RESPIRATORY MOTION ESTIMATION FROM SLOWLY ROTATING X-RAY PROJECTIONS

Rongping Zeng, Jeffrey A. Fessler, James Balter

EECS / Radiation Oncology Dept., The University of Michigan
rzeng@umich.edu, fessler@umich.edu, jbalter@umich.edu

ABSTRACT

As radiotherapy has become increasingly conformal, geometric uncertainties caused by breathing and organ motion have become an important issue. Accurate motion estimates may lead to improved treatment planning and dose calculation in radiation therapy. However, respiratory motion is 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 motion model from a set of projection views of the thorax that are acquired using a slowly rotating cone-beam CT scanner, such as a radiotherapy simulator. We use a conventionally reconstructed 3D thorax image, acquired by breath-hold CT, as a reference volume. We represent respiratory motion using a flexible parametric nonrigid motion model based on B-splines. The motion parameters are estimated by optimizing a regularized cost function that includes the squared error between the measured projection views and the reprojections of the deformed reference image. Preliminary 2D simulation results show that there is good agreement between the estimated motion and the true motion.

1. INTRODUCTION

Extensive research has been performed on estimating cardiac motion [1], aiding the diagnosis of cardiac function. However, there has been less work on building patient specific 4D respiratory motion models. Understanding thorax motion, especially the motion of tumors inside lungs, is important for radiation treatment planning. For example, knowledge of the space that tumors reach during a respiration cycle may guide delivery of the X-ray dose to focus more on tumors while sparing the normal adjacent tissue.

We propose a method 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 one minute 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.

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 use currently is the gradient descent method.

Instead of using projections, one may think that a feasible alternative would be to estimate respiratory motion by registering a set of thorax CTs taken under different breath-hold conditions. There are at least two shortcomings to this idea. First, unnaturally controlled breathhold states tend to be discontinuous, and some people (like lung cancer patients) cannot control their respiration well. Second, patients might be required to be scanned dozens of times all at once to generate enough views for each phase in the respiration cycle, and this long-time exposure to X-ray is undesirable. In contrast, in our method, patients breathe naturally and are only scanned for a few rotations, which overcomes those disadvantages.

The paper is organized as follows. First the theory is described, including the temporal motion model, similarity measure and optimization method. Then the preliminary simulation results are presented, followed by our conclusion and proposals for future work.

2. THEORY

2.1. Temporal Motion Model

Let $\{f_{t_1}, \cdots, f_{t_m}, \cdots, f_{t_M}\}$ denote the M-frame moving image sequence, where f_{t_m} is the image at time t_m. Assuming these images are all a deformation of the reference image f_r, then there exists a correspondence between f_{t_m} and f_r:

$$
\begin{aligned}
f_{t_m}(x,y) &= (\mathcal{W}_{\theta_m} f_r)(x,y) \\
&= f_r(T_x(x,y,\theta_m^{(x)}), T_y(x,y,\theta_m^{(y)})), \quad (1)
\end{aligned}
$$

where $\mathcal{W}_{\theta_m}$ is the warping operator controlled by parameter θ_m, $T_x(x,y,\theta_m^{(x)})$ and $T_y(x,y,\theta_m^{(y)})$ describe the deformation functions along x direction and y direction respectively, and vector θ_m is the parameter of the deformation functions. As described in [2], a smooth nonrigid spatial deformation can be modeled using spline functions as follows:

$$
T_x(x,y,t_m) = x + \sum_{i=1}^{I}\sum_{j=1}^{J} \theta_{m,ij}^{(x)} \, \beta\left(\frac{x}{\Delta_{\mathrm{x}}} - i\right) \beta\left(\frac{y}{\Delta_{\mathrm{y}}} - i\right)
$$

$$
T_y(x,y,t_m) = y + \sum_{i=1}^{I}\sum_{j=1}^{J} \theta_{m,ij}^{(y)} \, \beta\left(\frac{x}{\Delta_{\mathrm{x}}} - i\right) \beta\left(\frac{y}{\Delta_{\mathrm{y}}} - i\right),
$$

where $\beta(\cdot)$ is the spline function of degree n, $\theta^{(x)}$ and $\theta^{(y)}$ are spline coefficients, and Δ_{x} and Δ_{y} are constant scalars that control the width of the basis functions. We chose B-spline functions because of its advantages of good accuracy and analytical derivatives [3], but the theory generalizes.

Since the deformation of tissue over time should also be continuous, we adopt the following parametric temporal motion model:

$$
\begin{aligned}
T_x(x,y,t) &= x + \sum_{m=1}^{M}\sum_{i=1}^{I}\sum_{j=1}^{J} \theta_{mij}^{(x)} b\left(\frac{t}{\Delta_{\mathrm{t}}} - m\right) \\
&\qquad \beta\left(\frac{x}{\Delta_{\mathrm{x}}} - i\right) \beta\left(\frac{y}{\Delta_{\mathrm{y}}} - i\right)
\end{aligned}
$$

$$
\begin{aligned}
T_y(x,y,t) &= y + \sum_{m=1}^{M}\sum_{i=1}^{I}\sum_{j=1}^{J} \theta_{mij}^{(y)} b\left(\frac{t}{\Delta_{\mathrm{t}}} - m\right) \\
&\qquad \beta\left(\frac{x}{\Delta_{\mathrm{x}}} - i\right) \beta\left(\frac{y}{\Delta_{\mathrm{y}}} - i\right),
\end{aligned}
$$

where the temporal basis function $b(t)$ can take many forms, such as splines. We use the rectangular function currently for simplicity.

2.2. Measurement

As described before, the motion estimates are based on two set of measurements. One set is a motionless thorax image f_r, reconstructed from a conventional 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 projection views from a slowly rotating scanner:

$$
\{p_{\phi_1}, \cdots, p_{\phi_m}, \cdots, p_{\phi_M}\},
$$

where ϕ_i is the projection angle, and M is the total number of projections.

Let $\mathcal{A}_{\phi_m}$ denote the projection operator at angle ϕ_m, then

$$
\begin{aligned}
p_{\phi_m} &= \mathcal{A}_{\phi_m} f_{t_m} + \varepsilon_m \\
&= \mathcal{A}_{\phi_m} \mathcal{W}_{\theta_m} f_r + \varepsilon_m,
\end{aligned}
$$

where ε_m represents additive noise for the mth projection view. Based on the motion model defined in the previous section, the task of the motion estimation is to find the deformation parameters $\boldsymbol{\theta} = \{\theta_1, \cdots, \theta_m, \cdots, \theta_M\}$ using f_r from $\{p_{\phi_m}\}$. This is not a conventional image reconstruction problem, but is a kind of tomographic image registration problem.

2.3. Similarity Measure

The goal of our motion estimation is to find the deformation parameters $\boldsymbol{\theta}$ that make the calculated projections based on the deformed image f_r best match the measured projections. For simplicity, we focus here on the least-squared error metric:

$$
L(\boldsymbol{\theta}) = \frac{1}{2} \sum_{m=1}^{M} \left\| p_{\phi_m} - \mathcal{A}_{\phi_m} \mathcal{W}_{\theta_m} f_r \right\|^2,
$$

where $\mathcal{W}_{\theta_m} f_r$ is the warped f_r as shown in (1). The principles generalize to more complicated statistical models.

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

$$
\begin{aligned}
\hat{\boldsymbol{\theta}} &= \arg\min_{\boldsymbol{\theta}} \psi(\boldsymbol{\theta}) \\
\psi(\boldsymbol{\theta}) &= L(\boldsymbol{\theta}) + \lambda R(\boldsymbol{\theta}) \quad\quad (2)
\end{aligned}
$$

where $R(\boldsymbol{\theta})$ denotes the regularization function, and the parameter λ controls the trade-off between the similarity term and the regularity term.

Assuming that the motion is smooth temporally, we can regularize by using a temporal roughness penalty of the following form:

$$
R(\boldsymbol{\theta}) = \|\boldsymbol{C}\boldsymbol{\theta}\|^2,
$$

where $\boldsymbol{C}$ is a differencing matrix.

2.4. Optimization

For simplicity, we currently use the gradient descent method to minimize the cost function ψ as follows:

$$\boldsymbol{\theta}^{n+1} = \boldsymbol{\theta}^n - \alpha \nabla \psi(\boldsymbol{\theta}^n),$$

where α is the step size, and $\nabla \psi(\boldsymbol{\theta}^n)$ is the gradient of $\psi(\boldsymbol{\theta})$ evaluated at $\boldsymbol{\theta}^n$. Because of the linearity of the projection operator and the closed form of the B-spline functions, the gradients can be calculated analytically using the chain rule. The expressions are omitted due to space constraints.

A faster convergent algorithm, the preconditioned gradient descent (PGD) method, can also be used as the optimization algorithm, having the following iterative form:

$$\boldsymbol{\theta}^{n+1} = \boldsymbol{\theta}^n - \alpha P \nabla \psi(\boldsymbol{\theta}^n),$$

where P is a preconditioning matrix. The ideal preconditioning matrix P_0 would satisfy $P_0 H = I$, where H is the Hessian of $\psi(\boldsymbol{\theta})$, and I is the identity matrix. Since it is difficult to compute H^{-1}, we used the diagonal preconditioner $P = (\text{diag}\{H\})^{-1}$ to approximate H^{-1}.

3. SIMULATION RESULTS

Fig. 1 shows the reference image f_r, a 128×128 thorax phantom, used in our simulation [4]. We simulated a parallel-beam slow scanner to generate the data set $\{p_{\phi_m}\}_{m=1}^{m=M}$, shown in Fig. 2, consisting of $M = 15$ projections at angles uniformly spaced over $180°$. Each projection view had 160 radial samples. During the scanning process, the thorax phantom deforms according to a synthetic motion represented by the temporal motion model described in Section 2.1. Here the deformation of each frame is defined by cubic splines with a control grid of 2×2 points, meaning that there are 4×2 parameters for each frame. Since there are 15 frames, there are 120 parameters in this motion model. In this case there is no motion model mismatch, allowing quantitative comparisons with ground truth. Future studies will include more realistic motion models including motion model mismatch.

Fig. 3, Fig. 4, and Fig. 5 display the estimation results. We initialized $\boldsymbol{\theta}^0$ by random deviations from the true values. The deviations were randomly distributed between ± 8 pixels (1 pixel corresponds to about 3mm). As shown in Fig. 4, the estimated values of most parameters are fairly close to their true values. We also examine two points from the image (marked in Fig. 1) and compared the estimated displacements with their true vlues. The comparison shown in Fig. 5 illustrates that there is good agreement between the true movement and the movement found by the algorithm. The largest error is around 1 pixel. It is necessary to investigate this issue further to ascertain whether the algorithm can be improved, and if so, how to achieve the improvement.

4. CONCLUSION

In this paper, we proposed a nonrigid motion estimation method that uses a reference image and a sequence of projection views from a slowly rotating scanner. Cubic B-spline functions were applied as the basis of our parametric temporal motion model. Preliminary results presented in Section 3 illustrate the potential of this method. More simulation work will be undertaken in the near future, including, for example, experimenting on different time basis functions and penalty functions, testing the performance of the algorithm under noisy situations, and extending the current 2D implementation to the 3D case. Finally, the algorithm will be applied to real lung data.

5. REFERENCES

[1] J. C. McEachen, A. Nehorai, and J. S. Duncan, "Multiframe temporal estimation of cardiac nonrigid motion," *IEEE Tr. Im. Proc.*, vol. 9, no. 4, pp. 651–65, Apr. 2000.

[2] J. Kybic, P. Thevenaz, A. Nirkko, and M. Unser, "Unwarping of unidirectionally distorted EPI images," *IEEE Tr. Med. Im.*, vol. 19, no. 2, pp. 80–93, Feb. 2000.

[3] M. Unser, "Splines: A perfect fit for signal and image processing," *Signal Processing*, vol. 16, no. 6, pp. 22–38, Nov. 1999.

[4] G. Zubal, G. Gindi, M. Lee, C. Harrell, and E. Smith, "High resolution anthropomorphic phantom for Monte Carlo analysis of internal radiation sources," in *IEEE Symposium on Computer-Based Medical Systems*, 1990, pp. 540–7.
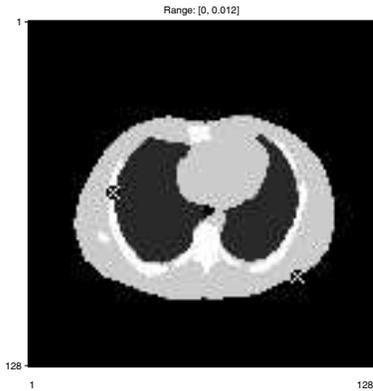
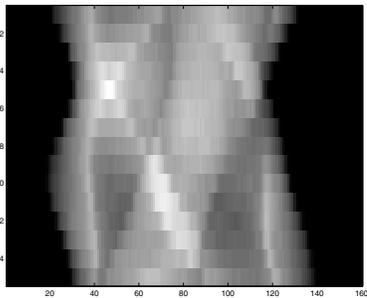**Fig. 1**. Image f, marked points are (100,90) and (32,64) respectively

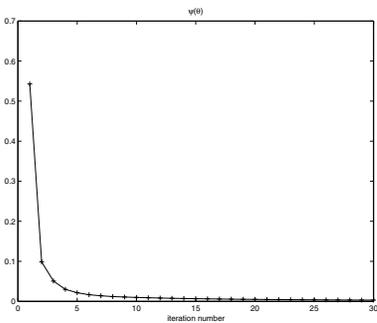**Fig. 2**. Projection $\{p_{\phi_m}\}_{m=1}^{15}$
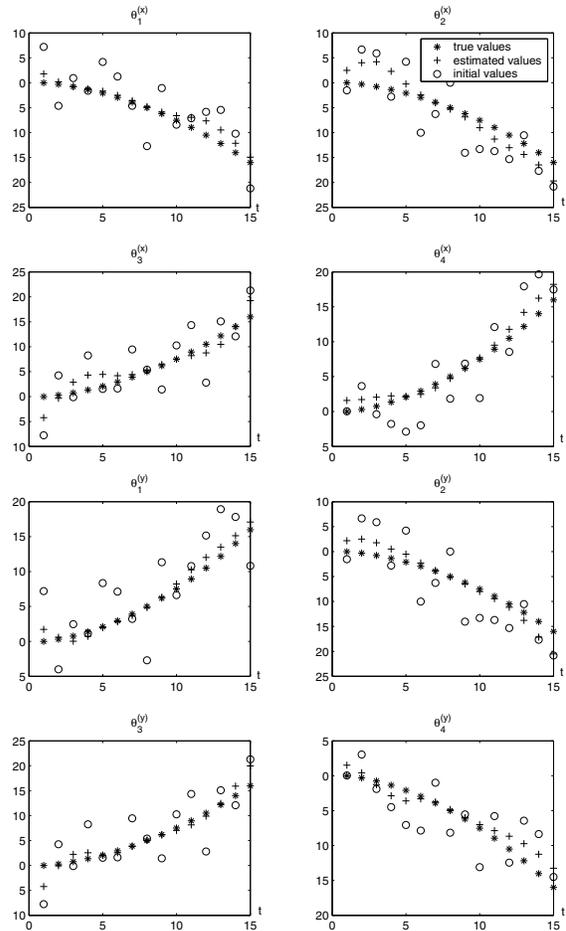

Fig. 3. Cost function

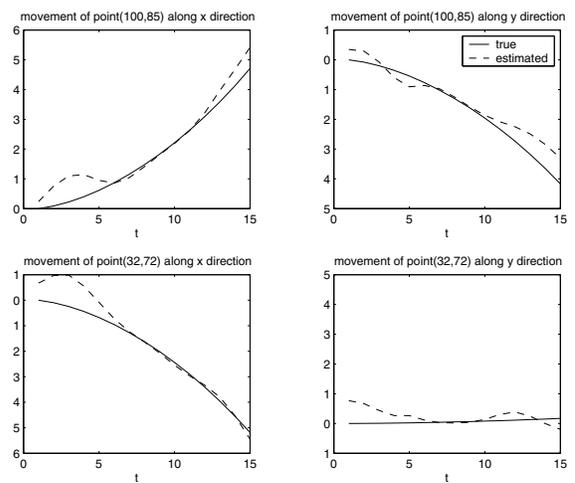**Fig. 4**. Comparison of the the estimated motion parameters with their true values

Fig. 5. Comparison of the true movement and the estimated movement of point(100,90),(32,64) as marked in Fig. 1