

Let x be a 1D vector. The multilevel 1D TF transform of x up to L levels is

$$W_1 x = \Bigg[\underbrace{[W_1x]_0^1\ [W_1x]_1^1\ [W_1x]_2^1}_{\text{Level:1}} \cdots \underbrace{[W_1x]_0^l\ [W_1x]_1^l\ [W_1x]_2^l}_{\text{Level:}l} \cdots \underbrace{[W_1x]_0^L\ [W_1x]_1^L\ [W_1x]_2^L}_{\text{Level:}L}\ x^L \Bigg], \tag{A2}$$

where the fine-to-coarse convolutions are defined by the following with $x^0 = x$:

$$x^l = w_0^l * x^{l-1}, \quad [W_1x]_0^l = w_0^l * x^l, \quad [W_1x]_1^l = w_1^l * x^l,$$
$$[W_1x]_2^l = w_2^l * x^l, \quad 1 \le l \le L. \tag{A3}$$

Here for the efficient multilevel decomposition, instead of downsampling x^l, we dilute the masks (A1) to $w^l{}_i$, e.g.,

$$w_0^l = \frac{1}{4}[1\ \underbrace{0 \cdots 0}_{2^l-1}\ 2\ \underbrace{0 \cdots 0}_{2^l-1}\ 1]. \tag{A4}$$

And its transpose ${W_1}^T$ is defined as

$$W_1^T(W_1x) = x^L + \sum_{l=1}^{L} \left([W_1x]_0^l + [W_1x]_1^l + [W_1x]_2^l\right). \tag{A5}$$

Consequently, the 1D TF norm is the sum of the L1 norm of the TF terms at multilevels with multifilters, i.e.,

$$\|W_1x\|_1 = \|x^L\|_1 + \sum_{l=1}^{L} \left\{ \left\|[W_1x]_0^l\right\|_1 + \left\|[W_1x]_1^l\right\|_1 + \left\|[W_1x]_2^l\right\|_1 \right\}, \tag{A6}$$

Also, it is worthy to mention that, in terms of the computation speed, the algorithm with TF regularization is in general faster than using the transform without left inverse, such as TV, since it would require additional computation of W^TW otherwise.

[a]Electronic mail: hao.gao@emory.edu

[1] J.-J. Sonke, L. Zijp, P. Remeijer, and M. van Herk, "Respiratory correlated cone beam CT," Med. Phys. **32**, 1176–1186 (2005).

[2] S. Kriminski, M. Mitschke, S. Sorensen, N. Wink, P. Chow, S. Tenn, and T. Solberg, "Respiratory correlated cone-beam computed tomography on an isocentric C-arm," Phys. Med. Biol. **50**, 5263–5280 (2005).

[3] L. Dietrich, S. Jetter, T. Tücking, S. Nill, and U. Oelfke, "Linac-integrated 4D cone beam CT: First experimental results," Phys. Med. Biol. **51**, 2939–2952 (2006).

[4] T. Li, L. Xing, P. Munro, C. McGuinness, M. Chao, Y. Yang, B. Loo, and A. Koong, "Four-dimensional cone-beam computed tomography using an on-board imager," Med. Phys. **33**, 3825–3833 (2006).

[5] L. Xing, B. Thorndyke, E. Schreibmann, Y. Yang, T. F. Li, G. Y. Kim, G. Luxton, and A. Koong, "Overview of image-guided radiation therapy," Med. Dosim. **31**, 91–112 (2006).

[6] T. Li, A. Koong, and L. Xing, "Enhanced 4D cone-beam CT with inter-phase motion model," Med. Phys. **34**, 3688–3695 (2007).

[7] S. Leng, J. Tang, J. Zambelli, B. Nett, R. Tolakanahalli, and G. Chen, "High temporal resolution and streak-free four-dimensional cone-beam computed tomography," Phys. Med. Biol. **53**, 5653–5673 (2008).

[8] Q. Zhang, Y. Hu, F. Liu, K. Goodman, K. E. Rosenzweig, and G. S. Mageras, "Correction of motion artifacts in cone-beam CT using a patient-specific respiratory motion model," Med. Phys. **37**, 2901–2909 (2010).

[9] X. Jia, Y. Lou, B. Dong, Z. Tian, and S. Jiang, "4D computed tomography reconstruction from few-projection data via temporal non-local regularization," Lect. Notes Comput. Sci. **6361**, 143–150 (2010).

[10] H. Gao, J. F. Cai, Z. Shen and H. Zhao, "Robust principal component analysis-based four-dimensional computed tomography," Phys. Med. Biol. **56**, 3181–3198 (2011).

[11] H. Gao, H. Yu, S. Osher, and G. Wang, "Multi-energy CT based on a prior rank, intensity and sparsity model (PRISM)," Inverse Probl. **27**, 115012 (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. **57**, 1517–1525 (2012).

[13] H. Wu, A. Maier, R. Fahrig, and J. Hornegger, "Spatial-temporal total variation regularization (STTVR) for 4D-CT reconstruction," Proc. SPIE **8313**, (2012).

[14] T. Goldstein and S. Osher, "The split Bregman algorithm for l_1 regularized problems," SIAM J. Imaging Sci. **2**, 323–343 (2009).

[15] J. F. Cai, S. Osher, and Z. Shen, "Split Bregman methods and frame based image restoration," Multiscale Model. Simul. **8**,337–369 (2009).

[16] H. Gao, "Fast parallel algorithms for the x-ray transform and its adjoint," Med. Phys. (in press).

[17] W. P. Segars, D. S. Lalush, and B. M. W. Tsui, "A realistic spline-based dynamic heart phantom," IEEE Trans. Nucl. Sci. **46**, 503–506 (1999).

[18] http://www.cirsinc.com/products/modality/18/dynamic-thorax-phantom/.

[19] https://sites.google.com/site/spatiotemporaltensorframelet.