

Robust principle component analysis based four-dimensional computed tomography

Hao Gao¹, Jian-Feng Cai¹, Zuwei Shen³, and Hongkai Zhao²

¹ Department of Mathematics, University of California, Los Angeles, CA 90095, USA

² Department of Mathematics, University of California, Irvine, CA 92697, USA

³ Department of Mathematics, National University of Singapore, 117543, Singapore

E-mail: haog@math.ucla.edu, cai@math.ucla.edu, matzuows@nus.edu.sg, zhao@math.uci.edu

Abstract. The purpose of this article for four-dimensional (4D) computed tomography (CT) is three-fold. (1) A new spatiotemporal model is presented from matrix perspective with the row dimension in space and the column dimension in time, namely, Robust PCA based 4DCT model (Robust Principle Component Analysis based 4D CT). That is, instead of viewing 4D object as a temporal collection of three-dimensional (3D) images and looking for local coherence in time or space independently, we percept it as a mixture of low-rank matrix and sparse matrix and explore the maximum temporal coherence of spatial structure among phases. Here the low-rank matrix corresponds to the “stationary” background or reference state in space over time, while the sparse matrix stands for the “moving” or “changing” components, (under some proper sparsifying transform,) e.g., tumors or organs in motion. Furthermore, this Robust PCA based 4DCT model can be applicable in other imaging problems for motion reduction or/and change detection. (2) A dynamic data acquisition procedure, i.e., a temporally spiral scheme, is proposed that can potentially maintain the similar reconstruction accuracy while using fewer projections of the data. The key point of this dynamic scheme is to reduce the total number of measurements and hence the radiation dose by acquiring complementary data in different phases without redundant measurements of the common background structure. (3) An accurate, efficient, yet simple-to-implement algorithm based on split Bregman method is developed for solving the model problem with the sparse representation in tight frames.

1. Introduction

Respiratory motion can degrade the image quality of computed tomography (CT), and consequently cause substantial errors in the dose delivery for thoracic and upper abdominal tumors in radiation therapy [1, 2]. With the time-resolved data acquisition, Four-dimensional (4D) CT possesses an unprecedented capability for accurate patient imaging and treatment planning in spite of organ/tumor motion [3-7].

Two methodologies for 4D CT algorithms exist. In the first one, different temporal phases (each phase corresponding to a 3D image) are essentially considered as independent phases in image reconstruction. For example, with an external respiratory signal for synchronization, the acquired projection data are binned into different phases according to amplitude or phase-angle sorting [8], after which the standard 3D reconstruction is carried out for each phase. There is no correlation between phases in reconstruction to this point. To alleviate the view-aliasing artifacts due to the reduced number of projections, the image registration based on a deformable model of respiratory motion can be used either in image space [9], or in data space with an artifact-free reference image [10]. The similar ideas also appear in other 4D imaging techniques [11], such as 4D positron emission tomography (PET) [12].

In contrast, in the second methodology, the “time” dimension is explicitly incorporated into the reconstruction algorithm. That is all the phases are treated as a single entity. An apparent reason for

this temporal fusion is that the images at different phases are intrinsically interconnected to each other due to some underlying physical or biological mechanism. As a result, this spatiotemporal synthesis feature is highly desirable in any truly 4D algorithm. For example, a spatiotemporal regularization via non-local means is utilized to enforce the temporal similarity between two consecutive phases in 4D CT [13]. A unified spatiotemporal strategy is also considered in 4D inverse planning for intensity modulated radiation therapy [14].

In this paper, we will present a different spatiotemporal model for 4D CT from matrix perspective, namely, Robust PCA based 4DCT model (Robust Principle Component Analysis based 4D CT), i.e., RPCA-4DCT model in abbreviation. That is, instead of viewing 4D object as a temporal collection of 3D images and looking for local coherence in time or space independently, we percept it as a mixture of low-rank matrix and sparse matrix, with the row dimension in space and the column dimension in time, and explore the maximum temporal coherence of spatial structure among phases. Here the low-rank matrix corresponds to the “stationary” background or reference state over time, while the sparse matrix stands for the “moving” or “changing” components, (under some proper sparsifying transform,) e.g., tumors or organs in motion. Here the image sparsity is enforced in wavelet tight frame domain rather than itself [15].

In addition, we will also introduce a dynamic data acquisition scheme to maximize the utility of the RPCA-4DCT model, and develop an efficient solution algorithm. Specifically, a temporally spiral scanning procedure can potentially maintain the similar reconstruction accuracy while using fewer projections of the data that are complementary at different phases to avoid redundant measurements of the common background structure; while being accurate, split Bregman method offers an extremely efficient, yet simple-to-implement strategy for solving a class of general l_1 -type problems, including the proposed RPCA-4DCT model [16, 17].

The RPCA-4DCT model is motivated by the recent work for data analysis in statistics, i.e., Robust Principle Component Analysis (RPCA) [18]. That is, with the data matrix consisting of a low-rank part and a sparse part, both can be (almost) exactly recovered by minimizing the sum of nuclear norm of the low-rank component and l_1 norm of the sparse component subject to certain assumptions (incoherence conditions). The similar models have been considered in several applications, such as video denoising [19], video surveillance [18], face recognition [18] and others [20-25]. A key difference of the RPCA-4DCT model from RPCA and most existing applications is that the available data here are tomographic measurements linked to the object through some ill-posed system matrix rather than directly from the object itself. Although the required incoherence conditions for guaranteeing the success of RPCA to exactly recover both low-rank and sparse matrices cannot be rigorously justified in such an illposed inverse problem, we will show in this paper that the RPCA-4DCT model indeed offers not only the improved overall image quality for 4D CT, but also a quite satisfactory decomposition into background and motion/change. Moreover, the RPCA-4DCT model can be augmented when in conjunction with tight frame transform, a dynamic data acquisition scheme pertinent to 4D CT, and split Bregman method. Finally, we remark that the proposed RPCA-4DCT model is a general model that can be potentially applicable in other imaging problems aiming at motion reduction or/and change detection besides 4D CT.

2. Models and Algorithms

2.1. Model

The 4D object to be imaged can be viewed as a temporal sequence of 3D images, i.e.,

$$X = \{x_j, j \leq n_t\}, \quad (2.1)$$

where X is the 4D object with temporal index j , and x_j corresponds to one of n_t phases, that is usually a piecewise-constantly discretized 3D image in space. Note that the respiratory cycle is generally assumed in order to acquire enough data for reconstructing each phase in the cycle [3]. However, this temporal periodicity is not necessary in the model formulation, thus is not assumed in (2.1). The

assumption we impose on the model in this study is fairly natural and practical, i.e., the temporal variation of X in space is “sparse” (with some proper sparsifying transform) with respect to a “stationary” background. Shortly, this assumption will be quantified as a matrix decomposition model, i.e., RPCA-4DCT model, with each component characterized in the proper norm.

The available data at each phase are

$$Y = \{y_j := A_j x_j + N_j, j \leq n_t\}. \quad (2.2)$$

Here y_j is assumed to be the x-ray transform of x_j with certain measurement noise N_j , and A_j corresponds to the system matrix that can be assembled according to line integrals in the image space between source-detector pairs [26]. Originally the system matrix should be independent of the index j since it is usually determined solely by the scanning geometry. In the following we will introduce a dynamic scanning strategy with the consequent dynamic system matrix to explore the possibility of 4D low-dose CT with fewer projections of data. On the other hand, the model and the algorithm discussed later also apply to other formulations of the system matrix, such as Fourier-based one [26].

2.1.1. Existing models. In 4D CT, one tries to reconstruct the 4D object X from its projection data Y . An apparent way is to reconstruct x_j solely from y_j for each phase, and then post-process x_j 's altogether for artifact reduction (if necessary), i.e.,

$$\begin{cases} x_j = \arg \min_{x_j} \|A_j x_j - y_j\|^2 + R(x_j), j \leq n_t \\ X = F(x_1, \dots, x_j, \dots, x_{n_t}) \end{cases}. \quad (2.3)$$

The first equation of (2.3) represents the solution of x_j through the minimization of a least-square data fidelity term and a regularization term R on the image x_j . Here the regularization is necessary for reducing the image artifact that may be due to the noise or the insufficient number of projections. This is a well-known iterative reconstruction strategy that is commonly used in algebraic reconstruction techniques (ART) when the system matrix is underdetermined, i.e., the number of data is less than the number of unknowns [26]. For comparison of models, L_2 regularization and total variation (TV) regularization [27] will be employed, i.e., the following with i as spatial index,

$$\|x_j\|^2 = \sum_i x_{ij}^2 \quad \text{and} \quad |\nabla x_j| = \sum_i \sqrt{(\partial_x x_{ij})^2 + (\partial_y x_{ij})^2 + (\partial_z x_{ij})^2}. \quad (2.4)$$

In the second equation of (2.3), F represents the post-processing, for example, to alleviate the view-aliasing artifacts or smooth the image variation between phases. A commonly used method is based on the deformable model [9]. Alternatively, when an artifact-free reference image is available, the image deformable model can also be used to “smooth” the data first, and then followed by the phase-wise reconstruction [10].

An immediate benefit of the model (2.3) is that the problem is computationally minimal in the sense that it is almost equivalent to solve a few 3D CT problems with some extra cost for pre-/post-processing. However, this model is fundamentally defective. That is the interplay between 3D images among different phases is considered as a pure image registration problem rather than a truly 4D reconstruction. As a result, the embedded features that are available only through a 4D model can never be revealed otherwise, e.g., through (2.3).

Therefore, one should incorporate, 3D images at different phases as a single entity, into the reconstruction model. The next question is: what should the model look like?

A natural thought is to consider the following,

$$X = \arg \min_X \sum_j [\|A_j x_j - y_j\|^2 + R(x_j)] + R_t(X), \quad (2.5)$$

where the reconstruction of x_j 's at different phases is performed simultaneously with an additional regularization term in time, i.e., $R_t(X)$. That is the data fidelity term is enforced at all phases, while the solution is regularized both spatially and temporally. Note that despite of the simultaneous consideration of all phases, the regularization is however independently carried out locally in space

and time. In the result section, we will adopt TV regularization in both space and time for comparison of models [28]. That is, with the alternative representation of X in pixels rather than phases, e.g., x_i consisting of all phases at the i th pixel,

$$X = \{x_i, i \leq n_s\}, \quad (2.6)$$

the temporal TV regularization is defined as

$$|\nabla_t x_i| = \sum_j |\partial_t x_{ij}|. \quad (2.7)$$

Then, the model (2.5) becomes the following with TV regularization in both space and time,

$$X = \arg \min_X \sum_j \|A_j x_j - y_j\|^2 + \lambda_1 \sum_j |\nabla x_j| + \lambda_2 \sum_i |\nabla_t x_i|. \quad (2.8)$$

Eq. (2.8) is indeed a way to model 4D CT as a 4D reconstruction problem. However, this model is still not very satisfactory in the sense that the spatiotemporal regularization is enforced “locally”, while in reality the 4D entity is a “global” mixture in time and space.

2.1.2. Robust PCA based 4DCT model. The major contribution of this work is to introduce a new spatiotemporal 4D model from matrix perspective, i.e., Robust PCA based 4DCT model or RPCA-4DCT model in short. That is X parameterized conventionally in space via (2.1) or in time via (2.6) can be represented in matrix with row dimension for spatial variable and column dimension for temporal variable

$$X = [x_1 \quad \cdots \quad x_j \quad \cdots \quad x_{n_t}]. \quad (2.9)$$

As mentioned earlier, the respiratory motion can be regarded as a sequence of 3D images with different temporal “sparse” deviations from a common “background”. Motivated by this observation, we consider the following natural low-rank and sparse decomposition of X ,

$$X = X_1 + X_2. \quad (2.10)$$

In (2.10), X_1 is the low-rank matrix component for modeling the stationary background of X . Please note that X_1 is assumed to resemble each other rather than to be constant in time, which can be naturally characterized as a low-rank matrix in mathematics. On the other hand, X_2 is the sparse matrix component for modeling the sparse deviation from the background X_1 . Here the sparsity can be the image itself or the image under some proper sparsifying transform that will be discussed shortly.

Accordingly, when the temporal change of images is sparse in the original representation, we are going to consider the following matrix minimization problem for RPCA-4DCT model

$$(X_1, X_2) = \arg \min_{(X_1, X_2)} \|A(X_1 + X_2) - Y\|^2 + \lambda_* \|X_1\|_* + \lambda_1 \|X_2\|_1, \quad (2.11)$$

where A represents a linear operator composed of system matrices $\{A_j\}$, the nuclear norm for penalizing the rank of the matrix X_1 (altogether for all phases) is defined as the sum of its singular values $\{\sigma_k\}$ with the regularizing parameter λ_* , and the l_1 norm for promoting the sparsity of X_2 (independently for each phase) is simply the absolute sum of its entries with the regularizing parameter λ_1 , i.e.,

$$\|X_1\|_* = \sum_k \sigma_k \quad \text{and} \quad \|X_2\|_1 = \sum_j (\sum_i |X_{2,ij}|). \quad (2.12)$$

Comparing with (2.8), the RPCA-4DCT model via (2.11) offers a unified treatment in time and space, while each 4D component is characterized via the proper norm, such as (2.12). The gain of the overall reconstruction quality via the RPCA-4DCT model will be apparent in the result section. In addition, the extraction of X_2 from the background X_1 is a unique feature of the RPCA-4DCT model with captured dynamic details that would sometimes be crucial but hardly recognizable even for human eyes.

On the other hand, there may be a concern that the computational cost would increase dramatically due to the simultaneous reconstruction of X at all phases. In Section 2.3, we will address this question

with an efficient algorithm, which shows that the collective optimization via (2.8) or (2.11) costs roughly the same as independent optimizations via (2.3).

2.1.3. Connection with prior works.

RPCA-4DCT model is motivated by RPCA for data analysis in statistics [18]. In [18], a model to recover principle component X_1 (modeled by a low-rank matrix) from data X with outliers X_2 (modeled by a sparse matrix), is converted to the following minimization problem when certain incoherence conditions are satisfied

$$(X_1, X_2) = \arg \min_{(X_1, X_2)} \|X_1\|_* + r \|X_2\|_1, \quad (2.13)$$

Subject to $X_1 + X_2 = Y$

where r is shown to be the following for the matrix with n_1 rows and n_2 columns, so that no tuning parameter is necessary,

$$r = \frac{1}{\sqrt{\max(n_1, n_2)}}. \quad (2.14)$$

2.1.4. *RPCA-4DCT model revisited.* Physical images usually have sparse structure under some carefully constructed dictionary, if not in the original representation. In this study, we find that the tight frame system derived from [15, 29] in general serves the purpose in terms of low-rank and sparse decomposition. As a result, the RPCA-4DCT model via (2.11) is revised as

$$(X_1, X_2) = \arg \min_{(X_1, X_2)} \|A(X_1 + X_2) - Y\|^2 + \lambda_* \|X_1\|_* + \lambda_1 \|WX_2\|_1, \quad (2.15)$$

where W represents the framelet analysis operator with $W^T W = I$. In this study, a multilevel tight framelet decomposition without downsampling under the Neumann (symmetric) boundary condition is used with piecewise linear framelets [30, 31].

A key difference of our RPCA-4DCT model (2.15) from RPCA (2.13) and most existing applications is that the available data set Y here is the tomographic data of X generated by some system matrix rather than directly from X itself. As a result, the required incoherence conditions for guaranteeing the success of RPCA cannot be rigorously justified in such an illposed inverse problem. Another difference is that, the sparsity of X_2 is enforced in the transform domain in RPCA-4DCT model, while it is in the original image domain in RPCA. Besides, the component X_2 here is for modeling the motion or the change which is sometimes crucial for CT, and the data/image noise is controlled by the data fidelity term; in contrast, in some applications of RPCA, the component X_2 is considered to be the noise that is of less interest.

The sparsity of images under tight framelets has been successfully used to solve many image restoration tasks including image denoising, image deblurring, image inpainting, etc (e.g. [17, 30, 31]). Most importantly, tight framelets are redundant systems, which lead to robust image representations. Therefore, partial loss and noise of the data in CT can be tolerated without adverse effects. Moreover, the filters in piecewise linear B-spline framelets are the 1st and 2nd discrete difference operators respectively, and the multiscale structure of the framelets enables their multilevel correspondences. These difference operators are well organized in a way such that they have the unitary property $W^T W = I$. Altogether, the piecewise linear B-spline framelet can provide more difference operators than traditional partial differential equation based methods such as TV regularization, and hence it can handle images with abundant structures.

Another benefit using tight framelet systems is the availability of more efficient numerical methods to solve the resulting minimization (2.15). We will use the split Bregman method to solve (2.15), and there is a system of linear equations to be solved in each iteration of the split Bregman method. Since A is usually under-determined and $W^T W = I$, it can be verified that the coefficient matrix of the resulting linear system has clustering eigenvalues. Therefore, the system of linear equation can

be solved efficiently by the conjugate gradient (CG) method, and CG gives the exact solution after only a few steps.

On the other hand, with *a priori* knowledge of the noise in (2.2), the RPCA-4DCT model (2.15) can be further augmented by characterizing the first term (the data fidelity term) with the appropriate norm. For example, if the data come with impulse noise, l_1 norm is particularly suitable for characterizing such outliers [32-35]. In this study, for simplicity, we assume Gaussian data noise and penalize the data fidelity term with L_2 norm.

Finally, we remark that when the considered 4D object changes drastically in time, it can be reformulated into a few overlapping 4D objects according to the properly chosen temporal windows. Consequently, the overall reconstruction is with respect to the weighted sum of 4D components [36].

2.2. Dynamic Scanning

One of the major practical concerns of CT is its ionizing radiation dose. For instance, it has been estimated that although CT studies constitute only 4% of all radiological procedures, they account for 40% of the radiation dose delivered [37]; furthermore, “CT could account for as much as 60% of manmade radiation exposures to Americans” [38]. Tremendous effort has been devoted in dose reduction [1, 2, 39-41]. For example, as a theoretically justified methodology, “interior tomography”, an internal region of interest (ROI) can be exactly reconstructed only from local projection data directly associated with this ROI, which would conventionally require the whole body x-ray illumination [41-43].

Here we propose a dynamic scanning procedure pertinent to 4D CT that can potentially offer further dose reduction through the reduced number of projections. This dynamic scanning is temporally “spiral” as illustrated in Figure 1.

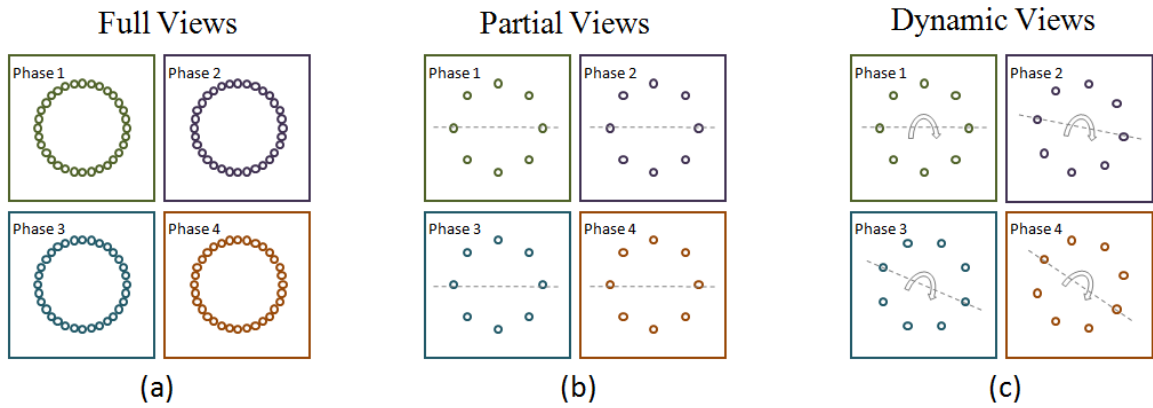


Figure 1. An illustration example on scanning procedures for 4D CT. (a) “Full Views” corresponds to full data acquisition with 32 projections for each phase; (b) “Partial Views” corresponds to partial data acquisition with 8 projections, which is temporally stationary in each phase; (c) “Dynamic Views” again with 8 projections is however dynamically variant among phases so that any view can be swept in some phase within a full dynamic cycle while avoiding redundant measurements of the common background structure at different phases. In this example one data acquisition cycle is synchronized with 4D images with 4 phases. Please note that no temporal periodicity of the images is assumed.

While Figure 1(a) is with full views of data (“Full Views”), Figure 1(b) is a temporally stationary data acquisition scheme with the reduced number of views that is stationary over time (“Partial Views”). In contrast, Figure 1(c) is a dynamic scanning procedure (“Dynamic Views”). In “Dynamic Views”, the positioning of acquired views differs between consecutive phases, however changes periodically so that each view is covered at least once during a dynamic period of scanning. In this

way, the redundant measurements of the common background structure at different phases can be avoided. Please note that we do not assume the temporal periodicity of images.

The potential advantage of the proposed ‘‘Dynamic Views’’ (Figure 1(c)), is that it can achieve the comparable image quality with less number of views than ‘‘Full Views’’ (Figure 1(a)), or provide better image quality with the same number of views than ‘‘Partial Views’’ (Figure 1(b)). A heuristic explanation is that the acquired views in one period can be fused into a dataset of full views so that the tangible data for each phase are from full views rather than partial views. What are missing in the data acquired from ‘‘Dynamic Views’’ comparing with ‘‘Full Views’’ mostly correspond to the redundant measurements of the stationary background.

Moreover, ‘‘Dynamic Views’’ is practically feasible as long as the object to be imaged can be regarded as a single temporal phase within the time period of each data acquisition. Next we will use the reported parameters in [3] to illustrate its applicability in 4D CT for respiratory motion. That is, assuming (1) the 4D object of interest is one respiratory cycle with 6s in time, (2) the scanner rotation time for full views is 1.5s, (3) a fraction of the entire views is to be used for each phase, i.e., 1/8th, then it is safe to consider the 4D model with up to 32 phases.

In contrast, in the standard scanning procedure, the over-sampled views of data are first acquired without pre-arrangement of the scanning according to phases, and then are binned into different phases according to amplitude or phase-angle sorting using an external respiratory signal [8]. Please note that the 4D objects to be imaged have to assume certain temporal periodicity in order to carry out the synchronized binning. Comparing with this standard scanning procedure, the proposed dynamic scanning scheme has the following apparent advantages: (1) no external respiratory signal is necessary, (2) the 4D objects do not have to be periodic, (3) most importantly, the carefully designed periodic data acquisition scheme provides the almost equivalent image quality as the full-view acquisition. On the other hand, it is synergetic with other potential scanning procedures for dose reduction, such as multi-source interior tomography [44].

2.3. Algorithm

In this section, we consider the solution of RPCA-4DCT model via the following optimization

$$(X_1, X_2) = \arg \min_{(X_1, X_2)} \frac{1}{2} \|A(X_1 + X_2) - Y\|^2 + \lambda_* (\|X_1\|_* + r \|WX_2\|_1), \quad (2.16)$$

where r is defined by (2.14), and λ_* is the only regularizing parameter to be determined.

Here we adopt an accurate, efficient, but simple-to-implement algorithm for solving the non-differentiable l_1 -type problems, such as (2.16), namely split Bregman method, that is essentially equivalent to augmented Lagrangian method, however was independently developed from a different perspective to improve the ROF model [16, 45]. In particular, the tight frame regularized split Bregman method will be implemented here [17]. That is (2.16) can be exactly solved through the following simple iterative scheme with $X_1^0 = X_2^0 = 0, f^0 = 0, d_1^0 = v_1^0 = 0$ and $d_2^0 = v_2^0 = 0$,

$$\left\{ \begin{array}{l} (X_1^{k+1}, X_2^{k+1}) = \arg \min_{(X_1, X_2)} \|A(X_1 + X_2) - Y + f^k\|^2 + \mu_* \|X_1 - d_1^k + v_1^k\|^2 + \mu_1 \|WX_2 - d_2^k + v_2^k\|^2, \\ d_1^{k+1} = \arg \min_{d_1} \frac{1}{2} \|X_1^{k+1} + v_1^k - d_1\|^2 + \frac{\lambda_*}{\mu_*} \|d_1\|_*, \\ d_2^{k+1} = \arg \min_{d_2} \frac{1}{2} \|WX_2^{k+1} + v_2^k - d_2\|^2 + \frac{r\lambda_*}{\mu_1} \|d_2\|_1, \\ v_1^{k+1} = v_1^k + X_1^{k+1} - d_1^{k+1}, \\ v_2^{k+1} = v_2^k + WX_2^{k+1} - d_2^{k+1}, \\ f^{k+1} = f^k + A(X_1^{k+1} + X_2^{k+1}) - Y. \end{array} \right. \quad (2.17)$$

The convergence of this iterative scheme when the variables are vectors has been established in [17]. Although the split Bregman method here by (2.17) is for matrix variables, the convergence can be obtained by mimicking the proofs in [17].

The first step of (2.17) corresponds to one iteration step in a typical differentiable L_2 minimization, and the solution is simply from its optimal condition. In implementation, A and W are regarded as linear operators rather than matrices. For efficiency, CG is utilized, in which only the evaluations of linear operators on X_j are necessary, such as $A_j X_j^l$ and WX_j^2 , without the explicit formulation and inversion of the whole system. As we have mentioned before, it can be verified that the considered matrix system has clustering eigenvalues. Therefore, CG gives the exact solution after only a few iterations. Please note that if only single iteration is used in CG, (2.17) can be viewed as a typical case of operator splitting methods [46-48]. However, it is found here that the conventional CG with a few iterations is more realistic in terms of reconstruction accuracy and convergence speed, which was also mentioned in [13].

The second step of (2.17) can be exactly solved by so-called singular value thresholding (SVT) [49]. That is

$$d_1^{k+1} = D_{\lambda^*/\mu^*}(X_1^{k+1} + v_1^k), \quad (2.18)$$

where the thresholding is with respect to singular values σ of the input matrix, i.e.,

$$D_\tau(X) := U \cdot \text{diag}(\max(\sigma - \tau, 0)) \cdot V^T, \text{ with } X = U \cdot \text{diag}(\sigma) \cdot V^T. \quad (2.19)$$

For this step, the major computational cost is from singular value decomposition (SVD), which can be expensive even in the overall scheme (2.17). To alleviate the potential heavy computational burden, a more efficient and sophisticated solution without SVD was proposed in [50]. However, it is sufficient to consider this step by SVT in this study since the number of columns of the matrix considered here (corresponding to the number of phases in 4D object) is small enough so that SVT (2.19) is computationally negligible overall.

Similarly, the solution to the third step of (2.17) is given by so called shrinkage formula, i.e.,

$$d_2^{k+1} = T_{r\lambda^*/\mu_1}(WX_2^{k+1} + v_2^k), \quad (2.20)$$

with

$$T_\tau(X) := \text{sgn}(X) \cdot \max(|X| - \tau, 0). \quad (2.21)$$

Note that the shrinkage formula (2.21) is a scalar operation for each entry of X , while SVT (2.19) is a global operation on X .

Regarding the parameters in (2.17), the followings are recommended

$$r = \frac{1}{\max(n_1, n_2)} \text{ and } \mu^* = \mu_1 = \lambda^*, \quad (2.22)$$

where n_1 (n_2) are the number of rows (columns) of the matrix X . Here the choice of r is supported by the rigorous analysis for RPCA [18] although the theory assumptions cannot be justified rigorously here due to the system matrix A . Furthermore, it is found that $\lambda^* \in [0.1, 1]$ generally provides the satisfactory performance in terms of both accuracy and speed.

In the proposed algorithm via (2.17), the dominant component is to solve L_2 problems in the first step. Due to its iterative nature, empirically it is not necessary to solve each CG step with very high accuracy in order for the whole loop to achieve fast convergence besides the fact that the system has a good condition number as mentioned before. We find that CG with 10 to 20 inner iterations is adequate for (2.17) to have the acceptable reconstruction accuracy within 50 outer iterations.

The similar split Bregman strategy as (2.17) can be used for solving other models in Section 2.1. Moreover, another immediate advantage of split Bregman method is that the computation cost for RPCA-4DCT model (2.15), is similar to the corresponding L_1 -type 4D models, such as (2.8) with TV regularization in both time and space, or a temporally-independent sequence of 3D models, such as (2.3) with TV regularization in space.

3. Results

The purpose here is twofold: (1) to compare different models, (2) to justify the proposed dynamic scanning.

To simplify the discussion, let “ L_2 ” (“TV”) be the 3D model that solve 4D CT by each individual phase, i.e., (2.3) with L_2 (TV) regularization; “TV+TVt” be the 4D model that solve 4D CT as a single entity, i.e., (2.8) with TV regularization in both space and time; the RPCA-4DCT model be the proposed matrix model that solve 4D CT as a single entity, i.e., (2.15) with low-rank component and sparse component (in tight frames). Please note that we do not compare here with the standard filtered backprojection (without regularizing solutions), which generally gives worse accuracy than “ L_2 ” when reconstructing with insufficient number of views.

In this proof-of-concept study, for simplicity, the spatial dimension is 2D rather than 3D, i.e., 128 by 128 spatial grid. Here 32 temporal phases are adopted for the justified reasons in Section 2.2. With the parallel scanning geometry, the length of the detectors with 256 detector pixels is equal to the side length of the spatial square domain, and “Full Views” consist of 256 projections. The reconstructions with three different data acquisition scheme (Figure 1) will be compared: “Full Views” corresponds to the use of all data; “Partial Views” corresponds to the use of 32 projections that is temporally stationary in phases; “Dynamic Views” corresponds to the use of 32 projections that is dynamically adjusted among phases so that any view is covered in some phase within a full dynamic period while avoiding redundant measurements of the common background. Here one data acquisition cycle is synchronized with 8 phases and the data are acquired with 4 cycles.

Except “ L_2 ”, which is differentiable and therefore can be solved with iterations involving only one step that is similar to the first step of (2.17), all aforementioned other models can be solved through split Bregman iterations similar to (2.17). As a result, since the major computational cost in all models is on the L_2 step similar to the first step in (2.17), the total computation time approximately only depends on the number of iterations. It is found that roughly 20-50 iterations together with the parameters specified by (2.22) are sufficient. In particular, all models except “ L_2 ” are similar in computational cost, while “ L_2 ” fails to achieve the satisfactory accuracy (comparing with other models) regardless of the number of iterations.

Two spatiotemporal phantoms are used for evaluation. Phantom 1 is utilized for evaluating the reconstruction with motion, while the purpose of Phantom 2 is for change detection.

Phantom 1 is to mimic a half respiratory cycle, which is based on the modified Shepp-Logan phantom that consists of piecewise constant regions (Figure 2). The temporal variations consist of (1) the intensity increase of the top circle (with a relatively large diameter), (2) the vertical movement of two central circles (with a relatively small diameter) apart of each other, and (3) the horizontal movement of two ellipses (with a relatively low contrast) apart of each other.

Phantom 2 is to model the case with small temporal variations, which can be even hardly recognizable by human eyes (Figure 3). It is based on a brain image and the temporal variations consist of the horizontal movement of two ellipses (with a relatively very low contrast) apart of each other.

3.1. Model comparison

In this section, a small fraction of full views is used, i.e., one eighth of the full data (32 projections). Specifically we adopt “Dynamic Views” as the data acquisition scheme (Figure 1(c)).

For Phantom 1, the result from the RPCA-4DCT model is shown in Figure 4, which clearly shows that the RPCA-4DCT model is able to not only recover images, but also provide automatic image decomposition into the background (that is mathematically low-rank) and the variation (that is sparse under tight frame transform). In contrast, the results from other models are shown in Figure 5. Since Phantom 1 is in favor of TV regularization due to its components in piecewise constants, “TV” and “TV+TVt” offer the equivalent image quality although the RPCA-4DCT model is slightly better in terms of reconstruction errors (Table 1).

On the other hand, the result from the RPCA-4DCT model for Phantom 2 is shown in Figure 6, which indicates again that the matrix model is not only superior in the overall image quality, but also able to capture these small features that would be impossible otherwise. For comparison, the results from other models are shown in Figure 7. With temporal TV regularization, “TV+TVt” provides better

overall image quality than “TV” both visually (Figure 7) and quantitatively (Table 1), which however is worse than the RPCA-4DCT model. The blurred or smoothed details are apparent for either “TV” or “TV+TVt”. Besides, neither clearly shows the temporal variations, which in contrast are available through the RPCA-4DCT model without extra computational cost.

Please note that although the decomposition via the RPCA-4DCT model is non-unique, the variations are clearly captured as in Figure 4 and 6. Furthermore, such a clear decomposition with captured details via the RPCA-4DCT model is usually not available through the post-processing of images from other models, such as the simple subtraction of images with respect to the first phase.

3.2. Scanning comparison

The reconstructions are performed on Phantom 2 with the RPCA-4DCT model to compare three different data acquisition schemes (Figure 1), i.e., “Full Views” with 256 projections for each phase, “Partial Views” with 32 projections for each phase that are temporally invariant in terms of positioning of projections, and “Dynamic Views” with 32 projections for each phase that are dynamically variant so that each view is available for some phase in a full scanning cycle. Here one data acquisition cycle is synchronized with 8 phases and the data are acquired with 4 cycles.

The reconstruction results for “Dynamic Views”, “Partial Views” and “Full Views” are presented in Figure 6, 8 and 9 respectively, and the quantitative reconstruction accuracy is summarized in Table 2. The results clearly show that “Dynamic Views” has a great potential for dose reduction since it offers a satisfactory accuracy with only a small fraction of the full data.

Acknowledgments

This work is partially supported by NSF grant DMS0811254.

Table 1. Reconstruction accuracy from various models (with “Dynamic Views”). The quantities are the quotient differences between the ground truth X^0 and the reconstructed images X , i.e., $\|X - X^0\| / \|X^0\|$ with $\|\cdot\|$ as L_2 norm.

| Phantom | L_2 | TV | TV+TVt | LR+TF |
|---------|-------|-------|--------|-------|
| 1 | 0.203 | 0.008 | 0.006 | 0.005 |
| 2 | 0.197 | 0.076 | 0.040 | 0.022 |

Table 2. Reconstruction accuracy from various data scanning schemes with the RPCA-4DCT model. The quantities are the quotient differences between the ground truth X^0 and the reconstructed images X , i.e., $\|X - X^0\| / \|X^0\|$ with $\|\cdot\|$ as L_2 norm.

| Phantom | Partial Views | Dynamic Views | Full Views |
|---------|---------------|---------------|------------|
| 2 | 0.187 | 0.022 | 0.005 |

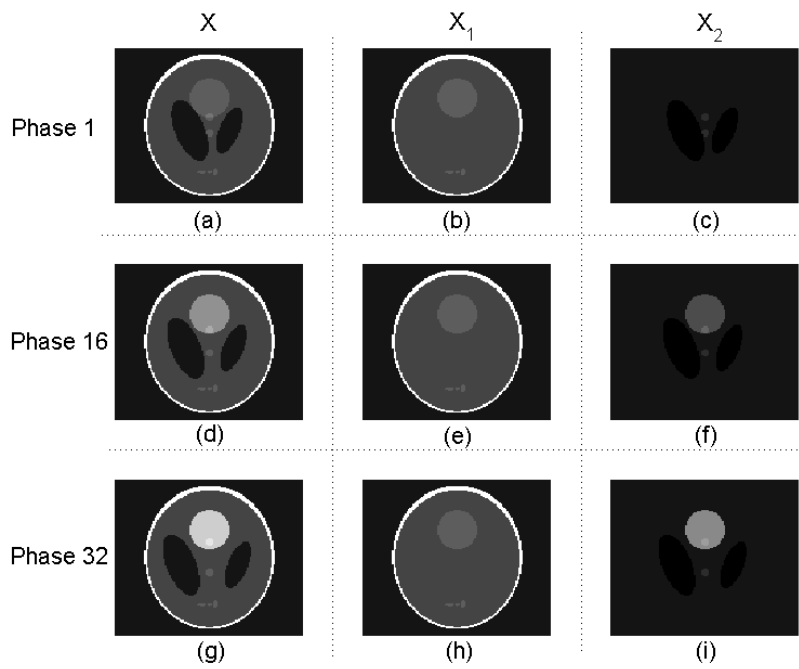


Figure 2. Phantom 2 for 4D CT. (a), (b) and (c) are the image X , the background of the image X_1 and the motion/change of the image X_2 respectively at Phase 1, i.e., $X=X_1+X_2$. Similarly, (d), (e) and (f) correspond to X , X_1 and X_2 at Phase 16, and (g), (h) and (i) correspond to X , X_1 and X_2 at Phase 32.

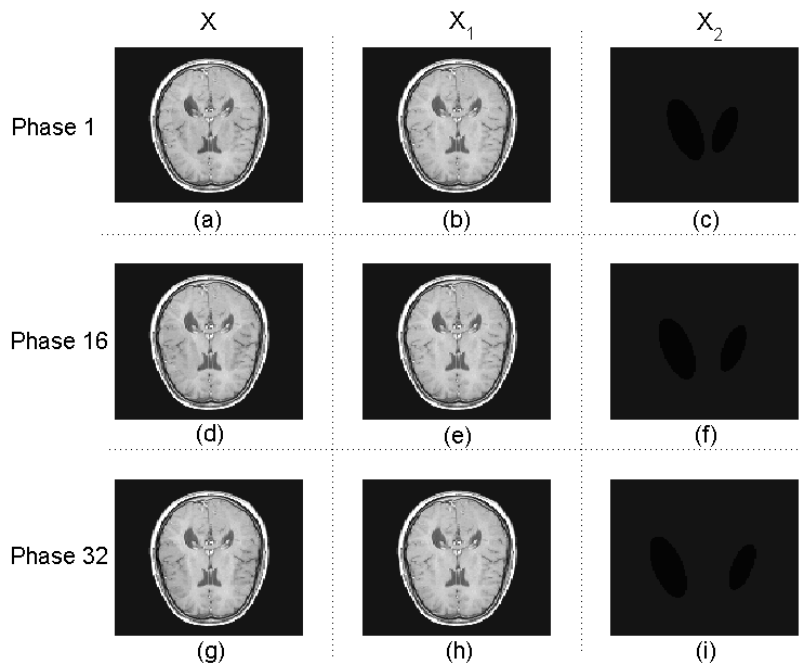


Figure 3. Phantom 2 for 4D CT. (a), (b) and (c) are the image X , the background of the image X_1 and the motion/change of the image X_2 respectively at Phase 1, i.e., $X=X_1+X_2$. Similarly, (d), (e) and (f) correspond to X , X_1 and X_2 at Phase 16, and (g), (h) and (i) correspond to X , X_1 and X_2 at Phase 32.

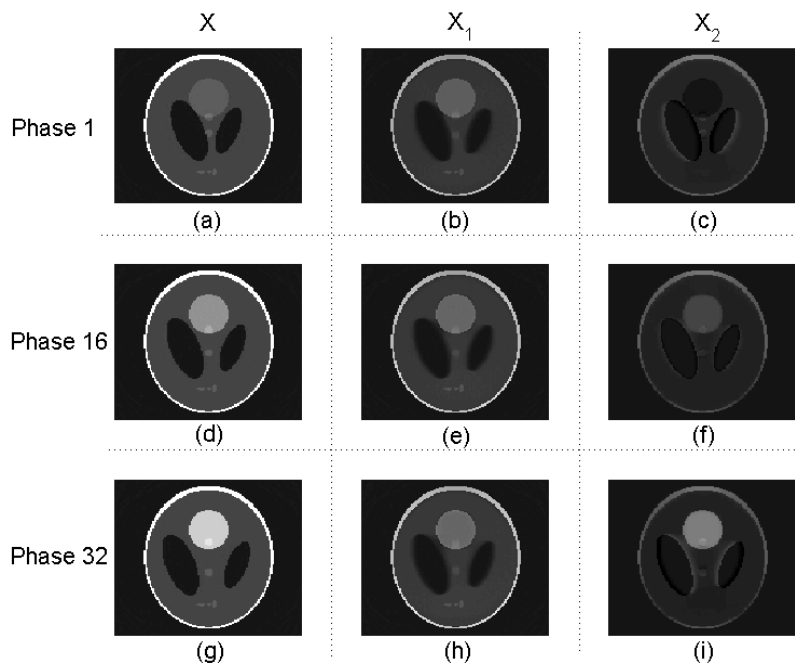


Figure 4. Reconstructed images with RPCA-4DCT model for Phantom 1. (a), (b) and (c) are the total image X , the low-rank component X_1 and the sparse component (in tight frames) X_2 respectively at Phase 1, i.e., $X=X_1+X_2$. Similarly, (d), (e) and (f) correspond to X , X_1 and X_2 at Phase 16, and (g), (h) and (i) correspond to X , X_1 and X_2 at Phase 32.

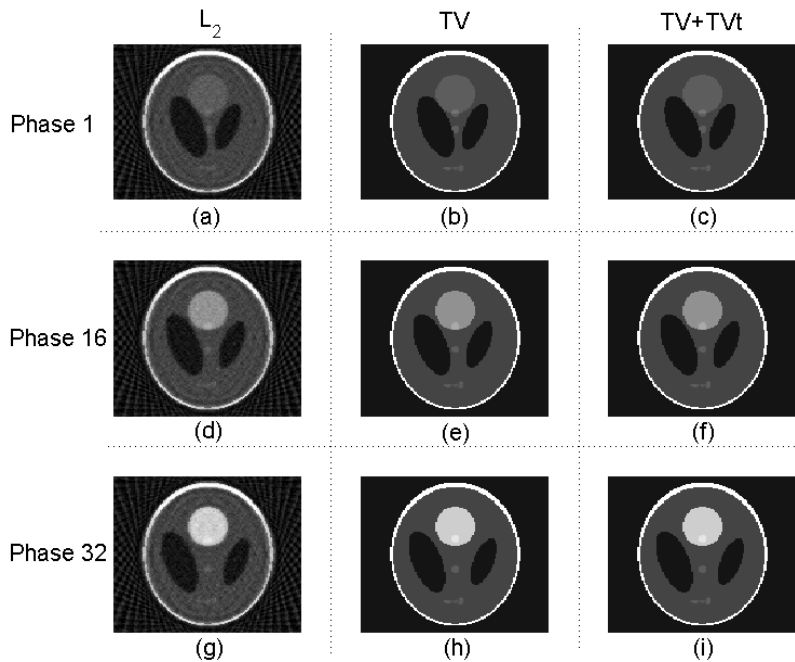


Figure 5. Reconstructed images with other various models for Phantom 1. (a), (b) and (c) are from “ L_2 ”, “TV” and “TV+TVt” respectively at Phase 1. Similarly, (d), (e) and (f) correspond to the above models at Phase 16, and (g), (h) and (i) correspond to the above models at Phase 32.

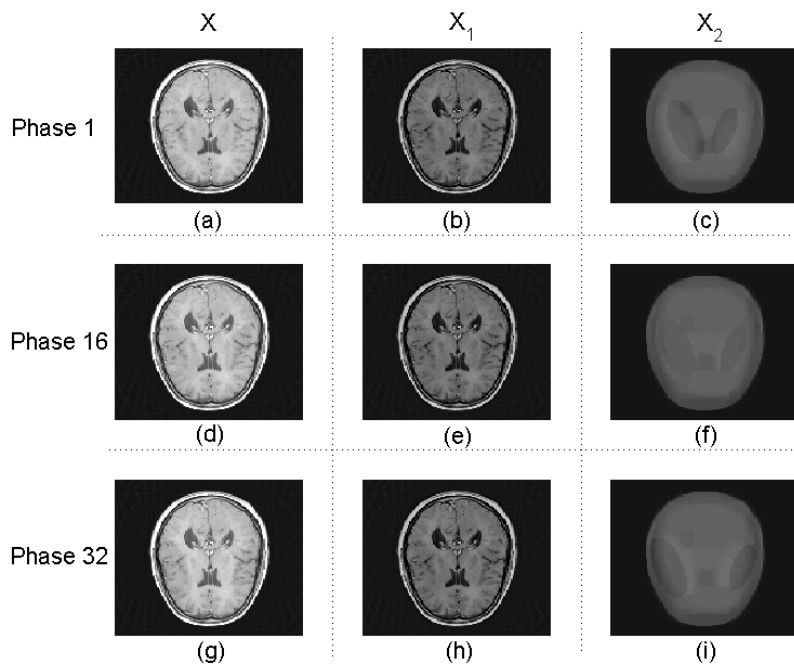


Figure 6. Reconstructed images with RPCA-4DCT model for Phantom 2. (a), (b) and (c) are the total image X , the low-rank component X_1 and the sparse component (in tight frames) X_2 respectively at Phase 1, i.e., $X=X_1+X_2$. Similarly, (d), (e) and (f) correspond to X , X_1 and X_2 at Phase 16, and (g), (h) and (i) correspond to X , X_1 and X_2 at Phase 32.

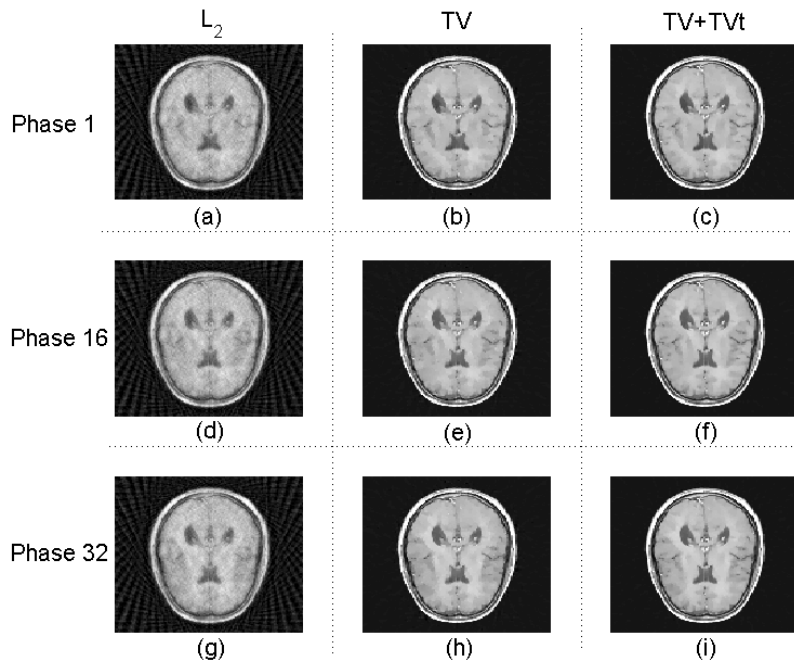


Figure 7. Reconstructed images with other various models for Phantom 2. (a), (b) and (c) are from “ L_2 ”, “TV” and “TV+TVt” respectively at Phase 1. Similarly, (d), (e) and (f) correspond to the above models at Phase 16, and (g), (h) and (i) correspond to the above models at Phase 32.

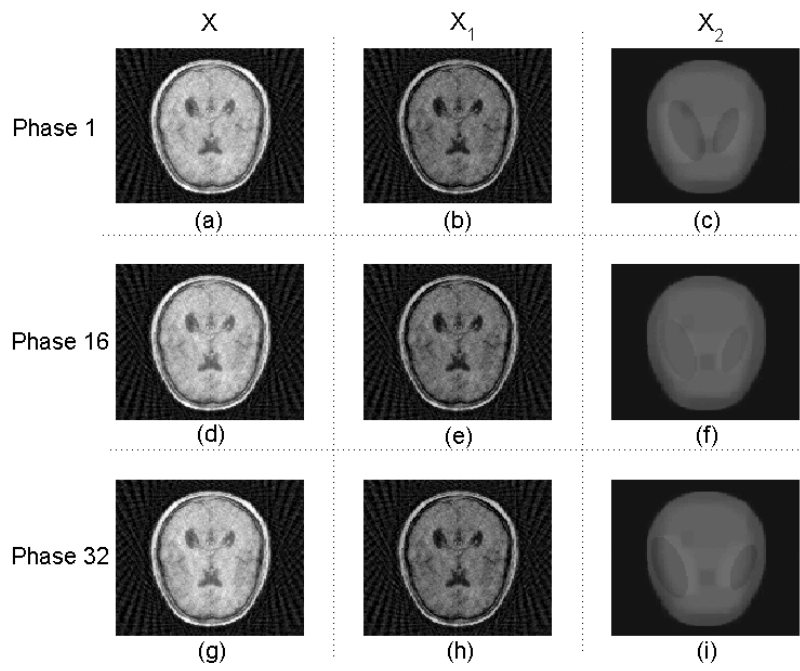


Figure 8. Reconstructed images with the RPCA-4DCT model for Phantom 2 with “Partial Views”.

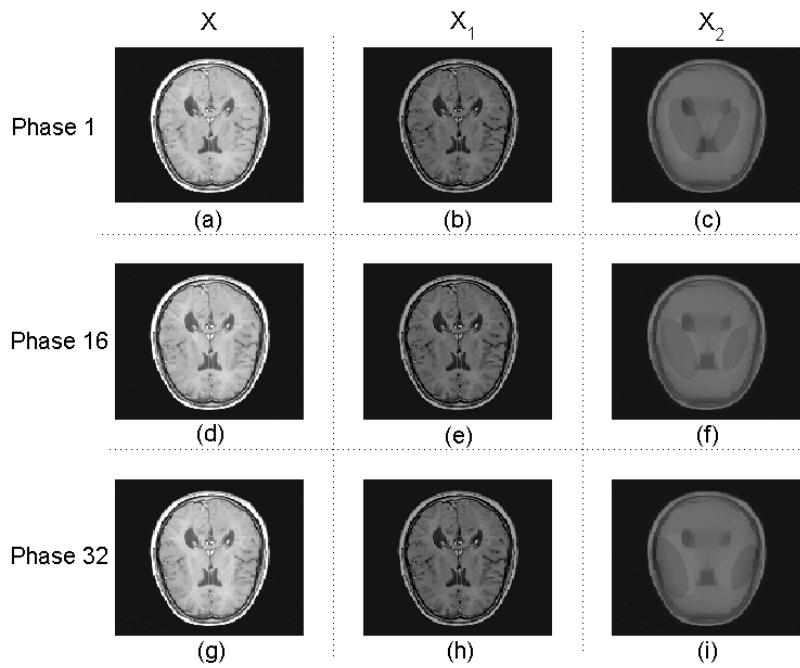


Figure 9. Reconstructed images with the RPCA-4DCT model for Phantom 2 with “Full Views”.

References

- [1] L. Xing, B. Thorndyke, E. Schreibmann, Y. Yang, T. F. Li, G. Y. Kim, G. Luxton, and A. Koong, "Overview of image-guided radiation therapy", *Med. Dosim.*, **31**, 91-112 (2006).
- [2] S. B. Jiang, J. Wolfgang, and G. S. Mageras, "Quality assurance challenges for motion-adaptive radiation therapy: gating, breath holding, and four dimensional computed tomography", *Int. J. Radiat. Oncol. Biol. Phys.*, **71**, S103-S107 (2008).
- [3] S. S. Vedam, P. J. Keall, V. R. Kini, H. Mostafavi, H. P. Shukla, and R. Mohan, "Acquiring a four-dimensional computed tomography dataset using an external respiratory signal", *Phys. Med. Biol.*, **48**, 45-62 (2003).
- [4] D. Low, M. Nystrom, E. Kalinin, P. Parikh, J. Dempsey, J. Bradley, S. Mutic, S. Wahab, T. Islam, G. Christensen, D. Politte, and B. Whiting, "A method for the reconstruction of four-dimensional synchronized CT scans acquired during free breathing", *Med. Phys.*, **30**, 1254-1263 (2003).
- [5] P. J. Keall, G. Starkschall, H. Shukla, K. M. Forster, V. Ortiz, C. W. Stevens, S. S. Vedam, R. George, T. Guerrero, and R. Mohan, "Acquiring 4D thoracic CT scans using a multislice helical method", *Phys. Med. Biol.*, **49**, 2053-2067 (2004).
- [6] E. Rietzel, T. Pan, and G.T. Chen, "Four-dimensional computed tomography: image formation and clinical protocol", *Med. Phys.*, **32**, 874-889 (2005).
- [7] T. Li, E. Schreibmann, B. Thorndyke, G. Tillman, A. Boyer, A. Koong, K. Goodman, and L. Xing, "Radiation dose reduction in four-dimensional computed tomography", *Med. Phys.*, **32**, 3650-60 (2005).
- [8] W. Lu, P. J. Parikh, J. P. Hubenschmidt, J. D. Bradley, and D. A. Low, "A comparison between amplitude sorting and phase-angle sorting using external respiratory measurement for 4D CT", *Med. Phys.*, **33**, 2964-2974 (2006).
- [9] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, "Non-rigid registration using free-form deformations: Application to breast MR images", *IEEE Trans. Med. Imag.*, **18**, 712-721 (1999).
- [10] T. Li, A. Koong, and L. Xing, "Enhanced 4D cone-beam CT with inter-phase motion model", *Med. Phys.*, **34**, 3688-3695 (2007).
- [11] E. Schreibmann, B. Thorndyke, T. Li, J. Wang, and L. Xing, "Four-dimensional image registration for image-guided radiotherapy", *Int. J. Radiat. Oncol. Biol. Phys.*, **71**, 578-586 (2008).
- [12] S. A. Nehmeh, Y. E. Erdi, K. E. Rosenzweig, H. Schoder, S. M. Larson, O. D. Squire, and J. L. Humm, "Reduction of respiratory motion artifacts in PET imaging of lung cancer by respiratory correlated dynamic PET: Methodology and comparison with respiratory gated PET", *J. Nucl. Med.*, **44**, 1644-1648 (2003).
- [13] X. Jia, Y. Lou, B. Dong, Z. Tian, and S. Jiang, "4D computed tomography reconstruction from few-projection data via temporal non-local regularization", *Lecture Notes in Computer Science*, **6361**, 143-150 (2010).
- [14] L. Lee, Y. Ma, Y. Ye, and L. Xing, "Conceptual formulation on four-dimensional inverse planning for intensity modulated radiation therapy", *Phys. Med. Biol.*, **54**, N255-N266 (2009).
- [15] A. Ron and Z. Shen, "Affine systems in $L_2(\mathbb{R}^d)$: the analysis of the analysis operator", *J. Funct. Anal.*, **148**, 408-447 (1997).
- [16] T. Goldstein and S. Osher, "The split Bregman algorithm for l_1 regularized problems", *SIAM J. Imaging Sci.*, **2**, 323-343 (2009).
- [17] J. F. Cai, S. Osher, and Z. Shen, "Split Bregman methods and frame based image restoration", *Multiscale Model. Simul.*, **8**, 337-369 (2009).
- [18] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?", *Technical report*, Stanford University (2009).

- [19] H. Ji, C. Liu, Z. Shen, and Y. Xu, "Robust video denoising using low rank matrix completion", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, San Francisco (2010).
- [20] Y. Peng, A. Ganesh, J. Wright, W. Xu, and Y. Ma, "RASL: robust alignment by sparse and low-rank decomposition for linearly correlated images", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, San Francisco (2010).
- [21] G. Liu, Z. Lin, and Y. Yu, "Robust subspace segmentation by low-rank representation", *Proceedings of the 26th International Conference on Machine Learning (ICML)*, Haifa, Israel (2010).
- [22] K. Min, Z. Zhang, J. Wright, and Y. Ma, "Decomposing background topics from keywords by principal component pursuit", *Proceedings of ACM International Conference on Information and Knowledge Management (CIKM)*, Toronto, Canada (2010).
- [23] G. Zhu, S. Yan, and Y. Ma, "Image tag refinement towards low-rank, content-tag prior and error sparsity", *Proceedings of ACM Multimedia*, Firenze, Italy (2010).
- [24] Z. Zhang, X. Liang, A. Ganesh, and Y. Ma, "TILT: transform invariant low-rank textures", *Proceedings of Asian Conference on Computer Vision*, Queenstown, New Zealand (2010).
- [25] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma, "Robust photometric stereo via low-rank matrix completion and recovery", *Proceedings of Asian Conference on Computer Vision*, Queenstown, New Zealand (2010).
- [26] T. Buzug, *Computed Tomography: From Photon Statistics to Modern Cone-Beam CT*, Springer-Verlag (2008).
- [27] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms", *J. Phys. D*, **60**, 259-268 (1992).
- [28] J. Weickert and C. Schnörr, "Variational optic flow computation with a spatio-temporal smoothness constraint", *Journal of Mathematical Imaging and Vision*, **14**, 245-255 (2001).
- [29] I. Daubechies, B. Han, A. Ron, and Z. Shen, "Framelets: MRA-based constructions of wavelet frames", *Appl. Comput. Harmon. Anal.*, **14**, 1-46 (2003).
- [30] A. Chai and Z. Shen, "Deconvolution: a wavelet frame approach", *Numer. Math.*, **106**, 529-587 (2007).
- [31] J. F. Cai, R. H. Chan, and Z. Shen, "A framelet-based image inpainting algorithm", *Appl. Comput. Harmon. Anal.*, **24**, 131-149 (2008).
- [32] S. Alliney, "Digital filters as absolute norm regularizers", *IEEE Trans. Signal Process.*, **40**, 1548-1562 (1992).
- [33] M. Nikolova, "Minimizers of cost-functions involving nonsmooth data fidelity terms. Application to the processing of outliers", *SIAM J. Numer. Anal.*, **40**, 965-994 (2002).
- [34] T. F. Chan and S. Esedoglu, "Aspects of total variation regularized L1 function approximation", *SIAM Journal on Applied Mathematics*, **65**, 1817-1837 (2005).
- [35] H. Gao and H. K. Zhao, "Multilevel bioluminescence tomography based on radiative transfer equation. Part 2: total variation and l1 data fidelity", *Opt. Express* **18**, 2894-2912 (2010).
- [36] H. Ji, Z. Shen, and Y. Xu, "Robust video restoration by joint sparse and low rank matrix approximation", Preprint (2010).
- [37] P.C. Shrimpton and S. Edyvean, "CT scanner dosimetry", *Br. J. Radiol.*, **71**, 1-3 (1998).
- [38] O. W. Linton, A. Fred, and F. A. Mettler, "National conference on dose reduction in CT, with an emphasis on pediatric patients", *Amer. J. Roentgenol.*, **181**, 321-329 (2003).
- [39] X. Pan, E. Y. Sidky, and M. Vannier, "Why do commercial CT scanners still employ traditional, filtered back-projection for image reconstruction?", *Inverse Probl.*, **25**, 123009 (2009).
- [40] B. P. Fahimian, Y. Mao, P. Cloetens, and J. Miao, "Low dose x-ray phase-contrast and absorption CT using Equally-Sloped Tomography", *Phys. Med. Bio.* **55**, 5383-5400 (2010).
- [41] Wang, G., Y. Ye, and H. Yu, *Interior tomography and instant tomography reconstruction from truncated limited angle projection data*. US Patent 7,697,658 B2 (2010).

- [42] Y. B. Ye, H. Y. Yu, Y. C. Wei, and G. Wang, "A general local reconstruction approach based on a truncated Hilbert transform", *Int. J. Biomed. Imaging*, Article ID 63634 (2007).
- [43] H. Kudo, M. Courdurier, F. Noo, and M. Defrise, "Tiny *a priori* knowledge solves the interior problem in computed tomography", *Phys. Med. Biol.*, **53** 2207-2231 (2008).
- [44] G. Wang, H. Y. Yu, and Y. B. Ye, "A scheme for multisource interior tomography", *Med. Phys.*, **36**, 3575-3781 (2009).
- [45] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, "An iterative regularization method for total variation-based image restoration", *Multiscale Model. Simul.*, **4**, 460-489 (2005).
- [46] P. L. Combettes and V. R. Wajs, "Signal recovery by proximal forward-backward splitting", *Multiscale Model. Simul.*, **4**, 1168-1200 (2005).
- [47] E. Hale, W. Yin, and Y. Zhang, "Fixed-point continuation for ℓ_1 -minimization: methodology and convergence", *SIAM J. Optim.*, **19**, 1107-1130 (2008).
- [48] X. Zhang, M. Burger, X. Bresson, and S. Osher, "Bregmanized nonlocal regularization for deconvolution and sparse reconstruction", *UCLA CAM Report*, 09-03 (2009).
- [49] J. F. Cai, E.J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion", *SIAM J. Optimiz.*, **20**, 1956-1982 (2010).
- [50] J. F. Cai and S. Osher, "Fast singular value thresholding without singular value decomposition", *UCLA CAM Report*, 10-24 (2010).