

Learning Methods for Lung Tumor Markerless Gating in Image-Guided Radiotherapy

Ying Cui
Dept. of Electrical and
Computer Engineering
Northeastern University
Boston, MA 02115
cui.yi@neu.edu

Jennifer G. Dy
Dept. of Electrical and
Computer Engineering
Northeastern University
Boston, MA 02115
jdy@ece.neu.edu

Gregory C. Sharp
Dept. of Radiation Oncology
Mass. General Hospital and
Harvard Medical School
Boston, MA 02114
gcsharp@partner.org

Brian M. Alexander
Dept. of Radiation Oncology
Mass. General Hospital and
Harvard Medical School
Boston, MA 02114
bmalexander@partners.org

Steve B. Jiang
Dept. of Radiation Oncology
University of California-San
Diego
La Jolla, CA 92093
sbjiang@ucsd.edu

ABSTRACT

In an idealized gated radiotherapy treatment, radiation is delivered only when the tumor is at the right position. For gated lung cancer radiotherapy, it is difficult to generate accurate gating signals due to the large uncertainties when using external surrogates and the risk of pneumothorax when using implanted fiducial markers. In this paper, we investigate machine learning algorithms for markerless gated radiotherapy with fluoroscopic images. Previous approach utilizes template matching to localize the tumor position. Here, we investigate two ways to improve the precision of tumor target localization by applying: (1) an ensemble of templates where the representative templates are selected by Gaussian mixture clustering, and (2) a support vector machine (SVM) classifier with radial basis kernels. Template matching only considers images inside the gating window, but images outside the gating window might provide additional information. We take advantage of both states and re-cast the gating problem into a classification problem. Thus, we are able to use the SVM classifier for gated radiotherapy. To verify the effectiveness of the two proposed techniques, we apply them on five sequences of fluoroscopic images from five lung cancer patients against the gating signal of manually contoured tumors as ground truth. Our five-patient case study shows that both ensemble template matching and SVM are reasonable tools for image-guided markerless gated radiotherapy with an average of approximately 95% precision in terms of delivered target dose at approximately 35% duty cycle.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

KDD '08, August 24–27, 2008, Las Vegas, Nevada, USA.
Copyright 2008 ACM 978-1-60558-193-4/08/08 ...\$5.00.

Categories and Subject Descriptors

H.2.8 [Database Applications]: Data Mining

General Terms

Algorithms, Design, Experimentation

Keywords

applied machine learning, clustering, classification, mixture model, support vector machine, image-guided radiotherapy

1. INTRODUCTION

Image Guided Radiation Therapy (IGRT) combines scanning and radiation equipment, to provide images of the patient's organs in the treatment position, at time of treatment, optimizing the accuracy and precision of the radiotherapy. Treatment errors related to respiratory organ motion may greatly degrade the effectiveness of conformal radiotherapy for the management of thoracic and abdominal lesions. This has become a pressing issue in IGRT. Respiratory gated radiotherapy holds promise to precisely deliver a lethal dose to the tumor, while minimizing the incidence and severity of normal tissue complications, for mobile tumors in the thorax and the abdomen [13]. Respiratory gating is a method of synchronizing radiation with respiration, during the imaging and treatment processes. In computer-driven respiratory-gated radiotherapy, a small plastic box with reflective markers is placed on the patient's abdomen. The reflective markers move during breathing, and a digital camera hooked up to a central processing unit monitors these movements in real time. A computer program analyzes the movements and triggers the scanner (simulation of treatment), or the treatment beam, always at the same moment of the respiratory cycle. With this technique, it is also possible to choose the respiratory phase: depending on its location, the tumor will be treated during inspiration or expiration so as to avoid exposure of critical organs. Figure 1 shows an imaging system mounted with two orthogonal x-ray tubes and fast amorphous silicon flat panels on the

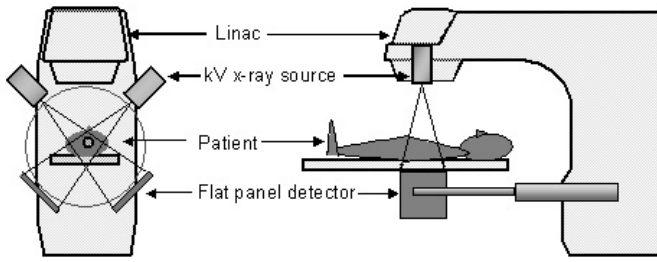


Figure 1: An example of an on-board imaging system for respiratory gated radiotherapy.

gantry of a medical linear accelerator (linac), which is used in respiratory gated radiotherapy.

Precise target localization in real-time is particularly important for gated radiotherapy. In gated radiotherapy, the radiation beam is held at a fixed position, and the beam is turned *ON* when the tumor is at the right position (in the gating window) and turned *OFF* when the tumor is outside the gating window. However, direct detection of the tumor mass in real-time during the treatment is often difficult. Various surrogates, both external and internal, are used to identify the tumor position. Depending on the surrogates used, we categorized the respiratory gating into external (optical) gating and internal (fluoroscopic) gating. During gated treatment, the internal or external surrogate signal is continuously compared against a pre-specified range of values, called the gating window. When the surrogate signal is within the gating window, a gating signal is sent to the linac to turn on the radiation beam.

External gating techniques rely on the correlation between tumor location and the external surrogates, such as markers placed on the patient’s abdomen [8, 19]. An example of an external surrogate is the small box with reflective markers placed on the patient’s abdomen described in the first paragraph above. The major weakness in external gating is the uncertainty in the correlation between the external marker position and internal target position [13]. Current internal gating uses internal tumor motion surrogates such as implanted fiducial markers, as established by the Hokkaido group [11, 18]. And it has been shown that internal surrogates can generate accurate gating signals. However, due to the risk of pneumothorax, the implantation of radiopaque markers in patients’ lungs will unlikely become a widely accepted clinical procedure [15, 1, 10]. Therefore, it is crucial to be able to perform accurately gated treatment of lung cancer without implanted markers.

The objective of this work is to generate precise and robust gating signals from fluoroscopic lung tumor images without markers via machine learning techniques. In a previous feasibility study, we applied a template matching method to generate gating signals for lung radiotherapy without implanted markers [3]. In this paper, we investigate two ways to improve markerless gated radiotherapy. Ensemble methods have been shown to increase the accuracy of weak classifiers [5]. Thus, we explore utilizing an ensemble of templates as our first approach. Secondly, template matching only considers images inside the gating window, but images outside the gating window might provide additional information for improving the precision in localizing the tumor. We assigned images inside the gating window as “ON” and those outside

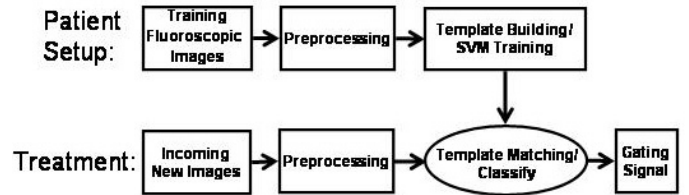


Figure 2: Block diagram for showing the process of the proposed clinical procedure for generating the gating signal.

as “OFF” classes and re-cast the gating problem into a classification problem. Then, as our second approach, we apply a support vector machine (SVM) classifier to gated radiotherapy. The framework of the proposed clinical treatment procedure is shown in Figure 2.

We verify the effectiveness of these approaches on five sequences of fluoroscopic images from five lung cancer patients against the gating signal of manually contoured tumors marked by a radiologist as ground truth. Our case study on these five patients shows that both ensemble template matching and SVM are reasonable tools for image-guided markerless gated radiotherapy with an average of approximately 95% precision in terms of delivered target dose at approximately 35% duty cycle. This paper presents a successful case study where machine learning and data mining algorithms are applied to a real world problem.

The rest of the paper is organized as follows. We start by describing the gating problem, how data is acquired and pre-processed in Section 2. Section 3 presents a detailed description of our ensemble template matching method. Section 4 re-frames the gating problem into a classification problem and provides a solution through SVM. To test our algorithms, Section 5 shows the evaluation metrics and validation results on five patient datasets. Finally, in Section 6, we conclude and outline future directions.

2. DATA ACQUISITION AND PRE-PROCESSING

In this section, we describe how data is acquired and pre-processed, the first two components during patient set-up in Figure 2.

2.1 Image Acquisition

In this study, the raw fluoroscopic image data we used comes from the system called the integrated radiotherapy imaging system (IRIS) [2], which consists of two pairs of gantry-mounted diagnostic x-ray tubes and flat panel imagers, shown in figure 3. The system can be used to acquire a pair of real-time orthogonal fluoroscopic images for lung tumor tracking.

2.2 Building Training Data

Before treatment, a sequence of orthogonal fluoroscopic images (approximately ten seconds in our experiments) are taken and used for patient set-up as training images. The tumor position in the gating window, where the treatment beam should be turned on, is identified either manually by a clinician or automatically by matching digitally reconstructed radiographs (DRRs) from the simulation 4D CT



Figure 3: The Integrated Radiotherapy Imaging System (IRIS), used as the hardware platform for the proposed gating technology in this paper.

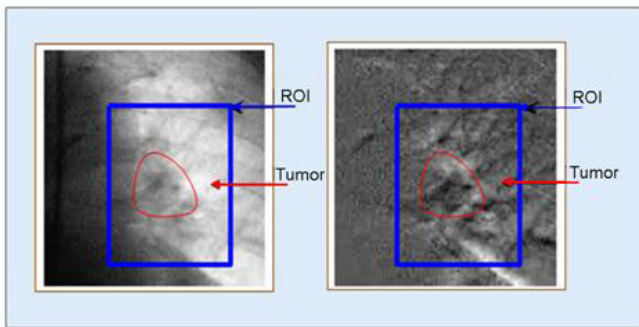


Figure 4: Tumor contour and the region of interest (ROI). Left: original fluoroscopic image. Right: motion-enhanced image.

scan [20]. In this investigation, the images were manually contoured. A rectangular region of interest (ROI) is then created in those images (see Figure 4). This ROI is set to be large enough to contain tumor motion in the training period.

Lung tumors move primarily due to the patient’s breathing. Typically, the gating window is set at the end-of-exhale (EOE) phase of the breathing cycle due to its longer duration and better stability. Figure 5 top left illustrates the tumor motion in the up-and-down direction as a function of time. Motion in the left-and-right direction is relatively very small (1–2mm) compared to the motion in the up-and-down direction (13–18mm). The tumor position is based on the centroid of the manually contoured image. Tumor in the lower positions correspond to the exhale phase, and those in the higher positions correspond to the inhale phase. A measure of radiation treatment efficiency is the gating duty cycle, the total time the beam is turned *ON* divided by the total time (beam is *ON* and *OFF*). Assuming that the de-

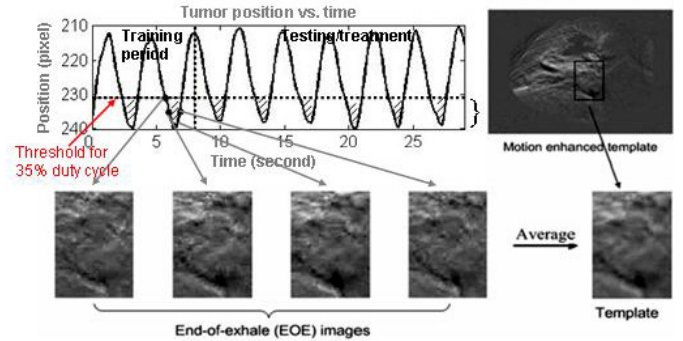


Figure 5: The top left figure is the breathing waveform represented by the tumor location. To the left of the vertical dotted line is the training period. To the right of the vertical line is the treatment or testing period. Under the horizontal dotted line (threshold corresponding to a given duty cycle) is the end of exhale phase. Bottom figures showed different end of exhale images during the training session, which are averaged to generate a single template.

sired gating duty cycle is given (in our experiments we set it at 35% and 50% which are typically used in radiotherapy), a corresponding threshold can be determined to define the gating window as shown by the horizontal line in Figure 5 top left. All the images in the gating window, i.e., with locations below this threshold, as shown in Figure 5, are labeled as EOE images.

2.3 Pre-Processing

We pre-process our images by first applying motion enhancement and then reduce the dimensionality by principal component analysis (PCA).

Motion Enhancement.

We apply a simple pre-processing technique called motion enhancement [12] on our training images. Given a sequence of images $I[t]$, where $t = 1, \dots, N$ is the sequence number. We compute the average image, $\sum_{t=1}^N I[t]/N$, and the motion-enhanced image (MEI) is the difference between the original image and the average image, $I[t] - \sum_{t=1}^N I[t]/N$. The intuition behind MEI is that the average captures the static structures and smears the moving structures, thus the difference will amplify the moving structures. Figure 4 shows the original fluoroscopic image of a tumor in an ROI together with a motion-enhanced view of it. We see that the tumor is clearer in the MEI.

Principal Components Analysis.

A typical EOE image has a size of 100×100 pixels. This leads to a dimensionality of size 10,000. To reduce the dimensionality, we apply principal component analysis (PCA) [14]. PCA finds a linear transformation, $Y = A^T X$, that projects the original high-dimensional data X with d dimensions to a lower dimensional data Y with q dimensions where $q < d$, such that the mean squared error between X and Y is as small as possible. X here is $d \times N$, where N is the number of data points, and A is a $d \times q$ matrix. The

solution is the transformation matrix A whose columns correspond to the q eigenvectors with the q largest eigenvalues of the data covariance. It also projects the high-dimensional dataset to the lower dimensional subspace in which the original dataset has the largest variance (i.e., restricts attention to those directions along which the scatter of the data points is greatest). PCA is applied only as a pre-processing step to clustering and to the support vector machine. It is not applied to template matching in this paper because during treatment, projecting each image to $\text{span}\{A\}$ is quite time consuming.

3. ENSEMBLE TEMPLATE MATCHING AND GAUSSIAN MIXTURE CLUSTERING

In this section, we describe our ensemble template matching method for gated treatment of lung cancer using fluoroscopy images. Before we proceed, we define our notations for clarity: R is the ROI in an incoming fluoroscopic image, T_i is the i th reference template (such as one of EOE templates in the EOE gating window), the sign \oplus represents correlation, s is the score used for generating the gating signal, i.e., $g = H(s - s_0)$, where s_0 is the threshold score, $H(x)$ is the Heaviside step function. $H(x) = 0$ for $x < 0$ and $H(x) = 1$ for $x \geq 0$. The gating signal $g = 1$ means beam *ON* while $g = 0$ means beam *OFF*.

In our previous work, we built a single EOE template by simply averaging all the motion-enhanced EOE training images, as shown in Figure 5. During treatment, we compute the correlation score between the reference template and each incoming MEI. Assuming that the image R and template T are of the same size $m \times n$, the normalized correlation coefficient (correlation score s) is defined as

$$s = \frac{\sum_m \sum_n (R_{mn} - \bar{R})(T_{mn} - \bar{T})}{\sqrt{(\sum_m \sum_n (R_{mn} - \bar{R})^2)(\sum_m \sum_n (T_{mn} - \bar{T})^2)}} \quad (1)$$

Here \bar{R} and \bar{T} are the average intensity values of the image R and template T , respectively. A high correlation score indicates that the incoming image is similar to the reference and gating is enabled.

In essence, the template building procedure is to generate the representative patterns defined in the EOE window. Then during matching, we use these representatives to recognize those images who have the same pattern and should be treated. Therefore, accurate representative templates are key points leading to accurate gating signals. We have noted that, only using the mean of EOE training images as the template, may not be enough. Our experiments showed that sometimes it leads to erratic gating signals. That is because the correlation score curve may not be smooth, causing the gating signal to be noisy at the transition regions.

3.1 Ensemble/Multiple Template Method

Inspired by the success of ensemble methods in improving the classification accuracies of weak base classifiers [5], we explore applying an ensemble or multiple template matching for gating. Hopefully, an ensemble of templates can smooth out the resulting gating signals. One way of generating an ensemble of templates is to set each EOE image as a template, and a correlation score is computed for each template. After a set of correlation scores are computed, it is necessary to choose an intelligent way to combine them to get

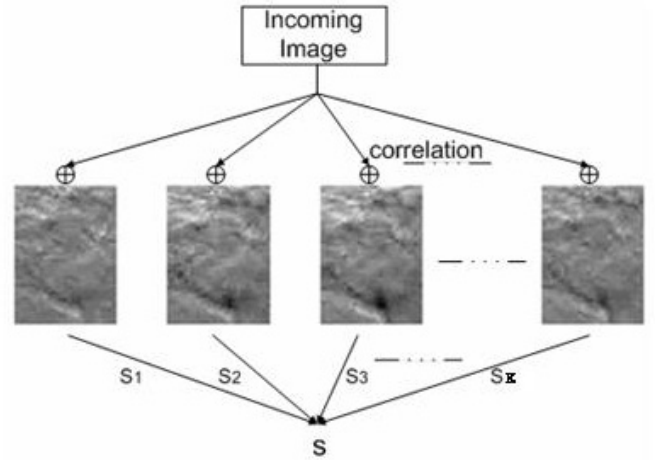


Figure 6: Ensemble/multiple template method. Here, each image is an end-of-exhale template. We match the incoming image with each template and get a set of correlation scores s_1, s_2, \dots, s_K . Then we apply a weighted average of these scores to generate the final correlation score s for gating.

a robust gating signal. There are several ways to combine the correlation scores (such as, taking the maximum score, taking the weighted average score). We found that applying a weighted average gave us the best results. We define these weights in subsection 3.3. This procedure is explained in Figure 6. Each image in the figure is an end-of-exhale template. We match the incoming image with each template and get a set of correlation scores s_1, s_2, \dots, s_K . Then we apply a weighted average of these scores to generate the final correlation score s for gating.

Although we can use the entire set of motion-enhanced images in the EOE gating window as reference templates, this approach is computationally expensive during treatment due to the need for computing several correlation scores. Many of the reference templates are very similar. Therefore, we want to find a way somewhere in between a single template method and using all the templates, which hopefully has the merits of both methods, i.e., computational efficiency and robustness. Here, instead of using all the templates, we would like to find a set of representative EOE templates. Ideally, these templates should carry all of the useful information of the original frames and discard the noise. Clustering methods are ideally suited to this task. Clustering algorithms group similar objects together and summarize each group with a representative template. We will use clustering to find a small set of templates and apply weighted averaging to combine the scores. We need to determine the number of templates and a method to find the clusters.

3.2 Finding Representative Templates by Clustering

To cluster the EOE templates, we apply Gaussian mixture clustering [16] to the PCA dimensionality reduced images. We denote the parameters of this model by Θ . In this model, we assume that each cluster comes from a multivariate Gaussian distribution, and our data (the image templates) come from a finite mixture of Gaussians. A Gaussian model as-

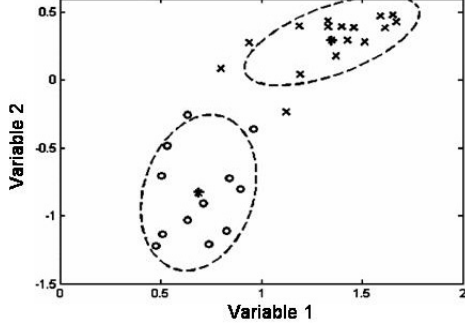


Figure 7: Scatter plot of our image data for patient 4 and 35% duty cycle in 2D with the clustering result. The "o" and "x" each represent different clusters, with the "*" and "*" symbol in bold and the covariances in ellipses.

sumes that there is a cluster template and the members of that cluster are variations of a template. Let X denote our data set which has d dimensions, and n denote the number of data points in X , we are trying to group the n data points into K clusters. Each of the K cluster is represented by its model parameters, $\theta_j = (\pi_j, \mu_j, \Sigma_j)$, where π_j is the prior probability (or mixture proportion) of cluster j , μ_j is the mean of cluster j and Σ_j is the covariance matrix of cluster j . More formally, we say that the probability of the data given the model is $P(X|\Theta) = \sum_{j=1}^K \pi_j P(X|\theta_j)$

$$P(X|\theta_j) = \frac{1}{\sqrt{(2\pi)^D |\Sigma_j|}} \text{EXP}\left\{-\frac{1}{2}(X - \mu_j)^T \Sigma_j^{-1} (X - \mu_j)\right\}$$

. To estimate the parameters, π_j , μ_j and Σ_j for each cluster, we apply the expectation-maximization algorithm (EM) [9]. The EM algorithm alternates through an expectation-step and a maximization-step until convergence. In the expectation-step, we estimate the cluster to which each image template belongs to given that the parameters (π_j , μ_j and Σ_j) are fixed. In the maximization-step, we estimate the parameters by maximizing the complete log-likelihood (the log-likelihood assuming we know the cluster memberships). The cluster means μ_j then become our representative templates.

To automatically determine the number of clusters, we apply the Bayesian information criterion (BIC) [17] score to penalize the log-likelihood function. We now maximize

$$\log P(X|\hat{\Theta}_{ML}) - (f/2)\log(n), \quad (2)$$

where f is the number of free parameters in the model. In our problem, we have

$$f = dK + (d(d+1)/2)K + (K-1) \quad (3)$$

to be estimated. We run the Gaussian mixture clustering from $K = 1$ to K_{max} ($K_{max} = 4$ in our experiments). Then pick the K with the largest BIC score. Note that if we do not add a penalty term, the log-likelihood increases as K increases. It can lead to a trivial result of picking K equal to the number of data samples (i.e., each data point will be

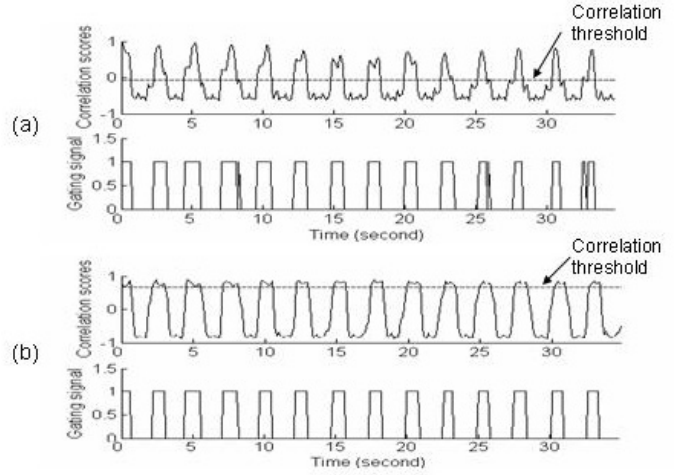


Figure 8: Results from different methods for an example patient. (a) single template method; (b) ensemble/multiple templates method with Gaussian mixture clustering. For each figure, the top curve is the correlation score and the bottom plot is the gating signal generated by the correlation score. Here we use 35% duty cycle.

considered as a cluster). A scatter plot of the clusters with their means and covariances is shown in figure 7.

3.3 Generating the Gating Signal

By the template clustering method, we can build a set of accurate representative templates. Accordingly, we will have a set of correlation scores for each incoming new image in the template matching step. Therefore, we need a way to combine the scores. As mentioned earlier, we are using a weighted average. Now the weights are just the prior probability π_j of each Gaussian mixture. We are performing a voting procedure where mixtures with more members have higher weights to vote for the final correlation score compared to mixtures with fewer members. We generate the gating signal based on the final correlation score. From the final correlation score of the training images, we determine a threshold that corresponds to the pre-set duty cycle. This threshold is then applied to the correlation scores calculated in real-time during treatment to generate the gating signal. When the score is above this threshold value, it indicates that the therapy beam should be enabled. Otherwise, the therapy beam should be turned off.

Figure 8 shows an example of gating signals generated by this ensemble template method and the previous single template method. We can see that the ensemble approach can achieve a reasonable gating signal and smooth correlation score, which coincides well with the smooth tumor motion. As demonstrated in this figure, through the voting process, the effect of errors caused by one template is compensated by the other templates. Thus the ensemble template method is less sensitive to noise compared to a single template.

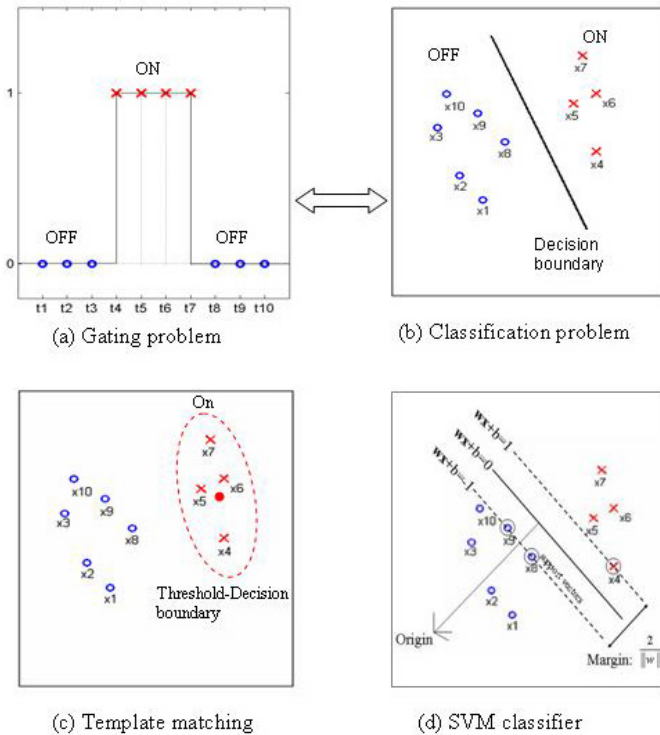


Figure 9: Re-cast the gating problem to a classification problem (a) and (b). (c) presents the decision boundary created by a single template matching and (d) displays decision boundary by an SVM classifier.

4. SUPPORT VECTOR MACHINES

Template matching as described in the previous section only look at images inside the gating window. However, instead of only using the images inside the gating window, can we also take advantage of images outside the gating window? Will this additional information be helpful to make decisions? This strategy is especially helpful for distinguishing images in the transition of the gating signal. We can measure how similar they are with the EOE images as well as how dissimilar they are from the non-EOE frames. Indeed, this can be viewed as a two-class classification problem. In this section, we re-cast the gating problem as a classification problem and present the support vector machine classifier as a solution.

Gating as a Classification Problem.

The goal for an automated method for gated radiotherapy is to decide when to turn the beam *ON* or *OFF*. This exclusivity condition provides us with a clue that the gating problem can actually be re-cast as a classification problem. We treat the image frames that correspond to a beam-on signal as one class and those that correspond to a beam-off signal as another class. For simplicity, we will call them beam-on and beam-off images. By this way, we reformulate the original gating problem into a two-class classification problem, as shown in Figure 9. In Figure 9a, ten time points on the gating signal are shown from time t_1 to t_{10} .

And each time point corresponds to a image frame, x_1 to x_{10} in Figure 9b. We represent each image frame as a vector, x_t . For this illustrative example, we project each x_1 to x_{10} onto two-dimensions as shown in Figure 1b. The set of images, $\{x_1, x_2, x_3, x_8, x_9, x_{10}\}$ are examples of beam-off images (class *OFF*) and $\{x_4, x_5, x_6, x_7\}$ are examples of beam-on images (class *ON*). The goal of classification is to build a classifier that outputs *ON* or *OFF* given a new input image x_t by learning parameters of this classifier from training examples such as x_1 to x_{10} .

The template matching method can be considered as a classifier that only takes advantage of the positive (*ON*) class. Correlation provides a measure of how close or similar each new image is from our template representing the *ON* class. The threshold is our decision boundary and turns the correlation score into a decision: scores higher than a threshold classify as *ON* and *OFF* otherwise. We select this threshold based on the desired duty cycle. In Figure 9c, a template would be the average of the *ON* examples shown in red dot in the center of the ellipse. And the threshold will be a fixed distance from this template. Note that a better way of automatically determining the decision boundary (threshold) is to consider both positive (*ON*) and negative (*OFF*) examples. In this paper, we build our classifier by looking at both *ON* and *OFF* training examples. In particular, we design a support vector machine (SVM) classifier to solve this gating problem. Figure 9d displays the decision boundary created by a linear SVM on this simple illustrative example data.

Support Vector Machine Classifier.

There are several possible classification algorithms for this task. Among them, SVM is one of the most popular learning methods for binary classification. SVM was originally designed by Vapnik [21]. It learns an optimal bound on the expected error, and finds an optimal solution as opposed to many learning algorithms that provides local optima (such as, neural networks [4]). The SVM objective is to find a boundary that maximizes the margin between two classes as well as separates them with the minimum empirical classification error. An SVM first projects instances into high dimensional space via kernels and then learns a linear separator that maximizes the margin between the two classes.

The SVM problem can be formulated as follows: suppose we have the training data $X = \{x_1, \dots, x_t\}$ and $\{y_1, \dots, y_t\}$ be the class labels of X . Without loss of generality, we assign class labels to take the value of either $+1$ or -1 . We want to have a large margin, and a small error penalty (slack variables ξ_i) for misclassifications, as shown in Figure 9d.

$$\begin{aligned} & \text{Minimize} && \|\bar{w}\|^2 + C(\sum_i \xi_i) \\ & \text{Subject to} && y_i(\bar{x}_i \cdot \bar{w} + b) \geq +1 - \xi_i \end{aligned} \quad (4)$$

Here, C is a user-defined parameter, where larger values mean higher penalty to errors. In addition, we apply the kernel trick to allow nonlinear decision boundaries. For the gating problem, we apply the radial basis function (RBF) kernel:

$$K(x, x') = \exp(-\gamma \|x - x'\|^2), \text{ for } (\gamma > 0).$$

During patient treatment (refer to Figure 2), each incoming image is pre-processed by projecting the pixels in the ROI to the reduced dimensional space by PCA as explained

in Section 2. The decision function will be of the following:

$$\begin{aligned}
 D(x) &= \operatorname{sgn}\left(\sum_{t=1}^n y_t \alpha_t \cdot K(x, x_t) + b\right), \\
 K(x, x_t) &= \exp(-\gamma \|x - x_t\|^2)
 \end{aligned}
 \tag{5}$$

Here y_t are the class labels for image vector t , and α_t and b are derived for a given C by solving the optimization problem described in equation 4, using quadratic programming [6]. The parameters γ and C of this function are learned during patient set-up or training. During treatment, these parameters are fixed. We create the gating signal based on the output of the decision function in equation 5 above with input y_t (the reduced-dimensional representation of our incoming fluoroscopy image at time t , i.e., the class label for image t).

5. EXPERIMENTS AND RESULTS

We collected fluoroscopic image data from five patients to evaluate the performance of the following two methods: (1) an ensemble template matching method where the representative templates are selected by Gaussian mixture clustering, and (2) a support vector machine (SVM) classifier with radial basis kernels. For each patient, we get a sequence of image frames by sampling at ten frames per second. Typically, each sequence contains 300–400 frames, corresponding to 30–40 seconds. A region of interest of around 100×100 pixels in size was selected to include the tumor positions at various breathing phases. The training period we used to build our templates consisted of two cycles, which corresponds to about 60–80 image frames.

To validate the algorithms, we compared our estimated gating signal with the reference gating signal. A radiation oncologist manually contoured the tumor in the first frame of the images. Then in the following frames, the contour was dragged to the correct places manually using a computer mouse by the radiation oncologist. The tumor centroid position in each image frame was calculated and used to generate the gating signal based on various duty cycles. Here, in our experiment, we used 35% and 50% duty cycles. Assuming t_0 is the total time (beam on and off), t_1 is the beam-on time (based on our estimation), t_2 is the correctly predicted beam-on time (true positive), we define the evaluation metrics: delivered target dose (TD) $TD = t_2/t_1$ and real duty cycle (DC) of the treatment $DC = t_1/t_0$.

5.1 Experiments by Ensemble Template

For the ensemble/multiple template matching method, we find the EOE training images in the first two to three breathing cycles (training period). We reduced the dimensions of these EOE images to two to three dimensions and perform clustering. We then computed the log-likelihood based on the Gaussian mixture model to get our BIC scores and determine the number of clusters. We obtained two or three (depends on which patient) clusters with this data. That means only two or three templates is enough for this application, which made our method more efficient compared to using all the EOE training images as templates. The final correlation score for a new incoming images is the weighted average of the scores from the reference templates with the prior probability of the mixture as weights.

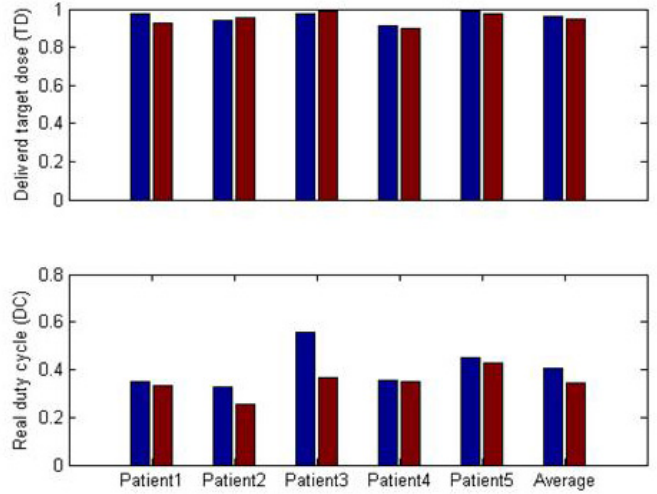


Figure 10: Experiment results in TD and DC for 35% proposed duty cycle. Blue bars: metric by SVM method. Red bars: metric by ensemble template matching method.

5.2 Experiments by Support Vector Machine

Image frames acquired during the set-up session is pre-processed and used to train the SVM classifier. We reduced the dimension from 100×100 to 1×50 for SVM. We then labeled the training images with +1 for beam-on images and -1 for beam-off images. We used LIBSVM [7] in our experiments. To train our SVM model, we applied a coarse-to-fine grid-search to determine the parameters γ and C for our radial basis function SVM (RBF-SVM) model. We found that trying exponentially growing sequences of γ and C is a practical method to identify the parameters ($\gamma = 10^{-10} 10^{-8} \dots 10^3$, $C = 10^{-5} 10^{-3} \dots 10^{10}$). Furthermore, to prevent overfitting in tuning the parameters, a ten-fold cross-validation procedure is performed on the training images to find a better model. Basically, the (γ, C) pair which provided the best ten-fold cross-validation accuracy on the training data is selected. We use the rest of the image data (data during treatment) as our testing samples. We pre-process the test data by PCA, and the predicted labels given by the SVM classifier serve as our estimated gating signals.

5.3 Results and Discussion

The experimental results are shown in Figures 10-11. In each figure, the upper figure shows the delivered target dose (TD) as a bar plot, with the SVM results in blue bars and the template-matching method results in red bars. Delivered target dose measures the true positive rate. The lower figure shows the real duty cycle (DC) in the same format, this is a measure of efficiency. For the proposed duty cycle of 35%, SVM achieves 95.8% average TD and 40.8% average DC, while the ensemble template matching method achieves 94.9% average TD and 34.7% average DC. When the proposed duty cycle equals 50%, SVM has an average TD of 98.4% and average DC of 53.1%, while ensemble template has an average TD of 97.6% and 49.5% DC.

Over all five patients we found that both methods are

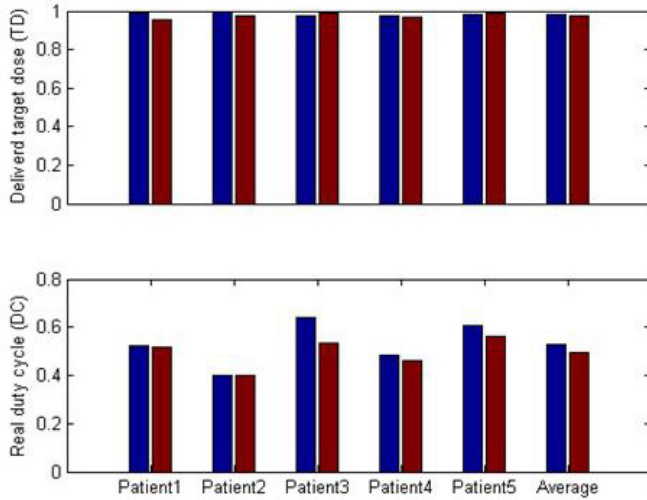


Figure 11: Experiment results in TD and DC for 50% proposed duty cycle. Blue bars: metric by SVM method. Red bars: metric by ensemble template matching method.

able to deliver most of the target dose correctly, both have a high TD, while SVM is more efficient than ensemble template. In general, SVM has an average DC that is 4%-6% longer. This is because SVM makes use of more information, both beam-on and beam-off images, while ensemble template matching only captures the information from the beam-on images. However, template matching has an advantage of detecting the aperiodic intra-fraction motion that occurs for some patients, although not seen in the current datasets. It is because the beam is only turned on when the correlation score is high, meaning the ROI of the incoming image is similar to the reference template and the tumor is at the right position. If the tumor drifts away or if shifts/rotations/deformations of the anatomy between patient setup and treatment happen, the correlation score should be always low and the beam will not be turned on. We then need to re-position the patient.

The time complexity for both methods is almost the same. We recorded the CPU time for executing both methods with a Pentium 4 Windows machine with 1G of RAM. The average CPU time for predicting the gating signal during testing is 0.06sec/frame by both methods. To obtain the gating signal for each time point, SVM needs to perform PCA and apply the reduced image vector as an input to the decision function, equation 7. For ensemble template matching, it needs to calculate the correlation between the original high dimensional motion-enhanced image with two or three templates (cluster means).

Figure 12 shows an example of the estimated gating signals. It plots the gating results by both methods. The top figure displays the predicted gating signal in red and reference signal in black for SVM. And the bottom figure shows the gating and reference signals for ensemble template matching. We can see that both gating signals coincide very well with the reference, and all errors occur at the edges.

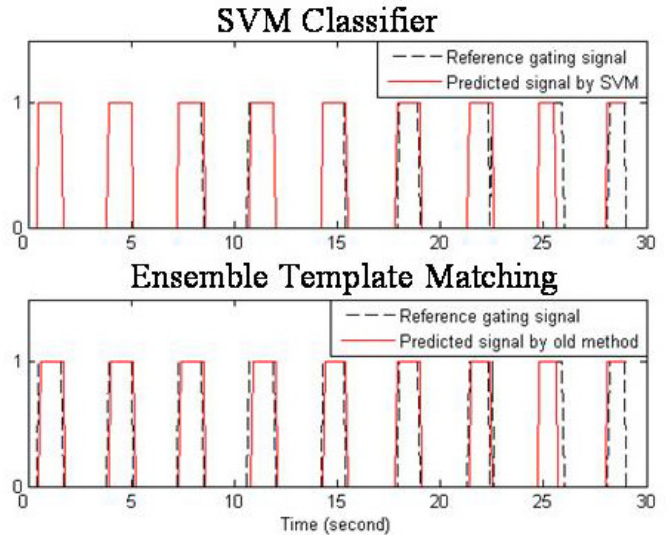


Figure 12: Example of estimated gating signals on patient 4 for proposed 35% duty cycle. Top: the predicted gating signal by SVM classifier. Bottom: the gating signal generated by ensemble template matching method.

6. CONCLUSIONS AND FUTURE WORK

This paper provides a case study where machine learning techniques have been successfully applied to an important real-world problem: gated radiotherapy. Working closely with the domain expert, we carefully selected the appropriate machine learning and data mining tools in developing the ensemble/multiple template matching method. Through our collaboration, we also provided our domain expert with a different view of the gating problem and re-cast it as a classification problem. Our study showed the feasibility of solving the gating problem by classification techniques. This provides us with wider recourses for gated radiotherapy. We can try other classification techniques, such as Bayesian classifier, neural networks, hidden Markov model, in our future work.

For our next step, we will (1) test the algorithms using more and longer patient data, (2) find a better way to get reference gating signal for validation, and (3) evaluate the dosimetric consequence of the current error level to see if there is a need to further lower the error rates. Then, we will consider clinical implementation of our methods.

Acknowledgments

The project is partially supported by an NCI grant (1 R21 CA110177 A 01A1) and by NSFcareer IIS-0347532, as well as by Varian Medical Systems, Inc.

7. REFERENCES

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

- [2] R. I. Berbeco, S. B. Jiang, G. C. Sharp, G. T. Chen, H. Mostafavi, and H. Shirato. Integrated radiotherapy imaging system (iris): design considerations of tumour tracking with linac gantry-mounted diagnostic x-ray systems with flat-panel detectors. *Phys Med Biol*, 49(2):243–55, 2004.
- [3] R. I. Berbeco, H. Mostafavi, G. C. Sharp, and S. B. Jiang. Towards fluoroscopic respiratory gating for lung tumours without radiopaque markers. *Phys Med Biol*, 50(19):4481–90, 2005.
- [4] C. M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, 1995.
- [5] L. Breiman. Bagging predictors. *Machine Learning*, 26(2):123–140, 1996.
- [6] S. A. J. Burges, C J C and B. Scholkopf. *Advances in Kernel Methods - Support Vector Learning*. MIT Press, Cambridge, USA, 1999.
- [7] C. C. Chang and C. J. Lin. Libsvm: a library for support vector machines. 2003. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [8] Q. Chen, M. Weinhaus, F. Deibel, J. Ciezki, and R. Macklis. Fluoroscopic study of tumor motion due to breathing: facilitating precise radiation therapy for lung cancer patients. *Med Phys*, 28(9):1850–6, 2001. TY - JOUR.
- [9] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal Royal Statistical Society, Series B*, 39(1):1–38, 1977.
- [10] P. R. Geraghty, S. T. Kee, G. McFarlane, M. K. Razavi, D. Y. Sze, and M. D. Dake. Ct-guided transthoracic needle aspiration biopsy of pulmonary nodules: needle size and pneumothorax rate. *Radiology*, 229(2):475–81, 2003.
- [11] T. Harada, H. Shirato, S. Ogura, S. Oizumi, K. Yamazaki, S. Shimizu, R. Onimaru, K. Miyasaka, M. Nishimura, and H. Dosaka-Akita. Real-time tumor-tracking radiation therapy for lung carcinoma by the aid of insertion of a gold marker using bronchofiberscopy. *Cancer*, 95(8):1720–7., 2002.
- [12] K. R. Jain, R and B. G. Schunck. *Machine Vision*. New York: McGraw-Hill, 1995.
- [13] S. B. Jiang. Radiotherapy of mobile tumors. *Semin Radiat Oncol*, 16(4):239–48, 2006.
- [14] I. T. Jolliffe. *Principal Component Analysis*. Berlin: Springer, 2002.
- [15] F. Laurent, V. Latrabe, B. Vergier, and P. Michel. Percutaneous ct-guided biopsy of the lung: comparison between aspiration and automated cutting needles using a coaxial technique. *Cardiovasc Intervent Radiol*, 23(4):266–72, 2000.
- [16] G. J. McLachlan and K. E. Basford. *Mixture Models, Inference and Applications to Clustering*. Marcel Dekker, New York, 1988.
- [17] G. Schwarz. Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464, 1978.
- [18] Y. Seppenwoolde, H. Shirato, K. Kitamura, S. Shimizu, M. van Herk, J. V. Lebesque, and K. Miyasaka. Precise and real-time measurement of 3d tumor motion in lung due to breathing and heartbeat, measured during radiotherapy. *Int J Radiat Oncol Biol Phys*, 53(4):822–34, 2002.
- [19] H. Shirato, T. Harada, T. Harabayashi, K. Hida, H. Endo, K. Kitamura, R. Onimaru, K. Yamazaki, N. Kurauchi, T. Shimizu, N. Shinohara, M. Matsushita, H. Dosaka-Akita, and K. Miyasaka. 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(1):240–7., 2003.
- [20] X. Tang, G. S. Sharp, and S. B. Jiang. Patient setup based on lung tumor mass for gated radiotherapy. *Med. Phys. (abstract)*, 33:2244, 2006.
- [21] V. Vapnik. *The nature of statistical learning theory*. Springer, New York, 1995.