

# Estimating 3-D Respiratory Motion From Orbiting Views by Tomographic Image Registration

Rongping Zeng\*, *Student Member, IEEE*, Jeffrey A. Fessler, *Fellow, IEEE*, and James M. Balter

**Abstract**—Respiratory motion remains a significant source of errors in treatment planning for the thorax and upper abdomen. Recently, we proposed a method to estimate two-dimensional (2-D) object motion from a sequence of slowly rotating X-ray projection views, which we called deformation from orbiting views (DOVs). In this method, we model the motion as a time varying deformation of a static prior of the anatomy. We then optimize the parameters of the motion model by maximizing the similarity between the modeled and actual projection views. This paper extends the method to full three-dimensional (3-D) motion and cone-beam projection views. We address several practical issues for using a cone-beam computed tomography (CBCT) scanner that is integrated in a radiotherapy system, such as the effects of Compton scatter and the limited gantry rotation for one breathing cycle. We also present simulation and phantom results to illustrate the performance of this method.

**Index Terms**—Aperiodicity penalty, *B*-spline, cone-beam projection views, respiratory motion.

## I. INTRODUCTION

GEOMETRIC uncertainties caused by respiratory motion complicate precision in radiotherapy treatment in the thorax and upper abdomen. Although active breath control (ABC) [1] can reduce this problem, it is very uncomfortable for patients to hold their breath, especially for those with lung cancer. Treating patients under free breathing conditions can be accomplished by constructing an internal target volume, in which the treatment plan is based on a larger margin encompassing the volume the tumor may reach during breathing [2]. Another method is respiratory-gated treatment [3]. In gated methods, treatment plan is based on the tumor position at a certain respiratory phase. During the treatment, the X-ray beam is turned on in the chosen phase and turned off otherwise. No matter what kind of methods are used, knowledge of how the patients' anatomy moves during respiration is required for designing a treatment plan.

To obtain a complete picture of the patients' anatomy at all times during breathing, intensive work has been dedicated to four-dimensional (4-D) CT imaging [4]–[8]. Here, 4-D refers

to a three-dimensional (3-D) spatial field plus a one-dimensional (1-D) respiratory index. Most current 4-D computed tomography (CT) imaging techniques use conventional scanners [4]–[6]. These methods often acquire multiple two-dimensional (2-D) slices at each table position, sort these slices into several respiratory phase bins, and then stack those slices that are within the same phase bins to form a 4-D model. Due to insufficient longitudinal coverage to image an entire volume during one breathing cycle, an assumption is made that a reproducible relationship exists between the internal motion and the “phase” of some external monitoring index. Inaccurate sorting due to imperfect correlation often leads to tissue discontinuity artifacts in the reconstructed CT volumes, especially at mid-inspiration or mid-expiration phases.

It is also feasible to image the thorax using cone-beam computed tomography (CBCT) scanners that are integrated on the gantry of the linear accelerator (linac) [7], [8]. Such scanners rotate slowly, with about 1 minute per full rotation. Therefore the collected projection views over a 360° scan span several breathing cycles. One way to obtain a motion-reduced thorax CT volume is to use only the projection views in the same breathing phase for image reconstruction, as described in [7]. However, assumptions of motion reproducibility remain a limitation and insufficient numbers of projections per breathing phase may result in severe artifacts in the reconstructed images. Another way is to use all the projection views and implement motion-compensated reconstruction. For example, Li *et al.* [8] obtained a motion model by coregistering the 4-D treatment planning CT images of the same patient and then incorporated the patient's motion model into the image reconstruction process. However, a motion model derived from 4-D CT data may not match the motion during a subsequent CBCT scan for two reasons. One is the possible changes of the patient's motion pattern through time, the other one is inaccurate spatial and temporal alignment between the 4-D CT data and the CBCT data.

We propose a different way to study the moving anatomy during breathing. The previous 4-D CT methods all attempt to directly reconstruct a sequence of static 3-D CTs through time. However, our method attempts to estimate the movement of patients' anatomy through time. With the estimated motion, a 3-D picture of the anatomy at any selected time point can be generated by deforming a static prior according to the estimated movement at that time. Our method utilizes an on-board imager (OBI)-based CBCT scanner. Although this type of scanner rotates slowly, it provides a large volume coverage (e.g., a 30 × 40 cm<sup>2</sup> detector) and a high temporal sampling rate (up to 15 projection views per second). Therefore, we have explored estimating respiratory motion from a sequence of cone-beam pro-

Manuscript received July 25, 2006; revised November 2, 2006. This work is supported in part by the National Institutes of Health under Grant P01 CA59827-06A1. Asterisk indicates corresponding author.

\*R. Zeng is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: rzeng@umich.edu).

J. A. Fessler is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: fessler@umich.edu).

J. M. Balter is with the Department of Radiation Oncology, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: jbalter@med.umich.edu).

Digital Object Identifier 10.1109/TMI.2006.889719

jection views acquired during free breathing, assuming that we also have available a static prior model of the anatomy, such as a breathhold planning CT. The basic idea of this approach is to model the motion as a time varying deformation of the prior and to estimate the motion parameters by maximizing the similarity between the modeled and measured projection views. We call this method deformation from orbiting views. Our method does not bin data over the respiratory cycle, hence, it reduces the tissue discontinuity artifacts in the aforementioned phase-sorting based 4-D CT imaging methods. Unlike the motion compensation method proposed in [8], our motion model is estimated from the CBCT data itself, hence, the possible inconsistencies of 4-D CT data do not exist.

We have described the general framework of deformation from orbiting views (DOVs) previously and presented preliminary 2-D simulation results [9]. Three-dimensional motion estimation from cone-beam projection views brings in several practical issues that were not addressed in our previous work, such as the limited gantry range for one breathing cycle and Compton scatter contamination in cone beam systems. For recently developed CBCT systems for radiotherapy, the angular span for one breathing cycle is around  $20^\circ$  to  $40^\circ$ . The drawback of this small angular range is that the projection views for one breathing cycle may be less informative about the motion along certain directions. To help overcome this limitation, we add pseudoperiodicity regularization to the estimator.

This paper starts with a description of the theory of DOV, including how the estimator is formed, how the penalty terms are designed and what strategies are used for this optimization problem. Simulation and experiment results are then presented and discussed.

## II. THEORY

The basic idea of DOV is to estimate respiratory motion from sequential X-ray cone beam projection views. This section is focused on the estimator for DOV. The estimator includes two kinds of terms, similarity terms and penalty terms. We first briefly review the relationship between the measured datasets, introduced in [9], and propose two possible similarity metrics. We then describe the penalty terms included in the estimator: a commonly used motion roughness penalty and an aperiodicity penalty specially designed for DOV. Afterward, the optimization strategies used for finding the estimates are given.

### A. The System Model

The proposed motion estimation method uses two sets of data. One is a reference thorax CT volume obtained from a conventional fast CT scanner under breathhold conditions, denoted  $f_{\text{ref}}(\mathbf{x}), \mathbf{x} \in \mathbb{R}^3$ . The other is a sequence of projection views of the same patient acquired at treatment time using a slowly rotating cone-beam system (1 min per rotation), denoted  $\mathbf{Y}_m, m = 1, \dots, M$  ( $M$  is the number of projection views). We establish the relationship between the two data sets  $f_{\text{ref}}(\mathbf{x})$  and  $\mathbf{Y}$  in this section.

We need to first address one concern about the slowly rotating cone-beam systems. Although the cone-beam scanners rotate slowly, the acquisition time of each projection view is short. For

example, recently developed systems can acquire 15 frames per second, which indicates that the imaging time for each frame is less than 0.067 s. We, therefore, assume that the respiratory motion is negligible within each single projection view.

Let the motion during the scan be denoted as  $\mathcal{T}_\theta(\mathbf{x}; t)$ , a time-dependent deformation controlled by parameters  $\theta$ . Since the projection views and the reference volume are all from the same patient, the ideal projection views  $\mathbf{g}_m$  can be related to  $f_{\text{ref}}$  in terms of the CT imaging principle through the motion as follows:

$$\mathbf{g}_m = \mathcal{A}_{\phi_m} f_{t_m} \quad (1)$$

$$f_{t_m}(\mathbf{x}) = f_{\text{ref}}(\mathcal{T}_\theta(\mathbf{x}; t_m)) \quad (2)$$

where  $\mathcal{A}_{\phi_m}$  denotes the X-ray projection [10] operator for projection angle  $\phi_m$ , and  $f_{t_m}$  is the deformed volume at time  $t_m$ . Combining (1) and (2), we obtain

$$\mathbf{g}_m = \mathcal{A}_{\phi_m} f_{\text{ref}}(\mathcal{T}_\theta(\mathbf{x}; t_m)). \quad (3)$$

However, in practice the projection views  $\mathbf{g}_m$  are estimated from the measured photon counts  $\mathbf{Y}_m$ , which are always degraded by noise, dominated by the Poisson effect [11]. For simplicity, we assume a monoenergetic model to describe the relationship between  $\mathbf{g}_m$  and  $\mathbf{Y}_m$  as follows:

$$Y_{m,n} \sim \text{Poisson}(I_{m,n} e^{-g_{m,n}} + S_{m,n}) \quad (4)$$

where  $I_{m,n}$  is a constant related to the incident X-ray intensity,  $S_{m,n}$  denotes the scatter contribution to  $Y_{m,n}$  and  $n$  is the detector element index. Then the projection views used for DOV can be estimated from  $\mathbf{Y}_m$  as follows:

$$\hat{g}_{m,n} = \log \left( \frac{I_{m,n}}{Y_{m,n} - \hat{S}_{m,n}} \right). \quad (5)$$

In (5),  $I_{m,n}$  can be measured by an air scan and  $\hat{S}_{m,n}$  is an estimate of the scatter contribution. There are a few popular ways to estimate scatter. One is to model the scatter as the convolution of a function with the primary counts. The function could be approximated by an exponential or Gaussian kernel [12]. Another way is to measure the scatter effect using a beam stop array [13]. One may also estimate the scatter by using the Monto-Carlo simulation method [14]. The DOV method can use any such scatter estimates.

We need to choose a deformation model to complete (3). Theoretically  $\mathcal{T}_\theta(\mathbf{x}; t)$  can be any suitable deformation model. We adopted a cubic  $B$ -spline based motion model as follows:

$$\mathcal{T}_\theta(\mathbf{x}; t) = \mathbf{x} + \sum_j \sum_i \theta_{j,i} \beta \left( \frac{t - \tau_j}{\Delta_t} \right) \beta \left( \frac{\mathbf{x} - \mathbf{x}_i}{\Delta_x} \right) \quad (6)$$

where  $\beta(\cdot)$  is the cubic  $B$ -spline function and  $\beta(\mathbf{x})$  the tensor product of cubic  $B$ -spline functions,  $\tau_j$  and  $\mathbf{x}_i$  the spatial and temporal knot locations,  $\Delta_x$  and  $\Delta_t$  control the width of the

spatial and temporal basis functions, respectively, and  $\theta$  the knot coefficients. There are two advantages of using a cubic  $B$ -spline model. One is that the small support of the cubic  $B$ -spline function eases the computation and optimization. The other is that the density of a  $B$ -spline control grid can be locally adjusted according to the characteristics of the signal to be fitted. For example, one can place more knots at regions where the signal changes faster and less knots otherwise. Based on the motion model (6), the estimation goal is to find the motion parameters  $\theta$  from the projection views  $\hat{g}_m$  and  $f_{\text{ref}}$ .

In terms of the relationship between  $g_m$  and  $f_{\text{ref}}$  described in (3), we construct an estimator of  $\theta$  as follows:

$$\hat{\theta} = \arg \min_{\theta} (D(\{\hat{g}_m\}, \{p_m(\theta)\}) + \beta_s R_s(\theta) + \beta_t R_t(\theta)) \quad (7)$$

where  $\{p_m(\theta)\} = \{\mathcal{A}_{\phi_m, f_{\text{ref}}}(\mathcal{T}_{\theta}(\mathbf{x}; t_m))\}$  is the modelled projection views of the warped reference volume,  $D(\cdot, \cdot)$  is a data fidelity term,  $R_s(\theta)$  is a motion roughness penalty term,  $R_t(\theta)$  is a temporal motion aperiodicity penalty term, and  $\beta_s$  and  $\beta_t$  are scalars that control the tradeoff between the three terms.

To sum up, the goal of DOV is to find the motion, a sequence of deformations through time, to match the projection views of the warped reference image to the measured projection views. It is essentially a registration problem. But unlike the traditional image registration problem, DOV works with the projection domain data and is thus more challenging. For example, a 3-D image registration task is to find a 3-D deformation field from two 3-D images, while DOV is tasked to find  $k$  3-D deformation fields from one 3-D image and  $n$  2-D projection views, where  $k \geq n$ . Evidently, DOV attempts to estimate more unknowns from less information. Thus, regularization is essential.

### B. Data Fidelity Term

This section elaborates on the data fidelity term in (7). We investigated the following two intensity-based metrics: sum of squared differences (SSD) and a correlation-based metric. Consequently, we call the estimators using those two metrics least-square (LS) estimator and correlation-based estimator.

• *SSD*: The expression of SSD is as follows:

$$\text{SSD}(\{\hat{g}_m\}, \{p_m(\theta)\}) = \frac{1}{2} \sum_{m=1}^M \|\hat{g}_m - p_m(\theta)\|^2. \quad (8)$$

This metric works well for registration of images from the same modality. This rule applies to DOV as well. To yield good estimates using this approach, the X-ray energies should be the same for imaging the static CT and for acquiring the cone-beam projection views. In addition to this, extra effort may be needed to correct the imaging artifacts such as Compton scatter effects, beam hardening effects, and presence of the radiotherapy table in the projection views (not present in the prior CT). The SSD represents “best case” performance when such effects are corrected. However, in practice it may be difficult to correct for all such effects, so the following correlation-base metric may be preferable.

• *Correlation-based metric*: In the correlation-based estimator, we used the negative-logarithm of the correlation coef-

ficient (NLCC) as the data fidelity metric. The expression is as follows:

$$\begin{aligned} \text{NLCC}(\{\hat{g}_m\}, \{p_m(\theta)\}) &= \sum_{m=1}^M -\ln(\text{cor}(\hat{g}_m, p_m(\theta))) \\ &= \sum_{m=1}^M \left( -\ln \sum_{n=1}^N (g_{m,n} - \bar{g}_m)(p_{m,n}(\theta) - \bar{p}_m(\theta)) \right. \\ &\quad \left. + \frac{1}{2} \ln \sum_{n=1}^N (g_{m,n} - \bar{g}_m)^2 \right. \\ &\quad \left. + \frac{1}{2} \ln \sum_{n=1}^N (p_{m,n}(\theta) - \bar{p}_m(\theta))^2 \right) \end{aligned} \quad (9)$$

where  $\bar{g}_m$  is the mean value of  $\hat{g}_m$  and  $\bar{p}_m(\theta)$  is the mean value of  $p_m(\theta)$ . In this data fidelity term, we use a logarithm to separate the numerator and denominators in the expression of the correlation coefficient, which simplifies the calculation of its gradient. Because the logarithm function is increasing, the logarithm step does not change the monotonicity of the correlation coefficient function. We negate the logarithm correlation coefficient because we are minimizing the cost function in the estimator (7).

Correlation-based metrics are suitable when the intensities of the images are linearly related. In X-ray imaging, the attenuation is larger when the X-ray energy is stronger. So we may expect the correlation-based estimator can perform well even if the energy spectra used for the conventional CT scanner and the cone-beam CT scanner are not identical.

### C. Penalty Design

This section elaborates on the penalty terms in (7).

• *Spatial and temporal motion roughness penalty*: The motion roughness penalty discourages rapidly changing breathing motion estimates that would be unrealistic. The spatial motion roughness can be measured qualitatively by the squared differences between the displacements of adjacent voxels, and the temporal motion roughness by the squared differences between the displacements of the same voxel at adjacent time points. To simplify this term, we replaced the displacement differences by the motion parameter differences. With this simplification, this term can be expressed mathematically as

$$R(\theta) = \frac{1}{2} \|\mathcal{C}\theta\|^2 \quad (10)$$

where  $\mathcal{C}$  is a differencing matrix, with a typical row having the form  $(\dots, 0, -1, 1, 0, \dots)$  for the first-order roughness penalty and  $(\dots, 0, -1, 2, -1, 0, \dots)$  for the second-order roughness penalty. It can be shown that the second-order differencing matrix has a very similar high-pass structure to that for penalizing displacements under a cubic  $B$ -spline deformation model. By including this penalty term, the optimization is guided toward a solution with a smoother breathing motion.

• *Aperiodicity penalty*: The aperiodicity penalty encourages similarity between deformation estimates that correspond to

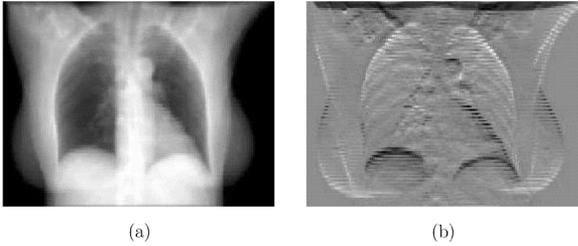


Fig. 1. (a) X-ray projection image and (b) its CC gradient image.

similar respiratory phases. This helps ensure temporal regularity. If the temporal knots are evenly spaced in each breathing period and each breathing period contains the same number of knots, then the deformation similarity can be quantified by the closeness of the coefficient values of knots that are located at similar respiratory phases, for the sake of simplicity. For example, in Fig. 5 (solid line) there are four breathing cycles, each containing five locally evenly spaced knots. Thus, every fifth knot corresponds to a similar phase, such as the knot group (1, 6, 11, 16), the knot group (2, 7, 12, 17), and so on. Based on this design, the aperiodicity penalty term also takes the form of (10), with the matrix  $\mathbf{C}$  having a typical row of  $(\dots, 0, -1, 0, \dots, 0, 1, 0, \dots)$ . The number of zeros between  $-1$  and  $1$  is related to the number of knots placed in each breathing period.

We add this penalty term to help overcome the limited gantry range for each breathing cycle. Current radiotherapy systems can rotate  $6^\circ$  per s, spanning around  $20^\circ$ – $40^\circ$  in one breathing cycle. Therefore, the measured projection views in one breathing cycle may poorly reflect the motion along certain directions. For example, if the gantry starts from  $0^\circ$  (anterior view), then the projection views in the first breathing cycle are less informative about the anterior–posterior (AP) motion, leading to poorer motion estimation accuracy along the AP direction in the absence of any other prior information. However, the projection views taken over the next breathing cycle can better capture the motion along AP direction. By using an aperiodicity penalty term, motion information contained in the adjacent breathing cycles can be “shared” to help compensate for the angular limitation.

We need a respiratory marker to determine the correspondences between the temporal knots for the aperiodicity penalty. We adopted and simplified the respiratory signal extraction method presented by Zijp’s [15]. The basic idea is to capture the superior–inferior (SI) transition of the diaphragm in the collected projection views. The method uses the following four steps.

*Step 1:* We applied a gradient filter (e.g.,  $h = [-1, 1]$ ) to each 2-D projection image along the Cranial-Caudal (CC) direction. This step is to emphasize the diaphragm-like transition feature in each projection image (Fig. 1).

*Step 2:* We took the absolute value of each gradient image then projected onto the CC axis (Fig. 2). The “image” formed by combining all the 1-D projections clearly shows some breathing pattern near the diaphragm region, while in the other regions there is no obvious intensity contrast (Fig. 3).

*Step 3:* The centroid of each 1-D projection was calculated and ordered in time. The formula for calculating the centroid of

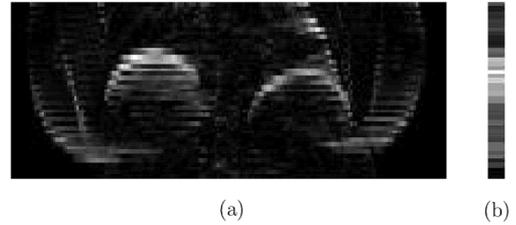


Fig. 2. (a) Absolute value of the gradient image and (b) its axial projection.

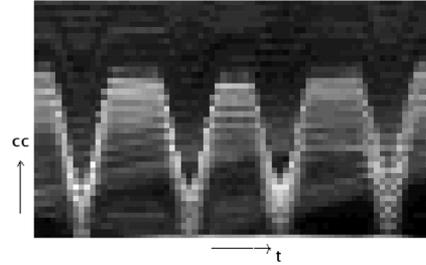


Fig. 3. Image formed by combining the 1-D axial projections. Each column corresponds to a single 1-D projection.

a 1-D signal  $s_n, n = 1, \dots, N$ , is

$$\text{centroid} = \frac{\sum_{n=1}^N n s_n}{\sum_{n=1}^N s_n}.$$

*Step 4:* The centroid signal was normalized and then smoothed by using a simple moving average filter.

As shown in Fig. 5, the estimated respiratory signal (dashed line) presents similar peak and valley patterns as that of the true respiratory signal (solid line). An advantage of this projection-view based method is that the resulting signal is related to internal anatomy positions, unlike external monitoring methods. We use this signal to decide the phase correspondence between temporal knots for calculating the aperiodicity penalty term. This is its only use here. Since this signal is not extremely important for the design of our motion model, a rough estimation of the breathing signal is sufficient for DOV.

#### D. Optimization Strategies

Because there is no analytical solution for (7), we used iterative methods to search for  $\hat{\theta}$ . In our previous work we used the Levenberg–Marquardt (LM) algorithm [16]. Although LM offers fast convergence, it is impractical for 3-D clinical data due to the computation of a large-sized Hessian. We then experimented with the Quasi-Newton (QN) algorithm, which approximates the Hessian by updating a preconditioning matrix. Based on our experiments, the approximation was not accurate enough to guide the optimization toward a correct direction for this optimization problem. We found that the preconditioned conjugate gradient (PCG) algorithm worked better than the other two algorithms.

The PCG algorithm does not use the gradient vector directly as its search direction. It modifies the gradient search directions so that the current search direction is conjugate to all the previous search directions. This modification ensures a more efficient search over the parameter space and hence converges

faster than the simple gradient descent algorithm. The updating scheme for each iteration  $n$  includes the following steps:

$$\begin{aligned}
 \mathbf{g}^{(n)} &= \nabla \psi(\boldsymbol{\theta}^{(n)}) \quad (\text{gradient}) \\
 \mathbf{p}^{(n)} &= \mathbf{P}\mathbf{g}^{(n)} \quad (\text{precondition}) \\
 \gamma_n &= \begin{cases} 0, & n = 0 \\ \frac{\text{real}(\langle \mathbf{p}^{(n)}, \mathbf{g}^{(n)} - \mathbf{g}^{(n-1)} \rangle)}{\text{real}(\langle \mathbf{p}^{(n-1)}, \mathbf{g}^{(n-1)} \rangle)}, & n > 0 \end{cases} \quad (11) \\
 \mathbf{d}^{(n)} &= -\mathbf{p}^{(n)} + \gamma_n \mathbf{d}^{(n-1)} \quad (\text{search direction}) \\
 \alpha_n &= \arg \min_{\alpha \in \mathbb{R}} \psi(\boldsymbol{\theta}^{(n)} + \alpha \mathbf{d}^{(n)}) \quad (\text{stepsize}) \\
 \boldsymbol{\theta}^{(n+1)} &= \boldsymbol{\theta}^{(n)} + \alpha_n \mathbf{d}^{(n)} \quad (\text{update}). \quad (12)
 \end{aligned}$$

We set  $\mathbf{P} = \mathbf{I}$ , which is actually the unpreconditioned case. The step size  $\alpha_n$  is often solved by line search. To save computation time, we used only one iteration of the Newton update to find a suboptimal step size  $\hat{\alpha}_n$  as follows:

$$\hat{\alpha}_n = \alpha_0 - \frac{\dot{\psi}(\alpha_0)}{\ddot{\psi}(\alpha_0)} \quad (13)$$

where the initial value  $\alpha_0$  is set to be zero to simplify the calculation. The gradient  $\mathbf{g}^{(n)}$ , the first derivative  $\dot{\psi}(\alpha_0)$  and the second derivative  $\ddot{\psi}(\alpha_0)$  can be found from (7) using the chain rule.

To accelerate the optimization procedure and to avoid local minima, we also applied a multiresolution technique [17].

### III. SIMULATION

This section presents our simulation results. The simulated datasets were generated based on several real clinical planning CTs and the geometry of a slowly rotating CBCT system, and thus should reflect sufficiently realistic conditions to illustrate the performance of this method. Furthermore, in the simulations absolute truth is known, permitting quantitative evaluation.

#### A. Simulation Setup

1) *Data Setup*: This section describes how we generated sequential cone-beam projection views of a moving CT volume using three breathhold treatment CT volumes of the same patient at different breathing phases (0%, 20%, 60% vital capacity above tidal exhale).

We selected the thorax CT at the end of exhale (0%) as our reference volume (Fig. 4), with  $192 \times 160 \times 60$  voxels and a voxel size of  $2 \times 2 \times 5 \text{ mm}^3$ . We then generated 70 cone-beam projection views of the warped reference volumes over a  $180^\circ$  rotation. (The warping process is described in the next paragraph.) The simulated cone-beam system had a flat-panel detector of  $180 \times 200$  elements of size  $4 \times 4 \text{ mm}^2$ . The source to isocenter distance and the isocenter to detector distance were 1000 mm and 500 mm, respectively. The gantry rotated  $6^\circ/\text{s}$  and spanned  $180^\circ$  over the four breathing cycles. We used a distance-driven method [18] to calculate the projection views. To simulate realistic projection views, scatter and Poisson noise were added according to the statistical property of projection views as described in (4). We first converted the projection views from attenuation to primary photon counts. The incident intensity  $I_{m,n}$  in (4) used for this conversion is  $10^6$  counts per ray [19]. We then applied a convolution method to generate the scatter counts, in

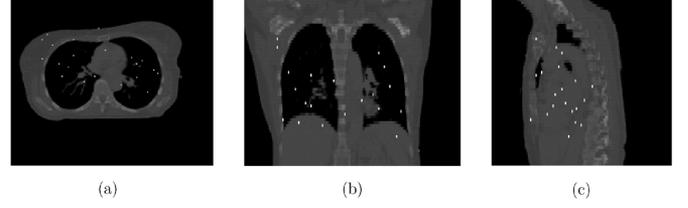


Fig. 4. Three views of the reference thorax CT volume. Points in the images are the projection positions on the three planes of the voxels randomly selected for accuracy plots. (a) Axial view. (b) Coronal view. (c) Sagittal view.

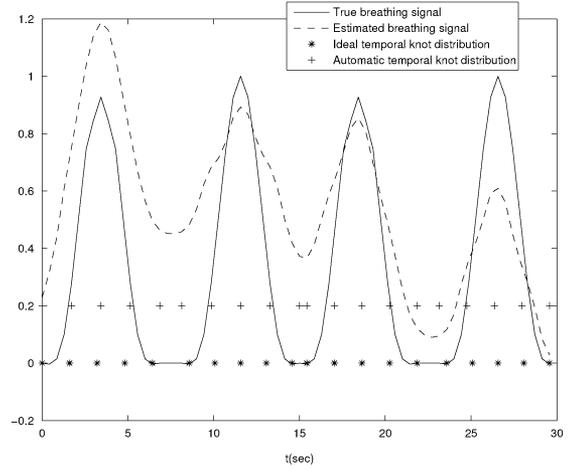


Fig. 5. Ideal temporal knot placement (“\*” line) and automatic temporal knot placement (“+” line).

which a normalized 2-D exponential kernel with a full-width at half-maximum of 4 cm [12] and a scatter to primary ratio (SPR) of 10% was convolved with the primary photon counts. In practice, the SPR may be higher. Finally, the scatter counts and the primary counts were added together and their sums were treated as parameters of the MATLAB function “poissrnd” to generate Poisson distributed noisy projection views.

The respiratory motion we simulated for generating the dynamic cone-beam projection views was based on the three breathhold CT volumes (0%, 20%, and 60% vital capacity). We first registered the 20% and 60% volumes to the 0% volume using a  $B$ -spline based deformation model. Then we selected some voxels near nonuniform regions such as the top surface of diaphragm and the intersections of bronchi, which have smaller registration errors. Afterward, we found the time points  $t_{20}$  and  $t_{60}$  corresponding to the 20% and 60% tidal volumes that best fit the SI displacement of the selected voxels into the following 1-D temporal motion model (14), [20],

$$z(t) = z_0 - a \cos^6(\pi t / \tau - \pi / 2) \quad (14)$$

where  $z_0$  is the SI position at exhale,  $a$  is the amplitude of the motion. Knowing the deformations at three time points and with the symmetry assumption between the motions of exhalation and inhalation, we performed temporal interpolation (separable for each of three directions) of the deformations at each voxel using the MATLAB function “csape” to form one cycle of temporally continuous breathing motion. Four breathing cycles with a total 30-s duration were simulated, each with different

breathing periods and amplitudes. The solid line in Fig. 5 shows the simulated respiratory signal.

### 2) Preparation for DOV:

- *Data preprocessing:* This step obtains the projection views  $\{\hat{g}_m\}$  from the measured photon counts using (5). We used a simple scatter estimate that was obtained by convolving the noisy photon counts with the same exponential kernel used for generating the scatter. For real cone-beam projection views, the scatter estimation should be more complex. In simulation, we deliberately used a simple scatter correction method so the scatter was incompletely corrected, as is the case in practice.

- *B-spline knot distribution:* The placement of B-spline control knots can be very flexible. It can be either a uniform distribution, or a nonuniform distribution. Theoretically, finer control grids enable more accurate approximation of a continuous signal. But in practice, due to the presence of noise, very fine control grids may overfit the noise. Furthermore, a finer control grid complicate optimization.

For our estimation, the control grid contained  $13 \times 9 \times 7$  spatial knots and 20 temporal knots, yielding 16 380 knots in total. The spatial control knots were evenly spaced in the thorax region, with knot spacings of 16, 16, and 10 voxels along the left-right (LR), AP, and SI direction, respectively. They were placed differently from the knot locations used for simulating the motion and with less density. The 20 temporal knots were distributed nonuniformly along the entire temporal axis, with the same number of knots evenly spaced in each active breathing period. The active breathing period is defined to be the interval from the beginning of inhale right to the end of exhale. A short rest interval follows each active breathing interval. Because the deformation during a rest interval would be very small with respect to the reference volume taken at the end of exhale, we did not place any temporal knots in this interval, reducing the number of parameters to be estimated. This nonuniform temporal knot placement facilitates establishment of the phase correspondence between knots as described in Section II-C.

- *Optimization setup:* For optimization, the motion 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. We also applied a multiresolution technique to accelerate the optimization procedure and to avoid local minima. We started the optimization from a downsampled-by-2 version of both the reference volume and the projection views, then used the coarser-scale result as an initialization for the next finer-scale optimization. It took about 65 iterations at the coarser level and 45 iterations at the finer level to converge. The total computation time was about 10 h using Matlab on a 3.4-GHz Pentium computer.

## B. Results and Discussion

In this section, we studied the effects of the temporal knot distribution, the aperiodicity penalty and the two similarity metrics on the DOV performance. We quantify DOV estimation accuracy using the means and standard deviations of the differences between the estimated and the true simulated displacements over the entire volume through all the time points.

TABLE I  
DEFORMATION ESTIMATION ACCURACY UNDER IDEAL TEMPORAL KNOT PLACEMENT

	LR	AP	SI
Mean error (mm)	0.11	0.04	0.16
STD deviation (mm)	0.63	0.81	1.83
MAX Abs error (mm)	6.70	10.82	19.46

1) *Effects of the Temporal Knot Placement:* We present two cases of results using the penalized LS estimator. One case uses an ideal temporal knot placement (“\*” signs in Fig. 5), based on the true respiratory signal. The other case was with automatic temporal knot placement (“+” signs in Fig. 5) according to the estimated respiratory signal from projection views. In the former case, since the true respiratory signal was used, the phase correspondences among the knots in adjacent breathing cycles were exact and thus the periodicity regularity term could accurately align the deformations at the same phases. The ideal case offers us a guideline on how well this proposed algorithm would perform. In the latter case, the peak intervals were detected automatically from the estimated breathing signal and temporal knots were spaced evenly in each peak intervals. Because of the mismatch between the estimated and true respiratory signals, offsets existed between the phases of the knots that were assumed to fall into the same breathing phases by the aperiodicity penalty term. This represents a practical case, where the ground truth of the respiratory signal is unavailable.

With the ideal temporal knot placement, the deformation estimation errors over the entire volume through time had nearly zero-mean Gaussian distributions. As can be seen from Table I, the standard deviations were less than 1 mm along the LR and AP direction and less than 2 mm along SI. These numbers indicate that most of estimation errors were very small. The standard deviation along the SI direction was almost twice of that along the LR and AP direction due to coarser reference image sampling in the SI direction. As an visual example of the estimation accuracy, we plotted the averaged motion curves of 20 randomly selected points (Fig. 4) in the thorax region in Fig. 6. This plot shows good agreement between the estimated and the true motion. Slightly larger deviations from the truth occur near the peaks of the second and third breathing cycles for the LR motion curve and near the peaks of the first and fourth breathing cycles for the AP motion curve. These deviations were expected since the projection views from those angles poorly captured the deformations along the LR or AP directions, respectively.

Table II lists the statistics of the deformation estimation errors with automatic temporal knot placement. Generally the estimated motion errors were slightly larger than those with the ideal temporal knot placement. Fig. 7 plots the true and estimated motion curves of the same 20 points as marked in Fig. 4. Unsurprisingly, the estimated motion curves also showed slightly larger deviation from the truth than those in the previous case. This degraded performance is mainly due to the phase offsets between knots. However, the aperiodicity penalty term did compensate for the insufficient of angular span per breathing cycle of the slowly rotating cone-beam scanner.

Comparison of the two results suggests that better temporal knot placement would improve the motion estimation accuracy.

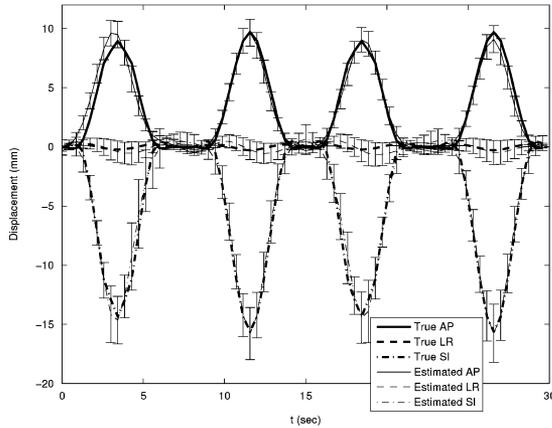


Fig. 6. Accuracy plot of the randomly selected 20 points under the optimization with ideal temporal knot placement. Thick lines represent the true motion curves averaged over the 20 points. Thin lines represent the estimated motion curves averaged over the 20 points. Error bars on the thin lines represent the standard deviations of the deformation estimation errors.

TABLE II  
ESTIMATION ACCURACY UNDER AUTOMATIC TEMPORAL  
KNOT PLACEMENT

	LR	AP	SI
Mean error (mm)	0.12	0.06	0.27
STD deviation (mm)	0.74	0.90	2.16
MAX Abs error (mm)	8.02	9.07	23.56

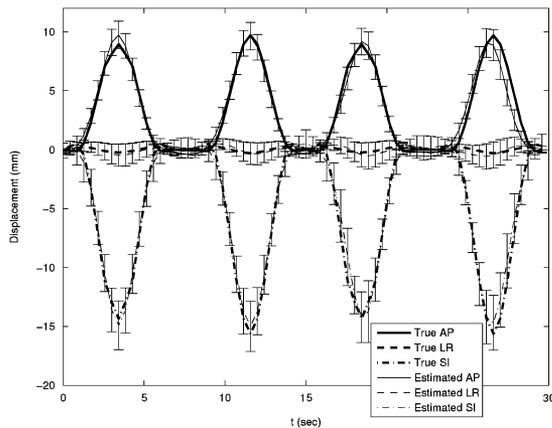


Fig. 7. Accuracy plot of the randomly selected 20 points under the optimization with automatic knot placement. Thick lines represent the true motion curves averaged over the 20 points. Thin lines represent the estimated motion curves averaged over the 20 points. Error bars on the thin lines represent the standard deviations of the deformation estimation errors.

Since the temporal knots are placed according to the respiratory signal, DOV would benefit from a better estimate of the respiratory signal.

Some large deformation errors did occur, even in the case of ideal temporal knot placement, e.g., a maximum absolute error of almost 10 mm along the LR direction. In examining the locations of the larger errors, we found that they tended to occur in image regions having nearly uniform intensities. Because deformations in those regions would exert only very slight changes on the projection views. So these errors are likely due to a lack of image structures, which is common for registration problems.

A second possible source of error is motion model mismatch, i.e., the respiratory motion could not be recovered fully by the  $B$ -spline motion model with the designed control grid. We did

TABLE III  
 $B$ -SPINE LEAST SQUARES FITTING ERROR UNDER  
IDEAL TEMPORAL KNOT PLACEMENT

	LR	AP	SI
Mean error (mm)	0.00	-0.01	0.00
STD deviation (mm)	0.15	0.25	0.57
MAX Abs error (mm)	3.81	4.02	9.48

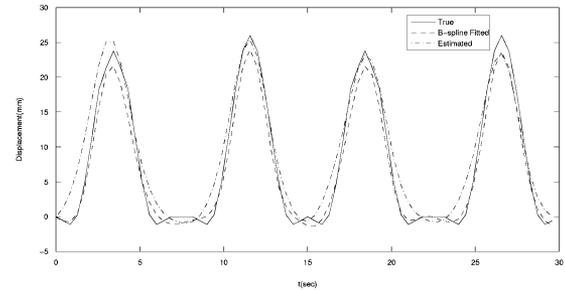


Fig. 8. AP motion curves of the voxel where the maximum  $B$ -spline fitting error along AP direction occurs: the true (solid line), the  $B$ -spline fitted (dashed line), and the estimated by DOV (dashed dotted line).

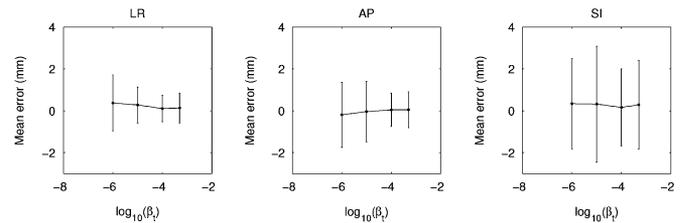


Fig. 9. Mean errors and standard deviations versus the aperiodicity penalty parameter  $\beta_t$ .

$B$ -spline least square fitting of the synthetic motion using the same control grid to examine how much error would result from the model mismatch alone. Table III gives the statistics of the  $B$ -spline approximation errors. Overall the approximation errors were very small, but there were also some relatively large errors. We examined the location where the largest AP motion fitting error occurred to see how well the DOV estimation performed at that voxel. Fig. 8 compares the estimated and the fitted AP motion curves of that voxel. These two curves are close to each other, indicating that the estimated motion was close to the optimum under the selected motion model at this voxel, which did happen to be in a nonuniform region.

2) *Effects of the Aperiodicity Penalty*: The aperiodicity penalty is important for DOV because of the limited gantry angles in one breathing cycle. A too small  $\beta_t$  may not sufficiently bring the motion information from the adjacent breathing cycles to compensate this limitation, while a too large  $\beta_t$  may subdue the role of the local motion information. This is a tradeoff. To study the impact of this term, we ran DOV using the penalized LS estimator with different  $\beta_t$  values and plotted the estimation accuracy in Fig. 9. As  $\beta_t$  increases from  $10^{-6}$ , the mean errors and the standard deviations in each direction tend to drop and then rise again after  $\beta_t$  is larger than  $10^{-4}$ .

3) *The Penalized LS and the Correlation-Based Estimator*: Table IV compares the estimation accuracies of the penalized LS and the correlation-based estimator. As can be seen from the table, the two estimators perform comparably when the intensity

TABLE IV  
ESTIMATION ACCURACY OF THE LS ESTIMATOR AND THE CORRELATION-BASED ESTIMATOR. MEAN ERRORS AND THE STANDARD DEVIATIONS WERE CALCULATED OVER THE WHOLE VOLUME THROUGH TIME

		LR	AP	SI
LS	Mean error (mm)	0.12	0.06	0.27
	STD deviation (mm)	0.74	0.90	2.16
cor	Mean error (mm)	0.11	0.06	0.20
	STD deviation (mm)	0.60	0.76	1.91

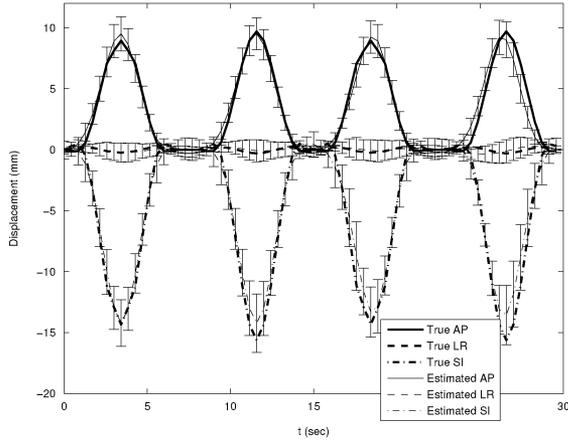


Fig. 10. Accuracy plot of the randomly selected 20 points using the correlation-based estimator with ideal temporal knot placement. Thick lines represents the true motion curves averaged over the 20 points. Thin lines represents the estimated motion curves averaged over the 20 points. Error bars on the thin lines represent the standard deviations of the deformation estimation errors.

of the modelled projection views matches those of the measured views. Fig. 10 draws the accuracy plot of the 20 points using the correlation-based estimator. This plot also resembles the accuracy plot of the penalized LS estimator in Fig. 7.

#### IV. PHANTOM EXPERIMENT

We realize that the realistic condition would be more complex than what can be simulated purely by computer, for example, the noise in the acquired data contains both Poisson and Gaussian, the residual scatter effect after correction may still be prominent. Therefore, we conducted a phantom experiment that is more realistic to further test the DOV performance.

##### A. Phantom

We used a partially deformable thorax phantom to test the performance of DOV, shown in Fig. 11. It is composed of a rigid frame and a compressible foam compartment inside, with some balls inserted. A rigid, flat plastic board is placed at the bottom of the phantom to simulate a diaphragm. This “diaphragm” is connected to a linear actuator through a piece of wood. Driven by the actuator, the “diaphragm” can move back and forth to compress and deform the material inside. The motion pattern of the “diaphragm” is controlled by the actuator. For this experiment, we used a motion profile with alternating amplitudes of 20 mm and 15 mm and alternating periods of 9 s and 6 s (Fig. 12).

##### B. Collecting Data

We first scanned the phantom using a conventional CT. The voltage of the X-ray tube for this CT was set to 120



Fig. 11. Picture of the movable phantom: 1) phantom; 2) diaphragm; 3) wood connector; 4) actuator.

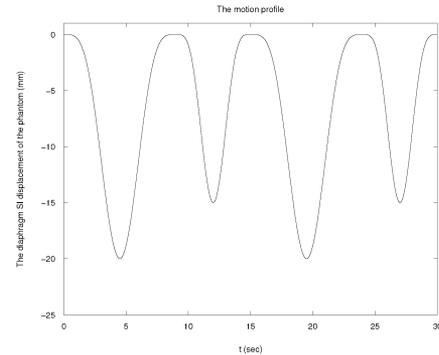


Fig. 12. Motion profile created for the actuator.

kv. We scanned the phantom in three motion states, with the “diaphragm” positioned at 0, 1, and 2 cm toward the neck. We named the three static volumes to be CT0, CT1, and CT2, respectively. The reconstructed volumes have a size of  $512 \times 512 \times 89$  with the voxel size of  $0.98 \times 0.98 \times 3$  mm<sup>3</sup>. CT2 was used as the reference volume for DOV. The other two were used as a measure of truth to evaluate the estimation accuracy of DOV.

Then we moved the phantom to a slowly rotating cone-beam system and started the actuator and took a 360° scan of the moving phantom. Manual laser alignment was performed to set up the phantom right before starting the cone-beam scan. But instead of placing the phantom at the correct setup position, we deliberately moved the phantom off about 1 cm along the axial direction to test DOV with setup errors. After completing the cone-beam scan of the phantom, we removed the phantom and collected a full cone-beam scan of the table. The table scan was used to normalize the measured photon counts of the phantom scan. For the cone-beam scanner, the voltage of the X-ray tube was set to 125 kv. The distance from the X-ray source to the detector was 1500.0 mm and to the isocenter was 1000.0 mm. The size of the 2-D flat-panel detector was  $397 \times 298$  mm<sup>2</sup>. The gantry rotated at 6° per second with a frame rate of about 11

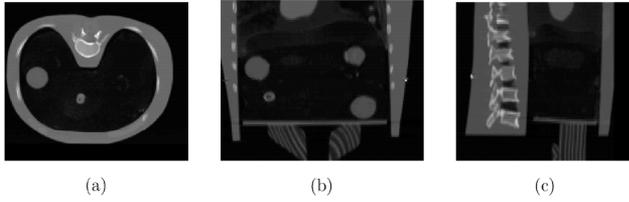


Fig. 13. Phantom CT. (a) Axial slice. (b) Coronal slice. (c) Sagittal slice.

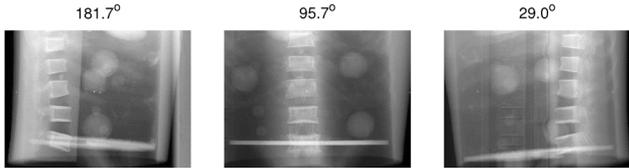


Fig. 14. Cone-beam projection views at angles 181.1°, 95.7°, 29.0°.

frames per second. Totally 668 views were collected for a full rotation.

### C. Preprocessing

We cropped the reference CT and downsampled it by 2 in the axial plane. After cropping and downsampling, the size of the reference CT was  $192 \times 180 \times 89$  and the voxel size was  $2.0 \times 2.0 \times 3 \text{ mm}^3$ . Fig. 13 shows three views of the CT volume. The intensity of the measured projections from the CBCT scanner are linear to the photon counts. We needed to convert the counts to attenuation. The conversion was done by taking logarithm of the table scan divided by the phantom scan, i.e.,

$$\text{attenuation} = \ln \frac{\text{the table scan}}{\text{the phantom scan}}. \quad (15)$$

An advantage of using the table scan rather than an air scan as a normalization factor is that the table artifact may be greatly removed from the phantom scan. However, the table artifact can not be totally removed, (as can be seen in Fig. 14), because the scatter and beam-hardening effects were different in the table scan and phantom scan. For the purpose of DOV, we only used the views in the first  $180^\circ$  interval and downsampled them by 4 in the temporal axis, so there were about 80 views spanned over  $180^\circ$  used by DOV. The projection views were truncated since the 2-D detector was not large enough to cover the whole width of phantom.

Before running DOV, we estimated the setup difference between the conventional CT and the CBCT system. Without setup correction, the estimated motion by DOV would compensate for the setup errors which do not belong to the real organ motion caused by breathing. Usually a rigid setup difference is assumed. It can be described by six parameters  $\{\phi_x, \phi_y, \phi_z, t_x, t_y, t_z\}$ : three rotation angles and three translations along each axis respectively. These parameters can be estimated by aligning the computed projection views of the reference volume to a few measured projection views. This method belongs to the field of 2-D–3-D registration which is commonly used for setup correction in radiotherapy [21], [22]. Usually the projection views used in those registrations do not include organ motion. However, this is not the case for the collected projection views in

TABLE V  
DOV ESTIMATION ACCURACY OF THE PHANTOM EXPERIMENT

	LR	AP	SI
Mean error (mm)	0.4	1.1	0.2
STD deviation (mm)	2.5	2.2	1.7

DOV. To meet the consistency requirement, the setup difference estimation used several projection views approximately corresponding to the motion phase of the reference volume. The correspondence can be established based on the extracted breathing signal from the diaphragm transition. Correlation-based metric was used in the registration. The estimated setup difference was 0.0001, 0.0038, 0.0061 rad; and 0.89, -0.02, 7.74 mm.

### D. Results

Since the energies of the X-ray tubes were different for acquiring the static reference CT and the cone-beam projection views, the intensities of the modelled projection views of the reference CT did not exactly match those of the measured projection views. Furthermore, other artifacts exist in the measured views such as the scatter and beam hardening effects. Using the penalized LS estimator would involve a mapping of the 120 kv reference CT to a 125 kv CT. To avoid this complexity, we chose the correlation-based estimator.

The precise motion of the interior of the phantom was unknown. To evaluate the estimated motion accuracy, we established a “ground truth” using the following landmark method. We located five landmarks in the reference volume CT2 and found their displacements at two motion phases (0 cm and 1 cm of the “diaphragm” movement) by registering CT2 to CT0 and CT1. The landmarks we selected were the centers of five balls inserted in the phantom. [Some balls can be seen in Fig. 13(b)]. We assumed the registration results to be true and compared the estimated motion of the landmarks at the three phases to the truth. The motion phase associated to each cone-beam projection view can be decided by the motion profile (Fig. 12). From the motion profile, we identified that  $t = 3.2, 17.5 \text{ s}$  corresponded to 0 cm position (CT0),  $t = 4.7, 9.3, 11.1, 15.7 \text{ s}$  corresponded to 1 cm position (CT1) and  $t = 6.8, 7.2, 7.5, 12.9, 13.2, 13.6, 21.1, 21.5, 21.8 \text{ s}$  corresponded to 2 cm position (CT2).

We calculated the mean errors and the standard deviations of the estimated deformations of the five landmarks at those 15 time points. The results were listed in Table V. In general, the errors were around the resolution of one voxel. Only the systematic error in the AP direction was slightly larger than expected. Explanation of this slightly larger systematic AP error requires further investigation. We realize that the established “ground truth” may be slightly rough. In the near future, we expect to design a finer ground truth to test the DOV performance, for example, by acquiring more static CTs to find a more precise phantom motion.

### E. Discussion

Although the estimation accuracy is expected to be improved further, the phantom experiment did illustrate the feasibility of the DOV principle, because this experiment reflected the realistic situations in the following three points of view. First,

the movable phantom imitated a real human thorax. Second, a slowly rotating cone-beam CT scanner integrated in a radiotherapy simulator was used to collect the projection views, in which real imaging artifacts existed such as the truncated views, the presence of radiotherapy table, Compton scatter and beam-hardening effects. Third, setup differences between the conventional CT scanner and the cone-beam CT scanner were also considered.

However, the motion pattern of the phantom we created may be simplified compared to the true breathing motion, which would be much more irregular in both amplitude and period. This irregularity may bring uncertainty in selecting the aperiodicity penalty parameter  $\beta_t$ . In our phantom experiment,  $\beta_t$  was set to be  $10^{-4}$ , the value that yielded the best estimation in the simulation, because the motion patterns of the phantom and the simulation were similar. In a real patient study, a monitor may be used to instruct patients to breathe in a more regular way to reduce the difficulty in selecting  $\beta_t$ .

Although the phantom motion pattern was simple, the spatial deformations of the phantom included intensive "sliding" between the edge of the "diaphragm" and the interior side of the body frame. This kind of deformation is challenging to fit by the pure  $B$ -spline model. In real patients, this "sliding" phenomenon would be somewhat reduced. Therefore we may expect better estimation accuracy in a real patient case.

Due to a large angle of the X-ray cone and the use of a 2-D detector, substantial scatter effects were present in the collected cone-beam projection views. In our current phantom experiment, a posterior scatter correction was skipped. This step will be implemented in the near future. We can build a beam stop array to obtain a scatter estimate. A beam stop array consists of small lead disks pressed into holes in an acrylic sheet. It can be placed right after the X-ray source collimator so the projection image reflects the scatter distribution at the positions where the lead disks are projected at, providing an estimate of the scatter [13]. These scatter images are acquired separately from the respiration correlated cone-beam projection views. We believe including a scatter correction can further improve the DOV accuracy.

The estimated deformation errors for the phantom experiment were largely around the resolution of one voxel. Smaller errors can be expected using a reference volume with a higher resolution. In this experiment, we applied a multi-resolution technique and used the downsampled-by-2 reference volume at our finest-level optimization. We did not increase the resolution to that of the originally acquired CT volume to save computation time. We realize that the long computation time will limit the usage of the DOV method in clinic. The bottleneck of the computation is evaluating (3) and (2), the 3-D deformed reference volumes and their projection views at each time points. Since most of the computation can be implemented separately view by view, the computation time can be reduced by parallel computing with a multiprocessor CPU.

## V. CONCLUSION

We propose DOV, a method to estimate respiratory motion from a sequence of slowly rotating X-ray cone-beam projection views. In this method, we adopt a  $B$ -spline motion model,

deform a breathhold thorax CT volume according to the motion model, and find the parameters of the motion model by optimizing the similarity between the measured projection views and the modeled projection views of the deformed reference volume. There are a few advantages of this method over the other 4-D CT imaging techniques. First, we do not assume any reproducibility between the internal motion and an external monitoring index, hence tissue discontinuity artifacts can be removed in the 4-D CT images generated by DOV. Although we use periodicity regularization in our cost function to compensate for the limited angular range over one breathing cycle, the regularization is different from and much weaker than the reproducibility requirement. Second, the  $B$ -spline motion model gives a continuous representation of the estimated motion once the motion parameters are solved. Third, motion is estimated from the on-board cone-beam projection data and can provide the latest update of the patient's motion pattern. We realize that the patients' anatomy may change after the latest reference volume is acquired. Since the deformable motion model may be able to describe modest changes in tumor position and shape given an appropriate knot distribution, even if there are slight changes on patients anatomy, we conjecture that the latest reference volume could still be used by DOV without degrading the performance dramatically. However, large deformations may either invalidate the reference model or increase degeneracy of motion estimation. The likely scenario is that a large average deformation near the reference state would lead to a repeat clinical simulation, thus updating the patient model for treatment as well as future motion estimation.

Our simulation and phantom experiment yielded encouraging results. Performance may be further improved by using a more robust similarity metric such as mutual information and by improving scatter correction. Better spatial knot placement may also be determined through the registration of two breathhold CT volumes. We also realize that real patient motion could not be totally characterized by simulations or physical phantoms. Therefore, pending research will focus on further improving the estimation accuracy and applying this method to clinical cone-beam data from patients undergoing radiation therapy treatments.

## ACKNOWLEDGMENT

The authors would like to thank B. De Man and S. Basu for permitting the use of their distance-driven projector. They would also like to thank D. Litzenberg and R. Kashani for their precious time in building the phantom and collecting data.

## REFERENCES

- [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. Radial. Oncol. Biol. Phys.*, vol. 44, no. 4, pp. 911–919, Jul. 1999.
- [2] H. A. Shih, S. B. Jiang, K. M. Aljarrah, K. P. Doppke, and N. C. Choi, "Internal target volume determined with expansion margins beyond composite gross tumor volume in three-dimensional conformal radiotherapy for lung cancer," *Int. J. Radial. Oncol. Biol. Phys.*, vol. 60, no. 2, pp. 613–622, 2004.
- [3] G. Starkschall, K. M. Forster, K. Kitamura, M. D. , A. Cardenas, S. L. Tucker, and C. W. Stevens, "Correlation of gross tumor volume excursion with potential benefits of respiratory gating," *Int. J. Radial. Oncol. Biol. Phys.*, vol. 60, no. 4, pp. 1291–1297, 2004.

- [4] 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.*, vol. 48, no. 1, pp. 45–62, Jan. 2003.
- [5] 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.*, vol. 30, no. 6, pp. 1254–1263, Jun. 2003.
- [6] T. Pan, T.-Y. Lee, E. Rietzel, and G. T. Y. Chen, "4-D-CT imaging of a volume influenced by respiratory motion on multi-slice CT," *Med. Phys.*, vol. 31, no. 2, pp. 333–340, Feb. 2004.
- [7] J.-J. Sonke, L. Zijp, P. Remeijer, and M. Herk, "Respiratory correlated cone beam CT," *Med. Phys.*, vol. 32, no. 4, pp. 1176–1186, Apr. 2005.
- [8] T. Li, E. Schreibmann, Y. Yang, and L. Xing, "Motion correction for improved target localization with on-board cone-beam computed tomography," *Phys. Med. Biol.*, vol. 51, no. 2, pp. 253–268, Jan. 2006.
- [9] R. Zeng, J. A. Fessler, and J. M. Baiter, "Respiratory motion estimation from slowly rotating X-ray projections: Theory and simulation," *Med. Phys.*, vol. 32, no. 4, pp. 984–991, Apr. 2005.
- [10] A. Macovski, *Medical Imaging Systems*. Upper Saddle River, NJ: Prentice-Hall, 1983.
- [11] H. Erdogan and J. A. Fessler, "Monotonic algorithms for transmission tomography," *IEEE Trans. Med. Imag.*, vol. 18, no. 9, pp. 801–814, Sep. 1999.
- [12] L. A. Love and R. A. Kruger, "Scatter estimation for a digital radiographic system using convolution filtering," *Med. Phys.*, vol. 14, no. 2, pp. 178–185, Mar. 1987.
- [13] R. Ning, X. Tang, and D. Conover, "X-ray scatter correction algorithm for cone beam CT imaging," *Med. Phys.*, vol. 31, no. 5, pp. 1195–1202, May 2004.
- [14] A. P. Colijn and F. J. Beekman, "Accelerated simulation of cone beam X-ray scatter projections," *IEEE Trans. Med. Imag.*, vol. 23, no. 5, pp. 584–590, May 2004.
- [15] L. Zijp, J.-J. Sonke, and M. Herk, "Extraction of the respiratory signal from sequential thorax cone-beam X-ray images," in *Proc. Int. Congr. Comp. Radiotherapy*, 2004.
- [16] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C*. New York: Cambridge Univ. Press, 1988.
- [17] P. Thevenaz and M. Unser, "Optimization of mutual information for multiresolution image registration," *IEEE Trans. Image Process.*, vol. 9, no. 12, pp. 2083–2099, Dec. 2000.
- [18] B. De Man and S. Basu, "Distance-driven projection and backprojection in three dimensions," *Phys. Med. Biol.*, vol. 49, no. 11, pp. 2463–2475, Jun. 2004.
- [19] P. J. La Riviere and D. G. Politte, "Penalized-likelihood sinogram smoothing for low-dose CT," *Med. Phys.*, vol. 32, no. 6, pp. 1676–1683, Jun. 2005.
- [20] A. E. Lujan, E. W. Larsen, J. M. Baiter, and R. K. Ten Haken, "A method for incorporating organ motion due to breathing into 3-D dose calculations," *Med. Phys.*, vol. 26, no. 5, pp. 715–720, May 1999.
- [21] J. Kim, J. A. Fessler, K. L. Lam, J. M. Baiter, and R. K. Ten Haken, "A feasibility study of mutual information based set-up error estimator for radiotherapy," *Med. Phys.*, vol. 28, no. 12, pp. 2507–2517, Dec. 2001.
- [22] J. Weese, G. P. Penney, P. Desmedt, T. M. Buzug, D. L. G. Hill, and D. J. Hawkes, "Fast voxel-based 2-D/3-D registration of fluoroscopy images and CT scans for image-guided surgery," *IEEE Trans. Inf. Technol. Biomed.*, vol. 1, no. 4, pp. 284–293, Dec. 1997.