

Estimating 3D Respiratory Motion from Orbiting Views

Rongping Zeng¹, Jeffrey A. Fessler¹, and James M. Balter²

rzeng@eecs.umich.edu, fessler@eecs.umich.edu, jbalter@med.umich.edu

Abstract—This paper describes a method for estimating 3D respiratory motion so as to characterize tumor motion. This method uses two sets of measurements. One is a reference thorax volume obtained from a conventional fast CT scanner under breath-hold condition. The other is a sequence of projection views of the same patient (acquired at treatment time) using a slowly rotating cone-beam system (1 minute per rotation) during free breathing. We named this method Deformation from Orbiting Views (DOV). Breathing motion over the entire acquisition period is estimated by deforming the reference volume through time so that its projections best match the measured projection views. The nonrigid breathing motion is described by a B-spline based deformation model. The parameters of this model are estimated by minimizing a regularized squared error cost function, using a conjugate gradient descent algorithm. Performance of this approach was evaluated by simulation. Results showed good agreement between the estimated and synthesized motion, with a mean absolute error of 1.63mm. Relatively larger errors tended to occur in uniform regions, which would not have significant effects on generating deformed volumes based on the estimated motion. The results indicate that it is feasible to estimate realistic nonrigid motion from a sequence of slowly rotating cone beam projection views.

Index Terms—4D CT, respiratory motion, cubic B-spline, cone-beam scanners.

I. INTRODUCTION

CONFORMAL radiotherapy requires that the delivery of X-ray is focused on the tumor area while sparing the normal adjacent tissue. However, geometric uncertainties caused by respiratory motion bring difficulties in conformal radiotherapy. Most current 4D (3D spatial + 1D breathing index) CT imaging techniques use conventional scanners to study respiratory motion [1]–[3]. These methods often acquire multiple 2D slices at each table position, sort these slices into several respiratory phase bins, and then stack those slices within the same phase bins to form 4D images. Due to the insufficient spatial coverage for imaging an entire volume for one breathing cycle, an assumption is made that there is a the reproducibility relationship between the internal motion and the “phase” of 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 acquire 4D data

using cone-beam CT scanners [4]. However, the limitation from motion reproducibility assumption still exists.

We propose a different way to study respiratory motion using a cone-beam scanner that is integrated into a radiotherapy system. This type of cone-beam scanner usually rotates slowly with about 1 minute per full rotation. This type of system provides large volume coverage (a 41×41 cm² detector) and a high temporal sampling rate (3-15 projection views per second). Therefore, it is possible to estimate respiratory motion from a sequence of cone-beam projection 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 deformation of the prior through time and to estimate the motion parameters by maximizing the similarity between the modeled and measured projection views. We call it deformation from orbiting views (DOV). This method reduces the motion reproducibility requirement of aforementioned 4D CT imaging methods.

II. THEORY OF DOV

The proposed motion estimation method uses two sets of measurements. One is a reference thorax CT volume obtained from a conventional fast CT scanner under breath-hold condition, denoted as $f_{\text{ref}}(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^3$. The other is a sequence of projection views of the same patient (acquired at treatment time) using a slowly rotating cone-beam system (1 minute per rotation), represented as \mathbf{g}_m , $m = 1, \dots, M$ (M is the number of projections). Although we allow the cone-beam scanners to rotate slowly, we require the acquisition time of each projection view to be short. For example, recently developed systems can acquire 15 frames per second, i.e., around 0.067 second per frame. We therefore assume that the respiratory motion is negligible within each single projection view. Let the motion be denoted as $\mathcal{T}_{\theta}(\mathbf{x}; t_m)$, a time-dependent deformation controlled by parameters θ . Since the projection views and the reference volume are all from the same patient, a mathematical relationship between \mathbf{g}_m and f_{ref} can be represented by the following equations, ignoring any noise and imaging artifacts,

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

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

where \mathcal{A}_{ϕ_m} denotes the projection operator from projection angle ϕ_m , and f_{t_m} is the deformed volume at time t_m .

This work is supported by NIH P01 CA59827-06A1.

¹Electrical Engineering and Computer Science Dept., The University of Michigan; ²Radiation Oncology Dept., The University of Michigan

However, in practice the measured projection views Y_m are always degraded by noise,

$$Y_{m,n} \sim \text{Poisson}(I_{m,n}e^{-g_{m,n}} + S_{m,n}), \quad (3)$$

where $I_{m,n}$ is a constant related to the incident X-ray energy, $S_{m,n}$ denotes the scatter contribution to $Y_{m,n}$ and n is the number of detector elements.

$\mathcal{T}_\theta(\mathbf{x}; t)$ can be any suitable deformation models. We adopted a cubic B-spline based motion model as follows,

$$\mathcal{T}_\theta(\mathbf{x}; t) = \mathbf{x} + \sum_j \sum_i \theta_{j,i} \beta\left(\frac{t - \tau_j}{\Delta_t}\right) \beta\left(\frac{\mathbf{x} - \mathbf{x}_i}{\Delta_{\mathbf{x}}}\right) \quad (4)$$

where $\beta(\cdot)$ is the cubic B-spline function and $\beta(\mathbf{x})$ the tensor product of cubic B-spline functions, τ_j and \mathbf{x}_i the spatial and temporal knot locations, $\Delta_{\mathbf{x}}$ and Δ_t control the width of the spatial and temporal basis functions respectively, and θ the knot coefficients. There are two advantages to use 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 B-spline control grid can be flexibly adjusted to meet the approximation accuracy of continuous signals. Based on the motion model (4), the estimation goal is to find the motion parameters θ from the projection views \mathbf{g}_m and f_{ref} .

Based on the relationship between \mathbf{g}_m and f_{ref} described in (1) and (2), we construct a regularized least square estimator of θ as follows,

$$\hat{\theta} = \arg \min_{\theta} (L(\theta) + \beta_s R_s(\theta) + \beta_t R_t(\theta)), \quad (5)$$

where $L(\theta) = \frac{1}{2} \sum_{m=1}^M \|\mathbf{g}_m - \mathcal{A}_{\phi_m} f_{\text{ref}}(\mathcal{T}_\theta(\mathbf{x}; t_m))\|^2$, a data fidelity term, $R_s(\theta)$ is a spatial motion roughness penalty term, $R_t(\theta)$ is a temporal motion aperiodicity penalty term, and β_s and β_t are scalars that control the trade-off between the three terms. $R_s(\theta)$ discourages wiggly spatial deformation, and $R_t(\theta)$ encourages similarity between deformations in similar respiratory phases. We refer to $R_t(\theta)$ as the ‘‘aperiodicity penalty’’, and we included this term to help overcome the limit angular range (20° to 40° for one breathing cycle) of a slowly rotating cone-beam scanner. 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. The aperiodicity penalty term can bring in some motion information contained in the adjacent breathing cycles to compensate for the angular limitation.

The aperiodicity penalty is a temporal regularity term. 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 approximated by the closeness of the coefficient values of knots that are located at similar respiratory phases. To determine the correspondence between the temporal knots, we need a respiratory signal. We adopted the idea presented in Zijp’s paper [5] and simplified their method. The basic idea is to capture the superior-inferior (SI) transition of the diaphragm in the collected projection views.

Since there is no analytical solution for (5), We used iterative method to search for $\hat{\theta}$. The conjugate gradient descent algorithm was chosen. The multi-resolution technique was also applied to speed up the optimization and to help avoid local minima.

III. SIMULATIONS AND RESULTS

A. Experiment setup

Performance of this approach was evaluated by simulation. The synthetic respiratory motion in the simulation was based on three breathhold thorax CTs at 0%, 20% and 60% tidal volume, with voxel size $0.2 \times 0.2 \times 0.5 \text{ cm}^3$. We selected the 0% breathhold CT (end-expiration) as the reference volume. 70 cone-beam projection views of the warped reference volume over a 180° rotation were generated in this simulation. The warping was according to a synthetic motion. In order to simulate a realistic respiratory motion, we first registered the three breathhold thorax CTs to find the deformations of 20% and 60% relative to 0%. Then we resolved the corresponding time phases for 20% and 60% tidal volume phases, by putting the SI displacements of some points into the following temporal breathing motion model [6],

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

where z_0 is the SI position at exhale, a the amplitude of the motion and T the period of breathing cycle. Knowing the deformations at three time points and with the symmetry assumption between the motion of exhalation and inhalation, we performed cubic spline interpolation of the deformations at each voxel to form one cycle of temporally continuous breathing motion. Four breathing cycles were simulated, a total duration of 30 seconds with a different breathing period for each cycle. The simulated cone-beam system had a flat-panel detector of 180×200 elements with each element area of $0.4 \times 0.4 \text{ cm}^2$. The source to isocenter distance and the isocenter to detector distance were 100cm and 50cm respectively. The gantry rotated from the anterior view with 6° per second and spanned 180° over the four breathing cycles. The simulated projection views were degraded by scatter effect and Poisson noise as described in the noisy model (3), where the scatter effect was generated using the convolutional method [7].

Before estimation, we removed the scatter effect first. We assumed the same convolutional kernel, convolved it with the noisy projection views and then subtracted the convolution result from the noisy projection views. So there was model mismatch between the estimated and the actual scatter since the convolution step was operated on the non-noisy projection views for generating the scatter effect while on the noisy projection views for removing the scatter effect. After removing scatter, we estimate the line integrals \mathbf{g}_m based on (3), i.e.,

$$\hat{g}_{m,n} = \log\left(\frac{I_{m,n}}{Y_{m,n} - S_{m,n}}\right), \quad (7)$$

The spatial control knots were evenly spaced in the thorax region. They were placed differently from the knot locations

used when simulating the motion, and with less density. The temporal knots were distributed non-uniformly along the entire temporal axis, with the same number of knots evenly spaced in each active breathing cycle.

B. Results and discussion

We present two groups of results. One case was with ideal temporal knot placement ('*' signs in Fig. 1), assuming the true respiratory phase signal is known; the other case was with automatic temporal knot placement according to the estimated respiratory phase signal from projection views ('+' signs in Fig. 1). The former one offers us a guideline on how well this proposed algorithm would perform under ideal situation, while the latter one represents a practical case, where the ground truth of the respiratory signal is unavailable.

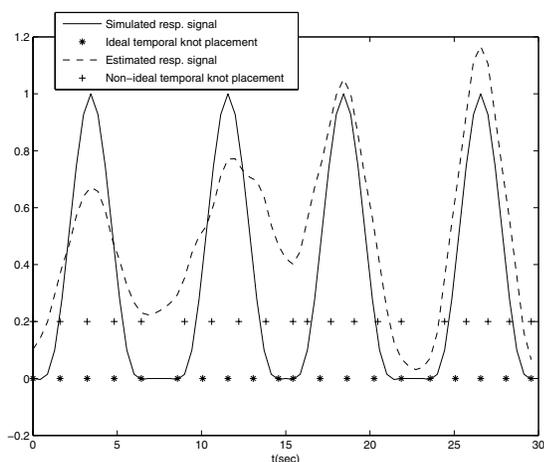


Fig. 1. Ideal temporal knot placement ('*' line) and automatic knot placement ('+' line)

- Case 1: ideal temporal knot placement

With the ideal temporal knot placement, the deformation estimation errors over the entire volume through time had a nearly zero-mean Gaussian distribution. As can be seen from Table I, the STD deviation were less than 1 mm along the LR and AP direction and less than 2mm along SI, similar for the root mean square (RMS) errors. These numbers indicated that most of estimation errors were very small. The RMS error along the SI direction was almost twice of that along the LR and AP direction due to a doubled image resolution in the SI direction. As an example of the estimation accuracy, we plotted the averaged motion curves of 20 randomly selected in the thorax region in Fig. 2. This plot showed a good agreement between the estimated and the true motion. Notice that slightly larger deviations from the truth presented near the peaks of the 2nd and 3rd breathing cycles for the LR motion curve and near the peaks of the 1st and 4th breathing cycles for the AP motion curve. These deviations were reasonable since the projection views from those projection angles poorly captured the thorax deformations along the LR or AP direction. But due to the periodicity regularization, the deficiency of motion info

due to limited angular span in one breathing cycle was greatly compensated.

TABLE I
DEFORMATION ESTIMATION ACCURACY UNDER IDEAL TEMPORAL KNOT PLACEMENT.

	LR	AP	SI
Mean error (cm)	0.013	0.009	0.011
STD deviation (cm)	0.069	0.083	0.179
RMS error (cm)	.064	.076	.164

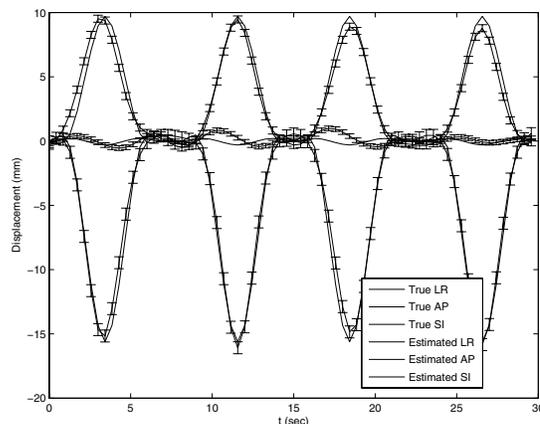


Fig. 2. Accuracy plot of the randomly selected 20 points under the optimization with ideal temporal knot placement. The Thick lines represents the true motion curves averaged over the 20 points. The thin lines represents the estimated motion curves averaged over the 20 points. Error bars on the thin lines represent the standard deviations of the deformation estimation errors.

However, there existed some large deformation errors, e.g. a maximum absolute error of 1cm in the LR direction. By examining the locations of the relatively large errors, we found that they tended to occur at the image regions with low contrast, which would not produce significant effect on generating deformed volumes based on the estimated deformations. As can be seen in Fig. 4 and Fig. 5, the differences between the true and predicted deformed volumes were quite small.

- Case 2: non-ideal temporal knot placement

In this case, the temporal knots were automatically placed based on the estimated respiratory signal. Because of the mismatch between the estimated and true respiratory signals, there were offsets between the phases of the deformations that were assumed to correspond to the same breathing phases by the aperiodicity penalty term.

The statistics of the deformation estimation errors are listed in Table II. Comparing the result with that in Table I, the deformation errors were generally larger than those under optimization with ideal temporal knot placement. Comparison of the true and estimated motion curves were plotted in Fig. 3. Unsurprisingly, the estimated motion curves also showed larger deviation from the truth than that in the previous case. This degraded performance was partially due to the larger B-spline approximation errors from motion model mismatch. Another reason was that there were offsets between the temporal knot

locations decided by automatic placement and ideal placement (Fig. 1). Periodicity regularization with these phase offsets resulted in less accuracy in the estimated motion, however, it did help overcome the insufficiency of angular span for one breathing cycle of the slowly rotating cone-beam scanners. Good placement of B-spline control knots is essential for this work. More investigation is needed on this issue.

TABLE II
ESTIMATION ACCURACY UNDER AUTOMATIC TEMPORAL KNOT
PLACEMENT.

	LR	AP	SI
Mean error (cm)	0.017	-0.001	0.014
STD deviation (cm)	0.0774	.1092	.2014
RMS error (cm)	.0740	.0995	.1875

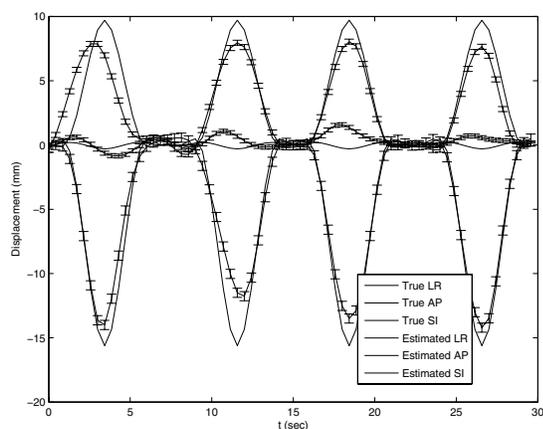


Fig. 3. Accuracy plot of the randomly selected 20 points under the optimization with automatic knot placement. The Thick lines represents the true motion curves averaged over the 20 points. The thin lines represents the estimated motion curves averaged over the 20 points. Error bars on the thin lines represent the standard deviations of the deformation estimation errors.

IV. CONCLUSION

We proposed DOV, a respiratory motion estimation method from a sequence of slowly rotating 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. The simulation yielded encouraging results, with the averaged deformation estimation errors around 1mm. Although there were some large deformation errors over the entire volume through time, but they tended to occur at uniform regions, which may not have significant meaning. In practice a region of interest should be identified by radiotherapists and the maximum errors in that region are more important, or we could set constraints on the maximum displacements of some voxels during optimization based on prior clinical information.

Currently we used a penalized least square estimator. Based on the statistical property of the projection views, we can

design a weighted least square estimator which may improve the estimation accuracy. We also need to investigate more on how to better place the spatial knots and temporal knots to reduce motion model mismatch. Finally, this method will be applied to real clinical cone-beam data.

REFERENCES

- [1] 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.
- [2] 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. Polite, 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. 1254–63, June 2003.
- [3] T. Pan, T.-Y. Lee, E. Rietzel, and G. T. Y. Chen, "4D-CT imaging of a volume influenced by respiratory motion on multi-slice CT," *Med. Phys.*, vol. 31, no. 2, pp. 333–40, Feb. 2004.
- [4] J.-J. Sonke, L. Zijp, P. Remeijer, and M. . Herk, "Respiratory correlated cone beam CT," *Med. Phys.*, vol. 32, no. 4, pp. 1176–86, Apr. 2005.
- [5] 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.
- [6] A. E. Lujan, E. W. Larsen, J. M. Balter, and R. K. Ten Haken, "A method for incorporating organ motion due to breathing into 3D dose calculations," *Med. Phys.*, vol. 26, no. 5, pp. 715–20, May 1999.
- [7] L. A. Love and R. A. Kruger, "Scatter estimation for a digital radiographic system using convolution filtering," *Med. Phys.*, vol. 14, no. 2, pp. 178–85, Mar. 1987.

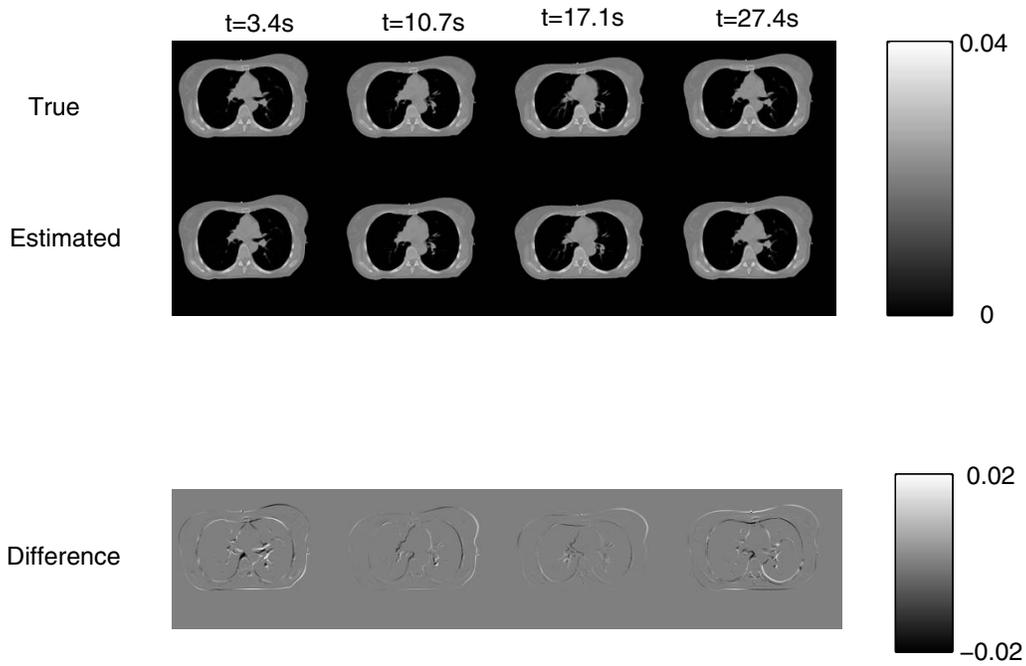


Fig. 4. Comparison of the true and estimated deformed volumes (axial view) at four time points under optimization with ideal temporal knot placement

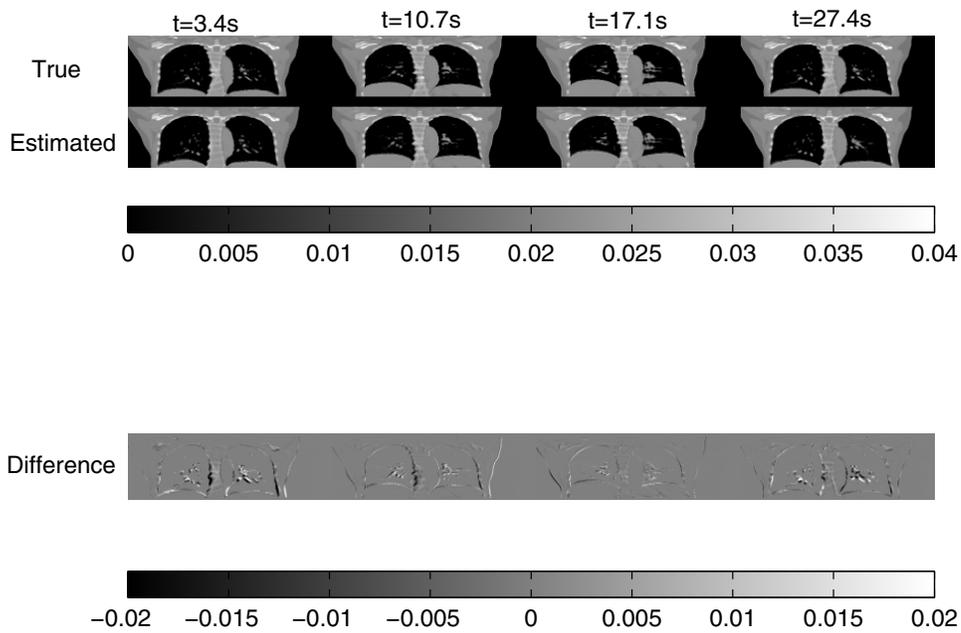


Fig. 5. Comparison of the true and estimated deformed volumes (axial view) at four time points under optimization with ideal temporal knot placement