SCIENTIFIC REPORTS

OPEN

Motion guided Spatiotemporal Sparsity for high quality 4D-CBCT reconstruction

Received: 19 May 2017
Accepted: 29 November 2017
Published online: 12 December 2017

Yang Liu, Xi Tao, Jianhua Ma, Zhaoying Bian, Dong Zeng, Qianjin Feng, Wufan Chen & Hua Zhang

Conventional cone-beam computed tomography is often deteriorated by respiratory motion blur, which negatively affects target delineation. On the other side, the four dimensional cone-beam computed tomography (4D-CBCT) can be considered to describe tumor and organ motion. But for current on-board CBCT imaging system, the slow rotation speed limits the projection number at each phase, and the associated reconstructions are contaminated by noise and streak artifacts using the conventional algorithm. To address the problem, we propose a novel framework to reconstruct 4D-CBCT from the under-sampled measurements—Motion guided Spatiotemporal Sparsity (MgSS). In this algorithm, we try to divide the CBCT images at each phase into cubes (3D blocks) and track the cubes with estimated motion field vectors through phase, then apply regional spatiotemporal sparsity on the tracked cubes. Specifically, we recast the tracked cubes into four-dimensional matrix, and use the higher order singular value decomposition (HOSVD) technique to analyze the regional spatiotemporal sparsity. Subsequently, the blocky spatiotemporal sparsity is incorporated into a cost function for the image reconstruction. The phantom simulation and real patient data are used to evaluate this algorithm. Results show that the MgSS algorithm achieved improved 4D-CBCT image quality with less noise and artifacts compared to the conventional algorithms.

Three-dimensional cone-beam computed tomography (3D-CBCT) has been widely used in image guided radiation therapy (IGRT)[1–3]. It can provide the volumetric information for tumor localization in IGRT. But for thoracic and upper abdominal regions, the 3D-CBCT image is often deteriorated by motion blur[4]. To address the problem, four-dimensional CBCT (4D-CBCT) imaging that incorporating temporal (phase) information on the basis of 3D-CBCT was proposed[5–7]. Compared to 3D-CBCT imaging, 4D-CBCT can provide the patient-specific respiratory motion information and multiple three-dimensional volumes to represent the different status in the breathing cycle[8–10]. For 4D-CBCT imaging, the respiratory signal is recorded or estimated and motion contained cone-beam projections are usually sorted into 8–10 subsets according to the respiratory signal. However, the gantry rotation speed and frame rate of the flat-panel imager limit the total number of cone-beam projections (usually 600~800), which result in relatively fewer projections in each respiratory phase. Consequently, the reconstructed CBCT images by using the conventional Feldkamp–Davis–Kress (FDK) algorithm[11] suffer from significant artifacts and noise[12,13]. In addition, the randomness of breathing would lead to the cone-beam projections bunched into several clusters, and the bunched sampling scheme will aggravate the noises and artifacts level in the reconstructed images[14].

To address this problem, various strategies have been proposed to improve the image quality of 4D-CBCT[15]. On one side, increasing the number of angular sampling, multiple-gantry rotation and slow-gantry rotation schemes were structured[7,13]. But these schemes will prolong the time of data acquisition and increase the risk of motion embracing. On the other side, lots of reconstruction algorithms[16–18] have been proposed to improve the quality of 4D-CBCT image, such like the total variation minimization based algorithms[15,19,20], the McKinnon–Bates (MKB) algorithm[21,22], the prior image constrained compressed sensing (PICCS) algorithm[23] and the auto-adaptive phase correlation (AAPC) algorithm[24]. As shown in the study by Frank Bergnera *et al.*, the reconstruction algorithms using an iterative scheme can remarkably reduce the 4D-CBCT specific artifacts[25]. Nevertheless, algorithms that use the full data set, at least for initialization, such as MKB and PICCS algorithm,

Guangdong Provincial Key Laboratory of Medical Image Processing, Southern Medical University, Guangzhou, Guangdong, 510515, China. Correspondence and requests for materials should be addressed to W.C. (email: chenwf@fimmu.com) or H.Z. (email: xinsier@smu.edu.cn)

are only a trade-off and achieve sub-optimal temporal resolution, of which the residual motion could still be found in the reconstructed 4D-CBCT images[26]. On the other hand, algorithms that only use the projections assigned to the current phase to reconstruct the final image can fully achieve the temporal resolution, such as the total variation minimization based algorithm. For these algorithm, when the projection number is very low or there exist small size and low-contrast objects, tiny structure may be usually erased with the piecewise smooth constraint.

In recent years, the image restoration or denoising algorithms based on block processing have shown increasing vitality. The earliest concept of 'patch' was proposed in Haralick's study on textural features for image classification[27]. In 1980, JS Lee pioneeringly used this concept for image enhancement[28]. Latterly, algorithms such like the dictionary learning based image reconstruction/restoration approaches[29], the nonlocal means filter[30] and the BM3D algorithm[31], *et al.*[32–34], have been successfully developed. Local patches often experience much less distortion than the global image and therefore it becomes easier to define the similarity between local patches. Low-rank and sparsity can be better reflected in patch based processing.

In this work, we propose to reconstruct 4D-CBCT volumes from the subset projections of current phase and incorporate the image domain phase-correlated information into the iterative procedure. Motived by the success of block based image restoration, we propose a Motion guided Spatiotemporal Sparsity (MgSS) to formulate the regularization for 4D-CBCT reconstruction. In this scheme, CBCT images of different phases are divided into small cubes (three-dimensional blocks) and the cubes are tracked with estimated motion vector fields through time (phase). After then, regional spatiotemporal sparsity is applied on the tracked cubes. Specifically, we recast the tracked cubes into four-dimensional matrix, and use the higher order singular value decomposition (HOSVD)[35] technique to analyze the regional spatiotemporal sparsity. Finally, a cost function is formulated with embedding the block based spatiotemporal sparsity. One simple but effective optimization algorithm was used for the cost function solution.

This paper is structured as follows: In the method part, we first present the flow chart of the proposed MgSS scheme and then give detailed introduction of each step. After this, we formulate the reconstruction framework that incorporates the MgSS scheme. In the experiment part, we exhibit the results of the NCAT phantom simulation data, 4DCT based simulation data and real patient data by using FDK, SART-TV and the MgSS algorithms. Lastly, in the discussion part, we simply discuss the superiority and limitation of the proposed algorithm.

Methods

Flow chart of the proposed MgSS algorithm. In this work, the proposed Motion guided Spatiotemporal Sparsity (MgSS) applying to 4D-CBCT reconstruction comprises the following steps;

(1) Estimate the motion maps for each voxel between adjacent phases of the 4D-CBCT images.
(2) Divide the CBCT sequence images into cubes (3D blocks), and track the cubes through time using the three-dimensional motion maps obtained in step 1.
(3) Stack the tracked cubes and apply regional spatiotemporal sparsity on the tracked cubes by framing the 4D-CBCT reconstruction problem to be a constrained optimization problem, and then optimize the problem to get the reconstructions.

In the following sections, we will introduce these steps in detail.

Motion Maps Estimation. In this study, we are not devoted to develop new algorithm to achieve the three-dimensional motion vector fields (3D-MVFs) between CBCT images of adjacent phases, but rather we use the Real-Time Image-based Tracker (RTIT)[36] toolbox to estimate the three-dimensional motion maps. The RTIT toolbox is an available open source with implementation of the optical flow based registration algorithms[37,38]. Assume that we obtain the initial/current 4D-CBCT estimation, with the RTIT, we can achieve the pixel based motion vector fields consisting of the changes in space coordinates that describe the distribution of the apparent motion velocities of intensity patterns in the sequences of images.

Cubes Tracking with MVFs. For the task of cubes tracking, we define the voxel space-time position as $u = (x,\ y,\ z,\ t)$, where x, y, z and t represent the spatial position of the voxel in the CBCT volumetric image and the temporal phase index, respectively. In MgSS, we fetch the cubes (3D blocks) from the images of the first phase and use the 3D-MVFs to track structurally similar cubes in the other phases. Specifically, the first phase image was initiated with highly overlapping cubes. In the extreme case, all the voxels in the first phase image can be defined as the central voxel of one cube, but this may result in huge computation burden. Thus in this study, rather than sliding by one voxel to every next, we use a step of $N_{step} = 2$ voxels to define the cubes. For the motion tracking, considering one cube $B(u_1) \in X_1^{N_b \times N_b \times N_b}$, here $u_1 = (x_1,\ y_1,\ z_1,\ t_1)$ indicates the central pixel position of the cubes with spatial position to be $(x_1,\ y_1,\ z_1)$ and temporal phase index to be t_1. By using Δu to denote the displacement of the central voxel in the cube between t_1 and t_2 phases, we can use $u_1 + \Delta u$ to track the center pixel location for cube in the t_2 phase. Considering $u_1 + \Delta u$ might be non-integral, the next block center voxel was taken as $u_2 = \{u_1 + \Delta u_2\}$, where "{}" is a rounding operation. Extracting the voxels centered from u_2, we can constitute a new cube $B(u_2)$ with the same block size. Following this schedule, $B(u_n)$ can be tracked with $u_n = \{u_1 + \Delta u_n\}$ using the above block-center-tracking method. A block cluster can be ulteriorly constructed: $\Theta_{MgSS} = [B(u_1), B(u_2), \ldots, B(u_{Nt})]$. Here the cluster Θ_{MgSS} is a four dimensional matrix with the size of $N_b \times N_b \times N_b \times N_t$. Based on the fact that not all the chest is moving during the CBCT scan, the tracked cubes for the static parts should be noisy blocks with same anatomy structures. As shown in Fig. 1, compared to the blocks of the same spatial position in the original phases, the tracked blocks exhibit more mutual similarity.

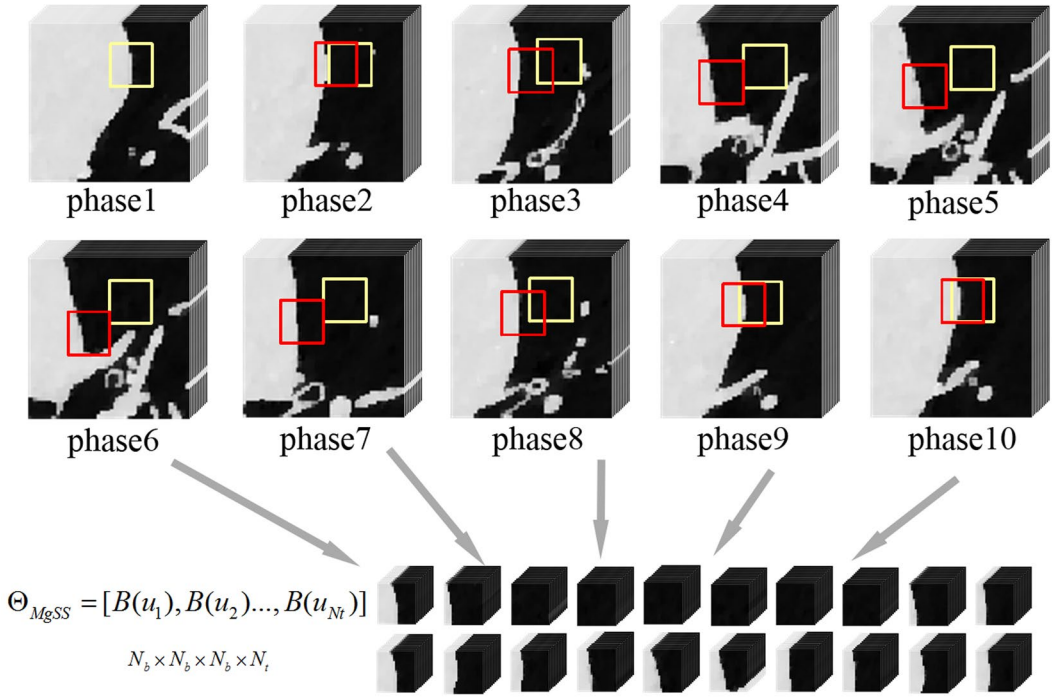

Figure 1. An example of blocks tracked (red squares) and not tracked (yellow squares) through all the frames.

Regional Spatiotemporal Sparsity. The technology about matrix rank sparsity has been successfully investigated in the field of dynamic image reconstruction, such as the k-t SLR method[39]. In some previous studies[40], matrix rank sparsity was often used to dispose the entire image. But the recent studies, such like Low-dimensional-structure self-learning and thresholding (LOST)[41] and compartment-based k-t principal component analysis[42] have indicated that the spatiotemporal sparsity and reconstruction quality could be further promoted by separating the entire image into blocks. Under the patch-based processing theory, decomposition of the tracked regions of dynamic datasets using the singular value decomposition (SVD) algorithm has been reported[43]. In Chen's work, blocks in the clusters were vectorized and each cluster Θ was first rearranged into a 2-D matrix $\widetilde{\Theta} \in \mathbb{C}^{N_s \times N_t}(N_s = N_b \times N_b)$. The SVD technique was then adopted to decompose the cluster: $\widetilde{\Theta}_{svd} = US_{svd}V^*$. If the cluster is truly spatiotemporal sparsity, a least number of significant values will be found in the singular matrix S_{svd}. For the standard SVD, the image blocks are manually vectorized and then the image structural properties in the spatial domain are ignored. To address the problem, the higher order singular value decomposition (HOSVD), which is able to directly decompose dynamic datasets into a multidimensional singular matrix rather than unfolding the blocks into column vectors, has been reported[35]. Also, in this work, we use the HOSVD to decompose the clusters due to its naturality and flexibility. By using the HOSVD technique, a four-dimensional tensor Θ can be decomposed as:

$$\Theta = S_{hosvd} \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)} \times_4 U^{(4)} \tag{1}$$

where $U^{(1)}$, $U^{(2)}$, $U^{(3)}$, $U^{(4)}$ are orthogonal matrices that contain the orthonormal vectors spanning the column space of the matrix Θ. Here, the symbol $\times_n$ stands for the n-th mode tensor product. The core tensor S_{hosvd} is not necessarily diagonal matrix, which implies that each dimension of the tensor can have a different rank. Generally, noise and artifacts can be attenuated by approximating the rank of the core tensor. However, the computation of the best rank approximation requires an iterative alternated least-square (ALS) algorithm and is quite time consuming[44]. Moreover, the approximation obtained from simple truncation has been proved to be in most cases quite similar to the optimal approximation[45]. Base on this, in our study, we apply a soft-thresholding strategy on the HOSVD coefficients to reduce noise and artifacts. The thresholding of the rank coefficients can be represented as follows:

$$\hat{S}_{hosvd} = H_\tau(S_{hosvd}) \tag{2}$$

where H_τ denotes the soft-thresholding operator with threshold τ. Then a new tensor can be synthesized by the inverse HOSVD transformation with truncated coefficients $\hat{S}_{hosvd}$ and the orthogonal matrices $U^{(1)}$, $U^{(2)}$, $U^{(3)}$, $U^{(4)}$:

$$\Theta^*_{hosvd} = \hat{S}_{hosvd} \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)} \times_4 U^{(4)} \tag{3}$$

The above operation is repeated for each reference cube, thus provides multiple estimates at the same coordinate. For this reason, the final estimates are aggregated by weighted averaging all the obtained block-wise estimates that overlapped at each voxel.

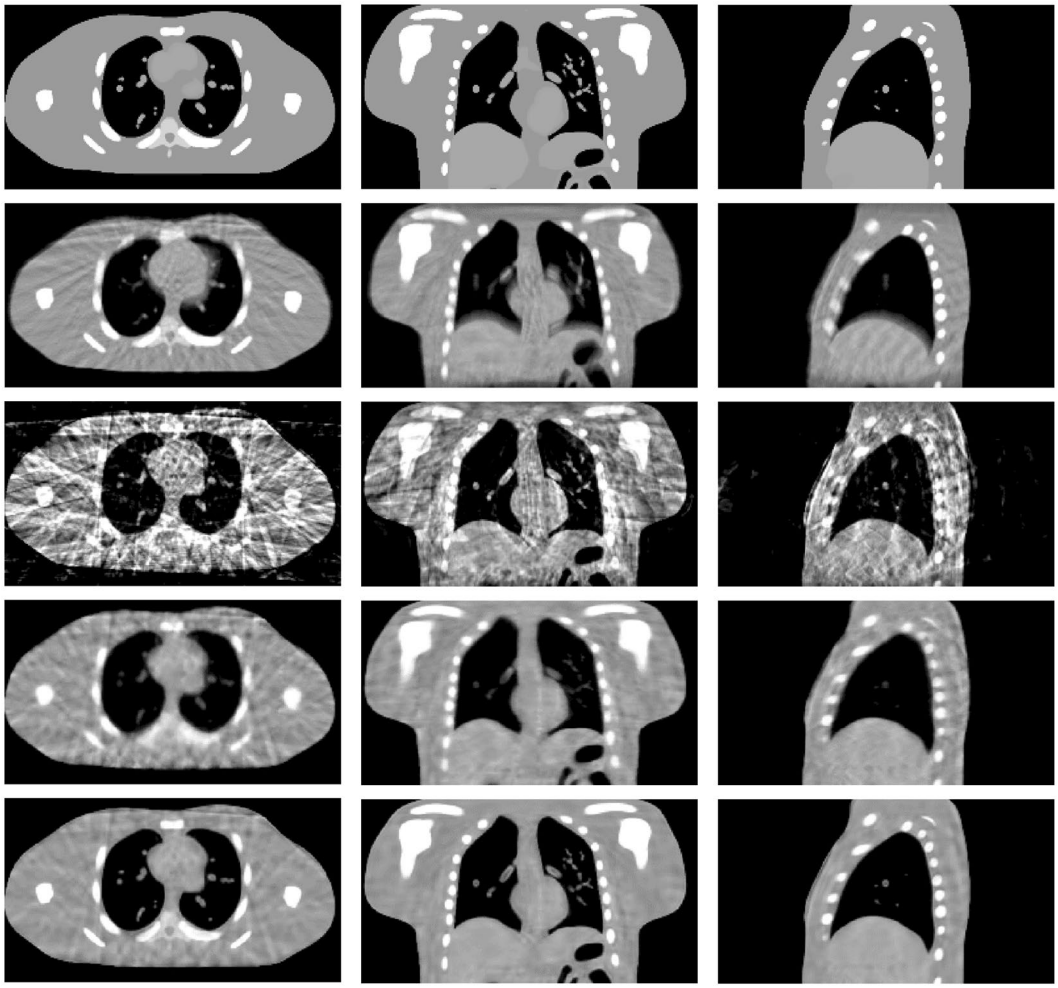

Figure 2. Results of 4D NCAT phantom with 21 projections for each phase. First row shows the begin-expiration phase of digital phantom. Second row shows the 3D-CBCT reconstructed from all projections by FDK. The third to last rows show 4D-CBCT images at the begin-expiration phase reconstructed by using FDK, SART-TV and proposed MgSS algorithm, respectively. The transverse, coronal, and sagittal planes have been shown in the first, second and third columns, respectively.

MgSS based 4D-CBCT reconstruction. To incorporate the motion tracking induced blocky spatiotemporal sparsity into the CBCT reconstruction framework, in this section, we formulate the following minimization scheme:

$$\begin{aligned} &\text{miniminze}_{f,u}\Phi(T_u \cdot f) \\ &s.t.\ \|Af - y\|^2 < \delta \end{aligned} \tag{4}$$

where f represents the estimated dynamic images, A is the CBCT imaging system matrix with elements of a_{ij}, y denotes the projection measurements. Operator Φ is the cube based sparsity penalties, while $T_u \cdot f$ denotes the dataset that includes a sequence of 4D matrix of the tracked cubes. Operator T_u includes two steps: (1) Track three-dimensional blocks with the motion vector, and (2) Rearrange the three-dimensional blocks into 4D matrix.

To optimize the problem in (4), in this work, inspired by the techniques used by Pan *et al.*[46], we present an efficient way to solve (4), which can be summarized as:

(1) SART[47] step:

$$f_j^{(n+1/2)} = f_j^{(n)} + \lambda\left(\sum_i\left[a_{ij}\frac{y_i - \sum_{n=1}^{N} a_{in} f_j^{(n)}}{\sum_{n=1}^{N} a_{in}}\right] / \sum_i a_{ij}\right) \tag{5}$$

Here λ is an over-relaxed factor.

(2) MgSS step: $f^{(n+1)} = W\left(\widetilde{\Theta}^*_{MgSS}(f^{(n+1/2)})\right)$ Here, $\widetilde{\Theta}^*_{MgSS}(f^{(n+1/2)})$ denotes the restored block-based clusters based on $f^{(n+1/2)}$ by using the HOSVD approach

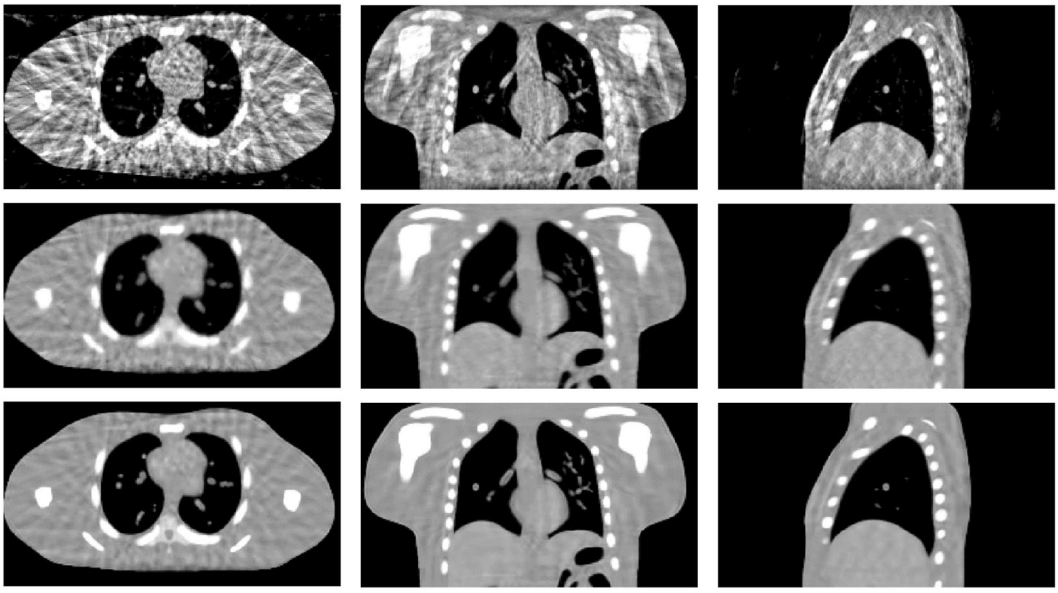

Figure 3. Reconstructions of 4D NCAT phantom with 31 projections for each phase. The first to third rows show 4D-CBCT at the begin-expiration phase reconstructed by FDK, SART-TV and proposed MgSS algorithms, respectively.

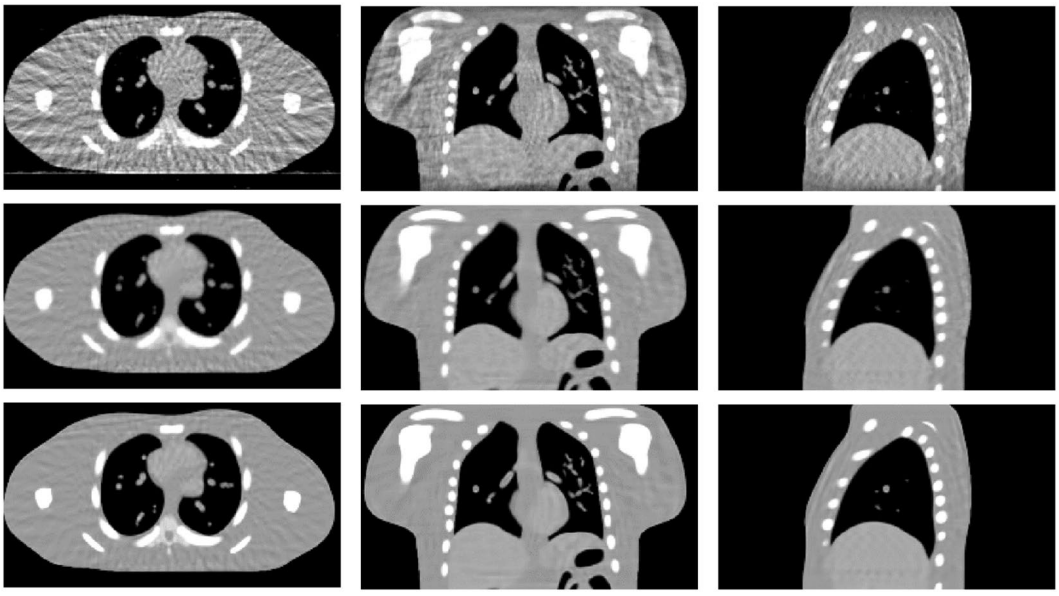

Figure 4. Reconstructions of 4D NCAT phantom with 51 projections for each phase. The first to third rows show 4D-CBCT at the begin-expiration phase reconstructed by FDK, SART-TV and proposed MgSS algorithms, respectively.

$$\Theta^*_{MgSS} = \hat{S}_{MgSS} \times_1 U^{(1)} \times_2 U^{(2)} \times_3 U^{(3)} \times_4 U^{(4)} \tag{6}$$

W is a weighted averaging operator indicates how the cubes are merged back into the image. The elements of W is the reciprocal of the times which one pixel was overlapped by different cubes.

In our implementation, the MgSS step is carried out after 10th iteration of SART step until $\left|f^{(n)} - f^{(n-1)}\right|^2 < \xi$ or reached the predetermined iterative number. Also one implicit motion vector field estimation step was performed using the intermediate reconstructions of SART.

Results

Data acquisition of digital simulation. *NCAT phantom based simulation.* In this work, the 4D NURBS-based Cardiac-Torso (NCAT)[48] phantom, which is capable of providing the realistic model of human anatomy

Views	Method	PhaseNumber (Tumor diameter:10 mm)									
		one	two	three	four	five	six	seven	eight	nine	ten
21	FDK	8.72E-03	8.77E-03	8.90E-03	9.02E-03	9.11E-03	9.14E-03	9.13E-03	9.04E-03	8.90E-03	8.72E-03
	SART-TV	2.16E-03	2.16E-03	2.17E-03	2.15E-03	2.16E-03	2.17E-03	2.17E-03	2.16E-03	2.18E-03	2.19E-03
	MgSS	9.11E-04	9.10E-04	9.09E-04	9.05E-04	9.07E-04	9.10E-04	9.09E-04	9.09E-04	9.15E-04	9.17E-04
31	FDK	8.33E-03	8.38E-03	8.53E-03	8.67E-03	8.77E-03	8.83E-03	8.82E-03	8.74E-03	8.59E-03	8.41E-03
	SART-TV	1.99E-03	1.93E-03	1.97E-03	1.98E-03	1.96E-03	1.98E-03	1.97E-03	1.98E-03	1.97E-03	1.98E-03
	MgSS	8.73E-04	8.74E-04	8.74E-04	8.70E-04	8.71E-04	8.71E-04	8.68E-04	8.67E-04	8.66E-04	8.69E-04
41	FDK	8.12E-03	8.07E-03	7.92E-03	7.79E-03	7.68E-03	7.62E-03	7.63E-03	7.71E-03	7.86E-03	8.04E-03
	SART-TV	1.80E-03	1.81E-03	1.82E-03	1.81E-03	1.83E-03	1.80E-03	1.79E-03	1.82E-03	1.81E-03	1.80E-03
	MgSS	8.11E-04	8.10E-04	8.11E-04	8.08E-04	8.11E-04	8.10E-04	8.10E-04	8.09E-04	8.11E-04	8.13E-04
51	FDK	8.00E-03	7.96E-03	7.82E-03	7.68E-03	7.58E-03	7.52E-03	7.53E-03	7.62E-03	7.76E-03	7.93E-03
	SART-TV	1.74E-03	1.75E-03	1.74E-03	1.73E-03	1.76E-03	1.74E-03	1.75E-03	1.76E-03	1.73E-03	1.74E-03
	MgSS	7.96E-04	7.93E-04	7.96E-04	7.92E-04	7.94E-04	7.93E-04	7.92E-04	7.89E-04	7.93E-04	7.95E-04

Table 1. rRMSE measures on the reconstructions of the FDK, SART-TV and MgSS reconstructions, respectively.

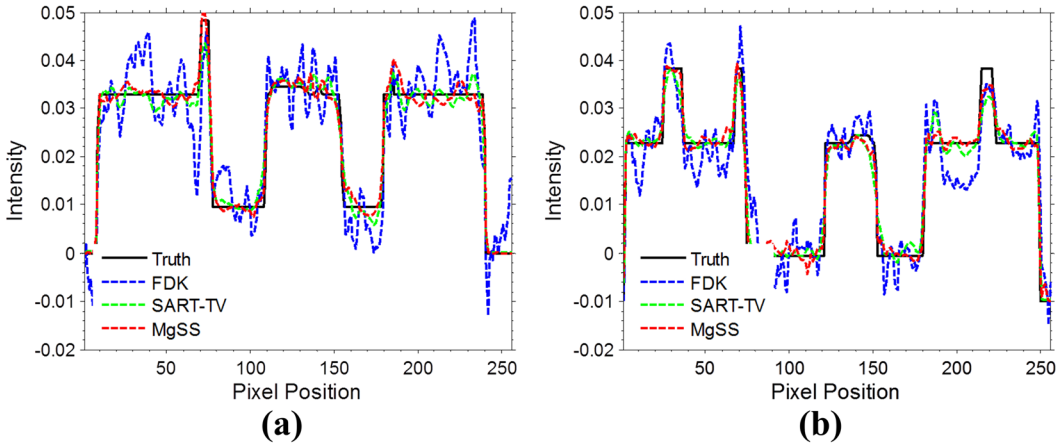

Figure 5. Horizontal profiles of reconstructions shown in Fig. 2. The profiles located at the pixel position x from 0 to 256 and y = 46 (**a**), y = 67 (**b**). The 'black line' generated from the NCAT phantom acts as the ground-truth for comparison.

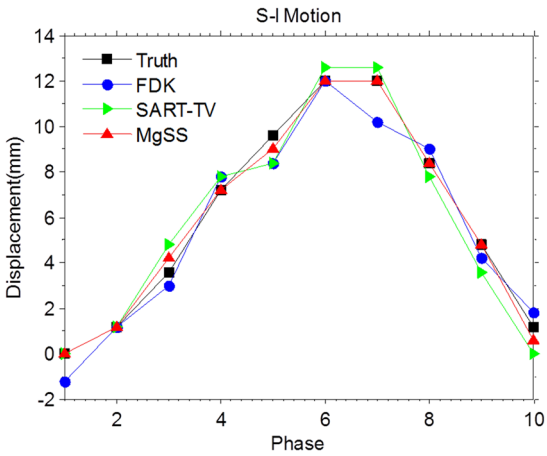

Figure 6. Motion trajectories extracted from the result of 4D-CBCT along the superior–inferior direction.

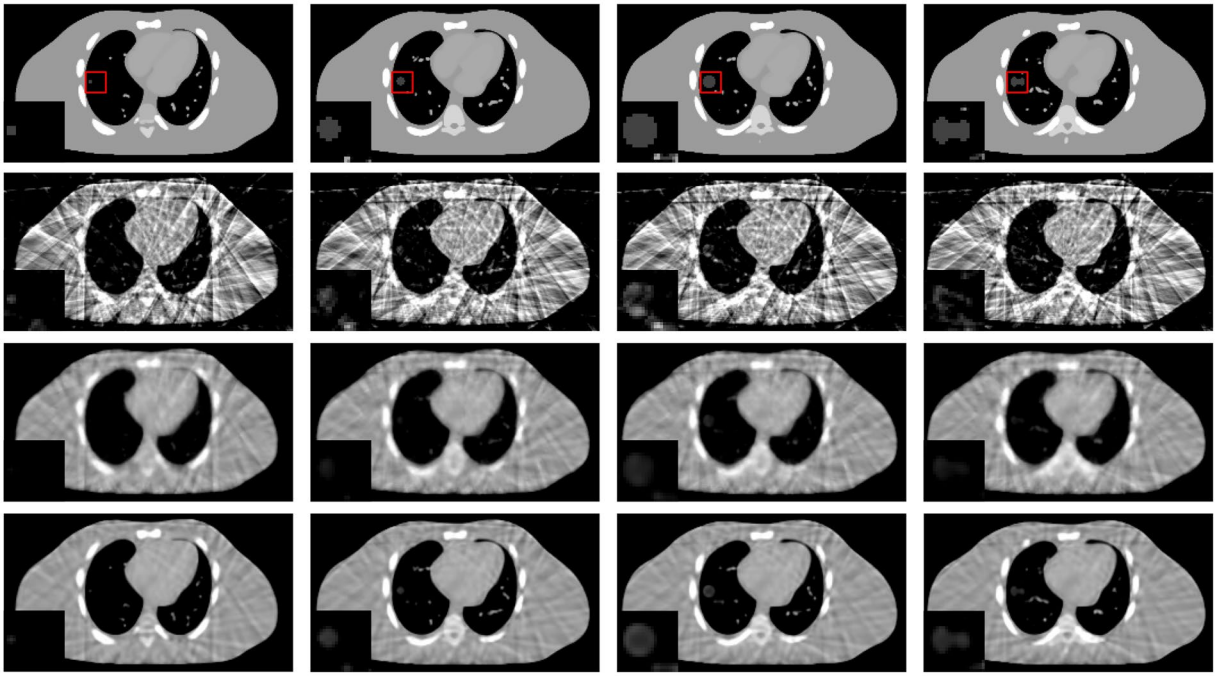

Figure 7. Reconstructions of 4D NCAT phantom with 21 projections. The first to the fourth columns show the transverse planes of 4D NCAT phantom with tumors of diameters: 6 mm, 16 mm, 22 mm and 28 mm, respectively.

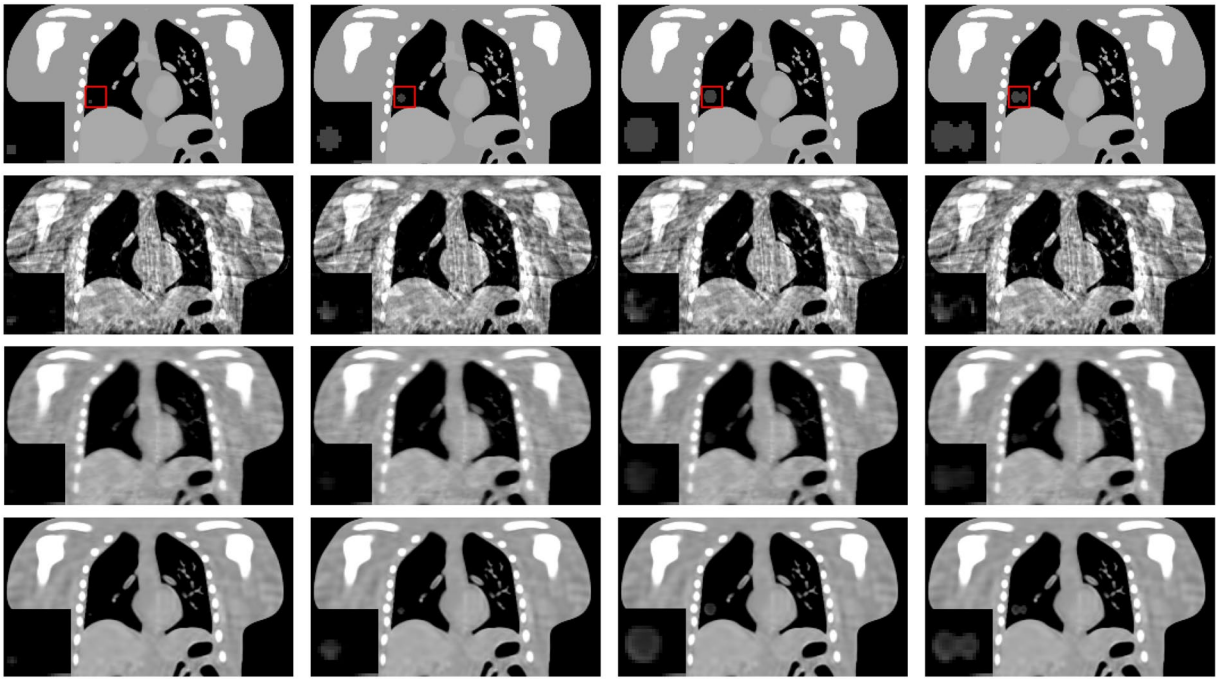

Figure 8. Reconstructions of 4D NCAT phantom with 21 projections. The first to the fourth columns show the coronal planes of 4D NCAT phantom with tumors of diameters: 6 mm, 16 mm, 22 mm and 28 mm, respectively.

and simulating cardiac and respiratory motion simultaneously, was used for data simulation. A dynamic phantom with ten respiratory phases and breathing period set to be 5 s was generated. The maximum diaphragm motion and the maximum chest anterior–posterior motion is 20 mm and 5 mm. We manually added several tumors with different sizes, contrast and shapes in the right lung field of phantom to test the robustness of the MgSS algorithm. The diameters of the spherical tumors were: 6 mm, 10 mm, 16 mm and 22 mm. The long diameter and short diameter of the non-spherical tumor is: 28 mm and 22 mm. For all simulations, the size of the digital phantom

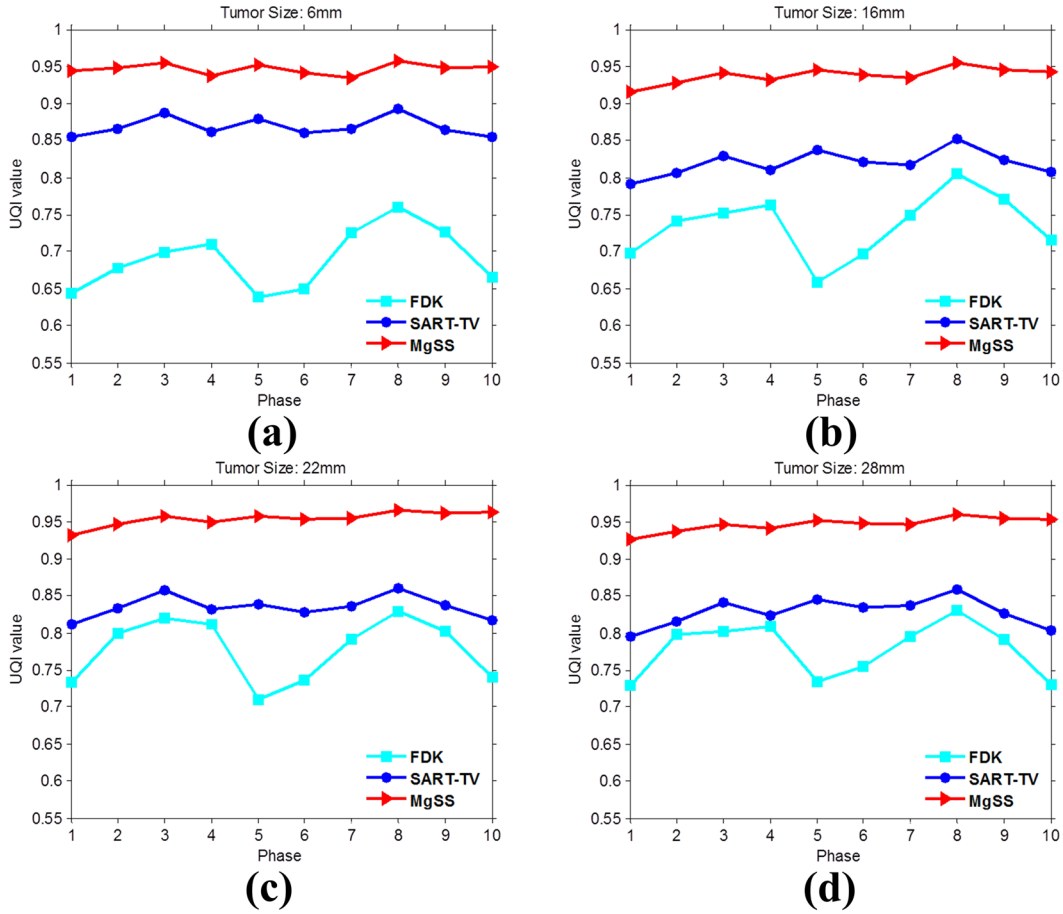

Figure 9. The UQI measures on the ROIs in Fig. 8 for ten phase.

is $256 \times 256 \times 150$, with voxel size of $2 \times 2 \times 2$ mm^3. The projections were generated by utilizing fast ray-tracing technique. The projection size of each angular view is 300×200 with detector pixel size 2×2 mm^2. In the process of simulation, projection views are evenly distributed over 360 degrees with the projection number of each phase range from 21 to 51. The noisy signal S_i at each detector bin i was simulated based on the Poisson noise model:

$$S_i = Poisson(I_0 \exp(-y_i)) + Normal(0, \sigma_e^2) \tag{7}$$

Here I_0 and σ_e^2 represent the incident x-ray intensity and the background noise, respectively. I_0 is set to be 2×10^6 and σ_e^2 is chosen to be 10.

4DCT based simulation. To further evaluate the performance of the MgSS algorithm, the 4DCT based digital phantom simulation was also performed. The 4DCT images were acquired on a 16-slice helical CT scanner (Brilliance Big Bore, Philips Medical Systems, Andover, MA). The three dimensional CT volumes at each phase were first interpolated to be isotropic data set with voxel size $1.0254 \times 1.0254 \times 1.0254$ mm^3. Then CB projections were computed from the reference 3D CT image using the projection matrix. The scan geometry was chosen according to the Varian On-Board Imager® and True Beam™ CBCT units.

Data acquisition of patient. The patient data was downloaded from an open data website (http://wiki.openrtk.org/index.php/ RTK/Examples/MCCBCTReconstruction). The cone-beam projections were acquired on the Elekta Synergy system. The clinical dataset consisted of 644 projections and were sorted to 10 phases based on the AS method[49]. The size of the digital flat panel is 512×512 with the pixel size of 0.8×0.8 mm^2. In our experiments, the isotropic reconstruction target resolution was set to be $1.2 \times 1.2 \times 1.2$ mm^3 on a $256 \times 256 \times 200$ matrix.

Evaluation metrics. To quantitative evaluate the performance of the proposed algorithm, we calculate the relative root mean square error (rRMSE) between the phantom images and the reconstructions. The rRMSE is defined as:

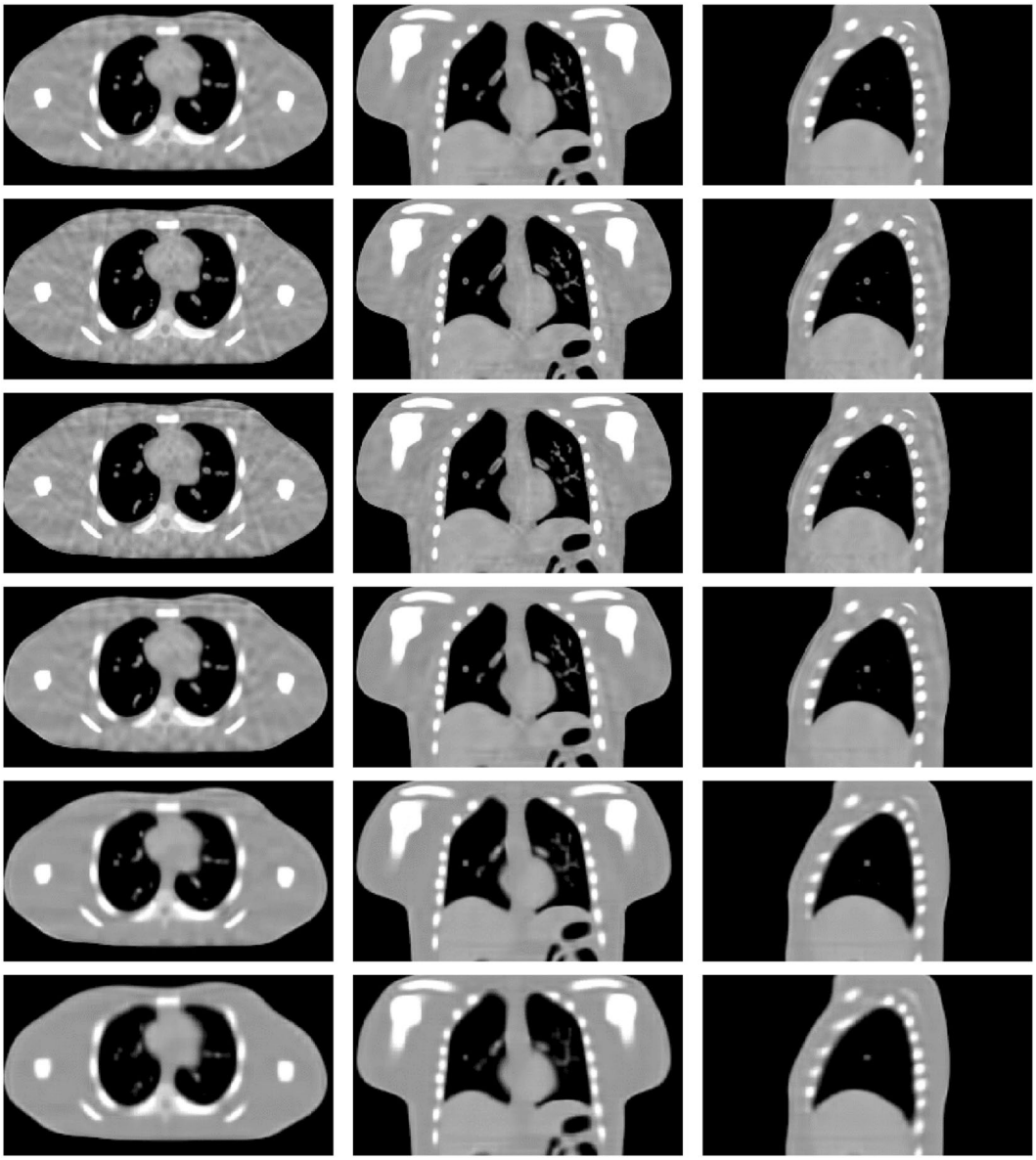

Figure 10. Ten phases reconstructions from -21 projections with block size set to be $5 \times 5 \times 5$, $7 \times 7 \times 7$, $9 \times 9 \times 9$, $13 \times 13 \times 13$, $17 \times 17 \times 17$, $23 \times 23 \times 23$.

$$\mathrm{rRMSE} = \sqrt{\sum_{m=1}^{Q}\left|f_m - f_{\mathrm{p},m}\right|^2 / \sum_{m=1}^{Q}\left|f_{\mathrm{p},m}\right|^2} \tag{8}$$

Here, f denotes the target image, f_p denotes the phantom image, and m is the voxel index.

The universal quality index[50] (UQI) index was utilized to conduct region of interest (ROI) based analysis by evaluating the degree of similarity between the reconstructed and the reference images. We select ROIs including the tumor and lung details within the reconstructed and reference images, the mean, variance and covariance of intensities in the ROIs can be respectively calculated as:

$$\overline{f} = \frac{1}{Q}\sum_{m=1}^{Q} f_m, \qquad \sigma^2 = \frac{1}{Q-1}\sum_{m=1}^{Q}(f_m - \overline{f})^2 \tag{9}$$

$$\overline{f}_{ture} = \frac{1}{Q}\sum_{m=1}^{Q} f_{ture,m}, \ \sigma_{ture}^2 = \frac{1}{Q-1}\sum_{m=1}^{Q}(f_{ture,m} - \overline{f}_{ture})^2 \tag{10}$$

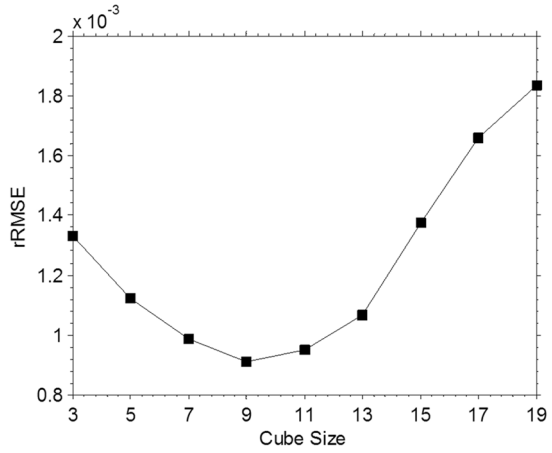

Figure 11. The averaged rRMSEs of ten phases reconstructions with different cube size.

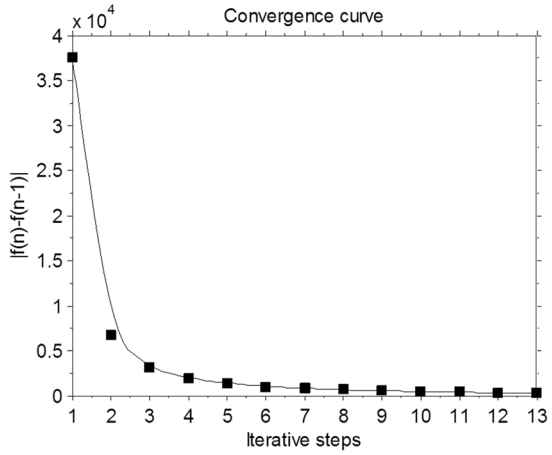

Figure 12. The convergence curve of the MgSS algorithm.

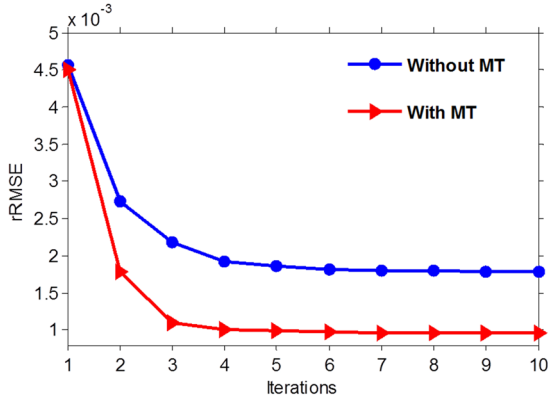

Figure 13. The rRMSE measures as a function of the number of iterations for the MgSS algorithm with and without motion tracking.

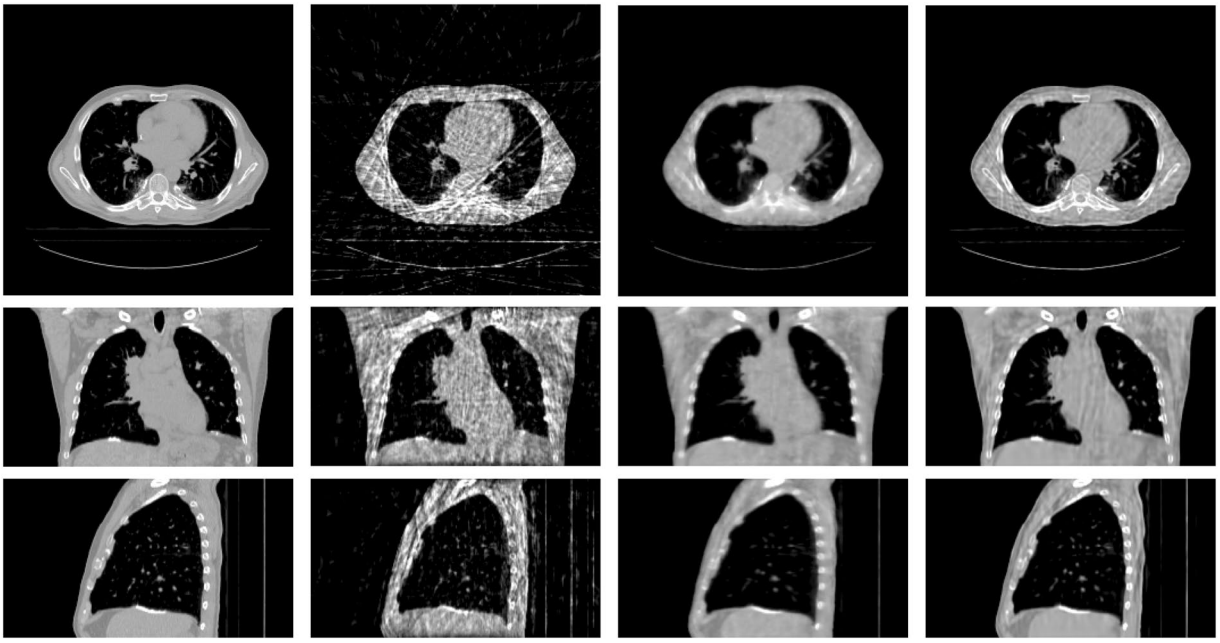

Figure 14. Results of realistic digital phantom with 21 projections for each phase. First column shows the begin-expiration phase of the 4DCT images. The second to last columns show 4D-CBCT images at the begin-expiration phase reconstructed by using FDK, SART-TV and proposed MgSS algorithms, respectively.

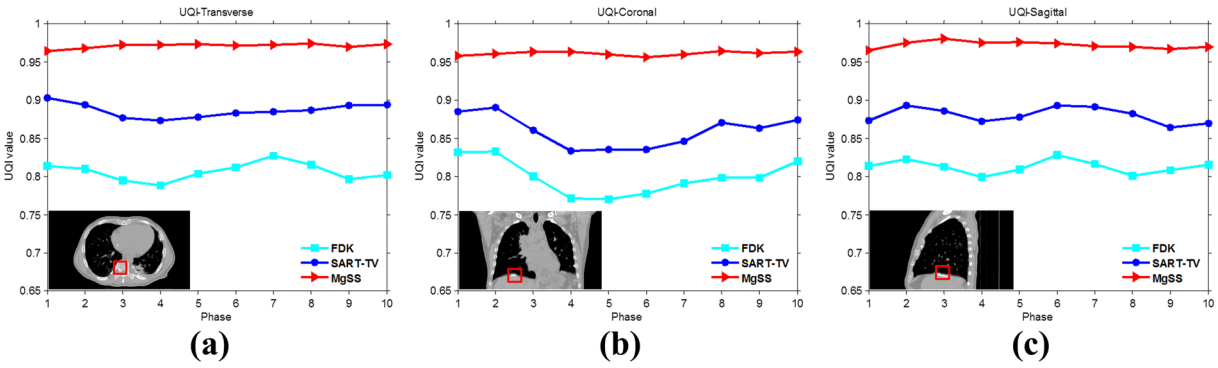

Figure 15. The UQI tests on ten phases reconstructions from -21views projection at the transverse, sagittal and coronal planes separately.

$$cov(f, f_{ture}) = \frac{1}{Q-1}\sum_{m=1}^{Q}(f_m - \overline{f})(f_{ture,m} - \overline{f}_{ture}) \tag{11}$$

the f_{ture} denotes the golden standard image, m is the voxel index, and Q denotes the number of voxels within the ROI. The UQI can be calculated as follows:

$$UQI = \frac{2cov(f, f_{ture})}{\sigma^2 + \sigma_{ture}^2}\frac{2\overline{f}_{ture}\overline{f}}{\overline{f}^2 + \overline{f}_{ture}^2} \tag{12}$$

UQI measures the intensity similarity between the two images, and its value ranges from zero to one. A UQI value closer to one suggests better similarity to the reference image.

Digital NCAT phantom study. *Visual inspection.* Figure 2 shows the results of the NCAT phantom reconstructed by using different methods at transverse, coronal, and sagittal planes for phase #1 with the projection number set to be 21. The columns one to three show the transverse, coronal, and sagittal images, respectively. First row in Fig. 2 shows the designed digital phantom images. The second row shows the results which were reconstructed by using the FDK algorithm from all projections. As we can see, the motion blurring artifacts are

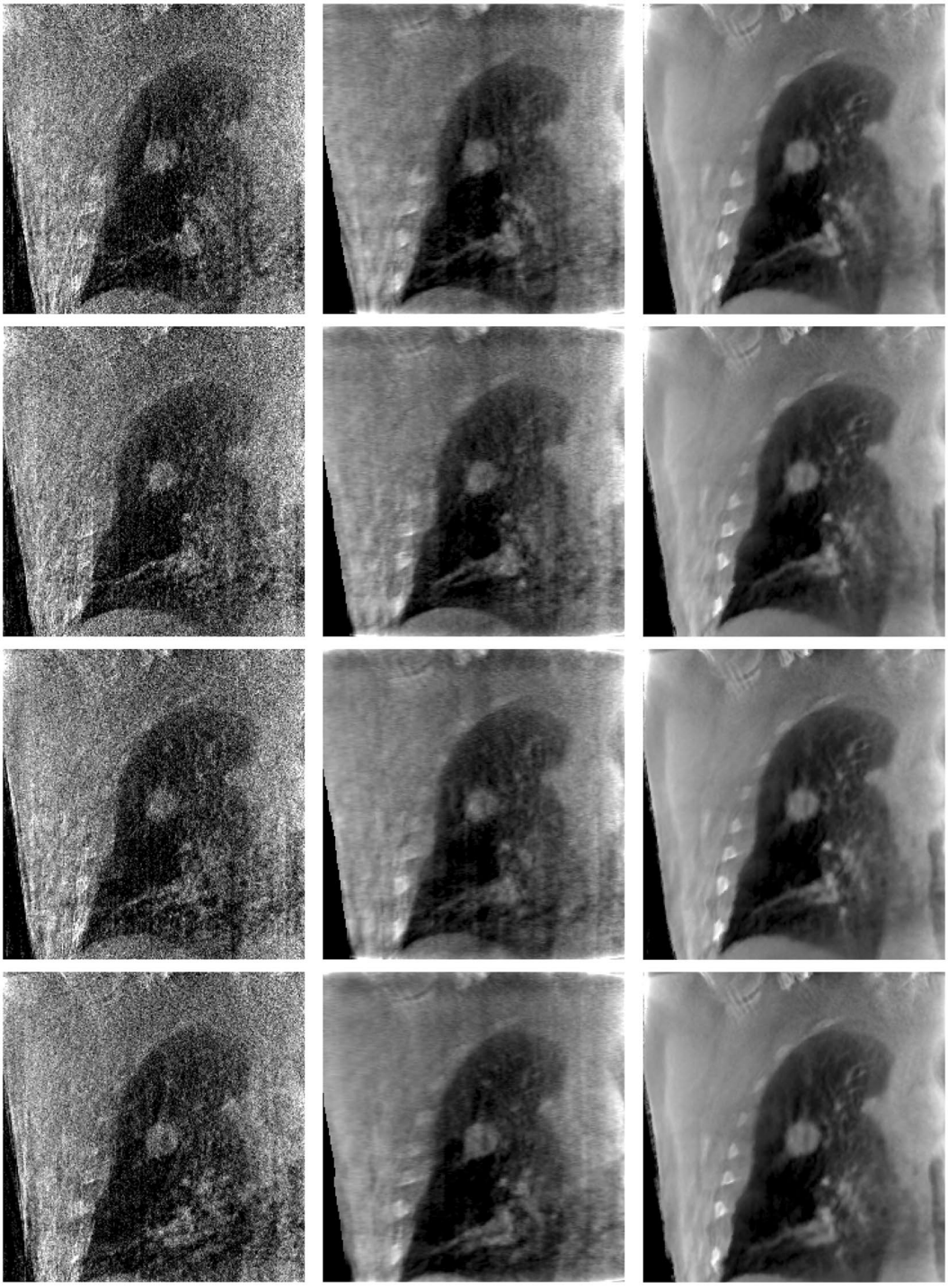

Figure 16. Results of patient data reconstructed by using different methods. The left to the right columns show the 4D-CBCT reconstructed by FDK, SART-TV and proposed MgSS algorithm, respectively.Each row shows images at the different phases in the breathing cycle: 20–30%, 40–50%, 60–70%, 80–90%.

obvious. The third row shows the phase correlated 4D-CBCT images reconstructed by the FDK algorithm, and we can see that the FDK reconstructions are full of noise and streak artifacts. For the SART-TV reconstructions, the fine structures inside the lung area are severely blurred though the view aliasing artifacts are suppressed. Compared to FDK and SART-TV algorithms, the proposed MgSS approach can yield images with superior quality. Moreover, Figs 3 and 4 demonstrate the reconstructions of -31 and -51 views, which further illustrate the gains of the proposed MgSS approach.

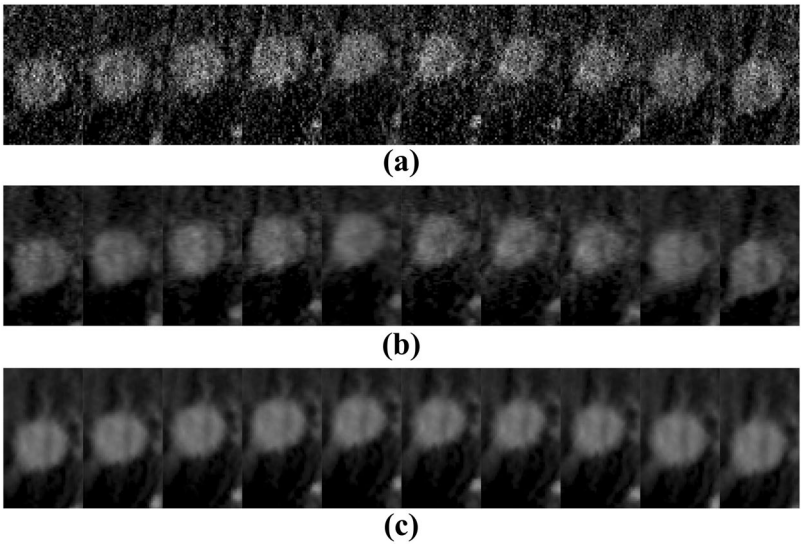

Figure 17. The zoomed tumor areas in the reconstructions of ten phases. The first to the third rows show the tumor image which were reconstructed by FDK, SART-TV, and MgSS, respectively.

Quantitative evaluation. Table 1 presents the results of rRMSE measures with the projection views range from 21 to 51 for each phase. The rRMSEs of ten phases reconstructions by using the proposed MgSS reduced by an average of 43% compared to those of FDK reconstructions. It can be obviously seen that the MgSS algorithm can achieve smaller rRMSE values compared to the other algorithms, which suggests the promising performance of the proposed MgSS approach.

Figure 5 illustrates the horizontal profiles of the transverse plane images in Fig. 2. It can be observed that the profiles obtained from the MgSS reconstructions match much better than the others, which suggest that the present MgSS approach achieves more noticeable gains than other approaches.

Motion trajectory accuracy. Because of the various applications of IGRT is interested in the tumor position information, we are devoted to extract the motion trajectories from different reconstruction schemes: FDK, SART-TV and the proposed MgSS algorithms. We define the motion trajectories which extracted from the NCAT phantom with high-contrast tumor as a reference. In this paper, we extract the movement information from the center of tumor. The motion trajectories have been showed in Fig. 6. As we can see, the motion information of tumor extracted from MgSS algorithm matches well with the reference trajectory, while the trajectories extracted from the FDK diverges the reference trajectory.

Low contrast lesion detection study. To test the robustness of the MgSS algorithm in reproducing the low contrast lesion, Figs 7 and 8 present the reconstructions of NCAT phantom with different tumor sizes and shapes at the transverse and coronal planes. The first row shows the designed phantom image used for visual comparison. The second to fourth rows show 4D-CBCT at the begin-expiration phase reconstructed by using FDK, SART-TV and proposed MgSS algorithms, respectively. It can be seen, the low contrast tumors in the NCAT phantom can't be completely rebuilt by the FDK and SART-TV algorithms. The morphology of tumors is partly destroyed. The phenomenon is particularly evident for tumors with small size. On the contrast, the MgSS algorithm can mostly recover the tumors with relatively complete morphological structures. Furthermore, Fig. 9 presents the UQI test on the ROIs shown in Fig. 8. All the results suggest that, with the introduction of phase-correlated information, the MgSS algorithm has better lesion detection ability than algorithms that only using the single phase information.

Parameter selection. In our method, there are two parameters need to be tuned: (1) the tracked cube size; (2) the soft-thresholding coefficients for the HOSVD processing. The cube size plays an important role in our work as it not only directly affects the accuracy of the results, but also affects the computation time. Traditionally, for the block-based techniques, the size of the cube is an empirical parameter specified by the user. A reasonable cube size can help us attend to the local structure characteristics while removing undesirable distortions.

To study the influence of the cube size on the proposed algorithm, we experimentally change the size of cube to get the associated reconstructions and calculate the rRMSEs between the reconstructions and the designed phantom image. We reconstruct the NCAT phantom from -21 projections with block size set from $3 \times 3 \times 3$ to $23 \times 23 \times 23$. Figure 10 shows the reconstructions of cube size to be $5 \times 5 \times 5$, $7 \times 7 \times 7$, $9 \times 9 \times 9$, $13 \times 13 \times 13$, $17 \times 17 \times 17$, $23 \times 23 \times 23$. As shown in Fig. 10, on one hand, boundary distortion could be observed when the cube size is too small. On the other hand, the image blur increases with the increase of the cube size. This phenomenon is due to the large cube contains too much structural information, the movement of central voxel in

the cube is not enough to describe the whole cube. Thus, to generate high quality 4D-CBCT image, appropriate size of the cube is needed. Moreover, as shown in Fig. 11, the averaged rRMSEs of ten phases reconstructions indicate that with the cube size range from $7 \times 7 \times 7$ to $11 \times 11 \times 11$, we can get relative smaller rRMSEs. In order to balance the reconstruction quality and computational time, in our other experiments, the cube size was fixed to $9 \times 9 \times 9$.

For the parameter of thresholding coefficients for the HOSVD processing, we choose with a threshold of $\sigma\sqrt{2\log p^2}$ to manipulate the coefficients of core tensor. This is also an experiential selection as used in other works[51]. Here, σ is defined as the standard deviation of a uniform region in the intermediary images during the iteration and p is the cube size.

Algorithm convergence. To validate and analyze the convergence of the present MgSS method, the $|f(\mathbf{n}) - f(\mathbf{n} - 1)|$ (absolute value of differences between two adjacent estimations) measuring on the entire to-be-reconstructed NCAT phantom image were performed. Figure 12 shows the $|f(\mathbf{n}) - f(\mathbf{n} - 1)|$ measures with respect to the number of iterations. Results show that the present MgSS algorithm can yield a steadily convergence solution.

Influence of motion tracking. To demonstrate the effect of the estimated motion fields on the reconstruction, we plot the rRMSE measures as a function of the number of iterations for the MgSS with and without motion tracking. Figure 13 shows the benefits of motion guidance, as MgSS with motion tracking can obtain reduced rRMSE as compared to MgSS without motion tracking. Figure 13 also illustrates the convergence of the proposed MgSS algorithm.

Realistic 4D CT based digital phantom study. Figure 14 shows the results of the realistic digital phantom reconstructed by using different methods at transverse, coronal, and sagittal planes for phase #1 with the projection number set to be 21. First column in Fig. 14 shows the 4DCT images using as the golden standard for comparison. The second column shows 4D-CBCT images reconstructed by the FDK algorithm. It can be seen that the FDK reconstructions are seriously contaminated by noise and artifacts, and some anatomical structures can't be clearly seen. For the SART-TV reconstruction, some fine structures have been erased though most of the view aliasing artifacts are suppressed. Compared to FDK and SART-TV algorithms, the proposed MgSS approach can yield images with superior quality. In addition, the UQI test between the reconstructions and golden standard image were calculated. Figure 15 shows the test results of then phases at the transverse, sagittal and coronal planes, separately. For then phases, the UQI values of MgSS reconstructions are always higher than 0.95, which suggests the promising performance of the proposed MgSS algorithm.

Patient study. Figure 16 shows the representative reconstructions of patient data by using the FDK, SART-TV and MgSS methods. Each row of Fig. 16 shows the image of 4D-CBCT at different phase in respiratory cycle: 20–30%, 40–50%, 60–70%, 80–90%. The first column in Fig. 16 shows the sagittal view of 4D-CBCT images reconstructed by conventional FDK. Because of the limited number of projections at each phase, severe noise and artifacts present in the FDK reconstruction. The second column of Fig. 16 shows the results of SART-TV algorithm. Although noise and artifacts have been remarkably reduced, details within the lung area and edges of bony structure can't be clearly seen. Last column of Fig. 16 shows the images reconstructed by MgSS method, from which we can observe noise is suppressed. The boundaries of bony structures as well as fine structures inside the lung are well preserved. To further illustrate the performance of our algorithm, zoomed images of the tumor areas in the then phasesof the reconstructions by using different reconstruction are presented (Fig. 17).

Discussion

In this work, we developed a MgSS algorithm to improve the image quality of 4D-CBCT. This algorithm is developed on the assumption that there exists high structural similarity between the images of neighbored phases, which is similar with Chen's work on dynamic MR reconstruction[43]. In Chen's work, two dimensional patches are tracked with estimated motion maps and then SVD was used to decompose the tacked cluster. For the standard SVD, the image blocks are manually vectorized and then the image structural properties in the spatial domain are ignored. For the proposed MgSS algorithm, the 3D-MVFs between phases were utilized to track the three dimensional cubes and then form the cluster Θ_{MgSS} with the size of $N_b \times N_b \times N_b \times N_t$. To preserve the structural properties, in this work, the higher order singular value decomposition which is able to directly decompose dynamic datasets into a multidimensional singular matrix rather than unfolding the cubes into column vectors, was used to process the four dimensional cluster Θ_{MgSS}. The MgSS algorithm can reveal the fine details shared by grouped blocks and preserve the essential unique features of each individual block. Results of digital phantom and patient data demonstrate the proposed approach can significantly suppress the view aliasing artifacts and noise.

One important step in the proposed MgSS algorithm is: cube based motion tracking. We utilize the Real-Time Image-based Tracker (RTIT)[36] toolbox to obtain the voxel-by-voxel displacement maps between 3D images of different phases. Specifically, we rely mainly on the displacement of the center voxel of cube in the current phase to determine the cube in the next phase. We assume the displacements of region are changing smoothly. Thus, the central voxel is sufficient to describe the cube motion. And in this case, the accuracy of 3D-MVFs estimation would play an important role in finding cubes with similar structures. In this work, the initial 3D-MVFs are generated after the first SART iteration, which may not so accurate. Notably, as the iteration goes on, the 3D-MVFs would be updated with intermediate reconstructions, and therefore both the image quality and the accuracy of the 3D-MVFs will be improved. Other approaches such like the 3D-MVFs initialized with those got from the 4D planning CT of the same patient can be considered. There is a very important parameter need to be tuned for the

proposed MgSS approach: the tracked cube size. A large size cube may include more details and local geometry, but on the other side, if the cube size is too large, the central voxel will be insufficient to describe the whole cube motion and also result in increasing computation time. Thus, the proper selection of the cube size is critical. In our studies, we change the block size from $3 \times 3 \times 3$ to $23 \times 23 \times 23$ and generate the reconstructions as shown in Fig. 10 and rRMSEs between the reconstructions and the phantom images. For visual inspection, too small or large cube size would lead to image blur or distortion. The rRMSE results suggest that when the cube size set to be $9 \times 9 \times 9$, we can get relative higher quality image with smaller rRMSE for the NCAT phantom. Also as illustrated in other studies of HOSVD technique, the appropriate selection of the cube size should base on the structural complexity of the target image and may be different for different images. But overall, in our work for lung CT imaging, we found the cube size set to be $9 \times 9 \times 9$ is robust enough to balance the structural information and computation time. Another issue of the proposed MgSS algorithm is the heavy computation burden. It takes about 30 minutes on a desktop computer (3.60 GHz Intel(R) i7 CPU with 8GB RAM) to run one iteration. On one hand, as mentioned above, we can short the computation time with a relatively small cube size as well as an elegant initialization. On the other hand, we can use a step of N_{step} pixels in transverse, coronal, and sagittal directions, respectively. Hence, the number of overlapping cubes is decreased by $1/N_{step}^3$. But with large N_{step} or large motion, gaps between moved blocks may appear, then a mask of the uncovered areas (gaps) on the n-th frame after block motion tracking can be detected and non-motion tracking is performed for the gap blocks to avoid potential additional gaps. Other techniques including using the graphics processing unit (GPU) can be also considered to short the computation time.

In summary, we have developed a MgSS algorithm to improve the image quality of 4D-CBCT. This method effectively utilizes the correlated information from other phases to reconstruct any particular phase of 4D-CBCT. By enforcing the regional spatiotemporal sparsity on the tracked cubes, noise and artifacts can be suppressed and the image quality of 4D-CBCT can be substantially improved.

References

1. Jaffray, D. A., Siewerdsen, J. H., Wong, J. W. & Martinez, A. A. Flat-panel cone-beam computed tomography for image-guided radiation therapy. *International Journal of Radiation Oncology* Biology* Physics* **53**, 1337–1349 (2002).
2. Oldham, M. *et al*. Cone-beam-CT guided radiation therapy: A model for on-line application. *Radiotherapy and oncology* **75**, 271. E1–271. E8 (2005).
3. Pouliot, J. *et al*. Low-dose megavoltage cone-beam CT for radiation therapy. *International Journal of Radiation Oncology* Biology* Physics* **61**, 552–560 (2005).
4. Chen, G. T., Kung, J. H. & Beaudette, K. P. Artifacts in computed tomography scanning of moving objects. *Semin Radiat Oncol* **14**, 19–26 (2004).
5. Sonke, J., Zijp, L., Remeijer, P. & van Herk, M. Respiratory correlated cone beam CT. *Medical physics* **32**, 1176–1186 (2005).
6. Li, T. *et al*. Four-dimensional cone-beam computed tomography using an on-board imager. *Medical physics* **33**, 3825–3833 (2006).
7. Lu, J. *et al*. Four-dimensional cone beam CT with adaptive gantry rotation and adaptive data sampling. *Medical physics* **34**, 3520–3529 (2007).
8. Ford, E. C., Mageras, G. S., Yorke, E. & Ling, C. C. Respiration-correlated spiral CT: a method of measuring respiratory-induced anatomic motion for radiation treatment planning. *Medical physics* **30**, 88–97 (2003).
9. Low, D. A. *et al*. A method for the reconstruction of four-dimensional synchronized CT scans acquired during free breathing. *Medical physics* **30**, 1254–1263 (2003).
10. Pan, T., Lee, T., Rietzel, E. & Chen, G. T. 4D-CT imaging of a volume influenced by respiratory motion on multi-slice CT. *Medical physics* **31**, 333–340 (2004).
11. Feldkamp, L. A., Davis, L. C. & Kress, J. W. Practical cone-beam algorithm. *JOSA A* **1**, 612–619 (1984).
12. Li, T., Koong, A. & Xing, L. Enhanced 4D cone-beam CT with inter-phase motion model. *Medical physics* **34**, 3688–3695 (2007).
13. Li, T. & Xing, L. Optimizing 4D cone-beam CT acquisition protocol for external beam radiotherapy. *International Journal of Radiation Oncology* Biology* Physics* **67**, 1211–1219 (2007).
14. Li, T., Schreibmann, E., Yang, Y. & Xing, L. Motion correction for improved target localization with on-board cone-beam computed tomography. *Phys Med Biol* **51**, 253–67 (2006).
15. Zhang, H. *et al*. Few-view cone-beam CT reconstruction with deformed prior image. *Med Phys* **41**, 121905 (2014).
16. Wang, J., Li, T. & Xing, L. Iterative image reconstruction for CBCT using edge-preserving prior. *Medical physics* **36**, 252–260 (2009).
17. Brock, R. S., Docef, A. & Murphy, M. J. Reconstruction of a cone-beam CT image via forward iterative projection matching. *Medical physics* **37**, 6212–6220 (2010).
18. Jia, X., Dong, B., Lou, Y. & Jiang, S. B. GPU-based iterative cone-beam CT reconstruction using tight frame regularization. *Physics in medicine and biology* **56**, 3787 (2011).
19. Sidky, E. Y. & Pan, X. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. *Physics in medicine and biology* **53**, 4777 (2008).
20. Sidky, E. Y. *et al*. Enhanced imaging of microcalcifications in digital breast tomosynthesis through improved image-reconstruction algorithms. *Medical Physics* **36**, 4920–32 (2009).
21. Kinnon, G. C. M. & Bates, R. H. T. Towards Imaging the Beating Heart Usefully with a Conventional CT Scanner. *IEEE Transactions on Biomedical Engineering* **BME-28**, 123–127 (1981).
22. Leng, S. *et al*. Streaking artifacts reduction in four-dimensional cone-beam computed tomography. *Medical Physics* **35**, 4649–59 (2008).
23. Leng, S. *et al*. High temporal resolution and streak-free four-dimensional cone-beam computed tomography. *Physics in Medicine & Biology* **53**, 5653–73 (2008).
24. Bergner, F. *et al*. Autoadaptive phase-correlated (AAPC) reconstruction for 4D CBCT. *Medical physics* **36**, 5695–5706 (2009).
25. Kida, S., Masutani, Y., Nakano, M., Imae, T. & Haga, A. Improvement of 4D Cone-beam CT image quality via iterative reconstruction method. *Ieice Technical Report* **112**, 73–75 (2012).
26. Bergner, F. *et al*. An investigation of 4D cone-beam CT algorithms for slowly rotating scanners. *Medical Physics* **37**, 5044–5053 (2010).
27. Haralick, R. M. Statistical and structural approaches to texture. *Proceedings of the IEEE* **67**, 786–804 (1979).
28. Lee, J. S. Digital Image Enhancement and Noise Filtering by Use of Local Statistics. *IEEE Transactions on Pattern Analysis & Machine Intelligence* **2**, 165–168 (1980).
29. Aharon, M., Elad, M. & Bruckstein, A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing* **54**, 4311–4322 (2006).

30. Buades, A., Coll, B. & Morel, J. M. A Review of Image Denoising Algorithms, with a New One. *Siam Journal on Multiscale Modeling & Simulation* **4**, 490–530 (2005).
31. Dabov, K., Foi, A., Katkovnik, V. & Egiazarian, K. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing* **16**, 2080–95 (2007).
32. Liu, C. & Freeman, W. T. A high-quality video denoising algorithm based on reliable motion estimation. *European Conference on Computer Vision* 706–719 (2010).
33. Dwivedi, A. & Shrivastava, S. K. An efficient and fast patch reordering approach for image denoising without losing structural information. *International Conference on Computer Communications* 1–6 (2014).
34. Parameswaran, S., Luo, E. & Nguyen, T. Q. Patch Matching for Image Denoising Using Neighborhood-based Collaborative Filtering. *IEEE Transactions on Circuits and Systems for Video Technology*, 1–1 (2016).
35. Costantini, R., Sbaiz, L. & Süsstrunk, S. Higher order SVD analysis for dynamic texture synthesis. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society* **17**, 42–52 (2008).
36. Ries, M. *et al*. Real-time 3D target tracking in MRI guided focused ultrasound ablations in moving tissues. *Magn Reson Med* **64**, 1704–12 (2010).
37. Zachiu, C., Senneville, B. D. D., Moonen, C. & Ries, M. A framework for the correction of slow physiological drifts during MR-guided HIFU therapies: Proof of concept. *Medical Physics* **42** (2015).
38. Zachiu, C., Papadakis, N., Ries, M., Moonen, C. & Denis, D. S. B. An improved optical flow tracking technique for real-time MR-guided beam therapies in moving organs. *Physics in Medicine & Biology* **60**, 9003–29 (2015).
39. Lingala, S. G., Hu, Y., Dibella, E. & Jacob, M. Accelerated Dynamic MRI Exploiting Sparsity and Low-Rank Structure: k-t SLR. *IEEE Transactions on Medical Imaging* **30**, 1042–1054 (2011).
40. Cai, J. F., Jia, X., Gao, H. & Jiang, S. B. Cine Cone Beam CT Reconstruction Using Low-Rank Matrix Factorization: Algorithm and a Proof-of-Principle Study. *IEEE Transactions on Medical Imaging* **33**, 1581–1591 (2014).
41. Akçakaya, M. *et al*. Low-dimensional-structure self-learning and thresholding: regularization beyond compressed sensing for MRI reconstruction. *Magn Reson Med* **66**, 756–67 (2011).
42. Vitanis, V. *et al*. High resolution three-dimensional cardiac perfusion imaging using compartment-based k-t principal component analysis. *Magn Reson Med* **65**, 575–87 (2011).
43. Chen, X., Salerno, M., Yang, Y. & Epstein, F. H. Motion-compensated compressed sensing for dynamic contrast-enhanced MRI using regional spatiotemporal sparsity and region tracking: Block low-rank sparsity with motion-guidance (BLOSM). *Magnetic resonance in medicine* **72**, 1028–1038 (2014).
44. Lathauwer, L. D., Moor, B. D. & Vandewalle, J. On the Best Rank-1 and Rank-(R 1, R 2, …, R N) Approximation of Higher-Order Tensors. *Siam Journal on Matrix Analysis & Applications* **21**, 1324–1342 (2000).
45. Pan, H., Huang, T. Z. & Ma, T. Two-step group-based adaptive soft-thresholding algorithm for image denoising. *Optik - International Journal for Light and Electron Optics* **127**, 503–509 (2016).
46. Jorgensen, J. S., Sidky, E. Y. & Pan, X. Quantifying admissible undersampling for sparsity-exploiting iterative image reconstruction in X-ray CT. *IEEE Trans Med Imaging* **32**, 460–73 (2013).
47. Andersen, A. H. & Kak, A. C. Simultaneous algebraic reconstruction technique (SART): a superior implementation of the ART algorithm. *Ultrasonic imaging* **6**, 81–94 (1984).
48. Segars, W. P. *et al*. Development and application of the new dynamic Nurbs-based Cardiac-Torso (NCAT) phantom (2001).
49. Zijp, L., Sonke, J. J. & Herk, M. V. Extraction of the Respiratory Signal from Sequential Thorax Cone-Beam X-Ray Images. In ICCR (2004).
50. Wang, Z., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society* **13**, 600 (2004).
51. Dabov, K., Foi, A. & Egiazarian, K. Image denoising with block-matching and 3D filtering. *Proceedings of SPIE - The International Society for Optical Engineering* **6064**, 354–365 (2006).

Acknowledgements

This work was partially supported by the National Natural Science Foundation of Chinaunder Grants (No. 81501466, No. 61571214, No. 81371544), the Natural Science Foundation of Guangdong Provinceunder Grants (No. 2015A030310018), theGuangzhou Science and Technology Project (No. 201710010099). The authors thank the RTK Group very much for the supplying of the opensource data of CBCT. The authors would also like to thank the anonymous reviewers for their constructive comments and suggestions that greatly improved the quality of the manuscript.

Author Contributions

Y.L. and H.Z. wrote the main manuscript text, Y.L., H.Z. and Z.B. did the whole experiments, X.T., D.Z., Q.F., J.M. and W.C. revised the manuscript. All authors reviewed the manuscript.

Additional Information

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2017