Respiratory motion estimation from slowly rotating x-ray projections: Theory and simulation

Rongping Zeng[a] and Jeffrey A. Fessler[b]
Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, Michigan 48109-2122

James M. Balter[c]
Department of Radiation Oncology, University of Michigan, Ann Arbor, Michigan 48109-0010

(Received 26 July 2004; revised 1 February 2005; accepted for publication 3 February 2005; published 18 March 2005)

Understanding the movement of tumors caused by respiratory motion is very important for conformal radiatherapy. However, respiratory motion is very difficult to study by conventional x-ray CT imaging since object motion causes inconsistent projection views, leading to artifacts in reconstructed images. We propose to estimate the parameters of a nonrigid, free breathing motion model from a set of projection views of the thorax that are acquired using a slowly rotating cone-beam CT scanner. This approach involves deforming a motion-free reference thorax volume according to the estimated parameters and comparing its projections to the corresponding measured projection views. The parameters are optimized by minimizing a regularized squared error cost function. Simulation results with a fan-beam geometry show good agreement between the estimated motion and the true motion, which supports the potential of this approach for estimating four-dimensional (three-dimensional spatial + temporal) respiratory motion. © *2005 American Association of Physicists in Medicine*. [DOI: 10.1118/1.1879132]

Key words: respiratory motion, cone-beam projection, regularized likelihood estimation, nonrigid deformation, B-spline

I. INTRODUCTION

Conformal radiotherapy requires that the delivery of radiation is focused on the tumor area while sparing the normal adjacent tissue. Although various methods have been developed for delivering precise radiotherapy, respiratory motion remains a significant source of error in treatment planning for the thorax and upper abdomen. Conventional techniques to deal with the motion effects often involve an increase in the margin of the target volume. To suppress respiratory motion, treatment can be done under breathhold condition assisted with breathing control devices.[1,2] But some patients, especially those with lung cancer, have difficulty holding their breath, so techniques for treatment in free breathing conditions are now being investigated.[3] In those techniques, consideration of organ motion is essential.

Because of the significance of motion effects in radiotherapy, a growing number of studies have been dedicated to four-dimensional (4D) CT imaging. Most work in this field focuses on the reconstruction of a CT volume with minimized motion artifacts. This work can be divided into three main categories: fast scanning, reconstruction for motion compensation, and gated image acquisition. In the first class, researchers endeavor to shorten scanner rotation times for data acquisition to reduce motion artifacts and improve temporal resolution.[4–6] In the second class, reconstruction algorithms for motion compensation are based on assumptions of *a priori* deformation model,[7–9] or based on the estimation or detection of motion using extra hardware.[10,11] In gated image acquisition techniques, devices are used to measure the breathing state either as a trigger signal to initiate the scan to acquire data at a certain breathing state,[12] or as a metric to sort the CT scans into bins of equivalent breathing states to form a volume.[13,14]

Although lots of work has been carried out related to 4D CT imaging, there have been fewer studies on building a 4D model for free breathing motion.[15,16] Naqa *et al.* proposed to estimate respiratory motion by registering several three-dimensional (3D) CT volumes at ordered breathing states, which are acquired using gated image acquisition techniques.[15] Usually, this technique requires multiple scans (about 15) for each table position. Such motion estimation methods are promising, but currently have two drawbacks. First, due to x-ray dose limitation, the motion model can only be estimated from a few CT volumes which may be insufficient to fit the respiratory motion of a whole breathing cycle. Second, when sorting the two-dimensional (2D) CT slices corresponding to the same breathing state to form a 3D image, the accuracy depends on the bin size of the breathing states and the accuracy of the device that measures the breathing states. Zijp *et al.* proposed a novel idea to sort the cone-beam x-ray projection images into several bins of breathing phases based on the position of the diaphragm in the projection images.[16] They detected the diaphragm by using image enhancement techniques and tracked the superior–inferior position of the diaphragm to generate a periodic one-dimensional breathing signal. Then 4D thorax volumes were reconstructed using the projections corresponding to the same breathing phase. The advantage of this method is that

no extra device is required. But one limitation of such methods is the assumption of periodicity of the respiratory motion.

In this work, we develop a different approach to estimate respiratory motion from two sets of measurements. One set is a sequence of projection views acquired from a slowly rotating cone-beam CT scanner, such as the type of imager that could be incorporated into a radiotherapy system. Usually such systems take about 1 min per rotation. During such a long acquisition period, patients would breathe freely, so the measured projection views capture information about respiratory motion. We also assume that we have available a motion-free reference volume, such as the reconstructed thorax images acquired by conventional breath-hold x-ray CT on a modern fast scanner. Our approach involves deforming this reference image according to the estimated motion parameters and comparing its projections to the corresponding measured projection views. In contrast to gated imaging techniques, our method requires only one to two scans instead of dozens of scans. Moreover, because the measured projection views correspond to the naturally continuous breathing state, the external state-measuring devices are unnecessary.

As is well known, estimation is an inverse procedure aimed at recovering unknown parameters from available measurements. Generally, for a nonlinear estimation problem, there are three main tasks: define a suitable system model, choose a good cost function, and select appropriate optimization algorithms. In our estimation problem, motion is defined by a parametric model based on B-splines. The cost function is the penalized least square error. The optimization algorithm we used is the Levenberg–Marquardt method.

The paper is organized as follows. Section II describes the problem and Sec. III describes the proposed estimation method, including the temporal motion model, similarity measure, and optimization method. Section IV presents our fan-beam simulation results. Finally, we set out our conclusion and future work.

II. PROBLEM STATEMENT

The following defines the estimation problem, beginning with a description of the x-ray projection operator, and then turning to an explanation of the measurements we will collect for the estimation problem.

A. X-ray projection operator

CT is a noninvasive imaging technique allowing the visualization of the internal structure of an object. In a CT system, the patient is placed between an x-ray source and an array of x-ray detectors. By rotating the source and the detector simultaneously around the patient, a large number of x-ray projections from different angles can be obtained during the data acquisition period. Ideally, each projection represents the summation or line integral of the attenuation coefficients of an object along a particular ray path, which can be represented mathematically as follows:

$$g_{m,n} = \int_{L_{\phi_m,n}} f_{t_m}(\boldsymbol{x})\,dl, \quad n = 1, \ldots, N, \quad m = 1, \ldots, M, \quad (1)$$

where $f_t(\boldsymbol{x})$, $\boldsymbol{x} \in \mathbb{R}^d$ denotes the attenuation coefficients of a d dimensional object ($d=2$ or 3) at time t, $g_{m,n}$ the projection value at the nth element of the detector from projection angle ϕ_m, N the number of detector elements, M the number of projection views, and $L_{\phi_m,n}$ the line between the x-ray source and the nth detector pixel at projection angle ϕ_m.

Let $\mathcal{A}_{\phi_m}$ denote the projection operator for angle ϕ_m, then Eq. (1) can be written simply as follows:

$$\boldsymbol{g}_m = \mathcal{A}_{\phi_m} f_{t_m}, \quad (2)$$

where $\boldsymbol{g}_m = (g_{m,1}, \ldots, g_{m,N})$. Ignoring beam hardening effects, the measurements $\boldsymbol{Y}$ from an x-ray detector are related to the projections by Beer's law:[17]

$$E[Y_{m,n}] = I_{m,n} e^{-g_{m,n}} + S_{m,n}, \quad n = 1, \ldots, N, \quad m = 1, \ldots, M, \quad (3)$$

where $E[\cdot]$ stands for the expectation of a random variable, $I_{m,n}$ is a constant related to the incident x-ray intensity, and $S_{m,n}$ denotes the scatter contribution to $g_{m,n}$.

B. Measurement model

The proposed method uses two sets of measurements. One set is a motionless thorax image f_{ref}, obtained with a conventional fast breath-hold CT scan. This image serves as a reference image and all the deformations are applied to this image. The other set is a sequence of M projection views from a slowly rotating scanner: $\{\hat{\boldsymbol{g}}_m\}$, $m=1, \ldots, M$, where ϕ_m is the projection angle at time t_m. The projection views are estimated from the measurements $\boldsymbol{Y}$ as follows:

$$\hat{g}_{m,n} = \log\left(\frac{I_{m,n}}{Y_{m,n} - S_{m,n}}\right), \quad (4)$$

where $S_{m,n}$ is a (possibly zero) scatter estimate.

Our method could also work with faster scanning, but recently developed cone-beam CT systems for radiotherapy usually have rotation speeds of approximately 6°/s. A full 360° rotation takes about 1 min. Patients breathe naturally during the scan. Although we allow the cone-beam scanner to rotate slowly, we require the acquisition time of each projection view to be short. For example, recently developed systems can acquire 15 frames/s, i.e., around 0.067 s/frame. We therefore assume that the respiratory motion is negligible within each single projection view.

The measured projections will be degraded by noise. For simplicity, we treat the x-ray detector outputs as independent Poisson random variables:[17]

$$Y_{m,n} \sim \text{Poisson}\{E[Y_{m,n}]\}, \quad n = 1, \ldots, N, \quad m = 1, \ldots, M. \quad (5)$$

We want to estimate a sequence of moving objects f_t using f_{ref} from $\{\hat{\boldsymbol{g}}_m\}_{t=1}^{M}$. This is not a conventional image reconstruction problem. Most reconstruction algorithms assume that the object is motionless during the whole data acquisi-

tion period, or assume simple motion models such as affine.[7–9] In our problem the object deforms while being scanned and the task is to estimate the motion by minimizing the difference between the measured projection views and the calculated projection views of the estimated moving objects. This is a kind of "tomographic image registration" problem.

III. THEORY

A. Temporal motion model

Let $\{f_t(\boldsymbol{x})\}$, $\boldsymbol{x} \in \mathbb{R}^d$, $t \in [0,T]$ denote the moving thorax volumes over scan duration T. We assume the moving volumes are all deformations of the reference thorax volume f_{ref} because they are from the same patient. Thus there exists geometric correspondence between f_t and f_{ref} as follows:

$$f_t(\boldsymbol{x}) = f_{\text{ref}}(\boldsymbol{\mathcal{T}}(\boldsymbol{x};\boldsymbol{\theta},t)), \tag{6}$$

where $\boldsymbol{\mathcal{T}}(\boldsymbol{x};\boldsymbol{\theta},t)$ is the unknown deformation function described by parameters $\boldsymbol{\theta}$. Usually, warping at each time point is described by d deformation functions, each along one dimension.

We assume the motion is nonrigid but smooth. We model the deformation function using a weighted sum of shifted basis functions,

$$\boldsymbol{\mathcal{T}}(\boldsymbol{x};\boldsymbol{\theta},t) = \boldsymbol{x} + \sum_{\tau=1}^{K} \sum_{\boldsymbol{i}} \boldsymbol{\theta}_{\tau,\boldsymbol{i}} b\left(\frac{t}{\Delta_t} - \tau\right) \beta\left(\frac{\boldsymbol{x}}{\Delta_x} - \boldsymbol{i}\right), \tag{7}$$

where Δ_x controls the width of the spatial basis function $\beta(\boldsymbol{x})$ and Δ_t controls the width of the temporal basis function $b(t)$. The general approach applies to any differentiable basis functions. Here we chose $b(t)$ to be a cubic B-spline,[18] and $\beta(\boldsymbol{x})$ to be the tensor product of cubic B-splines, i.e., $\beta(\boldsymbol{x}) = \Pi_1^d \beta(x_i)$, where $\boldsymbol{x} = [x_1, \ldots, x_d]'$, a d-dimensional spatial position vector. We chose B-splines for several reasons. B-splines offer good approximation of band-limited signals,[18] and B-splines have been used extensively for modeling nonrigid deformation.[19,20] The compact support of B-splines, and hence small overlap between knots, reduces the dependency between parameters thus makes the optimization problem easier to solve. Combining Eqs. (6) and (7), we can write the deformed object as

$$f_t = \mathcal{W}(\boldsymbol{\theta},t) f_{\text{ref}}, \tag{8}$$

where $\mathcal{W}(\boldsymbol{\theta},t)$ denotes the warping operator corresponding to Eq. (6).

Based on the above-defined motion model, the motion estimation problem is to find the deformation parameters $\boldsymbol{\theta}$. These parameters characterize both the temporal and spatial motion. After estimating $\boldsymbol{\theta}$, one can determine the motion of each point of the object at any time within the scanning period, because the motion model (7) is continuous with respect to the spatial variables $\boldsymbol{x}$ and the time variable t.

In Eq. (6), the deformation acts on a continuous-space reference image $f_{\text{ref}}(\boldsymbol{x})$. Since the actual reference image obtained from a CT scanner is a discrete image, interpolation is needed. We used uniform cubic B-splines for image-domain interpolation:

$$f_{\text{ref}}(\boldsymbol{x}) = \sum_{\boldsymbol{r}} c_{\boldsymbol{r}} \beta(\boldsymbol{x} - \boldsymbol{r}), \tag{9}$$

where the coefficients $c_{\boldsymbol{r}}$ are determined by prefiltering the reference CT volume appropriately.[18]

B. Estimation

We estimate the deformation parameters by minimizing the difference between the measured projection views and the calculated projection views of the deformed reference image. For simplicity, we focus here on the least-squared error metric:

$$L(\boldsymbol{\theta}) = \frac{1}{2} \sum_{m=1}^{M} \|\hat{\boldsymbol{g}}_m - \mathcal{A}_{\phi_m} \mathcal{W}(\boldsymbol{\theta},t_m) f_{\text{ref}}\|^2. \tag{10}$$

The principle generalizes to more complicated statistical models.

Depending on how many motion parameters one uses, for stable estimation it may be necessary to include regularization. We compute the estimate $\hat{\boldsymbol{\theta}}$ by minimizing the following regularized least-squares cost function:

$$\hat{\boldsymbol{\theta}} = \text{argmin}_{\boldsymbol{\theta}}\, \psi(\boldsymbol{\theta}),$$

$$\Psi(\boldsymbol{\theta}) = L(\boldsymbol{\theta}) + \alpha R(\boldsymbol{\theta}), \tag{11}$$

where $L(\boldsymbol{\theta})$ is described in Eq. (10), $R(\boldsymbol{\theta})$ denotes the regularization function, and the scalar α controls the trade-off between the similarity term and the regularity term. We select $R(\boldsymbol{\theta})$ to encourage smoothness of respiratory motion by using

$$R(\boldsymbol{\theta}) = \tfrac{1}{2} \|\boldsymbol{C}\boldsymbol{\theta}\|^2, \tag{12}$$

where $\boldsymbol{C}$ is a differencing matrix. The regularity term can be a combination of both temporal and spatial roughness penalties.

C. Optimization

General-purpose methods can be used to search for the values of parameters $\boldsymbol{\theta}$ that minimize the cost function $\psi(\boldsymbol{\theta})$. We chose the Levenberg–Marquardt method,[21] because it often offers fast convergence for least-squares problems. The iterative scheme is represented as follows:

$$\boldsymbol{\theta}^{n+1} = \boldsymbol{\theta}^n - (H + \lambda_n \,\text{diag}\{H\})^{-1} \nabla \psi(\boldsymbol{\theta}^n), \tag{13}$$

where $\nabla \psi(\boldsymbol{\theta}^n)$is the gradient of $\psi(\boldsymbol{\theta})$ evaluated at $\boldsymbol{\theta}^n$, H is an approximation of the Hessian matrix of $\psi(\boldsymbol{\theta})$, and λ_n is a positive tuning parameter. Usually, λ_n is initialized to be a modest value, say $\lambda_0 = 10$. After a successful step, it is decreased by a factor of 10; otherwise, it is increased by a factor of 10 for the next iteration. The gradient and Hessian are found using the chain rule. See the Appendix for explicit expressions. To save computation time, we update the Hessian only when the current iteration fails.

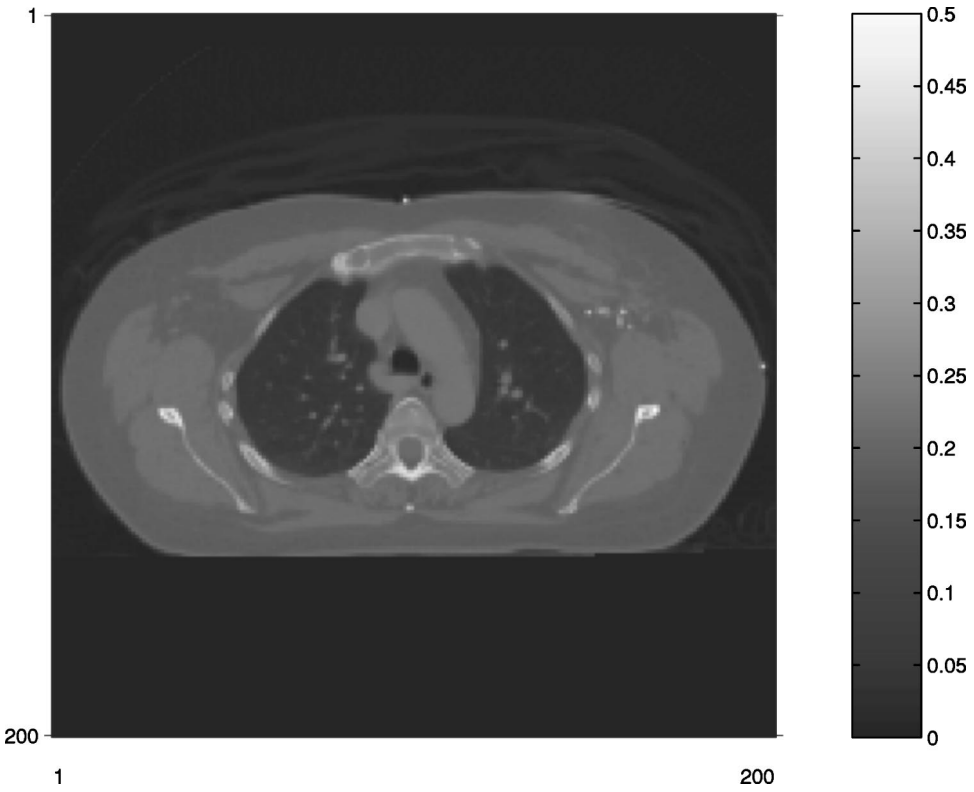

FIG. 1. The reference CT image f_{ref}.

IV. SIMULATION

This section presents our simulation results. For simplicity, we started with fan-beam (2D) geometry to study the performance of the proposed approach.

A. Experiment methods

Figure 1 shows the 200×200 reference image f_{ref} used in our simulation, one slice from a 3D breathhold thorax CT scan. The pixel size of this reference image is 1.96 mm. To form a sequence of moving images $\{f_{t_m}\}_{m=1}^{32}$, we warped f_{ref} according to the following synthetic motion functions:

$$\mathcal{T}_x(x,y,t) = x - 32 \sin^2\left(\frac{\pi t}{T}\right) \cos\left(\frac{\pi(x-x_0)}{X}\right) \times \exp\left[-8\left(\frac{2x}{X}-1\right)^2\right] \exp\left[-8\left(\frac{2y}{Y}-1\right)^2\right], \tag{14}$$

$$\mathcal{T}_y(x,y,t) = y + 16 \sin^2\left(\frac{\pi t}{T}\right) \cos\left(\frac{\pi(y-y_0)}{Y}\right) \times \exp\left[-8\left(\frac{2x}{X}-1\right)^2\right] \exp\left[-8\left(\frac{2y}{Y}-1\right)^2\right], \tag{15}$$

where X and Y denote the width and height of the reference image, respectively, x_0 and y_0 the coordinates of the upper-left pixel, and T the scan duration. The motion generated by these two equations are smooth and symbolically represents expansion and contraction of a breathing cycle. The real motion is usually smaller than what we generated. But since this is a 2D simulation, we have exaggerated the deformation to emulate the larger motion seen in 3D respiratory motion. After generating the deformed images, we computed 32 projection views for a fan-beam geometry with projection angles evenly spaced over 360°. The fan beam had 400 samples spaced by 2.60 mm, and the source to detector distance and isocenter to detection distance were 95 and 40.8 cm, respectively. We added the effects of Poisson noise as described in Sec. II B to produce the noisy data set $\{\hat{\boldsymbol{g}}_m\}_{m=1}^{32}$, shown in Fig. 2(b). The noise corresponds to 10^6 incident photons per detector element.[22] For this initial investigation, we simulated an imaging system that acquired 32 projection views over 360° over a single respiratory cycle ($T=4$ s). In practice multiple respiratory cycles will occur when collecting projection views over such a large angular range. In our future work we expect to include quasi-periodicity regularization terms to use projection views that span multiple respiratory cycles, but here we focus on the case of a single cycle for simplicity.

To illustrate the artifacts in reconstructed images caused by motion effects, we applied filtered backprojection reconstruction method to both the motion-free projections and the motion-corrupted projections (Fig. 2). Figure 2(d) shows blurring artifacts at the inner chest wall and degraded contrast of the mass between the lungs because of the inconsistent projection views caused by motion. Since our goal is to estimate respiratory motion rather than image reconstruction, we are not concerned about the streak artifacts present in both reconstructed images due to the small number of projection views.

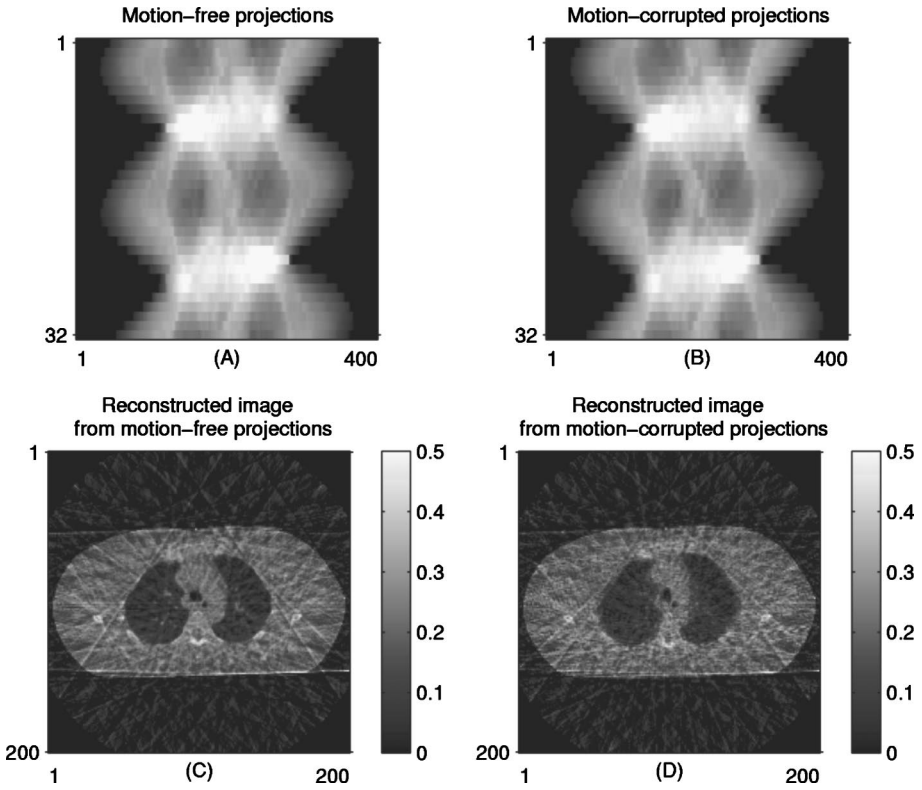

FIG. 2. Illustration of motion effects in image reconstruction. Motion-free projections (a) and the corresponding conventional fan-beam FBP reconstructed image (c). Motion-corrupted projections $\{\hat{\boldsymbol{g}}_m\}_{m=1}^{32}$ (b) and the corresponding reconstructed image (d).

For optimization, the deformation parameters were all initialized to be zero. We terminated the optimization algorithm when the absolute difference of the cost function value between the two most current iterations was less than a threshold. An important step for the estimation is to decide how to distribute the control knots. Generally, a finer control grid is favored for the motion that changes faster (less smoothly).[19] However, there is tradeoff. Coarse control grids may not be able to capture the real motion in detail; whereas overly fine control grids require more parameters, increase computation, and may overfit noise. We adjusted knot spacings manually, starting with a relatively coarse control grid, and then decreased the knot spacings until the optimizations with the two most recent control grids reached very similar results. This manual procedure finally arrived at a spatial control grid of 7×7 knots and a temporal grid of $K=5$ knots, with the knot spacings $\Delta_x=20$ pixels, $\Delta_y=20$ pixels, and $\Delta_t=8$ pixels. So the deformation model contains two sets of 245 parameters with each defining the deformation along x and y direction, respectively. We placed the knots in the image region where the support of each knot overlaps with the thorax. This placement helps prevent near zero-valued elements in the Hessian matrix, and also reduces computation relative to covering the entire square array. For regularization, we set $\alpha=5\times10^{-6}$ and used the second-order spatial and temporal roughness penalties with a typical row of the differencing matrix $\boldsymbol{C}$ in Eq. (12) having the form $(\ldots,0,-1,2,-1,0,\ldots)$.

B. Results

The Levenberg–Marquardt algorithm took about 15 iterations to converge. On a Pentium 2 GHz computer, each iteration consumed about 26 s without Hessian calculation and about 54 s otherwise. Fortunately the Hessian was updated only rarely. In fact here the computation of Hessian was needed only once at the beginning of the optimization. Figure 3 displays the transverse and longitudinal movement of a point located at the edge of the right lung. Figure 4 compares the motion vector of the image field at $t=2$ s. Both figures indicate good agreement between the estimates and the truth, even though there is model mismatch between the true deformation (14), (15) and our deformation model (7). The mean absolute error of the estimated motion for the whole image field was 0.23 mm and the maximum error was 1.94 mm. These errors are small compared to the synthetic motion, which has a maximum displacement of 15.14 mm and a mean absolute displacement of 1.11 mm. To examine the spatial distribution of errors, Fig. 5 shows the contours of the reference image superimposed upon an image of the difference between the estimated deformation and the true deformation (at time $t=2$ s). The relatively large errors tend to occur in image regions that lack structure, which may not have significant effect on generating deformed images based on the estimated motion. Figure 6 illustrates this viewpoint—only small discrepancies exist in the difference image [Fig. 6(c)] between the true and the estimated deformed images.

Since there was model mismatch between the B-spline motion model and the synthetic motion, we did B-spline least-squares fitting of the synthetic motion using the same control grid to investigate how much error would result from estimation alone. We found that the B-spline fitting rms error was 0.03 mm and the maximum absolute error 0.48 mm. So in this study model mismatch was a second source of error. Although using much finer control grids would reduce model

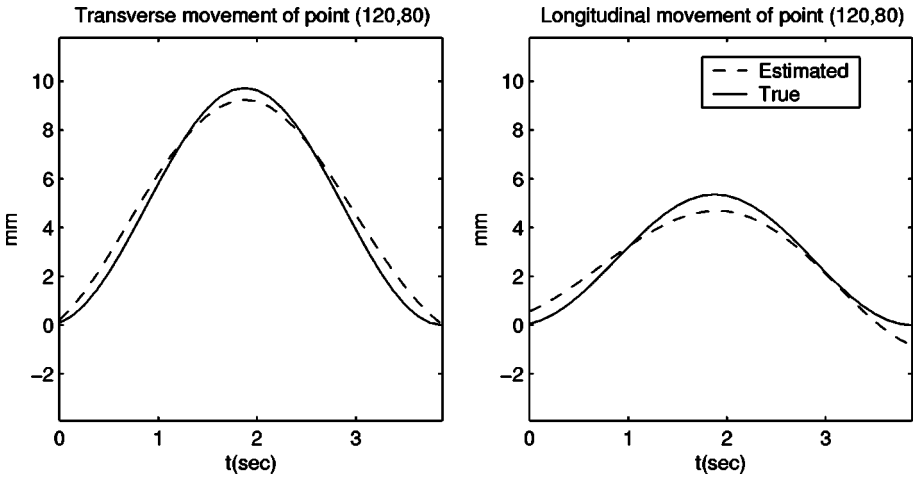

FIG. 3. Transverse (left panel) and longitudinal (right panel) movement for point (120, 80) located at the edge of the right lung.

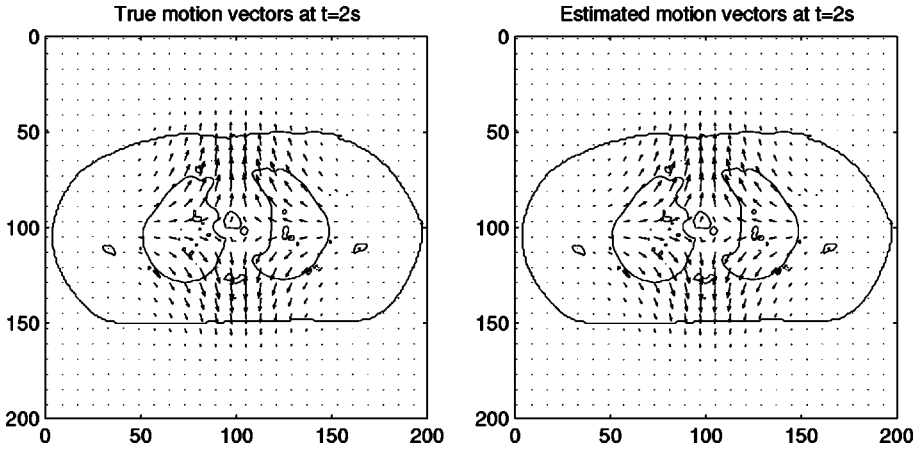

FIG. 4. True (left panel) and estimated (right panel) motion vectors at $t=2$ s (the time of maximum displacement).

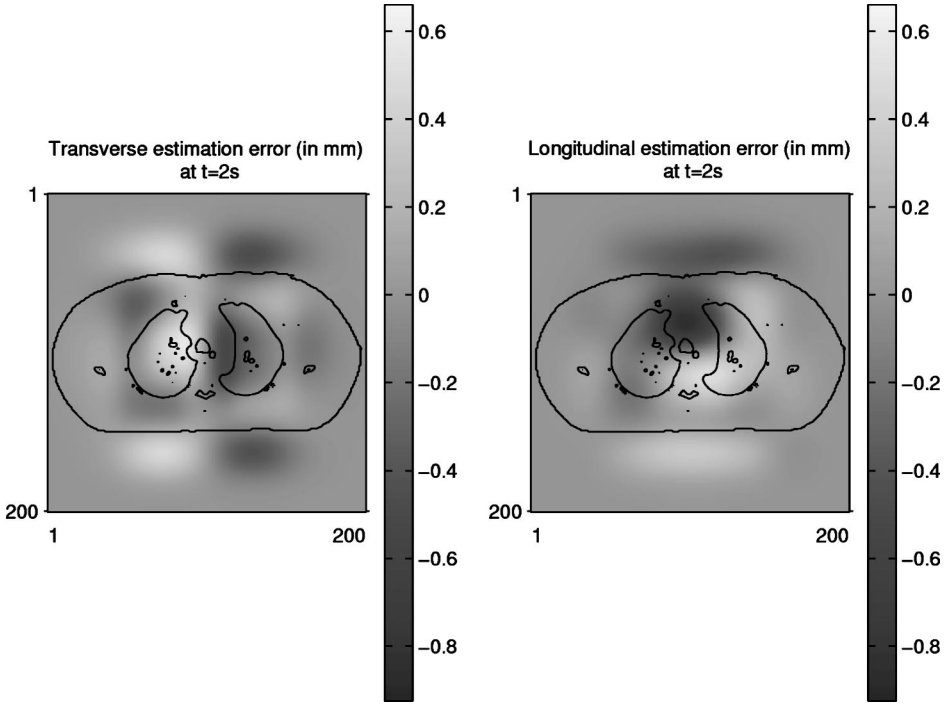

FIG. 5. Distribution of deformation estimation error, illustrated by superimposing the deformation estimation error image at $t=2$ s with the contour of the reference image. Transverse estimation error (left panel) and longitudinal estimation error (right panel).

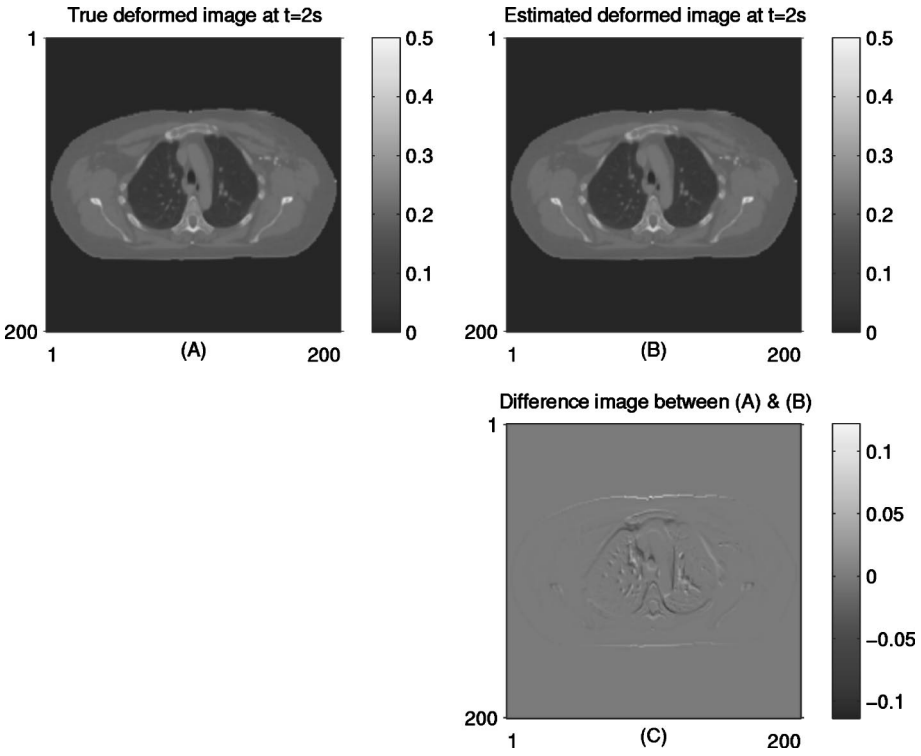

FIG. 6. True deformed image at $t=2$ s (top left panel), estimated deformed image at $t=2$ s (top right panel), and the difference image at $t=2$ s (bottom right panel).

mismatch error, more degrees of freedom associated with finer control grids would increase the number of local minima, degrading the performance of optimization algorithms.

V. DISCUSSION AND CONCLUSION

This paper described a method for estimating nonrigid motion from a motion-free reference image and a sequence of slowly rotating projection views of the moving object. Cubic B-spline functions were applied as the basis of our parametric temporal motion model. We used a regularized least-squares estimator of the motion parameters. For low-dose projection views, better performance may be achievable by applying the maximum-likelihood estimator for a Poisson statistic model. As proposed, the method does not require any external respiratory monitoring device. However, the approach could be generalized to use information from such devices, e.g., by adding appropriate terms to the cost functions.

Our initial investigation used a simulated fan-beam geometry. Since the theory described in Sec. III generalizes readily to 4D, the method should also be applicable to cone-beam projections of 3D thorax data and real respiratory motion, although the memory requirement and computation time will increase significantly. The parts that consume most of the computation time are the calculation of the gradient and Hessian in Eq. (10). Based on Eqs. (A1)–(A3), the complexity of the calculation of gradient is $O(MV)$, and that of the Hessian is $O(CMV)$, where V is the number of voxels, M the frame number of projection views, and C the number of knots of the deformation control grid. In the 4D case, computing the Hessian may require prohibitive memory and time, so a good approximation of the Hessian or alternate optimization algorithms not requiring a Hessian will be needed. We will conduct 4D studies in the near future, including experiments with real thorax data.

ACKNOWLEDGMENTS

This work is supported in part by NIH Grants Nos. P01 CA59827 and R01 CA60711.

APPENDIX: CALCULATION OF THE GRADIENT AND HESSIAN

We need to calculate the gradient and Hessian to implement the optimization algorithm described in Sec. III C. Explicit derivatives can be found using the chain rule.

There are two terms, $L(\boldsymbol{\theta})$ and $R(\boldsymbol{\theta})$, in the cost function $\psi(\boldsymbol{\theta})$. We first compute the partial derivatives of $L(\boldsymbol{\theta})$.

Let us introduce $\hat{f}_m = \mathcal{W}(\boldsymbol{\theta}, t_m) f_{\text{ref}}$ to denote the estimated deformed object at time t_m. Let $\theta_k^p (p=1,\ldots,d)$ be the coefficient of the kth knot in the pth dimension. Starting from Eq. (10), we obtain the first partial derivatives as follows:

$$\frac{\partial L}{\partial \theta_k^p} = -\sum_{m=1}^{M} \left\langle (\hat{\boldsymbol{g}}_m - \mathcal{A}_{\phi_m} \hat{f}_m), \mathcal{A}_{\phi_m} \frac{\partial \hat{f}_m}{\partial \theta_k^p} \right\rangle = -\sum_{m=1}^{M} \left\langle \mathcal{A}'_{\phi_m} (\hat{\boldsymbol{g}}_m - \mathcal{A}_{\phi_m} \hat{f}_m), \frac{\partial \hat{f}_m}{\partial \theta_k^p} \right\rangle, \tag{A1}$$

and the second partial derivatives

$$\frac{\partial^2 L}{\partial \theta_k^p \partial \theta_j^q} = \sum_{m=1}^{M} \left(\left\langle \mathcal{A}_{\phi_m} \frac{\partial \hat{f}_m}{\partial \theta_k^p}, \mathcal{A}_{\phi_m} \frac{\partial \hat{f}_m}{\partial \theta_j^q} \right\rangle - \left\langle (\hat{\boldsymbol{g}}_m - \mathcal{A}_{\phi_m} \hat{f}_m), \mathcal{A}_{\phi_m} \frac{\partial \hat{f}_m}{\partial \theta_k^p \partial \theta_j^q} \right\rangle \right), \tag{A2}$$

where $\langle a, b \rangle$ denotes the inner product of array a and b. We disregard the second derivative term in expression (A3) because are usual relatively small. Furthermore, $(\hat{\boldsymbol{g}}_m - \mathcal{A}_{\phi_m} \hat{f}_m)$ should be the random measurement errors and consequently the second derivative terms tend to cancel out after summation.[21] Therefore,

$$\frac{\partial^2 L}{\partial \theta_k^p \partial \theta_j^q} \approx \sum_{m=1}^{M} \left\langle \mathcal{A}_{\phi_m} \frac{\partial \hat{f}_m}{\partial \theta_k^p}, \mathcal{A}_{\phi_m} \frac{\partial \hat{f}_m}{\partial \theta_j^q} \right\rangle. \tag{A3}$$

We need to determine $\partial \hat{f}_m / \partial \theta_k^p$ to complete the calculation of Eqs. (A1) and (A3). We use the chain rule again. Using Eqs. (6) and (7), the expression for $\hat{f}_m(\boldsymbol{x})$ is

$$\hat{f}_m(\boldsymbol{x}) = \mathcal{W}(\boldsymbol{\theta}, t_m) f_{\text{ref}}(\boldsymbol{x}) = f_{\text{ref}} \left(\underbrace{\boldsymbol{x} + \sum_r \sum_{\boldsymbol{i}} \boldsymbol{\theta}_{r,\boldsymbol{i}} b\left(\frac{t_m}{\Delta_t} - \tau \right) \beta\left(\frac{\boldsymbol{x}}{\Delta_x} - \boldsymbol{i} \right)}_{\boldsymbol{x}'} \right). \tag{A4}$$

Let $\boldsymbol{\theta}_k$ correspond to the coefficient of the control knot located at $(\hat{\tau}, \hat{\boldsymbol{i}})$, where $\hat{\tau}$ is the temporal position and $\hat{\boldsymbol{i}}$ the spatial position, then from Eq. (A4)

$$\frac{\partial \hat{f}_m}{\partial \theta_k^p}(\boldsymbol{x}) = \frac{\partial \hat{f}_m}{\partial \theta_{\hat{\tau}, \hat{\boldsymbol{i}}}}(\boldsymbol{x}) = [\nabla f_{\text{ref}}|_{\boldsymbol{x}=\boldsymbol{x}'}]_p b\left(\frac{t_m}{\Delta_t} - \hat{\tau} \right) \beta\left(\frac{\boldsymbol{x}}{\Delta_x} - \hat{\boldsymbol{i}} \right), \tag{A5}$$

where $[\nabla f_{\text{ref}}|_{\boldsymbol{x}=\boldsymbol{x}'}]_p$, the pth element of the spatial gradient of the reference image evaluated at $\boldsymbol{x}'$, can be calculated from the interpolation model (9) as follows:

$$[\nabla f_{\text{ref}}|_{\boldsymbol{x}=\boldsymbol{x}'}]_p = \sum_{\boldsymbol{r}} c_{\boldsymbol{r}} [\nabla \beta(\boldsymbol{x}' - \boldsymbol{r})]_p, \tag{A6}$$

$$[\nabla \beta(\boldsymbol{x}' - \boldsymbol{r})]_p = \dot{\beta}(x_p) \prod_{t=1, l \neq p}^{d} \beta(x_l). \tag{A7}$$

The remaining calculation is that of the derivatives of $R(\boldsymbol{\theta})$. These can be found easily from Eq. (12) as follows:

$$\nabla R(\boldsymbol{\theta}) = C' C \boldsymbol{\theta}, \tag{A8}$$

$$\nabla^2 R(\boldsymbol{\theta}) = C' C. \tag{A9}$$

[a]Electronic mail: rzeng@eecs.umich.edu
[b]Electronic mail: fessler@eecs.umich.edu
[c]Electronic mail: jbalter@med.umich.edu

[1]J. W. Wong, M. B. Sharpe, D. A. Jaffray, V. R. Kini, J. M. Robertson, J. S. Stromberg, and A. A. Martinez, "The use of active breathing control (ABC) to reduce margin for breathing motion," Int. J. Radiat. Oncol., Biol., Phys. **44**, 911–919 (1999).
[2]V. M. Remouchamps, F. A. Vicini, M. B. Sharpe, L. L. Kestin, A. A. Martinez, and J. W. Wong, "Significant reductions in heart and lung doses using deep inspiration breath hold with active breathing control and intensity-modulated radiation therapy for patients treated with locoregional breast irradiation," Int. J. Radiat. Oncol., Biol., Phys. **55**, 392–406 (2003).
[3]P. J. Keall, V. R. Kini, S. S. Vedam, and R. Mohan, "Motion adaptive x-ray therapy: A feasibility study," Phys. Med. Biol. **46**, 1–10 (2001).
[4]H. Hu, "Multi-slice helical CT: Scan and reconstruction," Med. Phys. **26**, 5–18 (1999).
[5]R. A. Robb, E. A. Hoffman, L. J. Sinak, L. D. Harris, and E. L. Ritman, "High-speed three-dimensional x-ray computed tomography: The dynamic spatial reconstructor," Proc. IEEE **71**, 308–319 (1983).
[6]C. J. Ritchie, C. R. Crawford, W. Stanford, H. Anno, and Y. Kim, "Minimum scan speeds for suppression of motion artifacts in CT," Radiology **185**, 37–42 (1992).
[7]S. Roux, L. Desbat, A. Koenig, and P. Grangeat, "Exact reconstruction in 2D dynamic CT: Compensation of time-dependent affine deformations," Phys. Med. Biol. **49**, 2169–2182 (2004).
[8]C. R. Crawford, K. F. King, C. J. Ritchie, and J. D. Godwin, "Respiratory compensation in projection imaging using a magnification and displacement model," IEEE Trans. Med. Imaging **15**, 327–332 (1996).
[9]G. Wang and M. W. Vannier, "Preliminary study on helical CT algorithms for patient motion estimation and compression," IEEE Trans. Med. Imaging **14**, 205 (1995).
[10]W. Lu and T. R. Mackie, "Tomographic motion detection and correction directly in sinogram space," Phys. Med. Biol. **47**, 1267–1284 (2002).
[11]A. C. Dhanantwari, S. Stergiopoulos, and I. Iakovidis, "Correcting organ motion artifacts in x-ray CT medical imaging systems by adaptive processing. I. Theory," Med. Phys. **28**, 1562–1576 (2001).
[12]T. Flohr, B. Ohnesorge, H. Bruder, K. Stierstorfer, J. Simon, C. Suess, and S. Schaller, "Image reconstruction and performance evaluation for ECG-gated spiral scanning with a 16-slice CT system," Med. Phys. **30**, 2650–2652 (2003).
[13]D. A. Low, M. Nystrom, E. Kalinin, P. Parikh, J. F. Dempsey, J. D. Bradley, S. Mutic, S. H. Wahab, T. Islam, G. Christensen, D. G. Politte, and B. R. Whiting, "A method for the reconstruction of four-dimensional synchronized CT scans acquired during free breathing," Med. Phys. **30**, 1254–1263 (2003).
[14]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).
[15]I. El Naqa, D. A. Low, J. O. Deasy, A. Amini, P. Parikh, and M. Nystrom, "Automated breathing motion tracking for 4D computed tomography," in Proc. IEEE Nuc. Sci. Symp. Med. Im. Conf., 2003.
[16]L. Zijp, J-J. Sonke, and M. Herk, "Extraction of the respiratory signal from sequential thorax cone-beam X-ray images," in *Proc. Intl. Congr. Comp. in Radiotherapy*, 2004.
[17]A. Macovski, *Medical Imaging Systems* (Prentice-Hall, Englewood Cliffs, NJ, 1983).
[18]M. Unser, "Splines: A perfect fit for signal and image processing," Signal Process. **16**, 22–38 (1999).
[19]J. Kybic and M. Unser, "Fast parametric elastic image registration," IEEE Trans. Image Process. **12**, 1427–1442 (2003).
[20]D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: Application to breast MR images," IEEE Trans. Med. Imaging **18**, 712–721 (1999).
[21]W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C* (Cambridge University Press, New York, 1988).
[22]B. R. Whiting, L. J. Montagnino, and D. G. Politte, "Modeling x-ray computed tomography sinograms," 2001, Med. Phys. (submitted).