

Low-dose 4DCT reconstruction via temporal nonlocal means^{a)}

Zhen Tian

Department of Biomedical Engineering, Tsinghua University, Beijing 100084, China; Graduate School at Shenzhen, Tsinghua University, Shenzhen 518055, China; and Center for Advanced Radiotherapy Technologies, University of California San Diego, La Jolla, California 92093-0843

Xun Jia

Department of Radiation Oncology and Center for Advanced Radiotherapy Technologies, University of California San Diego, La Jolla, California 92093-0843

Bin Dong

Department of Mathematics, University of California San Diego, La Jolla, California 92093-0112

Yifei Lou

Department of Mathematics, University of California Los Angeles, Los Angeles, California 90095

Steve B. Jiang^{b)}

Department of Radiation Oncology and Center for Advanced Radiotherapy Technologies, University of California San Diego, La Jolla, California 92093-0843

(Received 6 September 2010; revised 6 January 2011; accepted for publication 6 January 2011; published 16 February 2011)

Purpose: Four-dimensional computed tomography (4DCT) has been widely used in cancer radiotherapy for accurate target delineation and motion measurement for tumors in the thorax and upper abdomen areas. However, its prolonged scanning duration causes a considerable increase of radiation dose compared to conventional CT, which is a major concern in its clinical application. This work is to develop a new algorithm to reconstruct 4DCT images from undersampled projections acquired at low mA s levels in order to reduce the imaging dose.

Methods: Conventionally, each phase of 4DCT is reconstructed independently using the filtered backprojection (FBP) algorithm. The basic idea of the authors' new algorithm is that by utilizing the common information among different phases, the input information required to reconstruct the image of high quality, and thus the imaging dose, can be reduced. The authors proposed a temporal nonlocal means (TNLM) method to explore the interphase similarity. All phases of the 4DCT images are reconstructed simultaneously by minimizing a cost function consisting of a data fidelity term and a TNLM regularization term. The authors utilized a modified forward-backward splitting algorithm and a Gauss–Jacobi iteration method to efficiently solve the minimization problem. The algorithm was also implemented on a graphics processing unit (GPU) to improve the computational speed. The authors' reconstruction algorithm has been tested on a digital NCAT thorax phantom in three low dose scenarios: All projections with low mA s level, undersampled projections with high mA s level, and undersampled projections with low mA s level.

Results: In all three low dose scenarios, the new algorithm generates visually much better CT images containing less image noise and streaking artifacts compared to the standard FBP algorithm. Quantitative analysis shows that by comparing the authors' TNLM algorithm to the standard FBP algorithm, the contrast-to-noise ratio has been improved by a factor of 3.9–10.2 and the signal-to-noise ratio has been improved by a factor of 2.1–5.9, depending on the cases. In the situation of undersampled projection data, the majority of the streaks in the images reconstructed by FBP can be suppressed using the authors' algorithm. The total reconstruction time for all ten phases of a slice ranges from 40 to 90 s on an NVIDIA Tesla C1060 GPU card.

Conclusions: The experimental results indicate that the authors' new algorithm outperforms the conventional FBP algorithm in effectively reducing the image artifacts due to undersampling and suppressing the image noise due to the low mA s level. © 2011 American Association of Physicists in Medicine. [DOI: [10.1118/1.3547724](https://doi.org/10.1118/1.3547724)]

Key words: 4DCT reconstruction, dose reduction, temporal nonlocal means, GPU

I. INTRODUCTION

Four-dimensional computed tomography (4DCT) has been widely used for treatment simulation in radiotherapy of tumors with respiratory motion,^{1–6} in that it can provide time-resolved volumetric images. Currently, there are two differ-

ent methods for 4DCT acquisition and sorting, namely, retrospective slice sorting^{1–6} and prospective sinogram selection.⁷ For the retrospective slice sorting method, the projection data are continuously acquired at each couch position for a time interval slightly longer than a full respira-

tory cycle, either in cine mode or in helical mode with a very low pitch.¹⁻⁶ Multiple slices corresponding to different acquisition time points are reconstructed at each couch position and then sorted into respiratory phase bins using various respiratory signals.^{1-6,8-10} For prospective sinogram selection method, the CT scanner is triggered by the respiratory signal for projection data acquisition.⁷ The projection data within the same phase bin are used to reconstruct CT slices corresponding to that breathing phase.

No matter which method is used, the reconstructed CT slices at different breathing phases represent a set of time-resolved volumetric images, which are called 4DCT images and can resolve the organ motions and reduce motion artifacts to a satisfactory extent.¹¹ However, the prolonged acquisition time results in a considerably increased radiation dose. For example, given the standard 4DCT parameters (140 kVp, 95 mA, and 0.5 s per rotation), the radiation dose of a 4DCT scan is about six times of that of a typical helical CT scan.¹² This fact has become a major concern in the clinical application of 4DCT and thus it is highly desirable to reduce its imaging dose.

Intuitively, the radiation dose can be reduced by lowering the mA s level and/or decreasing the number of acquired projections [decreasing the number of projections is not straightforward on currently available commercial CT scanners due to the use of continuous x-ray generation mode. However, technically, it is possible to modify the scanners to operate in high-frequency pulsed mode¹³⁻¹⁵ if there is a clinical need (such as the one suggested in this paper)]. However, these approaches in general will lead to amplified image noise and obvious streaking artifacts in the reconstructed 4DCT images if conventional filtered backprojection (FBP) algorithm is used. In the current FBP-based 4DCT reconstruction algorithm, each phase of the 4DCT images is reconstructed independently based on the acquired projection data associated with it. This method, though simple, completely neglects the highly temporal correlation between 4DCT images at successive phases, as apparently the same anatomical features exist in successive phases of 4DCT images with slight motion and deformation. It is expected that taking into account this information during reconstruction, one can reconstruct 4DCT images with high quality, even at the low dose contexts of low mA s level and/or under-sampled projection data.

Inspired by this idea, in this paper we propose a new 4DCT reconstruction algorithm by utilizing the aforementioned temporal correlation between images at successive phases. In our algorithm, CT slices corresponding to different phases are reconstructed simultaneously as opposed to independently in conventional FBP-type algorithms. In particular, the temporal regularization is imposed between successive phases via a temporal nonlocal means (TNLM) term to take the interphase correlation into account. Our idea of the TNLM function is inspired by the so-called nonlocal means (NLM) method¹⁶ originated in image processing field. The NLM method assumes that there are lots of repetitive structures contained in an image and thus utilizes the similar image features at different spatial locations in the same im-

age to constructively enhance each other.¹⁷ This assumption, though valid in many cases in image processing problems, may not hold in medical images. However, it is reasonable to believe that along temporal direction, there exist repetitive features in the 4DCT images. Thus, we propose the novel TNLM method which extends the original NLM method into the temporal domain by exploring the similarity between neighboring phases in the context of 4DCT reconstruction.

The rest of this paper is organized as follows. In Sec. II we will present the new 4DCT reconstruction method. The experimental results with this new method will be given in Sec. III. We then give discussion and conclusions in Sec. IV.

II. METHODS

II.A. Reconstruction model

Let us divide a respiratory cycle into N phases labeled by $i=1, 2, \dots, N$. The 4DCT image of phase i is denoted by a vector f_i . P_i is the projection matrix of phase i that maps the image f_i into a set of projections corresponding to various projection angles. The measured set of projections is denoted by a vector y_i . We attempt to reconstruct the 4DCT images by solving the following optimization problem:

$$\{f_i\} = \arg \min_{\{f_i\}} \sum_{i=1}^N \left\{ (P_i f_i - y_i)^T \Sigma^{-1} (P_i f_i - y_i) + \frac{\mu}{2} [J(f_i, f_{i-1}) + J(f_i, f_{i+1})] \right\}, \quad (1)$$

where the first term in the summation is a data fidelity term, ensuring that the projections of the reconstructed 4DCT image at each phase match the corresponding observed projections. The symbol T denotes the matrix transpose. The covariance matrix Σ is a diagonal matrix with its nonzero elements corresponding to the variance of the pixel values of the measured projection images.¹⁸ The second term in Eq. (1), $J(\cdot, \cdot)$, is the regularization term and the parameter μ adjusts the relative weight between the data fidelity term and the regularization term.

In this paper, we propose a new TNLM function as the temporal regularization imposed between neighboring phases to explore the interphase similarity. A periodic boundary condition along the temporal direction is assumed, e.g., $f_{N+1} = f_1$. For two images at different phases f_i and f_j , $J(f_i, f_j)$ is defined as

$$J(f_i, f_j) = \sum_x \sum_y [f_i(x) - f_j(y)]^2 \omega_{i,j}^*(x, y). \quad (2)$$

The weighting factor $\omega_{i,j}^*(x, y)$ is independent of $f_i(x)$ and $f_j(y)$ but defined according to the ground truth images $f_i^*(x)$ and $f_j^*(y)$ as

$$\omega_{i,j}^*(x, y) = \frac{1}{Z} \exp \left[- \frac{\|R_i^*(x) - R_j^*(y)\|_2^2}{h^2} \right], \quad (3)$$

where $R_i^*(x)$ denotes a square patch on the image f_i^* centering at a pixel x and $\|R_i^*(x) - R_j^*(y)\|_2^2$ is the L_2 -norm of the difference between $R_i^*(x)$ and $R_j^*(y)$. Z is a normalization param-

eter such that $\sum_y \omega_{i,j}^*(x,y) = 1$. h is a parameter that adjusts to what extent we would like to enforce the similarity between patches.

II.B. Optimization approach

To solve the optimization problem in Eq. (1), we implement a forward-backward splitting algorithm,^{19,20} where the solution to Eq. (1) can be obtained by alternatively performing the following two steps until convergence:

$$g_i^{(k)} = f_i^{(k-1)} - \frac{1}{\beta} P_i^T \Sigma^{-1} (P_i f_i^{(k-1)} - y_i), \forall i, \quad (4)$$

$$\begin{aligned} \{f_i^{(k)}\} &= \arg \min_{\{f_i\}} E_1[\{f_i\}] \\ &= \arg \min_{\{f_i\}} \sum_{i=1}^N \left\{ \|f_i - g_i^{(k)}\|_2^2 \right. \\ &\quad \left. + \frac{\mu}{2\beta} [J(f_i, f_{i-1}) + J(f_i, f_{i+1})] \right\}, \end{aligned} \quad (5)$$

where the superscript k is the index for iteration steps. g_i is an auxiliary vector and $\beta > 0$ is a constant introduced by the splitting algorithm. Note that Eq. (4) is actually one step of the gradient descent algorithm toward a problem minimizing an energy function $E_2[f_i] = (P_i f_i - y_i)^T \Sigma^{-1} (P_i f_i - y_i)$. The introduction of β actually controls the step size of the gradient descent algorithm for numerical stability purpose. In practice, it is found that by substituting this one step gradient descent with a conjugate gradient least square (CGLS) method²¹ for the minimization of $E_2[f_i]$, the overall convergence can be enhanced, although the convergence after this modification is not mathematically proven. The CGLS algorithm by itself is an iterative algorithm. In each iteration k , we use the images obtained from the last iteration, i.e., $f_i^{(k-1)}$ as an initial guess. Let us denote $P_{ni} = \Sigma^{-1/2} P_i$ and $y_{ni} = \Sigma^{-1/2} y_i$. The detailed implementation of this CGLS algorithm is performed as follows:

CGLS Algorithm:

Initialize: $m=0$, $u_i^{(0)} = f_i^{(k-1)}$, $r_i^{(0)} = y_{ni} - P_{ni} u_i^{(0)}$, $s_i^{(0)} = P_{ni}^T r_i^{(0)}$.

Do the Steps 1–5 for M_1 times.

1. $a_i^{(m)} = \|P_{ni}^T r_i^{(m)}\|_2^2 / \|P_{ni} s_i^{(m)}\|_2^2$
2. $u_i^{(m+1)} = u_i^{(m)} + a_i^{(m)} s_i^{(m)}$
3. $r_i^{(m+1)} = r_i^{(m)} - a_i^{(m)} P_{ni} s_i^{(m)}$
4. $b_i^{(m)} = \|P_{ni}^T r_i^{(m+1)}\|_2^2 / \|P_{ni}^T r_i^{(m)}\|_2^2$
5. $s_i^{(m+1)} = P_{ni}^T r_i^{(m+1)} + b_i^{(m)} s_i^{(m)}$
6. $g_i^{(k)} = u_i^{(M_1)}$

Here the superscript m is the iteration step for the CGLS algorithm. In practice, it is not necessary to carry out this CGLS algorithm very precisely in each outer loop k , since the purpose of this CGLS step is only to generate a better solution $g_i^{(k)}$ based on the input $f_i^{(k-1)}$. Therefore, the iteration steps of the CGLS algorithm M_1 is chosen to be a small integer, such as $M_1 = 3$.

To solve the subproblem in Eq. (5), let us first take functional variation of $E_1[\{f_i\}]$ with respect to $f_i(x)$. Note that the weighting factor $\omega_{i,j}^*(x,y)$ is a constant defined according to the ground truth images $f_i^*(x)$ and $f_j^*(y)$. We arrive at

$$\begin{aligned} \frac{\delta E_1}{\delta f_i(x)} &= 2(f_i - g_i^{(k)}) + \hat{\mu} \left(f_i - \sum_y f_{i-1}(y) \omega_{i,i-1}^*(x,y) \right) \\ &\quad + \hat{\mu} \left(f_i - \sum_y f_{i+1}(y) \omega_{i,i+1}^*(x,y) \right), \end{aligned} \quad (6)$$

where $\hat{\mu} = \mu / \beta$. By setting this variation to be zero and rearrange different terms, we obtain the optimality condition as

$$\begin{aligned} f_i &= \frac{1}{\hat{\mu} + 1} g_i^{(k)} + \frac{\hat{\mu}}{2\hat{\mu} + 2} \left[\sum_y f_{i-1} \omega_{i,i-1}^*(x,y) \right. \\ &\quad \left. + \sum_y f_{i+1} \omega_{i,i+1}^*(x,y) \right]. \end{aligned} \quad (7)$$

This equation leads to a Gauss–Jacobi type iteration scheme²² for solving the problem in Eq. (5)

Gauss–Jacobi iteration Algorithm:

Initialize $m=0$ and $\mu_i^{(0)} = g_i^{(k)}$. Do the Step 1 for M_2 times.

1. $u_i^{(m+1)} = 1 / (\hat{\mu} + 1) g_i^{(k)} + \hat{\mu} / (2\hat{\mu} + 2) [\sum_y u_{i-1}^{(m)} \omega_{i,i-1}^*(x,y) + \sum_y u_{i+1}^{(m)} \omega_{i,i+1}^*(x,y)]$
2. $f_i^{(k)} = u_i^{(M_2)}$

Again the integer m here denotes iteration step. It follows from Theorem 10.1.1 in the literature²² that such an iteration scheme converges for any $\hat{\mu} > 0$. The total number of iteration steps M_2 is chosen to be a small integer, such as $M_2 = 1$.

Moreover, since the weighting factor $\omega_{i,j}^*(x,y)$ defined according to the ground truth images $f_i^*(x)$ and $f_j^*(y)$ is not known beforehand, we update the weight during the iteration according to the latest available images $g_i^{(k)}$ and $g_j^{(k)}$. In other words, we choose square patches from the reconstructed images obtained in the last iteration [$R_{g_i^{(k)}}(x)$ and $R_{g_j^{(k)}}(y)$] to calculate the weight $\omega_{i,j}(x,y)$ instead of $\omega_{i,j}^*(x,y)$.

$$\omega_{i,j}(x,y) = \frac{1}{Z} \exp[-\|R_{g_i^{(k)}}(x) - R_{g_j^{(k)}}(y)\|_2^2 / h^2]. \quad (8)$$

Additionally, a simple truncation of negative pixel values is necessary after each iteration to ensure the positivity of the reconstructed images. The TNLM algorithm can be summarized as follows:

TNLM Algorithm:

1. Initialize $f_i^{(0)}$ for $i=1, \dots, N$ to be the image reconstructed by FBP with all projections from all N phases.
2. Use CGLS with initial value $\{f_i^{(k-1)}\}$ to get $\{g_i^{(k)}\}$.
3. Update weighting parameter $\omega_{i,i-1}$, $\omega_{i,i+1}$, using $\{g_i^{(k)}\}$.
4. Get $\{f_i^{(k)}\}$ using Gaussian–Jacobi algorithm.
5. Ensure image positivity: $f_i^{(k)} = 0$ if $f_i^{(k)} < 0$.
6. Go back to step 2 until convergence.

The advantage of this TNLM algorithm is straightforward. In the k th iteration, the algorithm first obtains a better solution $g_i^{(k)}$ using the CGLS algorithm based on the solution from previous step ($f_i^{(k-1)}$). Since this step does not contain any regularization on the solution, the obtained $g_i^{(k)}$ will be contaminated by noise and various artifacts. The following TNLM step, i.e., step 4, updates the solution according to the Gauss–Jacobi algorithm, yielding a new solution $f_i^{(k)}$. In par-

ticular, $f_i^{(k)}$ is a weighted average of the input image $g_i^{(k)}$ and the images at neighboring phases $g_{i-1}^{(k)}$ and $g_{i+1}^{(k)}$. Moreover, the updated image $f_i^{(k)}$ pixel value at x depend on $g_i^{(k)}$ in a local fashion, i.e., also at pixel x but in a nonlocal fashion on the neighboring phases, i.e., at other pixels y . For those nonlocal terms, the weight is automatically adjusted according to the similarity between the patches around x in phase i and the patches around y in neighboring phases. As such, any features that repetitively appear in successive phases, such as true anatomical structures, are preserved during the iteration. In contrast, those features that do not repeat, such as streaking artifacts, are suppressed.

II.C. Implementation issues

One disadvantage of our TNLM reconstruction algorithm is its large computational burden of searching for similarity. During the implementation, for each pixel x on image f_i , a square patch $R_i(x)$ centering at this pixel on the image is compared to patches centering at all pixels y on the neighboring phase images $f_{i\pm 1}$ for computing the weighting factor $\omega_{i,i\pm 1}(x,y)$. Suppose the patches used for computing this weighting factor are two-dimensional squares with d pixels in each dimension. Then the total complexity of the searching scheme between two images is in the order of $O(L^2 \times L^2 \times d^2)$, where L is the size of the 4D-CT slices in each dimension. However, this approach is neither computationally efficient nor necessary. In fact, for a patch at location x on phase i , the patches similar to it on neighboring phases must locate within a neighborhood of x due to the finite motion range of respiratory motion. Therefore, it is adequate to restrict the search for the similar patches only within a search window. In practice, we set this search window to be a square region with W pixels in each dimension. Since it is usually true that $W \ll L$, this searching window can reduce the computation load down to $O(L^2 \times W^2 \times d^2)$.

Another technique we use to speed up the calculation is removing the redundant calculations in computing the weighting factor. Note that the weights $\omega_{i,j}(x,y)$ and $\omega_{j,i}(y,x)$ are actually same before normalization. Reusing this factor in the Gauss–Jacobi algorithm instead of recomputing it can reduce the computational load by a factor of about 2.

Besides, we also implement the 4DCT reconstruction algorithm on an NVIDIA Tesla C1060 card to speed up the computation. This graphics processing unit (GPU) card has a total number of 240 processor cores (grouped into 30 multi-processors with 8 cores each), each with a clock speed of 1.3 GHz. It is also equipped with 4 GB DDR3 memory, shared by all processor cores. We simply have each GPU thread responsible for 1 pixel of the CT slices. Because of the large number of GPU threads, the computation efficiency can be considerably elevated.

III. EXPERIMENTAL RESULTS

We tested our reconstruction algorithm on a digital NCAT phantom at a thorax region.²³ The x-ray projections were simulated using Siddon's algorithm²⁴ in fan-beam geometry

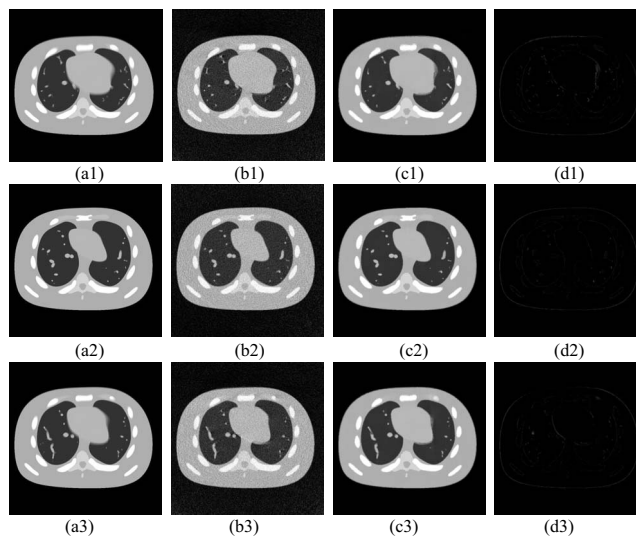


FIG. 1. The reconstruction 4DCT images for the 40% phase (end point of inhale) (top), the 70% phase (midpoint of exhale) (middle), and the 100% phase (end point of exhale) (bottom) with all projections at 20 mA s level (case 1). (a) Ground truth, (b) FBP, (c) TNLM, and (d) difference between ground truth and TNLM.

with an arc detector of 888 units and a spacing of 1.0239 mm. The source to detector distance is 949.075 mm and the source to rotation center distance is 541.0 mm. All these parameters mimic the realistic configuration of a GE Light-speed QX/I CT scanner. The gantry's rotation speed is set to be 0.5 s/rotation. The period of the respiratory cycle is 4 s. Our 4DCT reconstruction algorithm is tested on three low dose cases: Case 1: All projections acquired with low mA s protocol; case 2: Undersampled projections acquired with high mA s protocol; and case 3: Undersampled projections acquired with low mA s protocol.

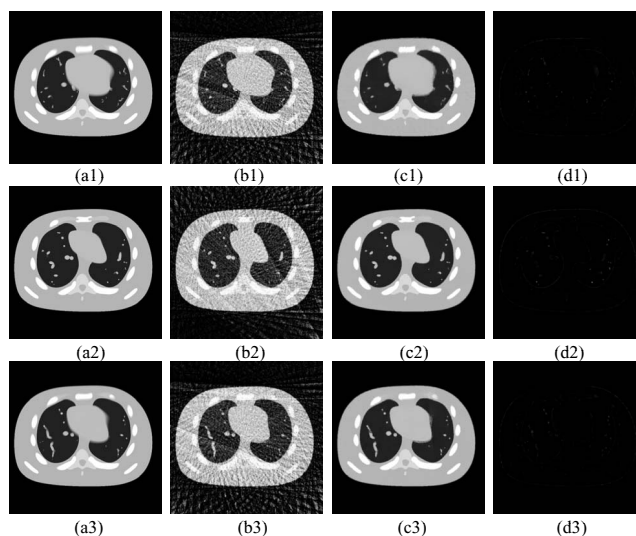


FIG. 2. The reconstruction 4DCT images for the 40% phase (end point of inhale) (top), the 70% phase (midpoint of exhale) (middle), and the 100% phase (end point of exhale) (bottom) with 500 projections at 100 mA s level (case 2). (a) Ground truth, (b) FBP, (c) TNLM, and (d) difference between ground truth and TNLM.

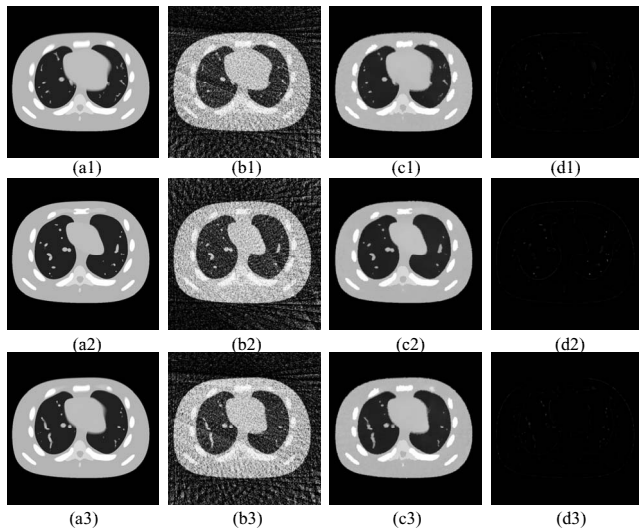


FIG. 3. The reconstruction 4DCT images for the 40% phase (end point of inhale) (top), the 70% phase (midpoint of exhale) (middle), and the 100% phase (end point of exhale) (bottom) with 500 projections at 20 mA s level (case 3). (a) Ground truth, (b) FBP, (c) TNLM, and (d) difference between ground truth and TNLM.

To simulate the noise-contaminated sinogram at low mA s situations, we add Gaussian noise signal to the noise free projection data, where the variance at a given entry i of the projection is taken as^{25,26}

$$\sigma_i^2 = \frac{1}{N_{0i}} \exp(\bar{p}_i), \quad (9)$$

where \bar{p}_i is the projection data value before adding noise. N_{0i} represents the average photon number just before entering the patient body, which is derived from the measurements at a certain mAs level. The variance computed as such also goes into the matrix Σ in Eq. (1).

In case 1, we first generated 4000 noise-contaminated projections at 20 mA s for ten phases and all of them are used for reconstruction. The reconstructed images at three phases, namely, the end of inhale, the middle of the exhale, and the end of exhale are shown in Fig. 1. The reconstruction results of the conventional FBP algorithm are also shown for comparison purpose. By visually inspecting, it is clear that our TNLM algorithm outperforms the FBP algorithm by greatly reducing the image noise. The blurring effect shown in the top and bottom row images is due to the residual heart motion within the respiratory phase bin. More discussion about this effect will be given in Sec. IV.

As for case 2, we generated 500 projections at 100 mA s to simulate the undersampling situation with high mA s

level. Case 3 is same as case 2 except the mA s is lowered down to 20 mA s to simulate the undersampling situation with low mA s level. The results of these two cases are shown in Figs. 2 and 3, respectively. Obvious streaking artifacts can be observed in the CT images reconstructed by the FBP algorithm due to undersampling [see Figs. 2(b)(1)–2(b)(iii) and Figs. 3(b)(1)–3(b)(iii)]. Moreover, the FBP results of case 3 are noisier than that of case 2 due to the low mA s level. On the other hand, our TNLM algorithm can still maintain the image quality to a satisfactory level in both cases. From the difference images between the ground truth image and TNLM result, it is found that the spatial resolution is degraded to a certain extent with the TNLM method and thus some small structures are lost, especially at the undersampling situation in cases 2 and 3.

To quantitatively evaluate the performance of our reconstruction algorithm in terms of maintaining image contrast and suppressing image noise, we calculate the contrast-to-noise ratio (CNR) and signal-to-noise ratio (SNR) which are defined as follows:

$$\text{CNR} = |\mu_s - \mu_b| / \sigma_b, \quad (10)$$

$$\text{SNR} = 10 \log_{10} \{ \|f - \bar{f}\|_2^2 / \|f - f^*\|_2^2 \}. \quad (11)$$

The CNR is defined on a given region of interest (ROI). μ_s is the mean value of the signal for the ROI, while μ_b and σ_b are the mean value and standard deviation of the nearby background. SNR is a quantity to measure the overall deviation of the reconstructed 4DCT image from the ground truth image. \bar{f} is the mean value of an image f and f^* is the ground truth image.

Table I lists the mean values of CNRs over ten phases of the two ROIs shown in Fig. 4. ROI1 is a small high contrast structure with the nearby dark lung region as its background. ROI2 is the vertebral body with the neighboring tissue as its background. In all three cases, our TNLM algorithm has improved the CNR by a factor of 3.9–10.2 depending on the ROIs and cases compared to the conventional FBP algorithm. In particular, for the FBP algorithm, all the CNRs of ROI2 are very small and thus the low contrast structures is hardly resolved, while for our TNLM algorithm, the corresponding CNRs are much larger. The results of CNRs illustrate that our algorithm is superior to the FBP algorithm in maintaining good contrast under low dose contexts. Table II lists the mean SNRs over ten phases, in which our TNLM algorithm has improved the SNRs by a factor of 2.1–5.9,

TABLE I. Mean CNR values over ten phases of the two ROIs in Fig. 4.

	Case1		Case2		Case3	
	ROI1	ROI2	ROI1	ROI2	ROI1	ROI2
FBP	6.83	1.80	4.09	1.13	2.56	0.69
TNLM	28.64	12.15	16.04	7.53	19.53	7.07

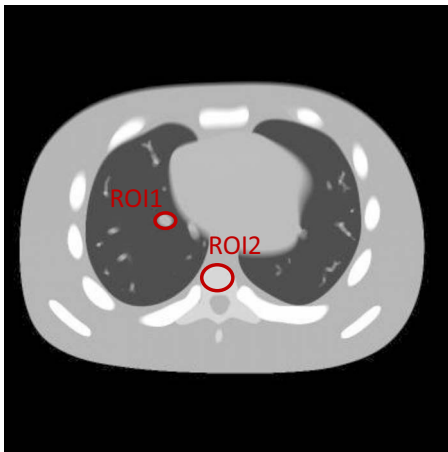


FIG. 4. Two ROIs used for CNR calculation.

indicating that our algorithm outperforms the FBP algorithm in effectively suppressing the image noise at low dose situation.

IV. DISCUSSION AND CONCLUSIONS

In this paper, we have presented a novel iterative 4DCT reconstruction algorithm via temporal regularization. The 4DCT images of different phases are reconstructed simultaneously by minimizing an energy function consisting of a data fidelity term and a temporal regularization term between every two neighboring phases. A temporal nonlocal means method is employed to take the temporal correlation of the 4DCT images into account. We utilized a modified forward-backward splitting algorithm to perform the optimization. The iterative reconstruction algorithm is implemented on a GPU platform to improve its efficiency. Our algorithm is tested on a digital NCAT phantom under low dose context by lowering the mA s level and/or decreasing the number of projections. The experimental results indicate that our algorithm performs much better than the conventional FBP algorithm in effectively reducing the image artifacts due to undersampling and suppressing the image noise due to the low mA s level. Specifically, the contrast-to-noise ratio has been improved by a factor of 3.9–10.2 and the signal-to-noise ratio has been improved by a factor of 2.1–5.9, depending on the cases. The total reconstruction time ranges from 40 to 90 s on a NVIDIA Tesla C1060 card (NVIDIA, Santa Clara, CA) for ten phases at a transverse slice. This reconstruction time may not meet the requirement for some clinical applications, yet the efficiency of our algorithm could be further increased by using some advanced technologies such as multigrid algorithm and multiple GPUs.

TABLE II. Mean SNR values over ten phases.

	Case1	Case2	Case3
FBP	10.49	6.03	3.39
TNLM	22.13	21.12	20.28

The blurred heart edges in some panels of Figs. 1–3 are caused by the residual heart motion within the respiratory phase bin. The breathing period used in this work is 4 s. Then, each of ten breathing phase bin covers 0.4 s, within which the respiratory motion is negligible but the heart motion is not. For example, we observe apparent blurring effect in Figs. 1(c)(i) and 1(c)(iii), but not Fig. 1(c)(ii); this is because the heart in Figs. 1(c)(i) and 1(c)(iii) is in the middle of systole/diastole, while in Fig. 1(c)(ii), it happens to be at the end of systole. The blurring effect also presents in the ground truth images, as shown in Figs. 1(a)(i) and 1(a)(iii), because we average over 400 images within a given breathing phase bin to produce the ground truth CT image at that phase. The blurring effect due to the residual heart motion also exists in the FBP images, which, however, is less visible due to the image noise and the streak artifact.

In this feasibility study, we have only quantitatively evaluated CNR and SNR. Other quantitative measures, such as spatial resolution of the images, are not considered in the current work. It is found that the spatial resolution is degraded to a certain extent with the TNLM method, especially at the undersampling situation in cases 2 and 3. We plan to conduct much more systematic and quantitative evaluations using a large set of representative patient images and a set of clinically relevant metrics in our future work.

ACKNOWLEDGMENTS

This work is supported in part by the University of California Laboratory Fees Research Program. The authors would like to thank NVIDIA for providing GPU cards for this project.

- ^{a)}Zhen Tian and Xun Jia have contributed equally to this work and should be considered co-first authors.
- ^{b)}Author to whom correspondence should be addressed. Electronic mail: sbjiang@ucsd.edu
- ¹E. Ford, G. Mageras, E. Yorke, and C. Ling, “Respiration-correlated spiral CT: A method of measuring respiratory-induced anatomic motion for radiation treatment planning,” *Med. Phys.* **30**, 88–97 (2003).
- ²P. Keall, G. Starkschall, H. Shukla, K. Forster, V. Ortiz, C. Stevens, S. Vedam, R. George, T. Guerrero, and R. Mohan, “Acquiring 4D thoracic CT scans using a multislice helical method,” *Phys. Med. Biol.* **49**, 2053–2067 (2004).
- ³D. Low, M. Nystrom, E. Kalinin, P. Parikh, J. Dempsey, J. Bradley, S. Mutic, S. Wahab, T. Islam, and G. Christensen, “A method for the reconstruction of four-dimensional synchronized CT scans acquired during free breathing,” *Med. Phys.* **30**, 1254–1263 (2003).
- ⁴T. Pan, T. Lee, E. Rietzel, and G. Chen, “4D-CT imaging of a volume influenced by respiratory motion on multi-slice CT,” *Med. Phys.* **31**, 333–340 (2004).
- ⁵E. Rietzel, T. Pan, and G. Chen, “Four-dimensional computed tomography: Image formation and clinical protocol,” *Med. Phys.* **32**, 874–889 (2005).
- ⁶S. Vedam, P. Keall, V. Kini, H. Mostafavi, H. Shukla, and R. Mohan, “Acquiring a four-dimensional computed tomography dataset using an external respiratory signal,” *Phys. Med. Biol.* **48**, 45–62 (2003).
- ⁷U. Langner and P. Keall, “Prospective displacement and velocity-based cine 4D CT,” *Med. Phys.* **35**, 4501–4512 (2008).
- ⁸R. Li, J. Lewis, L. Cervino, and S. Jiang, “4D CT sorting based on patient internal anatomy,” *Phys. Med. Biol.* **54**, 4821–4833 (2009).
- ⁹W. Lu, P. Parikh, J. Hubenschmidt, J. Bradley, and D. Low, “A comparison between amplitude sorting and phase-angle sorting using external respiratory measurement for 4D CT,” *Med. Phys.* **33**, 2964–2974 (2006).
- ¹⁰N. Wink, C. Panknin, and T. Solberg, “Phase versus amplitude sorting of

- 4D-CT data," *J. Appl. Clin. Med. Phys.* **7**, 77–85 (2006).
- ¹¹P. Keall, "4-dimensional computed tomography imaging and treatment planning," *Semin. Radiat. Oncol.* **14**, 81–90 (2004).
- ¹²J. van Sörnsen de Koste, S. Senan, C. Kleynen, B. Slotman, and F. Lagerwaard, "Renal mobility during uncoached quiet respiration: An analysis of 4DCT scans," *Int. J. Radiat. Oncol., Biol., Phys.* **64**, 799–803 (2006).
- ¹³B. Myagkov, E. Shikanov, and A. Shikanov, "Development and investigation of a pulsed x-ray generator," *At. Energy* **106**, 143–148 (2009).
- ¹⁴G. Yue, Q. Qiu, B. Gao, Y. Cheng, J. Zhang, H. Shimoda, S. Chang, J. Lu, and O. Zhou, "Generation of continuous and pulsed diagnostic imaging x-ray radiation using a carbon-nanotube-based field-emission cathode," *Appl. Phys. Lett.* **81**, 355–357 (2002).
- ¹⁵M. Kang, C. Park, C. Lee, J. Goo, and H. Lee, "Dual-energy CT: Clinical applications in various pulmonary diseases," *Radiographics* **30**, 685–698 (2010).
- ¹⁶A. Buades, B. Coll, and J. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.* **4**, 490–530 (2005).
- ¹⁷Y. Lou, X. Zhang, S. Osher, and A. Bertozzi, "Image recovery via non-local operators," *J. Sci. Comput.* **42**, 185–197 (2010).
- ¹⁸J. Wang, T. Li, Z. Liang, and L. Xing, "Dose reduction for kilovoltage cone-beam computed tomography in radiation therapy," *Phys. Med. Biol.* **53**, 2897–2909 (2008).
- ¹⁹P. Combettes and V. Wajs, "Signal recovery by proximal forward-backward splitting," *Multiscale Model. Simul.* **4**, 1168–1200 (2005).
- ²⁰E. Hale, W. Yin, and Y. Zhang, "Fixed-point continuation for l1-minimization: Methodology and convergence," *SIAM J. Optim.* **19**, 1107–1130 (2008).
- ²¹M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *J. Res. Natl. Bur. Stand.* **49**, 409–436 (1952).
- ²²G. Golub and C. Van Loan, *Matrix Computations* (Johns Hopkins University Press, Baltimore, 1996).
- ²³W. P. Segars, "Development and application of the new dynamic NURBS-based cardiac-torso (NCAT) phantom," Ph.D. dissertation, University of North Carolina, Chapel Hill, NC, 2001.
- ²⁴R. Siddon, "Fast calculation of the exact radiological path of a three-dimensional CT array," *Med. Phys.* **12**, 252–255 (1985).
- ²⁵T. Li, X. Li, J. Wang, J. Wen, H. Lu, J. Hsieh, and Z. Liang, "Nonlinear sinogram smoothing for low-dose x-ray CT," *IEEE Trans. Nucl. Sci.* **51**, 2505–2513 (2004).
- ²⁶J. Wang, H. Lu, Z. Liang, D. Eremina, G. Zhang, S. Wang, J. Chen, and J. Manzione, "An experimental study on the noise properties of x-ray CT sinogram data in Radon space," *Phys. Med. Biol.* **53**, 3327–3341 (2008).