Principal component reconstruction (PCR) for cine CBCT with motion learning from 2D fluoroscopy

Hao Gao,^{a)} Yawei Zhang, and Lei Ren

Department of Radiation Oncology, Duke University Medical Center, Durham, NC 27710, USA

Fang-Fang Yin

Department of Radiation Oncology, Duke University Medical Center, Durham, NC 27710, USA

Medical Physics Graduate Program, Duke Kunshan University, Kunshan, Jiangsu 215316, China

(Received 27 February 2017; revised 18 October 2017; accepted for publication 3 November 2017; published 11 December 2017)

Purpose: This work aims to generate cine CT images (i.e., 4D images with high-temporal resolution) based on a novel principal component reconstruction (PCR) technique with motion learning from 2D fluoroscopic training images.

Methods: In the proposed PCR method, the matrix factorization is utilized as an explicit low-rank regularization of 4D images that are represented as a product of spatial principal components and temporal motion coefficients. The key hypothesis of PCR is that temporal coefficients from 4D images can be reasonably approximated by temporal coefficients learned from 2D fluoroscopic training projections. For this purpose, we can acquire fluoroscopic training projections for a few breathing periods at fixed gantry angles that are free from geometric distortion due to gantry rotation, that is, fluoroscopy-based motion learning. Such training projections can provide an effective characterization of the breathing motion. The temporal coefficients can be extracted from these training projections and used as priors for PCR, even though principal components from training projections are certainly not the same for these 4D images to be reconstructed. For this purpose, training data are synchronized with reconstruction data using identical real-time breathing position intervals for projection binning. In terms of image reconstruction, with *a priori* temporal coefficients, the data fidelity for PCR changes from nonlinear to linear, and consequently, the PCR method is robust and can be solved efficiently. PCR is formulated as a convex optimization problem with the sum of linear data fidelity with respect to spatial principal components and spatiotemporal total variation regularization imposed on 4D image phases. The solution algorithm of PCR is developed based on alternating direction method of multipliers.

Results: The implementation is fully parallelized on GPU with NVIDIA CUDA toolbox and each reconstruction takes about a few minutes. The proposed PCR method is validated and compared with a state-of-art method, that is, PICCS, using both simulation and experimental data with the on-board cone-beam CT setting. The results demonstrated the feasibility of PCR for cine CBCT and significantly improved reconstruction quality of PCR from PICCS for cine CBCT.

Conclusion: With *a priori* estimated temporal motion coefficients using fluoroscopic training projections, the PCR method can accurately reconstruct spatial principal components, and then generate cine CT images as a product of temporal motion coefficients and spatial principal components. © 2017 American Association of Physicists in Medicine [https://doi.org/10.1002/mp.12671]

Key words: 4D CBCT, 4D CT, compressive sensing, image reconstruction, low-rank

1. INTRODUCTION

Four-dimensional (4D) CT images play an important role for thoracic and abdominal radiotherapy, including target delineation, treatment planning, and treatment delivery. The accuracy of image-guided radiotherapy (IGRT) highly depends on the degree of motion-induced artifacts in reconstructed 4D images. These artifacts cannot be reduced unless the temporal resolution is improved. This study aims at the reconstruction of 4D images with high-temporal resolution, namely cine CT, in the context of on-board cone-beam CT (CBCT),^{1–4} and the general methodology could be extended for diagnostic CT.

A major limitation of current on-board CBCT system for 4D imaging is its slow gantry rotation speed. Consequently, the acquired projections for each temporal phase are angularly sparse and undersampled. Moreover, there is a tradeoff between number of projections available to each image phase and number of phases to be reconstructed. Therefore, temporal resolution cannot be improved at no sacrifice of spatial image quality if each phase is reconstructed independently.

On the other hand, a crucial prior that can be used to improve 4D reconstruction is the temporal similarity or redundancy of 4D images, which is often enforced through some image regularization during the joint reconstruction of all 4D images. Such regularization methods can be roughly

classified into two categories. The first one is via the sparsity transform, such as reference image-based total variation (TV),⁵ spatiotemporal TV,^{6,7} tensor framelet,⁸ nonlocal mean,⁹ and nonlocal TV.¹⁰ The other one is via the low-rank model, such as nuclear norm^{11–13} and matrix factorization.¹⁴ The recent developments also involve regularization methods in the motion space, such as deformation vector field,^{15–17} PCA model,^{18,19} and 5D model.²⁰

In our previous work,¹⁴ the matrix factorization was utilized for an explicit low-rank representation of 4D images as a product of spatial principal components and temporal motion coefficients. However, the simultaneous reconstruction of spatial components and temporal coefficients was nonlinear and the image quality was relatively sensitive to reconstruction parameters. With the similar explicit low-rank factorization in motion space, the PCA model^{18,19} assumes that spatial components can be estimated *a priori* and reconstructs only temporal coefficients. However, spatial components from current 4D images to be reconstructed may not be accurately approximated by spatial components from prior 4D images, as spatial information is sensitive to data acquisition, such as modality change and geometric shift. In contrast, this work proposes using fluoroscopic training projections to learn temporal motion coefficients, and then only reconstructs spatial principal components using *a priori* estimated temporal coefficients, namely principal component reconstruction (PCR). Furthermore, the low-rank representation is in the image space instead of the motion space to avoid the nonlinearity in the motion space. The similar idea to fix temporal coefficients and only reconstruct spatial components was proposed for MRI, namely partial separability.^{21–24} However, due to the difference in imaging modality, the design of training data acquisition is different. In this sense, it is relatively straightforward for MRI to use fully sampled central k-space data for learning temporal coefficients by reconstructing a low-spatial resolution version of 4D images. Moreover, due to the use of the same dataset, temporal coefficients for high-spatial resolution 4D images to be reconstructed can be accurately estimated by these from coarse 4D images. In contrast, it is difficult in CT to use the same projection dataset for both the estimation of temporal coefficients and the reconstruction of spatial components, as coarse 4D images may not be satisfactorily reconstructed from these projections. In the following, we will provide the details for estimating temporal motion coefficients and then reconstructing spatial principal components.

2. METHODS

4D CBCT, especially cine CBCT with high-temporal resolution, is a sparse-view and undersampled image reconstruction problem. Without using any prior knowledge or temporal image regularization, the reconstruction quality is often unacceptable, for example, from the phase-independent image reconstruction method. Fortunately, 4D images exhibit a great deal of temporal similarity or redundancy, which can be used to improve the image reconstruction. The proposed

PCR method for cine CT consists of three essential components: (a) the degrees of freedom to be reconstructed are reduced by a low-rank representation of 4D phases using spatial principal components and temporal motion coefficients; (b) temporal coefficients are learned from training projections; (c) with temporal coefficients as priors, bilinear data fidelity becomes the linear data fidelity, and thus, spatial components can be conveniently reconstructed. Next, we will discuss these three points and the reconstruction algorithm in details.

2.A. Low-rank representation via matrix factorization

Let $X = \{x_t, t \leq T\}$ be a sequence of 4D images, where $x_t = \{x_{ijk}, i \leq N_x, j \leq N_y, k \leq N_z\}$ is a 3D image. When X is reshaped as a $N_x \cdot N_y \cdot N_z$ by T matrix with row dimension in space and column dimension in time, it can be represented as the product of spatial principal components $P = \{p_s, s \leq S\}$ and temporal motion coefficients $V = \{v_{st}, s \leq S, t \leq T\}$, that is,

$$X = PV. \quad (1)$$

Here, P is also reshaped as a $N_x \cdot N_y \cdot N_z$ by S matrix with $p_s = \{p_{ijks}, i \leq N_x, j \leq N_y, k \leq N_z\}$, and v_{st} weights the contribution from the s th principal component to the t th image, that is,

$$x_t = \sum_s p_s v_{st}. \quad (2)$$

Equation (1) is a low-rank representation of X , when the number of spatial components is smaller than the number of 4D images, that is, $S < T$. For example, $S = 15$ and $T = 50$ for the reconstructed results in the result session of this paper. Such a low-rank representation Eq. (1) via spatial principal components has been referred as matrix factorization mathematically¹⁴ or partial separability in the MRI field.^{21–24} The benefit of this low-rank representation is not only to reduce the degrees of freedom to be reconstructed but also a temporal regularization method to improve the reconstruction quality. Note that Eq. (1) is an explicit way of the low-rank regularization, in contrast with implicit low-rank regularization via nuclear norm.^{11–13} A distinct advantage of explicit low-rank representation is that the reconstruction quality can be significantly improved if temporal motion coefficients can be estimated *a priori* and then only spatial principal components need to be reconstructed.^{22–24}

2.B. Fluoroscopy-based motion learning

The key hypothesis for the proposed method to work is that temporal motion coefficients from projections can reasonably approximate these from 4D images in Eq. (1). Then, we can conveniently use these temporal coefficients from projections as priors and reconstruct only spatial principal components for 4D images.

However, acquired projections for image reconstruction may not be ideal for motion learning as these projections

involve the nonmotion changes due to the geometric variation caused by different projection angles. As a result, we propose to acquire the training projections at certain fixed angles for motion learning, that is, fluoroscopic training projections. That is, the training projections are acquired continuously for a few breathing periods at a fixed angle; then gantry is rotated to next angular position to acquire training projections for a few more breathing periods. In practice, these fluoroscopic projections can be taken during the gantry rotation to minimize the motion variation in time. For example, during 360-degree data acquisition, the gantry stops at 0, 90, 180, and 270 degrees for a few seconds each time to acquire fluoroscopic projections. Let N_v be the number of distinct angles. In the result section of this work, only 0 degree was utilized which is sufficient for both simulation and experimental studies in this proof-of-concept study.

Another issue to be addressed is that the breathing motion during training data acquisition is usually different from that during reconstruction data acquisition. To establish the one-to-one correspondence from reconstruction phases to training phases, we assume (a) some external equipment (e.g., Varian RPM system) is available to accurately measure real-time respiratory positions during both data acquisitions; (b) the range of respiratory positions during reconstruction data acquisition is a subset of the range of respiratory positions during training data acquisition. As a result, we can choose appropriate respiratory position intervals to group both training projections and reconstruction projections based on their corresponding respiratory positions, that is, to form T temporal intervals. Next, the training projections within each temporal interval are averaged to form $Y_0 = \{y_t, t \leq T\}$, from which temporal motion coefficients are learned. Here, the averaging is with respect to the number of projections within each temporal interval I_t , that is, $y_t = \sum_{k \in I_t} f_k / K$, $k \leq K$ for K projections f_k that belongs to I_t based on amplitude binning. Finally, these temporal coefficients learned from training projections will serve as priors for the reconstruction of spatial principal components.

That is, with one-to-one correspondence from Y_0 to X sorted by respiratory positions, Y is factorized as

$$Y_0 = P_y V_y. \quad (3)$$

Here, Y_0 is reshaped as a $N_a \cdot N_b \cdot N_v$ by T matrix with $y_t = \{y_{ijk}, i \leq N_a, j \leq N_b, k \leq N_v\}$, where N_a or N_b is the number of detector bins along a or b direction; P_y and V_y are spatial principal components and temporal motion coefficients from the training projection Y_0 . Note that V_y is derived from the training projection Y_0 [Eq. (3)], while V is from 4D images to be reconstructed [Eq. (1)]. Although P and P_y are not identical at all (e.g., Fig. 1), V and V_y are approximately the same or V is approximately a subset of V_y . This will be verified next [see Fig. 3(a)].

2.C. Reconstruction model

The idea of the proposed PCR method is recapped in Fig. 1. Once again, the key of PCR is that the temporal

FIG. 1. The PCR method for cine CBCT. The key hypothesis of PCR is that temporal motion coefficients from projections can reasonably approximate these from 4D images. Thus, fluoroscopic training data acquired at fixed angles (without geometric distortion) are used to learn temporal motion coefficients V_y , which is approximately equal to V for 4D images. Next, only spatial principal components P needs to be reconstructed using the priors V_y . Here both training data and reconstruction data are synchronized to T temporal phases after projection binning based on real-time respiratory positions. That is, by choosing appropriate respiratory position intervals, the training data are first grouped and averaged within each group to form Y_0 , which are then factorized into P_y and V_y ; using the same respiratory position intervals; the reconstruction data are also grouped into the same temporal phases as V_y , from which P is reconstructed. Last, the reconstructed 4D images X are simply the product of P and V_y . [Color figure can be viewed at wileyonlinelibrary.com]

motion coefficient V of 4D images is approximated by V_y from training projections. Then, we only need to reconstruct spatial principal components P . That is, we consider the following reconstruction model for PCR

$$\min_P \|AX - Y\|_2^2 + \lambda \|\nabla X\|_1, \text{ s.t. } X = P V_y. \quad (4)$$

In Eq. (4), the first term is the data fidelity term, where we explicitly enforce the low-rank regularization for X using the matrix factorization $X = P V_y$ with $S < T$. As a result, the data fidelity is bilinear in both P and V_y . However, with *a priori* estimated V_y , the data fidelity is now linear with respect to P only. If we were to reconstruct both P and V , the optimization problem is no longer convex due to the bilinear data fidelity and its nonlinearity may render the optimization algorithm to be inefficient, and dependent on reconstruction

parameters and initial guess, in addition to the nonexistence of a global minimum.

The second term of Eq. (4) is the image regularization term. Here, we choose $\lambda \|\nabla X\|_1$ to be the following spatiotemporal total variation regularization

$$\lambda \|\nabla X\|_1 = \sum_{(i,j,k,t)} \lambda_s \sqrt{\partial_i^2 X + \partial_j^2 X + \partial_k^2 X} + \lambda_t |\partial_t X|, \quad (5)$$

where $\nabla X = [\partial_i X \partial_j X \partial_k X \partial_t X]$ is the gradient operator with $\partial_{i,j,k,t}$ denoting the finite differencing operator along each Cartesian direction and temporal direction, respectively, λ_s the regularization parameter for spatial isotropic total variation and λ_t the regularization parameter for temporal total variation. The total variation regularization can be regarded as a special case of tensor framelet regularization,^{8,25,26} which can include higher order differencing at multilevel.

Regarding the choice of S , we empirically set $S = 15$ in our cine CBCT study. Note that this empirical choice is consistent with previous works for cine CBCT¹⁴ and cine MRI studies.²³ On the other hand, we calculated the curves of quantitative metrics at tumor positions when the number of principal components to be reconstructed in PCR varies from $S = 1$ to $S = 30$, for constant and shallow-deep breathing [Fig. 4(a)] using the simulation setting in the material section. The results are presented in Fig. 2, which suggests that the reconstructed image quality is reasonably stable with respect to the choice of number of spatial components to be reconstructed.

Regarding the choice of T , we set $T = 50$ here. Ideally temporal information is more accurate with more binning intervals or image phases. For example, one image phase per projection was considered for the 2D case.¹⁴ However, since all reconstructions are performed with complete GPU parallelization for fast reconstruction without memory exchange between GPU device and desktop host during the reconstruction, the GPU memory is the major limit for the number of

temporal frames that can be reconstructed in balance with desired spatial resolution and field of view.

Note that the PCR model can be regarded as a generalization of a popular state-of-art dynamic image reconstruction method, namely PICCS,⁵ in the sense that the averaged static background image used in PICCS for temporal regularization corresponds to the 1st spatial principal component of PCR [e.g., P_1 in Fig. 3(b)]. The fact that the temporal motion coefficient curve V_1 corresponding to P_1 [e.g., V_1 in Fig. 3(a)] is a constant verifies that P_1 is indeed an averaged background image. Generalizing beyond PICCS, with an explicit linear representation of 4D images X by multiple spatial principal components P in Eq. (4), PCR enforces the temporal similarity in high-order spatial components.

Next, we use a simulation study to illustrate the feasibility of the key hypothesis for the PCR method, that is, temporal motion coefficients V from temporal images can be approximated by temporal motion coefficients V_y from training projections, which then can be used for reconstructing spatial principal components P via Eq. (4).

The setting for the simulation study is given in the material section. The training projections were generated with constant breathing [Fig. 4(a)] for a single breathing period while the gantry angle was fixed at the zero degree along the anterior–posterior (AP) direction. The fine-resolution temporal images with 1.5 mm spatial resolution were averaged and downsampled to the coarse-resolution images with 3 mm spatial resolution. Then, V was extracted from the coarse-resolution images, and V_y was extracted from training projections. This was done through principal component analysis, after which both V and V_y were normalized with unity L2 norm for each spatial principal component. The first five temporal motion coefficients are plotted in Fig. 3(a), which indeed verifies the key hypothesis for the PCR method that V_y from training projections is a reasonable approximation of V for temporal images.

Based on this hypothesis, the PCR was performed to reconstruct 15 spatial principal components correspondingly

FIG. 2. Reconstructed results of quantitative metrics at tumor positions for PCR to study the dependence of the reconstructed image quality on the number of spatial principal components. [Color figure can be viewed at wileyonlinelibrary.com]

FIG. 3. Validation of PCR. The left figure (a) presents the comparison of 1st, 2nd, 3rd, 4th, and 5th temporal motion coefficients (V_1 – V_5) from ground truth images with these from training projections. Temporal coefficients from ground truth images are denoted by the “+” sign, while these from training projections are denoted by the “o” sign. The right figure (b) presents the comparison of spatial principal components from ground truth images with these reconstructed from PCR. The relative display windows are from 0% to 100%, 28%, and 19% for 1st, 2nd, and 3rd spatial principal components (P_1 – P_3), respectively. [Color figure can be viewed at wileyonlinelibrary.com]

FIG. 4. Real-time respiratory positions. (a) periodic breathing from simulation: constant breathing (red) and shallow-deep breathing (blue); (b) nonperiodic breathing from patients: respiratory positions (training breathing) during the acquisition of training fluoroscopy (red) and respiratory positions (reconstruction breathing) during the acquisition of reconstruction data (blue). Pattern 1: regular breathing; Pattern 2: breathing with baseline shift; Pattern 3: breathing with amplitude change; Pattern 4: breathing with both baseline shift and amplitude change. [Color figure can be viewed at wileyonlinelibrary.com]

using the estimated V_y . The reconstruction data were simulated using the procedure given in the material section. The ground truth of principal components was extracted corresponding to V from the coarse-resolution simulation images. Both the ground truth and reconstructed principal components are plotted in Fig. 3(b). By visual inspection of Fig. 3(b), with V_y estimated from the training projection, the principal components can be accurately reconstructed by the PCR method.

2.D. Solution algorithm

Here, we provide the solution algorithm for solving the proposed PCR model Eq. (4) with the following equivalent form

$$\min_P \|A(PV_y) - Y\|_2^2 + \lambda \|\nabla(PV_y)\|_1. \quad (6)$$

As noted earlier, Eq. (6) has linear data fidelity when V_y is estimated *a priori*. As a result, the optimization problem Eq. (6) is convex and can be efficiently solved. This benefit of explicit low-rank representation with estimated temporal motion coefficients for improved reconstruction quality has been emphasized for cine MRI.^{22–24}

In this work, the solution algorithm is developed for Eq. (6) using alternating direction method of multipliers (ADMM)²⁷ (also commonly known as split Bregman method²⁸), which is similar to our prior works^{11,12,25,26} that the readers may refer to for more details.

To deal with the nondifferentiable L1 norm, a dummy variable Z is introduced and the optimization problem is reformulated as the following constrained optimization problem

$$\min_{(P,Z)} \|A(PV_y) - Y\|_2^2 + \lambda \|Z\|_1, \text{ s.t. } \nabla(PV_y) = Z. \quad (7)$$

Note that $\lambda \|Z\|_1$ is defined by Eq. (5) and consists of an isotropic L1 norm in spatial variable and a L1 norm in temporal variable, that is,

$$\lambda \|Z\|_1 = \sum_{(i,j,k,t)} \lambda_s \sqrt{Z_i^2 + Z_j^2 + Z_k^2} + \lambda_t |Z_t|. \quad (8)$$

Furthermore, analytical formulas can be derived for solving $\lambda \|Z\|_1$, but not for $\lambda \|\nabla X\|_1$. This is exactly the motivation to introduce the dummy variable Z so that both the data fidelity and the image regularization can be decoupled and solved using analytical formulas.

This will become clear momentarily, as we start to alternate minimizing the following augmented Lagrangian L for Eq. (7) in P , Z , and U , respectively,

$$L(P, Z, U) = \|A(PV_y) - Y\|_2^2 + \lambda \|Z\|_1 + \mu \|\nabla(PV_y) - Z + U\|_2^2. \quad (9)$$

That is, the solution algorithm to Eq. (6) consists of the following three iterative steps derived from alternating minimization of Eq. (9)

$$P^{m+1} = \arg \min_P L(P, Z^m, U^m), \quad (10)$$

$$Z^{m+1} = \arg \min_Z L(P^{m+1}, Z, U^m), \quad (11)$$

$$U^{m+1} = U^m + \nabla(P^{m+1} V_y) - Z^{m+1}. \quad (12)$$

Due to the absence of nondifferentiable L1 norm, Eq. (10) is a standard differentiable L2 problem that can be solved analytically. Moreover, its solution should obey the following optimal condition

$$\begin{aligned} & [(A^T A + \mu \nabla^T \nabla)(P^{m+1} V_y)] V_y^T \\ & = [A^T Y + \mu \nabla^T (Z^m - U^m)] V_y^T. \end{aligned} \quad (13)$$

Here, we choose to solve Eq. (13) iteratively using linear conjugate gradient method, which does not require the explicit formulation of the system matrix for Eq. (13) and therefore is memory efficient.

On the other hand, Eq. (11) can be analytically solved by the following shrinkage formula S

$$Z^{m+1} = S(\nabla(P^{m+1} V_y) + U^m), \quad (14)$$

which consists of an isotropic shrinkage in spatial variables and a soft shrinkage in temporal variable.

$$\begin{aligned} S_s(Z_i, Z_j, Z_k) &= \frac{(Z_i, Z_j, Z_k)}{\sqrt{Z_i^2 + Z_j^2 + Z_k^2}} \\ &\max(\sqrt{Z_i^2 + Z_j^2 + Z_k^2} - \frac{2\lambda_s}{\mu}, 0), \end{aligned} \quad (15)$$

$$S_t(Z_t) = \text{sgn}(Z_t) \cdot \max(|Z_t| - \frac{2\lambda_t}{\mu}, 0). \quad (16)$$

Note that both shrinkages are component wise, that is, voxel-by-voxel operations.

3. MATERIALS

The breathing patterns considered in this work include: (a) two periodic breathing patterns, that is, constant breathing and shallow-deep breathing [Fig. 4(a)]; (b) four representative nonperiodic breathing patterns from patients [Fig. 4(b)]. The XCAT phantom together with these breathing patterns and default cardiac motion was used to simulate projection data and training fluoroscopy.

For periodic breathing patterns [Fig. 4(b)], the maximum amplitude of the shallow breathing was half of that of the deep breathing. The breathing period was a constant of 5 s. The reconstruction projection data were taken during a one-minute 360° gantry rotation, while the training projection data were taken during a half-minute scan at a fixed gantry angle along the anterior–posterior (AP) direction. Both reconstruction data and training fluoroscopy were taken at a constant rate of 600 projections per minute.

Four representative nonperiodic breathing patterns from patients [Fig. 4(b)] were selected: regular breathing (Pattern 1); breathing with baseline shift (Pattern 2); breathing with amplitude change (Pattern 3); and breathing with both baseline shift and amplitude change (Pattern 4). The averaged breathing periods were approximately 3 s for Pattern 1, 3, and 4, and 3.5 s for Pattern 2. Given original respiratory positions for each patient, one-minute respiratory positions were resampled at a constant rate of 800 positions per minute, which was used to generate a one-minute 360° gantry rotation of 800 projections. An additional portion of respiratory positions was resampled at the same rate for generating training AP fluoroscopy, that is, half-minute for Pattern 1, 3, and 4 and one-minute for Pattern 2. Here, the longer training period was needed for Pattern 2 to capture the baseline shift of Pattern 2. Note that each reconstruction or training projection was simulated using a XCAT volume corresponding to a unique respiratory position, for example, 800 XCAT volumes were generated to simulate 800 reconstruction projections, and additional 400 XCAT volumes were generated to simulate 400 training projections.

The source-to-detector distance was 150 cm, and the source-to-isocenter distance was 100 cm. Each 3D XCAT image for data generation was $256 \times 256 \times 200$ with 1.5 mm spatial resolution. The synthetic tumor that mimics the experimental phantom was added to the XCAT images. The tumor was of 3 cm diameter and had 4 cm motion range in the superior–inferior direction with its motion following corresponding respiratory positions. A stripe of 6 mm width was inserted at the center of the tumor to quantify the image resolution. The virtual detector was 150×150 with 4 mm pixel size. The reconstructed images were $128 \times 128 \times 100$ with 3 mm spatial resolution.

The experimental projection dataset was acquired from the Varian TrueBeam system. The same periodic respiratory positions [Fig. 4(a)] were programmed to the experimental phantom. The reconstruction projection data were taken during a one-minute 360° gantry rotation, while the training projection data were taken during a half-minute AP scan, both at a constant rate of 890 projections per minute. A stripe of ~3 mm width was at the center of the tumor position, which can be used to quantify the image resolution. Although the flat-panel detector was 1024×768 with 0.388 mm pixel size, the projection was averaged and downsampled to 1.522 mm pixel size and only the middle 256×100 portion was used for image reconstruction. The reconstructed images were $200 \times 200 \times 80$ with 1.5 mm spatial resolution.

The amplitude binning was used here to (a) sort projections to temporary phases based on respiratory positions and (b) correlate training temporal coefficients and reconstruction temporal coefficients. Equally spaced bins between minimal and maximal amplitudes were used to bin the projections. For periodic breathing, minimal and maximal amplitudes were the same as minimal and maximal respiratory positions, respectively. For nonperiodic breathing, the respiratory position range during training breathing was more than that during reconstruction breathing for Pattern 1 and 2, and less than

that during reconstruction breathing for Pattern 3 and 4. That is, some training phases were not available to be reconstructed for Pattern 1 and 2, while some reconstruction phases were not available to be trained for Pattern 3 and 4. To resolve this issue, (a) the minimum of maximal respiratory position during training and maximal respiratory position during reconstruction was set to be the maximal amplitude for binning, and the maximum of minimal respiratory position during training and minimal respiratory position during reconstruction was set to be the minimal amplitude for binning; (b) next, respiratory positions of reconstruction breathing larger than the maximal amplitude were set to be the maximal amplitude, and respiratory positions of reconstruction breathing smaller than the minimal amplitude were set to be the minimal amplitude during reconstruction.

The image reconstruction was performed on a GPU desktop with GeForce GTX TITAN X of 12 GB memory. The reconstruction code was fully parallelized using NVIDIA CUDA toolbox. It took approximately 1 min and 5 min for each reconstruction of 30 ADMM iterations with simulated and experimental data, respectively. Here, the choice of reconstructed image size and spatial resolution for simulation or experimental studies is mainly limited by the memory size of GPU.

The proposed PCR method was compared with PICCS, a state-of-art image reconstruction method.⁵ The reference image for PICCS was reconstructed with total variation regularization using all projections, and PICCS was solved using the similar ADMM-based solution algorithm as PCR. For either PICCS or PCR, μ was set as a percentage of the estimated condition number of $A^T A$ (i.e., the maximal diagonal element), $\lambda_s = 2\lambda_t/\mu$ and $\lambda_t = 2\lambda_s/\mu$ for the convenience of parameter tuning, and it was found that $\lambda_s = \lambda_t = \lambda$ provided satisfactory reconstruction results. Then, a sequence of image reconstructions was performed with several combinations of μ and λ that have equally spaced values in the logarithmic space. Last, the reconstructed images from each method with the best quality were selected to be presented here.

In this work, the image quality was quantitatively measured by contrast-to-noise ratio (CNR) and contrast recovery coefficient (CRC) as follows:

$$CNR = \frac{|m_t - m_b|}{\sqrt{\sigma_t^2 + \sigma_b^2}}, \quad (17)$$

$$CRC = \frac{|I_{\max} - I_{\min}|}{I_{\max} + I_{\min}}. \quad (18)$$

Both CNR and CRC were measured at the tumor positions. The mean m_t and the standard deviation σ_t were computed with respect to a target region inside the tumor (red-square ROI in Fig. 6), while m_b and the standard deviation σ_b were with respect to a background region outside the tumor (blue-square ROI in Fig. 6) that is adjacent to the target region. The averaged intensity $I_{\max}$ was computed inside the tumor stripe at the tumor center, while the averaged intensity $I_{\min}$ was computed close to the tumor stripe and inside the

tumor (Red stripe ROI in Fig. 6). Note that ROIs shown in Fig. 6 are specific to the displayed phase and moved rigidly across all phases along the AP direction based on the breathing pattern to match tumor position changes. Then, CNR and CRC were averaged across all temporal phases as quantitative metrics for reconstructed image quality.

4. RESULTS

4.A. Simulation results

The quantitative metrics for image quality evaluation (i.e., CRC and CNR) are plotted in Fig. 5 for all periodic and non-periodic breathing patterns (Fig. 4). These quantitative metrics suggest that PCR had improved reconstruction quality from PICCS. This is at no surprise since (a) PCR utilizes multiple spatial principal components while PICCS only has one, and (b) PCR utilizes the motion learning from fluoroscopy. In addition, to characterize reconstructed temporal changes, the spatial profile at tumor center (i.e., the entire profile of the AP direction along the red line in the sagittal view of Fig. 6) across all temporal phases is presented here for each pattern. The temporal plots suggest that the PCR is able to reconstruct the temporal changes well with significantly improved image quality from PICCS.

Representative results reconstructed from PICCS and PCR are presented in Figs. 6, 7, and 8, based on the CRC value from PICCS. Figure 6 presents the reconstructed results with the largest CRC value from PICCS (Phase 1 from Pattern 1); Fig. 7 presents the reconstructed results with the smallest CRC value from PICCS (Phase 30 from Pattern 1); Fig. 8 presents the reconstructed results with the median CRC value from PICCS (Phase 15 from Pattern 2, Phase 21 from Pattern 3, and Phase 22 from Pattern 4). By visual inspection, the reconstructed image quality from PCR was better than that from PICCS.

Note that by using the reference image (i.e., the static background image similar to the 1st spatial principal component) for temporal regularization, PICCS has an inevitable tendency to bias toward the reference image, which introduced undesirable artifacts, such as these streaking artifacts around the sternum (e.g., “Axial” from PICCS in Fig. 8) and the virtual stripes at tumor center [e.g., the temporal plot from PICCS in Fig. 5(a)]. In contrast, these artifacts were much less or not observed for PCR.

Noticeably, PCR was more robust in reconstructing the phases with relatively smaller number of projections, for example, Phase 30 from Pattern 1 as presented in Fig. 7. For this phase, PICCS suffered from the degraded image quality, although it reconstructed well at Phase 1 with relatively greater number of projections in Fig. 6. As a result, temporal plots from PICCS in Fig. 5 exhibit dark-striped discontinuities, which indicate that these phases with dark stripes were not reconstructed well (e.g., Fig. 7). In comparison, PCR is more robust than PICCS in reconstructing phases with fewer number of projections, for example, without dark stripes in temporal plots.

Fig 5. Reconstructed results of quantitative metrics at tumor positions from simulation data. [Color figure can be viewed at wileyonlinelibrary.com]

However, in contrast with random appearance of phases with insufficient number of projections (e.g., Fig. 6), the systematic appearance of phases with insufficient number of projections caused PCR to underperform as well, for example, the phases with smallest or largest amplitudes in shallow-deep breathing and Pattern 4 (see the temporal plots in Fig. 5). Such systematic appearance of insufficient number of projections is specific to breathing patterns. In this sense, it seems that constant and shallow-deep breathing are sufficient to study PCR, although both are periodic. That is, the constant breathing represents the situations without systematic appearance of insufficient number of projections, while the shallow-deep breathing represents the situations with systematic appearance of insufficient number of projections.

4.B. Experimental results

The quantitative curves of CRC and CNR are plotted in Fig. 9 for all periodic breathing patterns [Fig. 4(a)]. These quantitative metrics suggest that PCR had improved reconstruction quality from PICCS for experimental results.

Fig. 6. Reconstructed results (Phase 1) from simulation data with nonperiodic breathing of Pattern 1. The presented phase was selected with the largest CRC value for PICCS. [Color figure can be viewed at wileyonlinelibrary.com]

FIG. 7. Reconstructed results (Phase 30) from simulation data with nonperiodic breathing of Pattern 1. The presented phase was selected with the smallest CRC value for PICCS.

FIG. 8. Reconstructed results from simulation data with nonperiodic breathing of Pattern 2, 3, and 4. The presented phases were selected with the median CRC value for PICCS, that is, Phase 15 from Pattern 2, Phase 21 from Pattern 3, and Phase 22 from Pattern 4.

Representative results reconstructed from PICCS and PCR are presented in Figs. 10 and 11, based on the CRC value from PICCS. Figures 10 and 11 present the reconstructed results with the median CRC value from PICCS for constant and shallow-deep breathing, respectively, that is, Phase 5 from constant breathing and Phase 31 from shallow-deep breathing. By visual inspection, the reconstructed image quality from PCR was better than that from PICCS. For example, the 3 mm-resolution stripe at the tumor center is

FIG. 9. Reconstructed results of quantitative metrics at tumor positions from experimental data. [Color figure can be viewed at wileyonlinelibrary.com]

FIG. 10. Reconstructed results (Phase 5) from experimental data with constant breathing. The presented phase was selected with the median CRC value for PICCS.

FIG. 11. Reconstructed results (Phase 31) from experimental data with shallow-deep breathing. The presented phase was selected with the median CRC value for PICCS.

clearly visible from PCR, while it is blurred from PICCS in Figs. 10 and 11. Moreover, temporal changes at tumor center and across the strip in the AP direction are presented in Figs. 10 and 11. These temporal profiles further suggest that

the PCR had improved image quality for reconstructed temporal changes from PICCS.

5. CONCLUSIONS AND DISCUSSIONS

PCR is proposed to render the use of cine CT. The key of PCR is that temporal motion coefficients for 4D image phases can be estimated by temporal motion coefficients from fluoroscopic training projections. Then with *a priori* estimated temporal motion coefficients, the PCR reconstructs spatial principal components, based on which 4D images can be generated. Although this study focuses on on-board CBCT, the general methodology can be extended for planning CT as well. Simulation and experimental CBCT studies are performed to verify the feasibility of PCR for cine CT.

The reconstruction size in this work is limited by GPU memory, since all reconstructions are performed with complete GPU parallelization for fast reconstruction without memory exchange between GPU device and desktop host during the reconstruction. Alternatively, the reconstruction could be performed with partial GPU parallelization at the sacrifice of speed with memory exchange between GPU device and desktop host during the reconstruction, for example, with only parallelized projections and backprojections. In the latter case, the reconstruction size is no longer limited by GPU memory. However, we prefer fully parallelized GPU implementation in this work to save reconstruction time for this proof-of-concept study.

When the binned projections are not evenly distributed across temporal phases, the phases with fewer projections tend to have worse reconstruction quality. This problem commonly occurs for equally spaced amplitude binning. Note that this problem is not specific to certain reconstruction method. Although PCR can handle the random appearance of phases with insufficient number of projections (e.g., Fig. 7), it underperforms as well for the systematic appearance of phases with insufficient number of projections, for example, the phases with smallest or largest amplitudes in shallow-deep breathing and Pattern 4 (see the temporal plots in Fig. 5). The problem can be addressed by adaptive amplitude binning instead of equally spaced binning, that is, to choose intervals adaptively so that the number of projections per phase is roughly equal. However, a trade-off of adaptive binning for the shallow-deep breathing type is that for a fixed number of phases, it compromises the temporal resolution for the deep-breathing portion. As a result, it is beneficial for PCR to reconstruct as many phases as possible when keeping the number of spatial principal components the same, so that (a) the projections can be evenly distributed across all phases and (b) cine CT images can be dynamically accurate with high-temporal resolution.

The shallow-deep breathing case considered here can be regarded as a representative example of a nonperiodic breathing case (e.g., Pattern 4) in terms of the systematic appearance of phases with insufficient number of projections. However, this can be resolved through aforementioned adaptive projection binning. That is, the temporal resolution of reconstructed

phases needs to be balanced adaptively with projection number for each phase for optimized reconstruction quality.

ACKNOWLEDGMENT

This work is partially supported by NIH grant R01-CA184173.

^{a)} Author to whom correspondence should be addressed. Electronic mail: hao.gao@duke.edu.

REFERENCES

1. Sonke JJ, Zijp L, Remeijer P, van Herk M. Respiratory correlated cone beam CT. *Med Phys*. 2005;32:1176–1186.
2. Kriminski S, Mitschke M, Sorensen S, et al. Respiratory correlated cone-beam computed tomography on an isocentric C-arm. *Phys Med Biol*. 2005;50:5263.
3. Dietrich L, Jetter S, Tücking T, Nill S, Oelfke U. Linac-integrated 4D cone beam CT: first experimental results. *Phys Med Biol*. 2006;51:2939.
4. Li T, Xing L, Munro P, et al. Four-dimensional cone-beam computed tomography using an on-board imager. *Med Phys*. 2006;33:3825–3833.
5. Chen GH, Tang J, Leng S. Prior image constrained compressed sensing (PICCS): a method to accurately reconstruct dynamic CT images from highly undersampled projection data sets. *Med Phys*. 2008;35:660–663.
6. Ritschl L, Sawall S, Knaup M, Hess A, Kachelrieß M. Iterative 4D cardiac micro-CT image reconstruction using an adaptive spatio-temporal sparsity prior. *Phys Med Biol*. 2012;57:1517.
7. Wu H, Maier A, Fahrig R, Hornegger J. Spatial-temporal total variation regularization (STTVR) for 4D-CT reconstruction. In *SPIE Medical Imaging*, pp. 83133J-83133J. International Society for Optics and Photonics; 2012.
8. Gao H, Li R, Lin Y, Xing L. 4D cone beam CT via spatiotemporal tensor framelet. *Med Phys*. 2012;39:6943–6946.
9. Jia X, Lou Y, Dong B, Tian Z, Jiang S. 4D computed tomography reconstruction from few-projection data via temporal non-local regularization. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*. New York: Springer Berlin Heidelberg; 2010:143–150.
10. Liu J, Ding H, Molloy S, Zhang X, Gao H. TICMR: total image constrained material reconstruction via nonlocal total variation regularization for spectral CT. *IEEE Trans Med Imaging*. 2016;35:2578–2586.
11. Gao H, Cai JF, Shen Z, Zhao H. Robust principal component analysis-based four-dimensional computed tomography. *Phys Med Biol*. 2011;56:3181.
12. Gao H, Hengyong Y, Osher S, Wang G. Multi-energy CT based on a prior rank, intensity and sparsity model (PRISM). *Inverse Probl*. 2011;27:115012.
13. Chen GH, Li Y. Synchronized multiartifact reduction with tomographic reconstruction (SMART-RECON): a statistical model based iterative image reconstruction method to eliminate limited-view artifacts and to mitigate the temporal-average artifacts in time-resolved CT. *Med Phys*. 2015;42:4698–4707.
14. Cai JF, Jia X, Gao H, Jiang SB, Shen Z, Zhao H. Cine cone beam CT reconstruction using low-rank matrix factorization: algorithm and a proof-of-principle study. *IEEE Trans Med Imaging*. 2014;33:1581–1591.
15. Christoffersen CP, Hansen D, Poulsen P, Sørensen TS. Registration-based reconstruction of four-dimensional cone beam computed tomography. *IEEE Trans Med Imaging*. 2013;32:2064–2077.
16. Yan H, Zhen X, Folkerts M, et al. A hybrid reconstruction algorithm for fast and accurate 4D cone-beam CT imaging a). *Med Phys*. 2014;41:071903.
17. Wang J, Gu X. Simultaneous motion estimation and image reconstruction (SMEIR) for 4D cone-beam CT. *Med Phys*. 2013;40:101912.

18. Li R, Jia X, Lewis JH, et al. Real-time volumetric image reconstruction and 3D tumor localization based on a single x-ray projection image for lung cancer radiotherapy. *Med Phys*. 2010;37:2822–2826.
19. Ren L, Zhang Y, Yin FF. A limited-angle intrafraction verification (LIVE) system for radiation therapy. *Med Phys*. 2014;41:020701.
20. Liu J, Zhang X, Zhang X, et al. 5D respiratory motion model based image reconstruction algorithm for 4D cone-beam computed tomography. *Inverse Probl*. 2015;31:115007.
21. Liang ZP. Spatiotemporal imaging with partially separable functions. In: *2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. Arlington, VA: IEEE; 2007:988–991.
22. Zhao B, Haldar JP, Christodoulou AG, Liang ZP. Image reconstruction from highly undersampled (k, t)-space data with joint partial separability and sparsity constraints. *IEEE Trans Med Imaging*. 2012;31:1809–1820.
23. Christodoulou AG, Zhang H, Zhao B, Hitchens TK, Ho C, Liang ZP. High-resolution cardiovascular MRI by integrating parallel imaging with low-rank and sparse modeling. *IEEE Trans Biomed Eng*. 2013;60:3083–3092.
24. He J, Liu Q, Christodoulou A, Ma C, Lam F, Liang ZP. Accelerated High-Dimensional MR Imaging With Sparse Sampling Using Low-Rank Tensors; 2016.
25. Gao H, Qi XS, Gao Y, Low DA. Megavoltage CT imaging quality improvement on TomoTherapy via tensor framelet. *Med Phys*. 2013;40:081919.
26. Gao H, Guo M, Li R, Xing L. 4DCT and 4D Cone-Beam CT Reconstruction Using Temporal Regularizations. *Graphics Processing Unit-Based High Performance Computing in Radiation Therapy*; 2015: 63.
27. Boyd S, Parikh N, Chu E, Peleato B, Eckstein J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found Trends® Mach Learn*. 2011;3:1–122.
28. Goldstein T, Osher S. The split Bregman method for L1-regularized problems. *SIAM J Imaging Sci*. 2009;2:323–343.