2008 Seventh International Conference on Machine Learning and Applications

Tumor Targeting for Lung Cancer Radiotherapy Using Machine Learning Techniques

Tong Lin[1, 2], Laura Cervino[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, China*
[3]*Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, CA 92093, USA*
E-mail: sbjiang@ucsd.edu

Abstract

Accurate lung tumor targeting in real time plays a fundamental role in image-guide radiotherapy of lung cancers. Precise tumor targeting is required for both respiratory gating and tracking. Gating is considered as the current state of the art for precise lung cancer radiotherapy, which irradiates the tumor when it moves into a predefined gating window. Tracking seems to be a next-generation technique, and it operates in a more aggressive fashion by following the tumor position with radiation beam in real time. Existing methods for gating and tracking often rely on observed motion patterns of external surrogates or implanted fiducial markers. However, external surrogates suffer from certain degrees of inaccuracy, and implanted fiducial markers are in limited uses due to the risk of pneumothorax. Therefore, direct tumor targeting techniques without implanting fiducial markers are desired. Previous studies in fluoroscopic markerless targeting are mainly based on template matching methods, which may fail when tumor boundary is unclear in fluoroscopic images. In this paper, we propose a novel framework of markerless gating and tracking based on machine learning algorithms. Specifically, gating is treated as a two-class classification problem, which is solved by Principal Component Analysis (PCA) and Artificial Neural Network (ANN). Further, we formulate the tracking problem as a regression task, which employs the correlation between the tumor position and nearby surrogate anatomic features in the image. Four regression methods were tested in this study: 1-degree and 2-degree linear regression, artificial neural network (ANN), and support vector machine (SVM). Finally, we demonstrate the superb performance of the proposed markerless gating and tracking algorithms on 10 fluoroscopic image sequences of 9 patients. For gating, the target coverage (the precision) ranges from 90% to 99%, with mean of 96.5%. For tracking, the mean localization error is about 2.1 pixels and the maximum error at 95% confidence level is about 4.6 pixels (pixel size is about 0.5 mm).

1. Introduction

A major source of tumor targeting inaccuracy in hypofractionated lung cancer radiotherapy is from the respiration induced lung tumor motion (Jiang 2006). For some patients, the respiration tumor motion can be clinically significant (at the order of greater than 2–3 cm). Therefore, motion management techniques play a critical role to reduce the incidence and severity of normal tissue complications and to increase local control through dose escalation.

There are two promising techniques for managing tumor motion in radiotherapy: gating and tracking. Gating is considered as the current state of the art technique and has been adopted in clinical practice by

978-0-7695-3495-4/08 $25.00 © 2008 IEEE
DOI 10.1109/ICMLA.2008.143

some cancer centers. The goal of an automated gating algorithm is to decide when the tumor moves into a predefined small gating window, thus turning the beam on accordingly. Compared with gating which appears to be a passive targeting method, tracking operates in an aggressive fashion by following the tumor position with radiation beam in real time. Although technically challenging, tracking seems a promising next generation technique for motion management.

Directly detecting the tumor mass in real-time during the treatment is usually very difficult and unreliable. Previous studies often utilize observed motion patterns of various *external surrogates* or *implanted fiducial markers* to identify the tumor position. However, it has been observed that external surrogates may suffer from the accuracy problem. On the other hand, implanted fiducial markers are in limited uses due to the risk of pneumothorax. Therefore, tumor targeting techniques that rely on neither the external surrogates nor the implanted fiducial markers are desired.

Previous studies in *fluoroscopic markerless targeting* are mainly based on template matching methods (Berbeco et al 2005, Cui et al 2007, Cui et al 2007). However, the performance of template matching greatly degrades when the tumor boundary is unclear in fluoroscopic images. Additionally, building a set of representative templates is not an easy task. In this paper, we propose a novel framework for markerless gating and tracking based on machine learning algorithms. Specifically, gating is treated as a two-class classification problem, and tracking is formulated as a regression task. We demonstrate the superb performance of markerless gating and tracking algorithms on 10 fluoroscopic image sequences of 9 patients.

2. Methods and materials

2.1 Image data

To develop and evaluate the proposed algorithms, fluoroscopic image sequences (15 frames per second) for 9 lung cancer patients were acquired using an on-board x-ray imaging (OBI) system (Varian Medical Systems, Palo Alto, CA, USA). One patient had tumors in both left and right lungs. Therefore, 10 fluoroscopic sequences were used retrospectively. The average video length is about 40 seconds (*i.e.* 600 frames). For each patient, 15 seconds of fluoroscopic images (225 frames) at the beginning of the sequence are used as training data. In general, 15 seconds of fluoroscopic images include 3 to 4 breathing cycles. The remaining data are used for testing the proposed gating and tracking algorithms. The tumor positions are manually identified by expert observers to serve as the ground truth. All our algorithms are implemented on Matlab 7.3 platform.

2.2 Gating algorithm

We first downsample the original images with size 1024-by-768 to obtain 512-by-384 images. Then, a rectangular region of interest (ROI) (with size 100-by-100 pixels) is cropped to contain the target tumor motion and to exclude other irrelevant regions. The tumor positions in the ROI images are manually identified to serve as ground-truth. A gating window is set at the end-of-exhale phase, and the goal is to design algorithms that can automatically determine if the tumor is in the gating window. Since the tumor motion in the superior-inferior direction (y-direction) dominates, in this work we only set a threshold in the y-direction to generate the gating window (see figure 1).

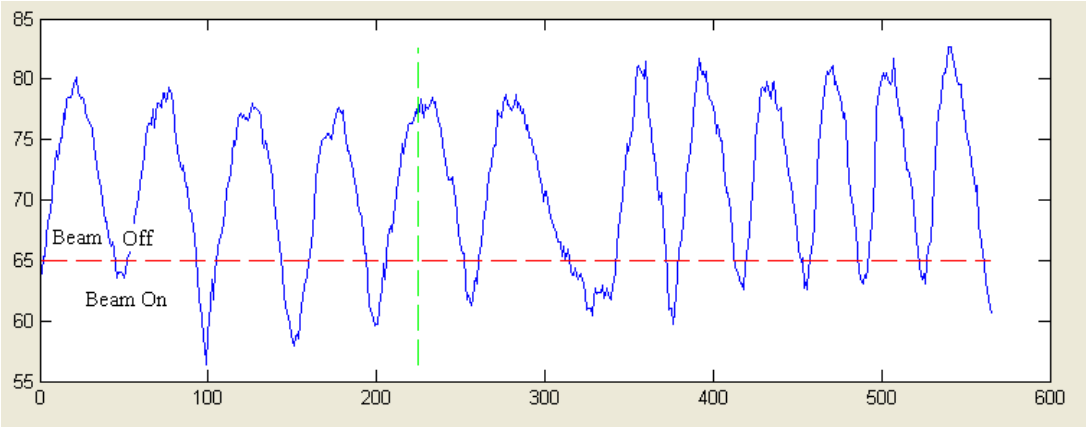

Figure 1: Divide data into training and testing subsets, and set a gating threshold in *y*-direction.

The first part of the proposed gating algorithm is *dimensionality reduction*, which aims at mapping the ROI images into a low-dimensional space. The purpose of dimensionality reduction is to reduce data amount and to extract significant features automatically. Five methods are tested in this work, including PCA (Turk et al 1991) and four nonlinear manifold learning methods (LLE, LTSA, Laplacian eigenmap, and diffusion maps) (Lin et al 2008). PCA is a classical linear dimensionality reduction method, which has been widely used in data analysis and pattern classification. However, PCA fails on curved nonlinear data sets since different branches of data points may be mixed to make distances disordered. Manifold learning methods seek to unfold the curved nonlinear data, while at the same time preserving certain properties (such as geodesic distances and angles). In this study, each 100-by-100 ROI image is considered as a 10,000-dimesional vector. In the training session, 225 frames of training data are firstly mapped into a 30-dimensional linear space by using PCA in order to reduce data amount greatly. Then these 30-dimensional data points are mapped into a 10-dimensional space by using all five dimensionality reduction methods further.

The second part is *classification* stage. In this work, artificial neural network (ANN) (Alpaydin 2004, Mitchell 1997) is used and compared with support vector machine (SVM) (Cui et al 2008). A neural network is an effective computational model for pattern classification and function approximation (or regression analysis). We employ a standard three-layer neural network with error backpropagation algorithm. Specifically, the input layer has 10 neurons to match the 10-dimensional input data after dimensionality reduction. There are 5 neurons in the hidden layer and only one neuron in the output layer.

2.3 Tracking algorithm

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 analysis techniques. The proposed algorithm consists of four main steps: 1) selecting surrogate windows; 2) extracting spatiotemporal patterns from the surrogate windows; 3) establishing regression between the tumor position and the spatiotemporal patterns; and 4) predicting tumor location using the established regression model. The first three steps are done using training image data before the treatment while the fourth step is done using the image data acquired during the treatment delivery in real time.

A few surrogate windows 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 a surrogate window is the same and the image

moves inside it. One window 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 windows 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. The diaphragm is closely related to the tumor motion in y-direction, and the nearby lung wall correlates to the tumor motion in x-direction (lateral). If the image quality is acceptable and the tumor itself has clear shape, the surrogate window placed on the tumor can also be helpful to predict the tumor position.

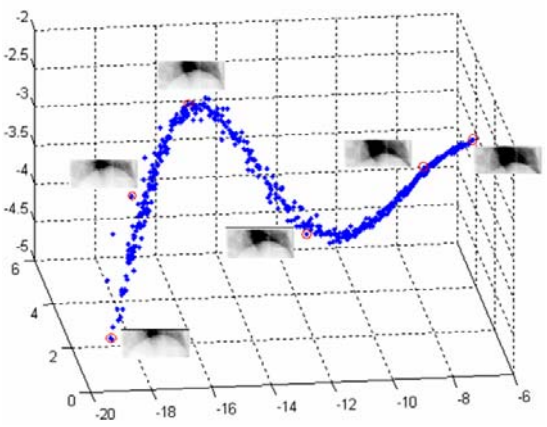

Figure 2: 3D embedding of the diaphragm images using PCA, and representative images are shown next to red circled points at different location in the 3D PCA space, representing different positions of the diaphragm.

The features in the selected surrogate windows are not tracked directly. Instead, we use the principal component analysis (PCA) to map each surrogate window to a low-dimensional space to get a compact coordinate representation (Turk and Pentland 1991). In our experiments, based on the eigenvalues, we choose to use the first three principal components to represent a surrogate window. The coordinate representation of three surrogate windows are denoted as (z_1, z_2, …, z_9), where (z_1, z_2, z_3) is for the first window, (z_4, z_5, z_6) is for the second, and (z_7, z_8, z_9) is for the last. Using the 3D representations in PCA embedding space, a surrogate window is reduced to a point, which follows a well defined trajectory. The more clear the anatomic feature in the window, the better defined the trajectory. The location of a surrogate window 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 2, 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.

The third step is to build regression model to predict the tumor position (x, y) based on the coordinate representations of surrogate windows (z_1, z_2, …, z_9).

3. Experimental results

3.1 Gating

In pattern classification, the performance can be measured with numbers of true positive (*tp*), false positive (*fp*), true negative (*tn*), and false negative (*fn*). In general, percentage measures can be used further, such as accuracy (=(*tp*+*tn*)/*all*), recall (=*tp*/(*tp*+*fn*)), and precision (=*tp*/(*tp*+*fp*)). The precision is also called target coverage (TC), which is of more clinical importance. For example, a TC of 80% means that 80% of the prescribed dose is delivered to the target and 20% delivered to the surrounding normal tissues. Figure 3 shows an example of the gating results.

We use CMU neural network implementation in C (Mitchell 1997) as our classification method, which is wrapped into a DLL that can be called from Matlab. For comparison, the support vector machine implementation Libsvm (Hsu et al 2007, Chang et al 2007) is employed. There are 10 image sequences used for testing the gating system. Each image sequence is divided into training data (175 frames), validation data (50 frames), and testing data (other

frames). We train the neural network 10 times to find the best initial network weights on the validation data set. Similarly, SVM needs to find the best combination of two parameters (RBF kernel parameter γ and penalty parameter C). The best SVM parameters are found by a complete grid-search on $C = 2^{-5}, 2^{-4} \cdots, 2^{14}, 2^{15}$ and $\gamma = 2^{-15}, 2^{-14} \cdots, 2^{14}, 2^{15}$.

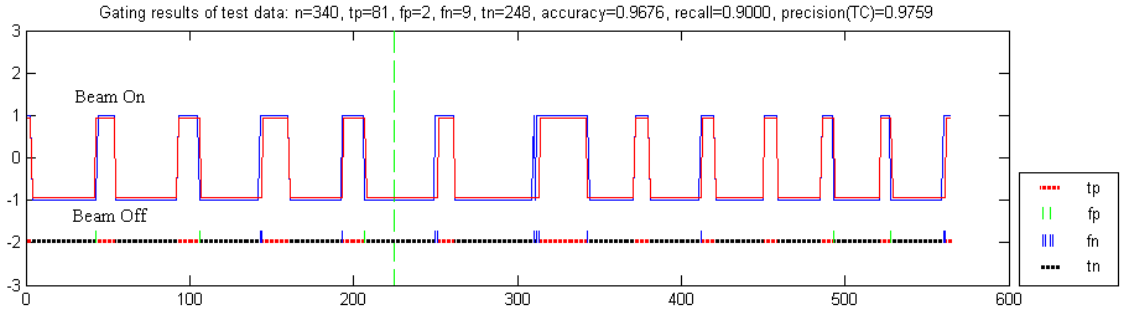

Figure 3: An example of gating signal showing four kinds of classification results: true positive (tp), false positive (fp), false negative (fn), and true negative (tn).

The experimental results of ANN and SVM show that on average PCA performs better than other 4 manifold learning methods, indicating that linear dimensionality reduction methods like PCA are suitable tools for the gating application. For one single patient, sometimes using certain manifold learning method can offer better precision (TC) than using PCA. On the other hand, ANN and SVM yield similar results if a same dimensionality reduction method is used. For example, using PCA and ANN yields an average result of accuracy **96.3**, recall **89.9**, and precision **97.8**. If using PCA and SVM, the average result is accuracy **94.9**, recall **84.9**, and precision **97.7**. From this comparison, we can see that ANN produces better accuracy and better recall than SVM.

Both ANN and SVM are implemented in C/C++. Training ANN is more efficient than training SVM since SVM needs to search the best combination of two parameters. The running time of training ANN 10 times is about 0.1521 second on an Intel Core 2 Duo 2.66 GHz Machine, while it takes 2.6332 seconds on training SVM to search the best parameters.

3.2 Tracking

In our experiments, we use Matlab neural network toolbox for ANN implementation and Libsvm for SVM implementation. Our results show that the 2-degree linear 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, $\overline{e}$ and the maximum tracking error at a 95% confidence level, e_{95}. The tracking results of four methods on 10 fluoroscopic videos show that the performance of all four regression methods is about the same, with ANN regression performing slightly better than others at $\overline{e}$ = 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.

4. Conclusion

In this work, we formulated the tumor targeting tasks (gating and tracking) as supervised learning problems, and the excellent experimental results were demonstrated on 10 fluoroscopic image sequences. In the future, we aims at improving the performance of our proposed tumor targeting system by following the three promising directions: (1) Powerful supervised learning algorithms with better accuracies can improve the performance of tumor targeting; (2) Temporal relationships should be incorporated to

make the gating and tracking results coherent in the temporal dimension; (3) Online learning methods can greatly adapt the tumor targeting system to each patient by allowing the correction of previous errors.

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.

References

Jiang S B 2006. Radiotherapy of mobile tumors. *Semin. Radiat. Oncol.* 16: 239-248.

Berbeco R L, Mostafavi H, Sharp G C, Jiang S B 2005. Towards fluoroscopic respiratory gating for lung tumours without radiopaque markers. *Phys Med Biol* 50: 4481-4490.

Cui Y, Dy J G, Sharp G C, Alexander B, Jiang S B 2007. Robust fluoroscopic respiratory gating for lung cancer radiotherapy without implanted fiducial markers. *Phys Med Biol* 52:741-755.

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(20): 6229-42.

Turk M and Pentland A 1991. Eigenfaces for Recognition,. *J. Cognitive Neuroscience*, 3(1): 71-86.

Lin T and Zha H 2008. Riemannian manifold learning. *IEEE Trans. Pattern Analysis and Machine Intelligence* 30(5):796-809.

Alpaydin E 2004. *Introduction to Machine Learning*. The MIT Press.

Mitchell T M 1997. Machine Learning. McGraw-Hill Press.

Cui Y, Dy J G, Sharp G C, Alexander B, Jiang S B 2008. Fluoroscopic gating without implanted fiducial markers for lung cancer radiotherapy based on support vector Machines. Manuscript.

Hsu C W, Chang C C and Lin C J 2007. A practical guide to support vector classification.

Chang C C and Lin C J 2007. LIBSVM : a library for support vector machines.