IOP PUBLISHING PHYSICS IN MEDICINE AND BIOLOGY

Phys. Med. Biol. **54** (2009) 981–992 doi:10.1088/0031-9155/54/4/011

Fluoroscopic tumor tracking for image-guided lung cancer radiotherapy

Tong Lin[1,2], Laura I Cerviño[1], Xiaoli Tang[1], Nuno Vasconcelos[3] and Steve B Jiang[1]

[1] Department of Radiation Oncology, University of California San Diego, La Jolla, CA 92093, USA
[2] Key Laboratory of Machine Perception (Ministry of Education), School of EECS, Peking University, Beijing 100871, People's Republic of China
[3] Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, CA 92093, USA

E-mail: sbjiang@ucsd.edu

Received 9 June 2008, in final form 28 November 2008
Published 16 January 2009
Online at stacks.iop.org/PMB/54/981

Abstract

Accurate lung tumor tracking in real time is a keystone to image-guided radiotherapy of lung cancers. Existing lung tumor tracking approaches can be roughly grouped into three categories: (1) deriving tumor position from external surrogates; (2) tracking implanted fiducial markers fluoroscopically or electromagnetically; (3) fluoroscopically tracking lung tumor without implanted fiducial markers. The first approach suffers from insufficient accuracy, while the second may not be widely accepted due to the risk of pneumothorax. Previous studies in fluoroscopic markerless tracking are mainly based on template matching methods, which may fail when the tumor boundary is unclear in fluoroscopic images. In this paper we propose a novel markerless tumor tracking algorithm, which employs the correlation between the tumor position and surrogate anatomic features in the image. The positions of the surrogate features are not directly tracked; instead, we use principal component analysis of regions of interest containing them to obtain parametric representations of their motion patterns. Then, the tumor position can be predicted from the parametric representations of surrogates through regression. Four regression methods were tested in this study: linear and two-degree polynomial regression, artificial neural network (ANN) and support vector machine (SVM). The experimental results based on fluoroscopic sequences of ten lung cancer patients demonstrate a mean tracking error of 2.1 pixels and a maximum error at a 95% confidence level of 4.6 pixels (pixel size is about 0.5 mm) for the proposed tracking algorithm.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Lung cancer radiotherapy is challenging even for patients with localized disease limited to the thorax. In-field recurrences are common, despite the use of concurrent chemotherapy and increasing doses of radiation. While tumor biology likely influences the aggressive clinical course of most lung cancers, targeting inaccuracies may also cause geographical miss, resulting in poor locoregional control with radiotherapy. A major source of targeting inaccuracy is from the respiration-induced lung tumor motion (Jiang 2006).

Two major classes of techniques have been developed to manage respiratory tumor motion. The main idea for the first class of techniques is to allow the tumor to move freely relative to the treatment beams and try to integrate the motion effect into the treatment plan (geometrical or dosimetric). These techniques include: (1) patient population based internal margin; (2) patient specific internal margin; (3) internal target volume (ITV) method; (4) IMRT optimization using a motion probability density function (PDF) (Trofimov *et al* 2005, Jiang 2006). The main idea for the other class of techniques is to 'freeze' the tumor motion relative to the treatment beams, which includes two categories: (1) control the tumor motion, using techniques such as *breath holding*, *forced shallow breathing* or *abdominal compression*; (2) allow free tumor motion but adjust the treatment equipment to maintain a constant target position in the beam's eye view when the beam is on, through *respiratory gating*, *beam tracking* or *couch-based motion compensation* (Jiang 2006).

The beam tracking technique follows the target dynamically with the radiation beam (Murphy 2004). It was first implemented in a robotic radiosurgery system (Adler *et al* 1999, Ozhasoglu *et al* 2000, Schweikard *et al* 2000, Murphy 2002, Murphy *et al* 2003). For linac-based radiotherapy, tumor motion can be compensated for using a dynamic multi-leaf collimator (MLC) (Keall *et al* 2001, 2005, Neicu *et al* 2003, Papiez 2003, Suh *et al* 2004, Papiez and Rangaraj 2005, Rangaraj and Papiez 2005, Webb 2005a, 2005b, Wijesooriya *et al* 2005, Neicu *et al* 2006). Beam tracking could be the best technique for stereotactic body radiation therapy (SBRT) due to its potentially high treatment efficiency and precision, although there are still many technical challenges, one of which is how to track the tumor location in real time with high precision.

Previous tumor tracking approaches can be roughly grouped into three categories: (1) deriving the tumor position based on external surrogates, such as the patient abdominal surface or lung volume (e.g., Jiang 2006). The relationship between the external surrogates and the internal tumor position can vary from day to day, and even during the treatment of the same day. Therefore, this approach is often considered to suffer from lack of accuracy. (2) Fluoroscopic tracking of radiopaque fiducial markers implanted inside or near the tumor (Shirato *et al* 2003, Tang *et al* 2007). The accuracy of this technology is better than 1.5 mm for tracking moving targets (Shirato *et al* 2000), which is much higher than the external surrogates approach. Another implementation of marker tracking is based on non-ionizing electromagnetic fields, using small wireless transponders implanted in human tissue (e.g., Balter *et al* 2005). However, no matter how marker tracking is realized, as long as the percutaneous marker implantation is involved, the clinical implementation of this technology in lung cancer radiotherapy is limited due to the risk of pneumothorax (Arslan *et al* 2002, Geraghty *et al* 2003). (3) Fluoroscopic tracking of the lung tumor without implanted fiducial markers. We have tested some conventional motion tracking methods, such as template matching, optical flow and active shape model (Cui *et al* 2007, Xu *et al* 2007, 2008) and achieved promising results when the tumor has reasonably high contrast and clear boundary in the images. However, these methods fail when the image quality is poor and the tumor is hard to identify in the images, which is often the case for lung cancer fluoroscopic images.

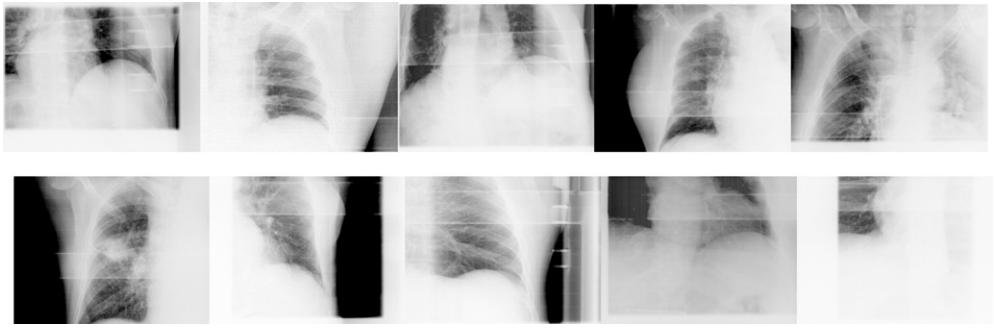

Figure 1. First frames of ten fluoroscopic image sequences used for testing the proposed tracking algorithm.

Tracking the lung tumor in fluoroscopic image sequences is a very challenging task because, compared to the colored video sequences that are often seen in the field of computer vision, the fluoroscopic image sequences have much less information (gray scale versus true color and less texture), and also in most cases the tumors do not have clear shapes. Additionally, the image quality is poor in many situations. Therefore, traditional tracking algorithms developed for computer vision applications may not be applied to fluoroscopic tumor images. In this paper, we propose a novel tumor tracking algorithm to solve this problem.

2. Methods and materials

2.1. Image data

To develop and evaluate the proposed algorithm, fluoroscopic image sequences for ten lung cancer patients were acquired at a rate of 15 frames per second using an on-board x-ray imaging (OBI) system (Varian Medical Systems, Palo Alto, CA, USA) and were used retrospectively. Figure 1 shows the first frames of these ten fluoroscopic image sequences. The average video length is about 40 s (i.e. 600 frames). For each patient, 15 s of fluoroscopic images (225 frames) at the beginning of the sequence are used as training data. The remaining data are used for testing the algorithm. The tumor positions are manually identified by expert observers to serve as the ground truth. In patients where the tumor was hard to identify visually in the fluoroscopic images, an anatomical structure nearby the tumor was considered instead of the tumor. Position of the tumor or anatomical structure (referred to as tumor hereafter) was identified using the center-of-mass of the tumor. The average diameter of the tumor or anatomical structure was 6.7 mm (range 4.8–13.9 mm), and they were mostly located in the middle or lower lobe. It was observed that the average motion range in the lateral and superior–inferior directions was 5.7 mm and 16.9 mm, respectively. All our algorithms are implemented on Matlab 7.3 platform.

2.2. Outline of the method

The proposed algorithm is based on the observation that the motion of some anatomic features in the images (called *surrogates*) may be well correlated to the tumor motion. The correlation between the tumor position and the motion pattern of surrogates can be captured by regression

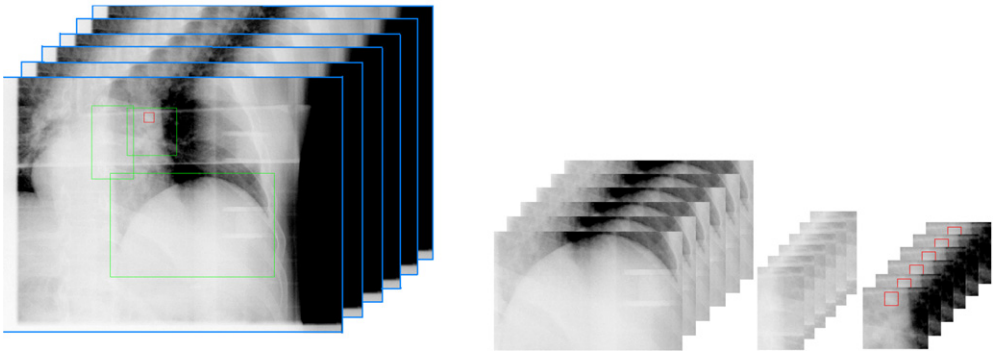

Figure 2. Left: the selection of surrogate ROIs (green rectangular regions—the tumor area is marked in red) on the first frame of the fluoroscopic image sequence. Right: the selected three surrogate ROIs.

analysis techniques. The proposed algorithm consists of four main steps: (1) selecting surrogate regions of interest (ROIs); (2) extracting spatiotemporal patterns from the surrogate ROIs; (3) establishing regression between the tumor position and the spatiotemporal patterns; and (4) predicting the tumor location using the established regression model. In a clinical setting, the first three steps would be performed using training image data before the treatment while the fourth step would be performed in real time using the image data acquired during treatment delivery. In this retrospective analysis, the first three steps are done using the first 225 frames in each sequence, and the fourth step is done with the remaining frames.

A few surrogate ROIs are created in the first frame of the training image sequence, which are assumed to be more or less correlated with the tumor motion. In the remaining training frames, the location of every surrogate ROI is the same while the image moves inside it. One ROI can be placed to contain the diaphragm, if visible in the image, which usually has a strong correlation with tumor motion in the superior–inferior (SI) direction (y-direction in this proposal). Other ROIs can contain any visible moving anatomic structures such as the lung boundary or even the tumor itself. In our preliminary experiments, only three surrogates are selected and placed on the diaphragm, the lung boundary and the tumor itself, as shown in figure 2. The diaphragm is closely related to the tumor motion in the y-direction, and the nearby lung wall correlates to the tumor motion in the x-direction (lateral). If the image quality is acceptable and the tumor itself has clear shape, the surrogate ROI placed on the tumor can also be helpful to predict the tumor position.

The features in the selected surrogate ROIs are not tracked directly. Instead, we use the principal component analysis (PCA) to map each surrogate ROI intensity map to a low-dimensional space to get a parametric representation (Alpaydin 2004). In our experiments, based on the obtained eigenvalues, we choose to use the first three principal components to represent a surrogate ROI. This 3D parametric representation of three surrogate ROIs is denoted as $(z_1, z_2, \ldots, z_9)$, where (z_1, z_2, z_3) is the parametric representation for the first ROI, (z_4, z_5, z_6) is for the second and (z_7, z_8, z_9) is for the last. Figure 3 shows the 3D representations of the three surrogate ROIs in PCA embedding space. We can see that a surrogate ROI is reduced to a point in 3D PCA space, which follows a well-defined trajectory. We can also see that the clearer the anatomic feature in the ROI, the better defined the trajectory. The location of a surrogate ROI selected on the first training frame is fixed for the remaining frames while the image moves inside it, leading to the trajectory in PCA space. This is illustrated in figure 4,

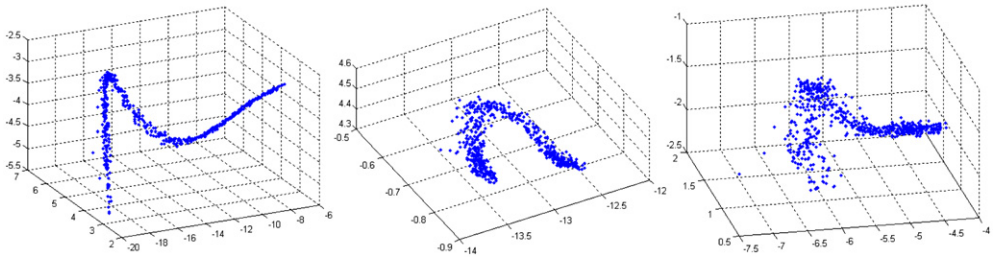

Figure 3. 3D representations of the three surrogate ROIs in the PCA embedding space. Left: ROI containing the diaphragm; middle: ROI containing the lung boundary and right: the ROI containing the tumor itself.

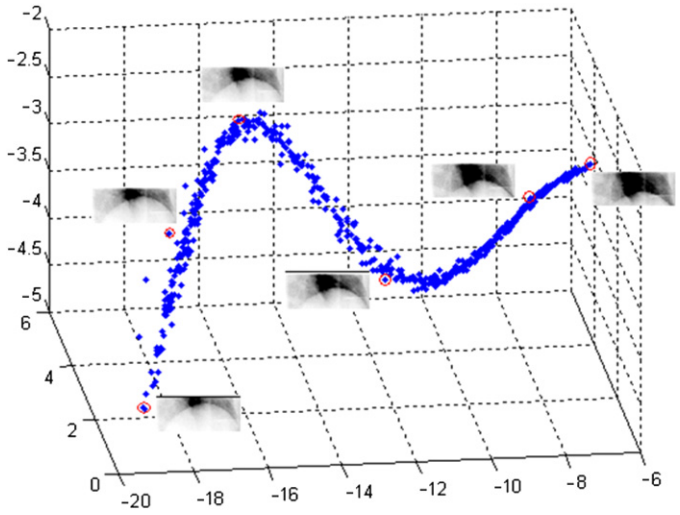

Figure 4. 3D embedding of the diaphragm images using PCA, and representative images are shown next to red circled points at different location in the 3D PCA space, representing different positions of the diaphragm.

where representative images are shown next to red circled points at different parts of the trajectory in 3D PCA space. As the diaphragm moves up and down, the corresponding point in 3D PCA space moves along the trajectory from one end to the other.

Figure 5 illustrates the curve representations of tumor positions (x, y) and 3D coordinates of three surrogate ROIs in PCA space. We can observe that there is a strong correlation between tumor positions and surrogate coordinates. The third step mentioned earlier is to build a regression model to predict the tumor position (x, y) based on the parametric representations of surrogate ROIs $(z_1, z_2, \ldots, z_9)$.

2.3. Principal component analysis (PCA)

PCA is a technique to reduce dimensionality with minimum loss of information. In this analysis, the dimension is the size (number of pixels) in the ROIs, that is, if an ROI consists of 50×50 pixels, the dimension is $n = 50 \times 50 = 2500$. Let $\mathbf{\Sigma} = \frac{1}{N} \sum_{i=1}^{N} (\mathbf{x_i} - \bar{\mathbf{x}})(\mathbf{x_i} - \bar{\mathbf{x}})^T$ be

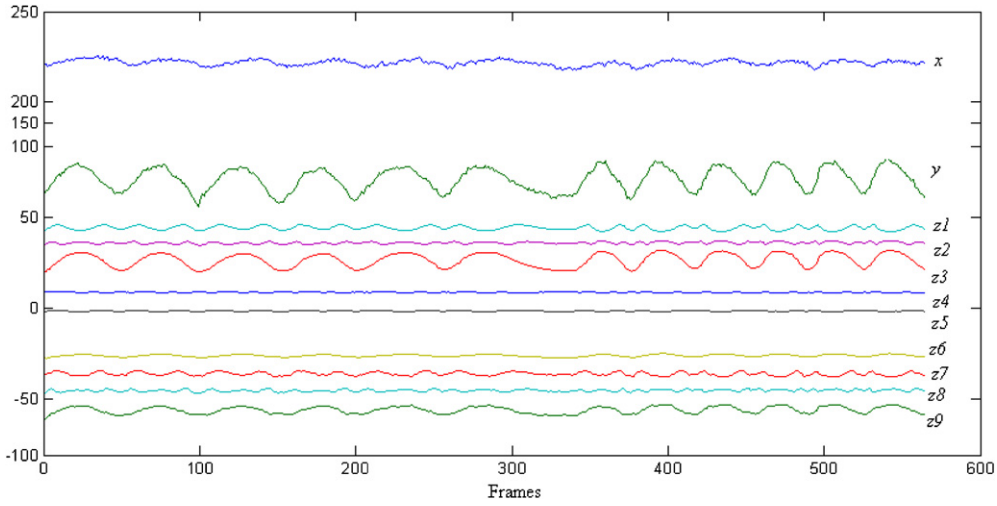

Figure 5. Curve representations of tumor positions (x, y) and 3D coordinates of three surrogates ($z_1, z_2, \ldots, z_9$) in PCA space.

the $n \times n$ data covariance matrix, where $\mathbf{x_i}$ is an n-dimensional data point (a vector containing the pixel intensities for the ROI at time point i), $\bar{\mathbf{x}}$ is the sample mean and N is the number of images in the sequence. Let $\mathbf{U} = [\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_d}]$ be the matrix with unit eigenvectors corresponding to the d largest eigenvalues $\Lambda = \mathrm{diag}[\lambda_1, \lambda_2, \ldots, \lambda_d]$ of the covariance matrix, such that $\mathbf{\Sigma U} = \mathbf{U\Lambda}$ (each eigenvector $\mathbf{u_i}$ has dimension $n \times 1$). Then we obtain the PCA projection $\mathbf{Y} = \mathbf{U^T\tilde{X}}$, where $\mathbf{\tilde{X}} = \mathbf{X}(\mathbf{I} - \mathbf{e}\mathbf{e}^T/N)$ is the zero mean version of the original data, $\mathbf{X}$ is an $n \times N$ matrix with the N vectors $\mathbf{x_i}$, and $\mathbf{Y}$ is a point in the PCA space representing the image in the surrogate ROI coordinates. Here $\mathbf{e}$ is a column vector of all ones. Geometrically, $\mathbf{U} = [\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_d}]$ forms a set of incomplete orthonormal basis for the data space, and the original data points are most spread out in these basis directions.

2.4. Regression analysis

In order to predict the tumor position, a regression model between the tumor position and the parametric representation of ROIs in PCA space needs to be built. Four regression methods are tested in this work, including linear regression, two-degree polynomial regression, artificial neural network (ANN) and support vector machine (SVM). Note that ANN and SVM can be used for both regression and classification. In regression, target outputs can be any real numbers, whereas in classification, target outputs have to be integers to indicate different classes.

2.4.1. Linear regression. Suppose we want to predict the target y from an n-dimensional input vector $z = [z_1, z_2, \ldots, z_n]$ with a linear model $y = a_0 + \sum_{i=1}^{n} a_i z_i$. We can set up a linear model in matrix forms: $\mathbf{y} = Z\boldsymbol{\alpha} + \boldsymbol{\varepsilon}$, where $\mathbf{y}$ is an m-by-1 vector of observations, $\mathbf{Z}$ is an $m \times p$ matrix of predictors, $\boldsymbol{\alpha}$ is a $p \times 1$ vector of parameters and $\boldsymbol{\varepsilon}$ is an $m \times 1$ vector of random disturbances. The parameter vector $\boldsymbol{\alpha}$ can be computed with the least-squares solution $\boldsymbol{\alpha} = (Z^T Z)^{-1} Z^T \mathbf{y}$ or with QR decomposition of $\mathbf{Z}$.

2.4.2. Two-degree polynomial regression. In this case we predict the target y from an n-dimensional input vector $z = [z_1, z_2, \ldots, z_n]$ with a two-degree polynomial model $y = a_0 + \sum_{i=1}^{n} a_i z_i + \sum_{i=1, j \geqslant i}^{n} a_{ij} z_i z_j$. This can be treated as a linear predictor where the dimension of the input is increased to $z = [z_1, z_2, \ldots, z_n, z_1 z_1, z_1 z_2, \ldots, z_1 z_n, z_2 z_2, z_2 z_3, \ldots, z_2 z_n, \ldots, z_n z_n]$. The procedure is then the same as with the linear model.

2.4.3. The artificial neural network (ANN). A typical neural network has three layers: input layer, hidden layer and output layer (Alpaydin 2004). The output values of the output neurons represent target labels (discrete numbers) for classification or target values (real numbers) for regression. In our regression application, there are two output neurons to provide prediction results of the tumor position (x and y coordinates). Each hidden neuron emits an output $y_j = f(w_{j0} + \sum_{i=1}^{n} z_i w_{ji})$, where $[z_1, z_2, \ldots, z_n]$ is the n-dimensional input sample, w_{ji} denotes the input-to-hidden weights at the hidden neuron j from the input neuron i and $f(\cdot)$ is a nonlinear activation function. A commonly used activation function is a sigmoid function $f(x) = 1/(1 + \exp(-ax))$, where a is a slope function. The relationship between an output neuron and hidden neurons can be defined similarly. The weights of this network can be initialized to zeros or to random numbers, and can be trained by the most general *backpropagation* method. Backpropagation attempts to minimize the square error between the desired outputs and the actual outputs, and the minimization problem can be efficiently solved by *stochastic gradient descent* algorithm.

2.4.4. Support vector machine (SVM). (Drucker *et al* 1997, Burges 1998, Cristianini and Shawe-Taylo 2000, Alpaydin 2004). Given a set of data points $\{(\mathbf{z_1}, y_1), (\mathbf{z_2}, y_2), \ldots, (\mathbf{z_n}, y_n)\}$, SVM can be used for classification ($y_i \in \{1, -1\}$) or regression ($y_i \in R^1$). In classification applications, SVM seeks to maximize the margin between two classes of samples

$$\max \frac{2}{\|\mathbf{w}\|^2} \text{ subject to } y_i(\mathbf{w}^{\mathbf{T}} z_{\mathbf{i}} + b) \geqslant 1, \qquad i = 1, 2, \ldots, n.$$

This is equivalent to the minimizing problem

$$\min \tfrac{1}{2} \|\mathbf{w}\|^2 \text{ subject to } y_i(\mathbf{w}^{\mathbf{T}} z_{\mathbf{i}} + b) \geqslant 1, \qquad i = 1, 2, \ldots, n.$$

In the case where the data is not linearly separable, the classification task can be solved by introducing some nonnegative slack variables

$$\min \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^{n} \xi_i \text{ subject to } y_i(\mathbf{w}^{\mathbf{T}} z_{\mathbf{i}} + b) \geqslant 1, \qquad \xi_{\mathrm{i}} \geqslant 0, \qquad i = 1, 2, \ldots, n,$$

where $C > 0$ is the penalty parameter of the error term. In the nonlinear case, SVM utilizes a nonlinear mapping $\phi(\cdot)$ to map the data into a higher dimensional space. The optimization problem is reduced to

$$\min \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^{n} \xi_i \text{ subject to } y_i(\mathbf{w}^{\mathbf{T}} \phi(z_{\mathbf{i}}) + b) \geqslant 1, \qquad \xi_{\mathrm{i}} \geqslant 0, \quad i = 1, 2, \ldots, n.$$

$K(z_{\mathbf{i}}, z_{\mathbf{j}}) = \phi(z_{\mathbf{i}})^T \phi(z_{\mathbf{j}})$ is called the kernel function, and four basic kernels are commonly used:

(1) Linear: $K(z_{\mathbf{i}}, z_{\mathbf{j}}) = z_{\mathbf{i}}^T z_{\mathbf{j}}$.

(2) Polynomial: $K(z_{\mathbf{i}}, z_{\mathbf{j}}) = \left(\gamma z_{\mathbf{i}}^T z_{\mathbf{j}} + r\right)^d, \gamma > 0$.

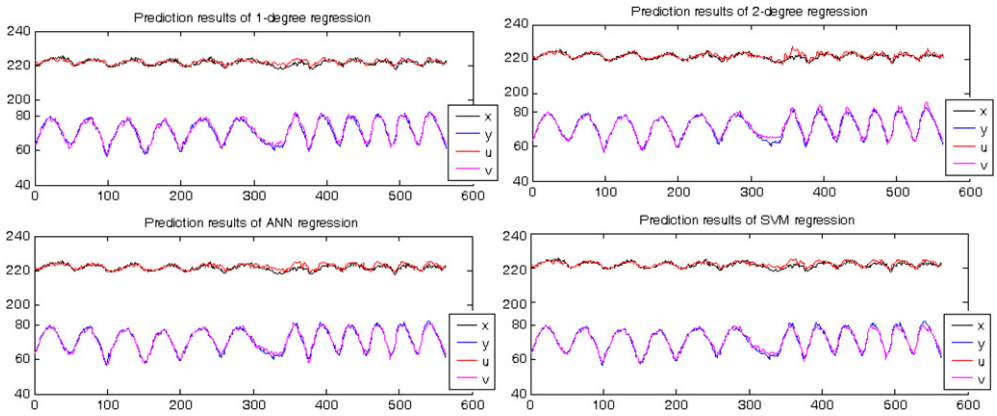

Figure 6. Tracking results of four regression methods for patient 1 as a function of frame index number. (x, y) are the ground truth (in pixels) and (u, v) are the predicted results.

(3) Radial basis function (RBF): $K(z_\mathbf{i}, z_\mathbf{j}) = \exp(-\|z_\mathbf{i} - z_\mathbf{j}\|^2)$, $\gamma > 0$.
(4) Sigmoid: $K(z_\mathbf{i}, z_\mathbf{j}) = \tanh\left(\gamma z_\mathbf{i}^T z_\mathbf{j} + r\right)$.

Here r, d, and γ are kernel parameters.

For regression tasks, the standard form of support vector regression is

$$\min_{\mathbf{w},b,\xi,\xi^*} \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{n}\xi_i + C\sum_{i=1}^{n}\xi_i^*,$$
$$\text{subject to } \mathbf{w}^\mathbf{T}\phi(z_\mathbf{i}) + b - y_i \leqslant \varepsilon + \xi_\mathbf{i},$$
$$y_i - \mathbf{w}^\mathbf{T}\phi(z_\mathbf{i}) - b \leqslant \varepsilon + \xi_\mathbf{i}^*,$$
$$\xi_\mathbf{i} \geqslant 0, \quad \xi_\mathbf{i}^* \geqslant 0, \quad i = 1, 2, \ldots, n.$$

3. Experimental results

In our experiments, we use Matlab neural network toolbox for ANN implementation and Libsvm for SVM implementation (Chang and Lin 2007, Hsu *et al* 2007). Figure 6 shows tracking results of the four regression methods for patient 1. We can see that all four methods perform better on the first 225 frames of training data than on the remaining testing data. The two-degree polynomial regression may suffer from the over-fitting problem, as the tracking errors for the testing data are much higher than that of the training data. For comparison, we compute the mean tracking errors, $\bar{e}$ and the maximum tracking error at a 95% confidence level, e_{95}. Table 1 summarizes the tracking results of the four methods on ten fluoroscopic videos. We can see that the performance of all four regression methods is about the same, with ANN regression performing slightly better than others at $\bar{e} = 2.1$ pixels and $e_{95} = 4.6$ pixels (the pixel size is about 0.5 mm). It is worth noticing that ANN is also more robust than other methods, with the maximum e_{95} of 6.5 pixels, while for the other three methods the maximum e_{95} at least doubles this value.

We observed the time performance of our algorithm. For this study, three ROIs were selected with a total of 9794 pixels. The average time needed to project the ROIs onto the PCA space and to obtain the tumor position was 8.2 ms in a PC with 1.80 GHz CPU and 1.97 GB

Table 1. Tracking results for the four regression methods. $\bar{e}$ is the mean localization error, and e_{95} is the maximum localization error at a 95% confidence level. Pixel size is 0.486 mm for patients 2, 8 and 10, and 0.518 mm for the rest.

Patient	Moving range (x, y) in pixels	One-degree linear regression $\bar{e}$, e_{95}	Two-degree linear regression $\bar{e}$, e_{95}	ANN regression $\bar{e}$, e_{95}	SVM regression $\bar{e}$, e_{95}
1	10, 25	1.7, 3.2	1.8, 3.9	1.9, 3.4	1.8, 3.6
2	7, 28	1.9, 3.7	1.9, 4.5	1.3, 2.9	1.7, 3.9
3	9, 29	2.4, 5.5	2.4, 7.2	2.3, 6.2	2.2, 5.8
4	25, 41	6.0, 12.9	6.1, 14.2	4.0, 6.5	6.9, 11.6
5	8, 28	1.4, 2.9	1.4, 3.4	1.7, 4.5	1.2, 2.5
6	11, 32	1.5, 3.2	1.2, 2.6	2.3, 4.5	0.9, 2.0
7	12, 37	1.6, 3.2	1.7, 3.9	1.8, 3.8	1.4, 3.0
8	10, 38	1.4, 2.8	1.6, 4.1	2.4, 5.7	1.5, 3.3
9	8, 24	1.5, 3.1	1.3, 2.6	1.3, 3.1	1.4, 2.7
10	11, 51	2.5, 6.8	2.6, 7.6	2.3, 5.8	2.8, 13.8
Average	11.1, 33.3	2.2, 4.7	2.2, 5.4	2.1, 4.6	2.2, 5.2

of RAM. The time required to find the estimated tumor position in a new acquired image is in the same order of magnitude as the time achieved by Moser *et al* (2008) of 6.4 ms, which is short enough for real-time tracking.

4. Discussion and conclusions

We have proposed a novel tumor mass tracking algorithm without implanted fiducial markers. Three surrogate ROIs are selected, and then the correlation between the tumor position and the surrogate representations can be captured by regression analysis techniques. Experimental results demonstrated the feasibility of the proposed tracking algorithm, which can be clinically used for respiratory gating or beam tracking in the future.

The tumor tracking technique proposed in this paper has been compared with a previous markerless tracking technique, the multiple template technique (Cui *et al* 2007). The tumor localization results obtained here are comparable with the results of template matching. We have tracked tumors which are visible to the observers or some visible anatomical structures when the tumor is not distinguishable. Although the performance of both methods is comparable, we believe that the method presented here will perform better than template matching when tracking a non-distinguishable tumor. In that case, we would expect the template matching technique, or any other methods that directly track the tumor, to fail; however, with the current approach, as long as good surrogates are selected in some of the ROIs, we would still be able to derive the tumor position.

The results presented in this paper indicate great promise for the development of algorithms that can track lung tumors in real time, in a non-invasive way, with high accuracy. Nevertheless, the achievement of these goals requires a number of further developments. First, the current work can be expanded to design regression functions capable of taking the smoothness of PCA trajectories (rather than single point locations) into account for the prediction of tumor positions. Second, surrogate ROIs should be selected automatically in an optimal way. Right now, we empirically selected three surrogate ROIs. There is no guarantee that the manually selected surrogates are the best ones for a particular patient, in terms of their

predictive power with respect to the tumor position. Since the motion of surrogate features is used to predict tumor motion, the quality of tumor tracking is very likely to improve if the ROIs are carefully selected.

There is also a need to determine the number of surrogates on which to base the prediction. We have used three ROIs in the paper. There is, however, no reason to believe that this is the optimal number of surrogates. In contrast, as is well known in statistics, improved predictions can typically be obtained with recourse to the law of large numbers, by taking averages over large numbers of measurements. In the context of fluoroscopic image tracking, this advises the adoption of regression functions based on a large number of surrogate ROIs. However, the amount of training data required to estimate the regression function grows exponentially with the number of variables of the regression problem. This number is the product of the number of surrogate ROIs with the number of PCA coefficients per ROI. When too large a number of ROIs are used, the training requirements become practically infeasible. This implies that there is a trade-off between the prediction accuracy and dimensionality problem. Finding the optimal value for this trade-off requires a careful selection of both the number of surrogates and the number of coefficients per surrogate. We have, so far avoided dimensionality problems by resorting to three surrogates and three PCA coefficients. Once again, there is no reason to believe that those are optimal numbers which will lead to the best possible predictions. A better strategy is to rely on statistical learning techniques to determine which coefficients are most informative for the prediction, and limit the regression problem to these.

The technique presented here finds the regression between tumor and surrogate representation based on knowledge of the tumor position ground truth of the training set. The clinical implementation of this method has to be well thought in order to reduce treatment time. The regression could potentially be performed before the first treatment from a fluoroscopic sequence uniquely acquired for this purpose. However, PCA is sensitive to the tumor size and position, so if the tumor changes size or relative position with respect to the chosen surrogates, this regression needs to be re-calculated. These issues will be addressed in the future work.

One may note that all the image sequences used in this work are only from anterior–posterior (AP) views. This is because tumors in images from other views, such as lateral view, are often much more difficult to identify even by expert observers. Thus for those images it is impossible to develop ground truth to test the developed algorithm. In the future work, we plan to use images with carefully implanted fiducial markers as the ground truth in order to test our algorithms in different views.

Finally one should note that due to the nature of fluoroscopic images, only 2D information of the tumor position can be extracted from them. Although superior–inferior motion is normally the largest one of all motions, and it is captured in fluoroscopic images, of the other two motions (lateral and anterior–posterior) only one will be captured by fluoro. In order to achieve complete 3D tracking, either a good correlation between the third coordinate and the planar image is obtained, for example from a 4D-CT scan of the patient, or another imaging method is used for the third coordinate. This will also be addressed in the future work.

Acknowledgments

The project is partially supported by an NCI grant (1 R21 CA110177 A 01A1), a National Science Foundation of China (NSFC) Grant 60775006, National Key Basic Research Program of China (NKBRP) Grant 2004CB318005. The authors would like to acknowledge Dr Ying Cui for her help.

References

Adler Jr J R, Murphy M J, Chang S D and Hancock S L 1999 Image-guided robotic radiosurgery *Neurosurgery* **44** 1299–306 (discussion 1306–7)

Alpaydin E 2004 *Introduction to Machine Learning* (Cambridge, MA: MIT Press)

Arslan S, Yilmaz A, Bayramgurler B, Uzman O, Nver E and Akkaya E 2002 CT- guided transthoracic fine needle aspiration of pulmonary lesions: accuracy and complications in 294 patients *Med. Sci. Monit.* **8** CR493–7

Balter J M, Wright J N, Newell L J, Friemel B, Dimmer S, Cheng Y, Wong J, Vertatschitsch E and Mate T P 2005 Accuracy of a wireless localization system for radiotherapy *Int. J. Radiat. Oncol. Biol. Phys.* **61** 933–7

Burges C J C 1998 A tutorial on support vector machines for pattern recognition *Data Mining Knowl. Discov.* **2** 121–67

Chang C C and Lin C J 2007 LIBSVM: a library for support vector machines http://www.csie.ntu.edu.tw/∼cjlin/papers/libsvm.pdf

Cristianini N and Shawe-Taylo J 2000 *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods* (Cambridge: Cambridge University Press)

Cui Y, Dy J G, Sharp G C, Alexander B and Jiang S B 2007 Multiple template-based fluoroscopic tracking of lung tumor mass without implanted fiducial markers *Phys. Med. Biol.* **52** 6229–42

Drucker H, Burges C J, Kaufman L, Smola A and Vapnik V 1997 Support vector regression machines *Advances in Neural Information Processing Systems-9* ed M C Mozer, M I Jordan and T Petsche (Cambridge, MA: MIT Press) pp 155–61

Geraghty P R, Kee S T, McFarlane G, Razavi M K, Sze D Y and Dake M D 2003 CT-guided transthoracic needle aspiration biopsy of pulmonary nodules: needle size and pneumothorax rate *Radiology* **229** 475–81

Hsu C W, Chang C C and Lin C J 2007 A practical guide to support vector classification http://www.csie.ntu.edu.tw/∼cjlin/papers/guide/guide.pdf

Jiang S B 2006 Radiotherapy of mobile tumors *Semin. Radiat. Oncol.* **16** 239–48

Keall P J, Joshi S, Vedam S S, Siebers J V, Kini V R and Mohan R 2005 Four-dimensional radiotherapy planning for DMLC-based respiratory motion tracking *Med. Phys.* **32** 942–51

Keall P J, Kini V R, Vedam S S and Mohan R 2001 Motion adaptive x-ray therapy: a feasibility study *Phys. Med. Biol.* **46** 1–10

Moser T, Biederer J, Nill S, Remmert G and Bendl R 2008 Detection of respiratory motion in fluoroscopic images for adaptive radiotherapy *Phys. Med. Biol.* **53** 3129–45

Murphy M J 2002 Fiducial-based targeting accuracy for external-beam radiotherapy *Med. Phys.* **29** 334–44

Murphy M J 2004 Tracking moving organs in real time *Semin. Radiat. Oncol.* **14** 91–100

Murphy M J, Chang S D, Gibbs I C, Le Q T, Hai J, Kim D, Martin D P and Adler J R Jr 2003 Patterns of patient movement during frameless image-guided radiosurgery *Int. J. Radiat. Oncol. Biol. Phys.* **55** 1400–8

Neicu T, Berbeco R, Wolfgang J and Jiang S B 2006 Synchronized moving aperture radiation therapy (SMART): improvement of breathing pattern reproducibility using respiratory coaching *Phys. Med. Biol.* **51** 617–36

Neicu T, Shirato H, Seppenwoolde Y and Jiang S B 2003 Synchronized moving aperture radiation therapy (SMART): average tumour trajectory for lung patients *Phys. Med. Biol.* **48** 587–98

Ozhasoglu C, Murphy M J, Glosser G, Bodduluri M, Schweikard A, Forster K M, Martin D P and Adler J R 2000 Real-time tracking of the tumor volume in precision radiotherapy and body radiosurgery—a novel approach to compensate for respiratory motion *Proc. 14th Int. Conf. on Computer Assisted Radiology and Surgery (CARS 2000), San Francisco, CA, USA* pp 691–6

Papiez L 2003 The leaf sweep algorithm for an immobile and moving target as an optimal control problem in radiotherapy delivery *Math. Comput. Modelling* **37** 735–45

Papiez L and Rangaraj D 2005 DMLC leaf-pair optimal control for mobile, deforming target *Med. Phys.* **32** 275–85

Rangaraj D and Papiez L 2005 Synchronized delivery of DMLC intensity modulated radiation therapy for stationary and moving targets *Med. Phys.* **32** 1802–17

Schweikard A, Glosser G, Bodduluri M, Murphy M J and Adler J R 2000 Robotic motion compensation for respiratory movement during radiosurgery *Comput. Aided Surg.* **5** 263–77

Shirato H *et al* 2003 Feasibility of insertion/implantation of 2.0-mm-diameter gold internal fiducial markers for precise setup and real-time tumor tracking in radiotherapy *Int. J. Radiat. Oncol. Biol. Phys.* **56** 240–7

Shirato H *et al* 2000 Physical aspects of a real-time tumor-tracking system for gated radiotherapy *Int. J. Radiat. Oncol. Biol. Phys.* **48** 1187–95

Suh Y, Yi B, Ahn S, Kim J, Lee S, Shin S and Choi E 2004 Aperture maneuver with compelled breath (AMC) for moving tumors: a feasibility study with a moving phantom *Med. Phys.* **31** 760–6

Tang X, Sharp G C and Jiang S B 2007 Fluoroscopic tracking of multiple implanted fiducial markers using multiple object tracking *Phys. Med. Biol.* **52** 4081–98

Trofimov A, Rietzel E, Lu H M, Martin B, Jiang S, Chen G T and Bortfeld T 2005 Temporo-spatial IMRT optimization: concepts, implementation and initial results *Phys. Med. Biol.* **50** 2779–98

Webb S 2005a The effect on IMRT conformality of elastic tissue movement and a practical suggestion for movement compensation via the modified dynamic multileaf collimator (dMLC) technique *Phys. Med. Biol.* **50** 1163–90

Webb S 2005b Limitations of a simple technique for movement compensation via movement-modified fluence profiles *Phys. Med. Biol.* **50** N155–61

Wijesooriya K, Bartee C, Siebers J V, Vedam S S and Keall P J 2005 Determination of maximum leaf velocity and acceleration of a dynamic multileaf collimator: implications for 4D radiotherapy *Med. Phys.* **32** 932–41

Xu Q, Hamilton R J, Schowengerdt R A, Alexander B and Jiang S B 2008 Lung tumor tracking in fluoroscopic video based on optical flow *Med. Phys.* **35** 5351–9

Xu Q, Hamilton R J, Schowengerdt R A and Jiang S B 2007 A deformable lung tumor tracking method in fluoroscopic video using active shape models: a feasibility study *Phys. Med. Biol.* **52** 5277–93