4D Cone-beam CT reconstruction using a motion model based on principal component analysis

David Staub^{a)}

Department of Radiation Oncology, Virginia Commonwealth University, Richmond, Virginia 23298

Alen Docef

Department of Electrical and Computer Engineering, Virginia Commonwealth University, Richmond, Virginia 23284

Robert S. Brock, Constantin Vaman, and Martin J. Murphy

Department of Radiation Oncology, Virginia Commonwealth University, Richmond, Virginia, 23298

(Received 28 July 2011; revised 31 October 2011; accepted for publication 31 October 2011; published 29 November 2011)

Purpose: To provide a proof of concept validation of a novel 4D cone-beam CT (4DCBCT) reconstruction algorithm and to determine the best methods to train and optimize the algorithm.

Methods: The algorithm animates a patient fan-beam CT (FBCT) with a patient specific parametric motion model in order to generate a time series of deformed CTs (the reconstructed 4DCBCT) that track the motion of the patient anatomy on a voxel by voxel scale. The motion model is constrained by requiring that projections cast through the deformed CT time series match the projections of the raw patient 4DCBCT. The motion model uses a basis of eigenvectors that are generated via principal component analysis (PCA) of a training set of displacement vector fields (DVs) that approximate patient motion. The eigenvectors are weighted by a parameterized function of the patient breathing trace recorded during 4DCBCT. The algorithm is demonstrated and tested via numerical simulation.

Results: The algorithm is shown to produce accurate reconstruction results for the most complicated simulated motion, in which voxels move with a pseudo-periodic pattern and relative phase shifts exist between voxels. The tests show that principal component eigenvectors trained on DVs from a novel 2D/3D registration method give substantially better results than eigenvectors trained on DVs obtained by conventionally registering 4DCBCT phases reconstructed via filtered backprojection.

Conclusions: Proof of concept testing has validated the 4DCBCT reconstruction approach for the types of simulated data considered. In addition, the authors found the 2D/3D registration approach to be our best choice for generating the DV training set, and the Nelder-Mead simplex algorithm the most robust optimization routine. © 2011 American Association of Physicists in Medicine. [DOI: 10.1118/1.3662895]

Key words: 4D cone-beam CT, motion modeling, principal component analysis, projection matching

I. INTRODUCTION

Time-dependent computed tomography (4DCT) has become an important tool for assessing and managing intrafraction motion of tumors and other structures of interest during radiotherapy.¹ For imaging of moving objects its advantages over 3DCT include the ability to describe tumor and organ motion, as well as the reduction of motion artifacts in reconstructed images.² Conventional 4D cone-beam CT (4DCBCT) reconstruction requires subdividing the CBCT projection data into time bins and then reconstructing each CT bin via filtered backprojection (FBP).^{2–6} Conventional 4DCBCT suffers from four important problems, due to both the drawbacks of projection binning as well as the inherent deficiencies of CBCT reconstruction via FBP. They consist of the following: (1) Poor image quality. Image quality is impaired by streak artifacts resulting from projection under-sampling due to binning.^{7–10} In addition, 4DCBCT inherits

the deficiencies of 3D CBCT including loss of CT number fidelity, increased noise, and reduced contrast as compared with fan-beam CT (FBCT).^{11,12} (2) Low temporal resolution. Generally around ten images spanning one or two breathing cycles are reconstructed with 4DCBCT. This results in a temporal resolution of around 500–1000 ms, which is significantly worse than 4D fan-beam CT (4DFBCT), and is much less than the actual frame rate at which the cone-beam projections are acquired (around ten images per second).¹⁰ (3) Incomplete motion information. Conventional 4DCBCT reconstruction yields a set of images that cover a single time averaged breathing cycle. For the purposes of accurate margin planning, tumor motion characterization, motion PDF creation, etc., it would be desirable to observe the anatomic motion over many breathing cycles. Changes in breathing pattern and outlier motion states will not be captured by conventional time averaged reconstruction.^{13–17} (4) Unknown

voxel trajectories between reconstructed images. Displacement vector fields (DVs) establishing the correspondence between voxels in different breathing phases are useful for a variety of important clinical tasks such as “contour propagation, dose accumulation, treatment adaptation, dosimetric evaluation, and 4D optimization.”¹⁸ With conventional 4DCBCT, deformable image registration (DIR) must be performed between phases of the image time series in a separate step to acquire these DVs. This introduces further complications to the clinical workflow and imparts errors to the voxel trajectories from both the image reconstruction and DIR processes.

Our 4DCBCT reconstruction strategy is to combine prior knowledge in the form of a prior fan-beam CT (FBCT) with the projection data to eliminate the dependence on full reconstruction via backprojection. Instead, we animate the FBCT with a parameterized motion model and then adjust the model parameters until the animated FBCT optimally emulates the motion recorded by the CBCT projections. This strategy addresses each of the conventional 4DCBCT reconstruction problems listed above: (1) Direct FBP reconstruction of the CBCT is avoided. The image quality and CT number fidelity of FBCT are preserved in the animated 4DCBCT. (2) A CT is reconstructed for each projection to which the motion model is compared; therefore, the temporal resolution can be as high as the frame acquisition rate of the raw CBCT projection data-set. (3) Time averaging of projection data is avoided and the reconstructed image sequence will cover as many breathing cycles as are present in the raw CBCT projection data-set. (4) The motion model intrinsically contains the voxel correspondences (trajectories) between different images in the reconstructed CT time series. The DVs linking the reconstructed images are an integral part of the reconstruction algorithm itself, thus no separate registration process adding extra error and time is needed to produce them.

Our method has two additional benefits. One is the reduction in storage space required for a 4DCBCT reconstructed with our algorithm. Instead of storing the full set of reconstructed images and corresponding DVs, one only needs to store the optimized motion model and the prior FBCT. A desired CT or DV from the time series can then be quickly reconstructed on the fly using a machine of modest computing power. In addition, as in the work of Li *et al.*^{19,20} described below, our motion model can be used for real-time tracking of moving structures such as tumors.

Motion modeling has been previously employed as a method of addressing the problems of 4DCT reconstruction. McClelland *et al.* designed a 4DFBCT reconstruction algorithm that created separate 4D motion models for each slab or couch position. B-splines were used to register each individual time series of slabs, and the resulting time series of control points were fit to a respiratory trace to model their temporal evolution. Finally, the motion models for separate slabs were combined to form a motion model covering the entire imaging volume.²¹ A similar approach was employed by Yang *et al.* who used their “5D motion model” to fit the motion described by deformable registration of 4DFBCT

slabs.²² Zhang *et al.* reconstructed and registered the phases of a 4DFBCT and performed principal component analysis (PCA) on the resultant displacement vector fields (DVs) to linearly correlate the DV motion with a small number of variables on a voxel by voxel scale. They chose the two eigenmodes with the largest variance as their variables and modeled the temporal evolution of the two eigenmodes using the patient breathing trace and the patient breathing trace time-shifted by three data points.²³

Another class of algorithms for 4DCBCT constrains a DV motion model by using it to deform a static prior CT and then adjusting the model parameters until digitally reconstructed radiographs (DRRs) taken through the deformed prior match the projections in a 4DCBCT data-set. We will call this iterative forward projection matching (IFPM). Zeng *et al.* used such a projection matching approach to constrain a DV motion model consisting of the tensor product of three spatial and one temporal B-splines.²⁴ Docef and Murphy²⁵ used a model consisting of the tensor product of three spatial B-splines to represent the DV and modulated the control points in time using a parameterized function of the patient breathing trace. Li *et al.*^{19,20} reconstructed a 4DFBCT via backprojection, deformably registered the sequential phases to a reference phase to obtain deformation vector fields (DVs), and then calculated the eigenmodes of the DVs to form a PCA model similar to that of Zhang *et al.*²³ They then deformed the reference phase from the 4DFBCT using the PCA model and matched individual DRRs to individual CBCT projections to estimate the eigenmode coefficients (or principal coefficients) separately for each time step in the raw 4DCBCT data-set.¹⁹ Their principal aim was to reconstruct the trajectory of a lung tumor. More recently, they have further explored the mathematical underpinnings of the PCA model and compared with the 5D lung motion model of Low *et al.*^{26,27}

The principal challenge in any approach combining a motion model and a prior image is in achieving a physically realistic model with a tractable number of parameters that is easily generalizable to any imaging subject. The 4D DV B-spline model of Zeng *et al.*,²⁴ while completely general, would require many thousands of control points to provide a reasonable time resolution for the duration of a typical CBCT exam. The method of Docef and Murphy²⁵ greatly reduces the number of control point parameters but requires assumptions about the relative phase and amplitude of the individual displacement vectors. We have therefore adopted a new strategy using a motion model based on the principal motion eigenmodes of breathing, obtained via PCA. The principal components are trained on a conventionally binned 4DCBCT of the patient, whereby they “learn” the relative amplitudes and phases of motion among the moving image elements without recourse to a specific physical model.

Our projection matching approach can be considered a hybrid of that of Docef and Murphy²⁵ and Li *et al.*^{19,20} As in Li, we use PCA of a representative set of DVs to relate the motion amplitude of each voxel to a small set of variables.

However, like Docef, we use a parameterized function of the breathing trace to model the temporal evolution of the principal coefficients. Furthermore, we fit the motion model to the entire set of CBCT projection radiographs simultaneously. This has the benefit of greatly reducing the number of parameters to be optimized compared with Zeng and Docef, and constraining the temporal behavior of the model as compared with Li. An additional difference between our implementation and that of Li *et al.*²⁰ is our use of the 4DCBCT data-set to generate the training DVFs as opposed to a previously acquired 4DFBCT, although they do mention that possibility. We expect, and they comment, that a motion model based on the most current training DVFs, i.e., those obtained from the raw 4DCBCT data-set itself, will be more accurate than a model based on training DVFs from a 4DFBCT that might be days or weeks old.

The success of our forward iterative projection matching method for constructing a 4DCBCT depends on two fundamental elements: (1) the DRR calculation to match the projection radiographs, and (2) the PCA-based motion model. In a prior study by Brock *et al.*²⁸ we investigated the sensitivity of 3DCBCT construction via iterative forward projection matching to the degree of mismatch between the DRRs and the CBCT projection images. In this study, we have focused on the sensitivity of IFPM 4DCBCT construction to the underlying PCA motion model. We have identified and analyzed three principal factors impacting the motion model. The first is the accuracy of the eigenvectors, which must be calculated from a preliminary estimate of the 4DCBCT. Li *et al.*^{19,20} calculated their eigenvectors from the sequential registration of a previously acquired backprojected and binned 4DFBCT. However, this injects three sources of error into the PCA analysis—image artifacts due to the 4D binning process, the possibility of producing outdated eigenvectors that do not accurately characterize the motion at the time of the cone beam projections, and inaccuracies and approximations contributed by the image registration process. In order to ensure that our motion model is current we do not use a previously acquired 4DFBCT to get the DVF training set; instead we bin the projections in the 4DCBCT data-set itself and register the binned data. The most obvious way to achieve these registrations is to perform a conventional FBP reconstruction on each bin and then deformably register the bins to a reference image. However, this approach will introduce registration errors due to the artifacts and inaccuracies inherent in backprojected CBCT. In an attempt to eliminate the deficiencies due to backprojected CBCT, we have used our previously developed iterative forward reconstruction method for CBCT (Refs. 25 and 28) to obtain the DVFs used to train the PCA model. In our analysis, we will compare the accuracy of this PCA model to one obtained from backprojected and deformably registered 4DCBCT images.

The second factor is the complexity of a subject's breathing motion, which determines the spectrum of eigenmodes needed for a physically realistic motion model. In a prior study, we showed how this spectrum evolves from a single eigenmode for pure sinusoidal breathing at a single phase,

into multiple modes as phase differences appear in the relative motion of individual voxels and the breathing becomes less regular.²⁹ In this study, we show how the increasing complexity of breathing impacts the accuracy of our reconstructed 4DCBCT.

The third factor is the target dataset used to constrain the model. We compare the accuracy of the reconstructed 4DCT motion model when the targeted fitting data is the simulated DVFs, the CTs driven by the simulated DVFs, or, in the realistic case, the simulated projection radiographs.

II. METHODS AND MATERIALS

II.A. Numerical simulation of 4DCBCT data

Rather than use actual CBCT data for our proof of concept testing, we numerically simulated our 4DCBCT data. This afforded us control over the motion complexity that we model and provided us a gold standard with which to evaluate our results at every step of the reconstruction process. A numerical 4DCBCT data-set was produced as follows:

- (1) We constructed a numerical CT phantom, also referred to as the source CT, in the form of a cubic volume cropped from a patient pelvis CT.
- (2) An initial DVF was created with three spatial components, with the amplitudes varying in space according to a three dimensional Gaussian function that was maximized at the center of the CT.
- (3) A time series of DVFs was produced by scaling the vectors in the initial DVF by the amplitude of a breathing trace sampled at successive time points. We produced DVF time series of different complexity by varying the breathing trace used to scale the initial DVF and/or introducing relative voxel-by-voxel temporal phase shifts between individual displacement vectors. Listed from simplest to most complex the four types of simulated motion were:
 - (a) The initial DVF scaled by a synthetic sine wave breathing trace with no phase shifts between displacement vectors.
 - (b) The initial DVF scaled by an actual patient breathing trace with no phase shifts between vectors. The patient breathing trace was one of those observed by Murphy and Dieterich³⁰, and is comparable to a RPM breathing trace.
 - (c) The initial DVF scaled by a synthetic sine wave breathing trace with multiple phase shifts between displacement vectors. Three separate phase shifts are used for the multiple shift case. They are distributed by scanning across rows of vectors and applying the different shifts in sequence, repeating the pattern every three vectors.
 - (d) The initial DVF scaled by an actual patient breathing trace with multiple phase shifts between vectors (same patient breathing trace as described above).
- (4) The time series of DVFs was applied to the source CT (our static prior image), producing a time series of deformed CTs.

(5) DRRs at different simulated gantry angles were cast through the time series of deformed CTs to produce a rotating set of projections taken through moving anatomy. This became our simulated 4DCBCT data-set. In our CBCT geometry, the CTs were composed of a $64 \times 64 \times 64$ matrix of $1 \times 1 \times 1$ mm voxels, and a flat panel detector was simulated as a 256×256 matrix of 1×1 mm pixels.

Generating our 4DCBCT in this manner not only produced the requisite set of rotating projections of our moving object but also gave us the true CTs underlying each projection and the true DVFs linking those CTs to the source CT. These are used in our analysis of the factors contributing to reconstruction accuracy.

II.B. Algorithm workflow

Our algorithm works to optimize the free parameters of the DVF motion model by producing an optimal match between the actual projections in the 4DCBCT data-set and simulated projections (DRRs) taken through the source CT animated with the motion model.

The process requires two data sets—a prior FBCT of the patient, which we designate as the source image, and a set of 2D CBCT projection images, which we designate as the target image set. The FBCT is assumed to be acquired at breathhold; the CBCT projections are acquired serially at 10–30 frames per second during free breathing. The algorithm workflow is then as follows:

- (1) The motion model generates a time series of DVFs using the current model parameter values.
- (2) This DVF time series is applied to the source CT, producing a time series of deformed CTs.
- (3) DRRs at appropriate gantry angles are cast through the deformed CTs to produce a time series of DRRs.
- (4) The simulated DRRs are compared with the target projections from the 4DCBCT data-set.
- (5) If the matching between the DRRs and projections is determined to be fully optimized, iteration stops and the algorithm is finished. Otherwise the motion model parameters are updated to produce better matching, and the algorithm returns to step 1.

II.C. Motion model description

Our motion model is based on the fundamental assumption that the time-changing displacements of all the anatomical elements of the patient can be completely described by their relative amplitudes and phases, and that these relative amplitude and phase relationships remain approximately constant for the duration of the CBCT scan. Under this assumption, we expect that the motion of each voxel will track with the motion of a single reference point, to within their relative amplitude and phase shift factors. If the motion of the reference point is represented by a measured breathing trace $b(t)$, then we assume that the motion of the other voxels will track with $b(t)$. This forms the basis for our breathing model.

To represent the model mathematically, let $\mathbf{u}(\mathbf{x}, t)$ be a single displacement vector at voxel position $\mathbf{x}$ and time t . Collect all of the displacement vectors $\mathbf{u}$ at time t into a feature vector $\mathbf{DVF}(t)$. The model constructs the temporal sequence of DVFs from an appropriately weighted sum of orthonormal basis vectors $\mathbf{e}_j$ spanning the space of DVF vector components. The weights $w_j(t)$ are modulated in proportion to the time-dependent breathing amplitude measured at one particular spatial point, as in following equation:

$$\mathbf{DVF}(t) = \sum_{j=1}^M w_j(t) * \mathbf{e}_j. \quad (1)$$

The basis vectors contain all the relative amplitude and phase information for the motion. The weighting factors accommodate the absolute amplitude information by modulating the measured breathing amplitude.

The most desirable basis is one in which the variance of the motion data is maximally concentrated along the first few basis vectors and becomes zero or negligible for all other basis vectors. In addition, we desire that the motion data vary independently along each basis vector in relation to all other basis vectors. Then, we must only determine weights for the limited set of important basis vectors and need not worry about any interdependence between them. Such a basis is represented by the principal components of the breathing motion. Although the principal component basis vectors are not necessarily stochastically independent they are maximally decorrelated, which serves as a reasonable approximation of independence. We determine this optimal basis by performing PCA on a training set comprised of a temporal sequence of DVFs derived from the breathing motion of the imaging subject. The training set must be of sufficient quality to characterize the phase and amplitude relationships present between voxels during patient breathing. Once the training set is acquired its covariance matrix is calculated and the M eigenvectors of the covariance matrix are determined using PCA. These eigenvectors, or principal components, are the orthonormal basis vectors $\mathbf{e}_j$ in Eq. (1). They capture the relative amplitude and temporal phase relationships of all of the individual displacement vectors. The weights, or principal coefficients, are modeled in time using the breathing trace $b(t)$ as a template so that, per our assumption above, the displacement vectors will approximately track with the observed movement of the breathing reference point. The idea of creating a complete spatial motion model by driving a set of principal component vectors by the movement of a single reference point has been independently investigated by Li *et al.*,²⁷ who also find that it is a reasonable approach. We relate each principal coefficient to the breathing trace through three adjustable parameters: an amplitude term a_j , a time shift τ_j , and a baseline shift term c_j . Mathematically the model can be expressed as

$$\mathbf{DVF}(\mathbf{p}, t_i) \cong \mathbf{DVF}_{av} + \sum_{j=1}^M pc_j(t_j) * \mathbf{e}_j \quad (2)$$

where

$$pc_j(t_i) = a_j * b(\tau_j + t_i) - c_j \quad (3)$$

The rationale for this simple temporal parameterization of the principal coefficients (and our underlying assumption about the relative movement of voxels) is found in the behavior of actual 4DCT principal coefficients, as shown for example in Vaman *et al.*,²⁹ where the first two principal coefficients track the breathing cycle to within an amplitude, phase, and offset factor as in Eq. (3).

The displacement vector field $\mathbf{DVF}(\mathbf{p}, t_i)$ represents the DVF generated with model parameters $\mathbf{p}$ mapping the source CT to the anatomy at time t_i . $\mathbf{DVF}_{av}$ is the mean of the DVF training set. The vectors $\mathbf{e}_j$ are the PCA eigenvectors. The coefficient $pc_j(t_i)$ is the value at time t_i of the j 'th principal coefficient corresponding to eigenvector $\mathbf{e}_j$. The $3*M$ fitting parameters a_j , τ_j , c_j for all principal coefficients form the contents of parameter vector $\mathbf{p}$. The factor $b(\tau_j + t_i)$ is the value of the breathing trace sampled at time $\tau_j + t_i$. This parameterization has been observed to provide a reasonable representation of the temporal behavior of principal coefficients derived directly from the principal component analysis of a set of empirical breathing DVFs, as discussed below.

The model utilizes M different eigenvector-principal coefficient pairs. In practice, it has been found that the first two or three PCA eigenvectors are usually sufficient to accurately describe the breathing motion in the training set.^{19,20,23,29} In the tests of our fitting algorithm a motion model utilizing five, seven, or ten eigenvectors produced no improvement in fitting accuracy over the three eigenvector model.

II.D. Generating the DVF training set for PCA

The DVF training set serves as the input to the PCA that produces the basis of eigenvectors, which is the heart of the motion model. Each DVF in the training set serves as a snapshot of the phase and amplitude relationships between voxel displacements at a particular time. For the motion model to perform well it is important that the training DVFs be reasonably free of artifacts and errors and sufficiently numerous to capture the essential features of the patient's breathing. For the tests reported here we have generated our DVF training set using four different methods:

- (1) Populate the training set with 16 DVFs sampled directly from the known DVFs of the simulated numerical data-set. This provides an idealized training set that can be used as the gold standard for the PCA eigenvectors.
- (2) Sample a subset of 16 CTs from the known CTs of the simulated numerical data-set and deformably register them to the source CT. This introduces the effect of deformable registration inaccuracy and the resulting DVF errors on the PCA eigenvectors.
- (3) Create idealized sets of "binned" cone-beam projections by sampling 16 CTs from the known CTs and casting 100 projections through each. Then directly register the 2D projections in each bin to the source CT via iterative forward projection matching.²⁸ Here, the deformable registration is constrained by 2D projection matching rather than 3D CT matching.

- (4) Bin the projections as above, perform a conventional FBP reconstruction on each of the phase bins, and register the reconstructed phases to the source CT. This adds the effect of CBCT back-projection artifacts to the uncertainties introduced by DIR in methods 2 and 3.

For this proof of concept study, we have implemented an idealized projection binning procedure for the IFPM and backprojection methods. By doing this we sidestep the complications of intrabin motion that can be expected to occur in real life. Our goal is to test the two methods under identical conditions and discover the best possible performance we can expect from each when those conditions are favorable. This provides us with a baseline to compare against in future studies when more realistic conditions are simulated.

II.E. Testing algorithm performance

We have tested our algorithm under a variety of conditions. The first was the complexity of the simulated breathing motion present in the target data-set. We varied the breathing motion by using either a fully periodic or pseudo-periodic breathing trace, and by introducing relative phase shifts in the voxel motion. The scenario where the simulated breathing motion is fully periodic and no phase shifts are present represents the simplest case, and pseudo-periodic motion with multiple phase shifts the most complex.

The second condition we tested was the method in which the DVF training set is generated. We used the four methods outlined in Sec. II D.

The third condition tested was the fitting target. The fact that our simulated test data-set contains not only projections but also their known underlying DVF and CT time series means that we can fit the motion model directly to the known DVFs, or to their associated CTs, or, in the most realistic case, to the CBCT projection radiographs. In this manner, we can identify the effects of each proxy connecting the motion model to the fitting target data on the success of the motion model optimization. As an example, the algorithm workflow recast to treat the true DVF time series as the fitting endpoint would look like the following:

- (1) The motion model generates a time series of DVFs using the current model parameter values.
- (2) The full set of simulated DVFs is compared against the known DVFs from the numerical data-set.
- (3) If the matching is determined to be fully optimized, iteration stops and the algorithm are finished. Otherwise the motion model parameters are updated to produce better DVF matching, and the algorithm returns to step 1.

Finally, we can change the optimization algorithm used to navigate to the best possible set of motion model parameters. We have tested the Levenberg–Marquardt, Polak–Ribiere conjugate gradient, BFGS, and Nelder–Mead downhill simplex algorithms. All optimization algorithms are implemented using the GNU Scientific Library (GSL), a C library containing various routines for scientific

FIG. 1. (a) A slice through the source CT, (b) a slice through a known deformed CT, (c) a slice through the deformed CT reconstructed using our method, (d) for comparison, a slice through the same deformed CT reconstructed by casting 100 projections through the known deformed CT and reconstructing with filtered backprojection. The simulated motion type for this example uses a sine waveform with no phase shifts. DRRs were used as the matching target and the training set was generated using the 2D/3D registration method for generating c). The grid overlaying the images was added to make the registration process more robust by adding additional high contrast features.

computing. The library is available freely online, where accompanying documentation is easily accessible.³¹ The Nelder-Mead simplex algorithm has proven to be most robust; therefore, it is used for all reported results. The N-M simplex algorithm initializes a simplex with vertices randomly placed around an initial guess in the parameter space. It then updates the worst vertex using simple geometric transformations. Effectively, as the algorithm iterates the simplex crawls across the solution space until it reaches a minimum, at which point it contracts until the RMS distance of the vertices to the simplex center falls below a threshold value. Once this stage is reached iteration terminates and the coordinates of the best vertex are returned as the optimal set of parameters.

III. RESULTS

Figure 1 displays an example slice of the source CT, the known (i.e., ground truth) deformed CT phase, the CT phase reconstructed by our proposed method, and the CT phase reconstructed via conventional 4D CBCT filtered backprojection. The results presented in Figs. 2–4 represent the success of our algorithm at reconstructing the known DVF and CT time series under variations of simulated breathing motion complexity, PCA training set generation method, and target data (i.e., fitting endpoint). The root mean square error (RMSE) is used as a metric to analyze the reconstruction quality. For DVF reconstruc-

tion, the RMSE is in mm and for CT reconstruction it is in CT number intensity units (i.u.). Vector components in the simulated DVFs range from about 0 to 6 mm (CT voxels are $1 \times 1 \times 1$ mm), while CT number intensity units in the simulated CT time series range from 0 to about 3000 i.u. Each data point in Figs. 2–4 represents the average of the RMSE calculated for each reconstructed DVF or CT at every time point considered by the fit. The data in Figs. 2–4 were obtained by fitting 15 time steps spanning about two breathing cycles. This provided a close up look at the algorithm's performance over a short, closely sampled time interval. However, to ensure that the performance of our algorithm did not degrade when we considered more extended periods of patient breathing we also fit data at 100 time points covering about ten patient breathing cycles for the most realistic fitting scenario, i.e., motion driven by the patient breathing trace with multiple phase shifts and projections as the fitting target. For all training set generation methods, including more time points spanning a larger number of breathing cycles improved the accuracy of the fit. For training set methods 1–4 the reduction in reconstruction error of the extended fit over the limited fit in terms of percent DVF RMSE/CT RMSE were 19%/15%, 12%/8%, 9%/7%, and 5%/5%, respectively.

Figure 5 displays the eigenvalue spectra for PCA performed on error free DVFs generated with different simulated motion complexity. Both the sine and patient single-phase motion types produce only one nonzero eigenvalue. The sine multiphase motion type produces two nonzero eigenvalues, and the patient multiphase motion type produces three nonzero eigenvalues (the third is small and difficult to see on Fig. 5).

Figure 6 demonstrates the possible mismatch between the parameterized function of the breathing trace as given by Eq. (3) and the temporal trace of the principal coefficient that it represents. When data-sets are generated using any but the patient multiphase motion type the breathing trace can perfectly model the principal coefficient traces in time using the idealized eigenvectors of method 1. This can be shown by projecting the known DVF time series into the relevant eigenvectors to directly produce principal coefficient traces, and then fitting the parameterized function of the breathing trace as in Eq. (3) directly to the principal coefficient traces. However, principal coefficient traces for data generated with the patient multiphase motion type cannot be perfectly represented with the parameterized function of the breathing trace, even with ideal eigenvectors. Figure 6 compares the known principal coefficient traces of patient multiphase data with their best representations when modeled with Eq. (3).

IV. DISCUSSION

Conventional backprojection, algebraic, and statistical reconstruction methods for CBCT must recover all of the Hounsfield numbers from the projection data. This places maximal demands on the amount of sampled projection data and the accuracy of the physical processes incorporated in the reconstruction process, resulting in well-known CBCT

FIG. 2. Variation of DVF (top) and CT (bottom) reconstruction error with training set generation method for different combinations of simulated motion complexity and fit target.

deficiencies. Phase-binned reconstruction for 4DCBCT only makes the problems worse. We have presented here a new reconstruction approach that sidesteps these problems. Our strategy of deforming a prior FBCT image to match CBCT projection images not only relaxes the demands on projection data sampling through the use of prior information, it obviates the need to recover Hounsfield numbers directly from the CBCT projections. Furthermore, the use of a motion model to animate the prior image allows us to obtain a temporal portrait of anatomical motion that is limited only by the time resolu-

tion and the duration of the CBCT scan, rather than by the practical considerations of conventional 4DCT phase bins. The resulting gain in 4DCBCT image quality is clearly seen in Fig. 1. Our 4DCBCT reconstruction concept was subjected to tests designed to address three questions: (1) What is the best way to generate the DVF training set for PCA? (2) How is the algorithm's performance influenced by the complexity of the breathing motion? (3) How does each proxy connecting the motion model to the output DRRs affect the accuracy with which the optimal model parameters can be determined? We

FIG. 3. Variation of DVF (top) and CT (bottom) reconstruction error simulated motion complexity for different combinations of training set generation method and fit target.

investigated two practical means of obtaining a DVF training set—the 2D/3D forward projection matching approach (method 3 from Sec. II D), and the conventional backprojected 3D/3D registration approach (method 4 from Sec. II D)—and compared them to the idealized methods 1 and 2. The two idealized methods represent the gold standards. Figure 2 clearly shows that the 2D/3D approach produces better results than the backprojected 3D/3D method. Furthermore, results from the 2D/3D method do not differ substantially from those obtained by directly registering a subsample of the known CTs to the source CT (method 2 from Sec. II D). As expected, avoiding registration altogether by using a subsam-

ple of the known DVFs (method 1 from Sec. II D) provides the best result. When the results from Fig. 2 for each PCA training method are averaged together they produce average RMSEs of 0.012 mm for the DVFs and 8.56 i.u. for the CTs from method 1, and 0.053 mm/21.65 i.u., 0.093 mm/ 24.91 i.u., 0.278 mm/57.30 i.u. for methods 2, 3, and 4, respectively. From this we see that method 4 produces a roughly threefold increase in RMSE over method 3, while method 3 introduces a less substantial increase in error over method 2. One point of interest is that while the average DVF RMSE of method 3 is nearly twice that of method 2, the CT RMSE is only slightly higher. This is due to the fact that large differences in the

FIG. 4. Variation of DVF (top) and CT (bottom) reconstruction error with fit target for different combinations of simulated motion complexity and training set generation method.

DVF often cause negligible changes in the resulting deformed CT, especially when the CT contains objects that are more or less homogeneous. Exceptions can occur at object boundaries, where small DVF changes may produce large intensity changes in edge voxels of the deformed image. Finally, we note that even areas of the numerical phantom that appear homogenous tend to have a voxel-to-voxel standard deviation of around 8–10 i.u., which is typical of a fan-beam CT. Average CT RMSEs from the first 3 methods fall within 2 or 3 factors of this number, while the RMSE of method 4 is greater by a

factor of 5 or 6. In our comparison of the DVF training set generation methods we chose to implement an idealized image binning procedure for methods 3 and 4. Our purpose was to eliminate any residual motion from the bins and ensure that each contained an adequate projection count in order to compare both approaches under ideal circumstances. In the real world, projections will be binned according to the patient breathing trace, which will result in intrabin residual motion and irregular numbers of projections per bin. We expect that under these conditions the iterative forward projection

FIG. 5. Eigenvalue spectra for PCA performed on DVFs generated with different simulated motion complexity. The DVFs are directly sampled from the sets of known DVFs.

matching method will further distance itself from the backprojection method. The IFPM method has been shown to produce accurate results with far fewer projections than are required for a FBP reconstruction of decent quality.²⁸ This allows for narrower bin widths which will in turn reduce residual motion

within the bin. In addition, if bins are narrower, then more bins can be used without redundancy, allowing for a more densely populated DVF training set that should lead to more accurate PCA eigenvectors. This presents an interesting trade-off problem for future study.

FIG. 6. Parameterized breathing trace fit to known principal coefficient traces for the first (top) and second (bottom) eigenmodes. The known principal coefficient traces were produced by taking DVF time series generated with the patient multiphase motion type and projecting them into the first and second eigenvectors. Eigenvectors were trained on a set of DVFs sampled from the known DVF time series.

The second concern was the performance of the algorithm when faced with different degrees of simulated motion complexity. We have broken this down into three subareas of inquiry: (1) How many eigenmodes are sufficient to create an accurate motion model? (2) How is convergence affected when more eigenmodes, and thus more free parameters, are included to model greater motion complexity? And (3) Can our breathing trace parameterization accurately represent known principal coefficient traces as motion complexity increases?

Our expectation based on our previous experience²⁹ as well as that of others^{19,20,23} was that three eigenmodes would be sufficient for an accurate model. Figure 5 illustrates the eigenvalue spectra for the four different complexities of simulated motion when PCA is computed on an ideal training set (i.e., a set of DVFs modulated directly by the breathing signal). The two single-phase motion types each generate only one nonzero eigenmode, the sine multiphase type generates two nonzero modes, and the patient multiphase type generates three non-negligible modes. This is what one would expect from the principal components analysis of time-varying DVFs. For a single-phase, sinusoidal motion, the covariance matrix of the DVFs has rank one, and thus only one nonzero eigenvector. For a sinusoidal motion with different voxel phases, the covariance matrix has rank two and there are only two nonzero eigenvectors. For more complex motions, the covariance matrix has higher rank (cf., reference 29 for the proof). For these idealized cases, three eigenmodes will be sufficient to fully model any of these motion types. However, when imperfect DVF training data are used from methods 2–4 a tail appears in the eigenvalue spectra extending nonzero values all the way to mode 15 (there are 16 DVFs per training set, which allows for a maximum of 15 nonzero eigenmodes). Even so, we have observed that in terms of DVF or CT error there is no benefit gained by modeling more than three modes.

In terms of convergence of the optimization routine, including more eigenmodes does not appear to have any effect. In the cases tested the algorithm reduced the objective function to the same value when fitting three modes as it did when fitting four or more modes.

To examine the effectiveness of our breathing trace parameterization it is useful to look at the bottom-most trend of Fig. 3. This trend plots the RMSE for fits to motion data of varying complexity under the most ideal conditions, i.e., the known DVFs are used as the fitting target and method 1 is used to generate the DVF training set. The first three data-points are more or less flat, with an increase in RMSE occurring only for patient multiphase motion complexity. This is due to the fact that pseudo-periodic motion with multiple phases cannot be perfectly modeled by the parameterized function of the breathing trace, while the simpler three motion types can. Figure 6 displays the parameterized breathing trace fit directly to principal coefficient traces for the first and second breathing modes of patient multiphase type data. Figure 6 shows that the breathing trace fit is close to the known principal coefficient traces but cannot represent them perfectly. This limitation becomes a greater issue when multiple eigenmodes contribute equally to

the overall variance of the motion data. However, we have previously observed only one dominant mode with one or two much weaker modes in a data-set tracking patient breathing²⁹ and have designed our simulated motion to match. Thus, the increase in fitting error is small when going from the sine multiphase motion type to the patient multiphase motion type, as illustrated by Fig. 3 where trend averaged RMSEs are 0.124 mm/31.61 i.u. and 0.126 mm/35.52 i.u. for sine and patient multiphase motion types, respectively.

Our third question concerns the compression of DVF data down to the DRR level at which matching to known data is performed. This gives the optimization algorithm less information to constrain the model parameters. When starting from parameter values initialized to zero, markedly worse results

Fig. 7. (a) Slice taken through the optimization landscape with DVFs used as the matching target. (b) A slice through the same section of parameter space but with DRRs used as the matching target. The data for this fit was generated using sine motion with no phase shifts and DVFs sampled from the known set of DVFs for the training set.

can occur when the algorithm uses DRRs as target data vs when it uses DVFs. However, this problem can be avoided by using a good starting guess for the fit parameters, as evidenced by the results in Fig. 4. One can get a good starting set of model parameters by projecting the training set into the eigenvectors to get time averaged principal coefficient traces and then fitting these to the corresponding time averaged breathing trace amplitudes. Figure 4 plots the RMSE as the matching target changes from the known DVF time series to the intermediate CT time series and finally to the DRR time series. Averaged across all trends the RMSE values are 0.099 mm/25.76 i.u., 0.107 mm/26.97 i.u., and 0.122 mm/31.20 i.u. for the DVF, CT, and DRR matching cases, respectively.

In addition, the effects of target data compression were a deciding factor in our choice of optimization algorithm. Our initial pick was the Levenberg–Marquardt algorithm, which offered fast convergence in few iterations for the DVF fitting case, but failed at the CT and DRR matching levels. At the DVF matching level, the optimization landscape is relatively smooth and contains a steep and well pronounced global minimum. However, when CT and DRR matching are introduced the sharp global minimum becomes a more gently sloping basin and shallow local minima appear as ripples in the texture of the optimization landscape. Examples of the different optimization landscapes encountered when using the DVFs or the DRRs as the matching targets are presented in Fig. 7. Both the Polak–Ribiere conjugate gradient method and the BFGS algorithm provided at least partial success, but the Nelder–Mead simplex algorithm, our only non-gradient method, proved to be the most dependable.

Our motion model requires the optimization of $3 \times M$ free parameters, where M is the number of eigenmodes incorporated. For our simulations M was equal to 3, giving us 9 free parameters. For comparison, the B-spline models of Docef and Murphy²⁵ and Zeng *et al.*²⁴ both contain thousands of free parameters while the PCA model of Li *et al.*^{19,20} requires only M free parameters, where they also set M equal to 3. However, unlike our model and that of Docef and Murphy and Zeng *et al.* their model contains no temporal dependence and therefore the fitting process must be repeated at every time step.

In terms of algorithm efficiency, when fitting data simulated with the most realistic conditions around 200 iterations are generally required for convergence of the Nelder–Mead simplex algorithm. On an Intel core i7 CPU running at 2.8 GHz with 12 GB of memory and considering 15 time steps in the fit this process generally takes between 5 and 10 min. On our initial GPU implementation of the code (GeForce GTX480) the same trial finishes in about 1–1.5 min.

V. CONCLUSIONS

We have developed a novel 4DCBCT reconstruction concept in which a prior FBCT is animated by a parameterized motion model such that it reproduces a set of CBCT projections. The resulting 4DCT has the Hounsfield number fidelity of a FBCT and the time resolution of the individual CBCT projections. It avoids the artifacts and other deficiencies of

conventional CBCT reconstruction, as well as the time averaging and motion blurring of conventional 4DCT. The motion model is based on the principal eigenmodes of breathing that can be detected in a preliminary 4DCBCT reconstructed via more conventional means. The present study has investigated several issues that can affect the accuracy and robustness of this model. In our proof of concept testing we have shown that our motion modeling algorithm reconstructs the CTs and DVFs underlying a set of projection data with good accuracy for the simulated motion types considered. Our results indicate that the quality of the DVF training set used for the PCA motion model is important and that, of the two practical training set generation methods tested, the 2D/3D forward iterative projection matching approach is clearly superior to conventional backprojection and registration. In addition, we have found that the Nelder–Mead simplex algorithm provides the most robust optimization.

In future work, we will update our DRR code to include physically realistic features such as scatter and test against actual phantom and patient projections. In addition, we will test different binning approaches to mitigate the effects of residual motion blurring on the DVF training set.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Jun Lu for assistance in reconstructing 4DCBCT images. This work was supported in part by NCI Grant No. R01CA123299. The authors report no conflicts of interest.

^{a)}Author to whom correspondence should be addressed. Electronic mail: david.staub32@gmail.com

¹P. Keall, “4-dimensional computed tomography imaging and treatment planning,” *Semin. Radiat. Oncol.* **14**, 81–90 (2004).

²S. 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).

³L. Dietrich *et al.*, “Linac-integrated cone beam CT: First experimental results,” *Phys. Med. Biol.* **51**, 2939–2952 (2006).

⁴M. J. Fitzpatrick, G. Starkschall, J. A. Antolak, J. Fu, H. Shukla, P. J. Keall, P. Klahr, and R. Mohan, “Displacement-based binning of time-dependent computed tomography image data sets,” *Med. Phys.* **33**, 235–246 (2006).

⁵T. Li, L. Xing, P. Munro, C. McGuinness, M. Chao, Y. Yang, B. Loo, and A. Koong, “Four-dimensional cone-beam computed tomography using an on-board imager,” *Med. Phys.* **33**, 3825–3833 (2006).

⁶J. J. Sonke, L. Zijp, P. Remeijer, and M. v. Herk, “Respiratory correlated cone-beam CT,” *Med. Phys.* **32**, 1176–1186 (2005).

⁷T. Li, A. Koong, and L. Xing, “Enhanced 4D cone-beam CT with inter-phase motion model,” *Med. Phys.* **34**, 3688–3695 (2007).

⁸Z. Qi and G.-H. Chen, “Extraction of tumor motion trajectories using PICCS-4DCBCT: A validation study,” *Med. Phys.* **38**, 5530 (2011).

⁹S. Rit, J. W. H. Wolthaus, M. v. Herk, and J.-J. Sonke, “On-the-fly motion-compensated cone-beam CT using an a priori model of the respiratory motion,” *Med. Phys.* **36**, 2283 (2009).

¹⁰S. Leng, J. Tang, J. Zambelli, B. Nett, R. Tolakanahalli, and G. H. Chen, “High temporal resolution and streak-free four-dimensional cone-beam computed tomography,” *Phys. Med. Biol.* **53**, 5653–5673 (2008).

¹¹R. Ning, X. Tang, and D. Conover, “X-ray scatter correction algorithm for cone beam CT imaging,” *Med. Phys.* **31**, 1195–1202 (2004).

¹²J. H. Siewerdsen and D. A. Jaffray, “Cone-beam computed tomography with a flat-panel imager: Magnitude and effects of x-ray scatter,” *Med. Phys.* **28**, 220–231 (2001).

- ¹³F. Bergner, T. Berkus, M. Oelhafen, P. Kunz, T. Pa, R. Grimmer, L. Ritschl, and M. Kachelriess, "An investigation of 4D cone-beam CT algorithms for slowly rotating scanners," *Med. Phys.* **37**, 5044–5053 (2010).
- ¹⁴J. Eom, X. G. Xu, S. De, and C. Shi, "Predictive modeling of lung motion over the entire respiratory cycle using measured pressure-volume data, 4DCT images, and finite-element analysis," *Med. Phys.* **37**, 4389–4400 (2010).
- ¹⁵J. R. McClelland, J. M. Blackall, S. Tarte, A. C. Chandler, S. Hughes, S. Ahmad, D. B. Landau, and D. J. Hawkes, "A continuous 4D motion model from multiple respiratory cycles for use in lung radiotherapy," *Med. Phys.* **33**, 3348 (2006).
- ¹⁶J. J. Sonke and J. Belderbos, "Adaptive radiotherapy for lung cancer," *Semin. Radiat. Oncol.* **20**, 94–106 (2010).
- ¹⁷S. Dieterich, K. Cleary, W. D'Souza, M. Murphy, K. H. Wong, and P. Keall, "Locating and targeting moving tumors with radiation beams," *Med. Phys.* **35**, 5684–5694 (2008).
- ¹⁸A. Knopf, C. Bert, E. Heath, S. Nill, K. Kraus, D. Richter, E. Hug, E. Pedroni, S. Safai, F. Albertini, S. Zenklusen, D. Boye, M. Sohn, M. Soukup, B. Sobotta, and A. Lomax, "Special report: Workshop on 4D-treatment planning in actively scanned particle therapy—recommendations, technical challenges, and future research directions," *Med. Phys.* **37**, 4608–4614 (2010).
- ¹⁹R. Li, X. Jia, J. H. Lewis, X. Gu, M. Folkerts, C. Men, and S. B. Jiang, "Real-time volumetric image reconstruction and 3D tumor localization based on a single x-ray projection image for lung cancer radiotherapy," *Med. Phys.* **37**, 2822–2826 (2010).
- ²⁰R. Li, J. H. Lewis, X. Jia, X. Gu, M. Folkerts, C. Men, W. Y. Song, and S. B. Jiang, "3D tumor localization through real-time volumetric x-ray imaging for lung cancer radiotherapy," *Med. Phys.* **38**, 2783–2794 (2011).
- ²¹J. R. McClelland, S. Hughes, M. Modat, A. Qureshi, S. Ahmad, D. B. Landau, S. Ourselin, and D. J. Hawkes, "Inter-fraction variations in respiratory motion models," *Phys. Med. Biol.* **56**, 251–272 (2011).
- ²²D. Yang, W. Lu, D. A. Low, J. O. Deasy, A. J. Hope, and I. E. Naqa, "4D-CT motion estimation using deformable image registration and 5D respiratory motion modeling," *Med. Phys.* **35**, 4577–4590 (2008).
- ²³Q. Zhang, A. Pevsner, A. Hertanto, and Y.-C. Hu, "A patient-specific respiratory model of anatomical motion for radiation treatment planning," *Med. Phys.* **34**, 4772–4781 (2007).
- ²⁴R. Zeng, J. A. Fessler, and J. M. Balter, "Estimating 3-D respiratory motion from orbiting views by tomographic image registration," *IEEE Trans. Med. Imaging* **26**, 153–163 (2007).
- ²⁵A. Docef and M. J. Murphy, "Reconstruction of 4D deformed CT for moving anatomy," *Int. J. CARS* **3**, 591–598 (2008).
- ²⁶D. A. Low, P. J. Parikh, W. Lu, J. F. Dempsey, S. H. Wahab, J. P. Hubenschmidt, M. M. Nystrom, M. Handoko, and J. D. Bradley, "Novel breathing motion model for radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* **63**, 921–929 (2005).
- ²⁷R. Li, J. H. Lewis, X. Jia, T. Zhao, W. Liu, S. Wuenschel, J. Lamb, D. Yang, D. A. Low, and S. B. Jiang, "On a PCA-based lung motion model," *Phys. Med. Biol.* **56**, 6009–6030 (2011).
- ²⁸R. S. Brock, A. Docef, and M. J. Murphy, "Reconstruction of a cone-beam CT image via forward iterative projection matching," *Med. Phys.* **37**, 6212–6220 (2010).
- ²⁹C. Vaman, D. Staub, J. Williamson, and M. Murphy, "A method to map errors in the deformable registration of 4DCT images," *Med. Phys.* **37**, 5765–5776 (2010).
- ³⁰M. J. Murphy and S. Dieterich, "Comparative performance of linear and nonlinear neural networks to predict irregular breathing," *Phys. Med. Biol.* **51**, 5903–5914 (2006).
- ³¹M. Galassi, J. Theiler, J. Davies, B. Gough, R. Priedhorsky, G. Jungman, M. Booth, J. O. Tähtinen, T. Walter, F. Rossi, C. Perassi, S. Jaroszewicz, N. Darnis, J. H. Stover, I. Alxneit, T. Keskitalo, L. Johnson, P. Alken, R. Ulerich, P. Holoborodko, and P. Gonnet, "GNU Scientific Library Reference Manual," 3rd edition (2009), <http://www.gnu.org/software/gsl/>.