Published in final edited form as:

Phys Med Biol. 2012 August 7; 57(15): 4771–4786. doi:10.1088/0031-9155/57/15/4771.

Real-Time Tumor Motion Estimation Using Respiratory Surrogate via Memory-Based Learning

Ruijiang Li^{1,*}, John H Lewis², Ross I Berbeco², and Lei Xing¹

¹Department of Radiation Oncology, Stanford University, Stanford, CA, USA

²Department of Radiation Oncology, Brigham and Women's Hospital, Dana-Farber Cancer Institute and Harvard Medical School, Boston, MA, USA

Abstract

Respiratory tumor motion is a major challenge in radiation therapy for thoracic and abdominal cancers. Effective motion management requires an accurate knowledge of the real-time tumor motion. External respiration monitoring devices (optical, etc.) provide a noninvasive, non-ionizing, low-cost, and practical approach to obtain respiratory signal. Due to the highly complex and nonlinear relations between tumor and surrogate motion, its ultimate success hinges on the ability to accurately infer the tumor motion from respiratory surrogates. Given their widespread use in the clinic, such a method is critically needed. We propose to use a powerful memory-based learning method to find the complex relations between tumor motion and respiratory surrogates. The method first stores the training data in memory and then finds relevant data to answer a particular query. Nearby data points are assigned high relevance (or weights) and conversely distant data are assigned low relevance. By fitting relatively simple models to local patches instead of fitting one single global model, it is able to capture highly nonlinear and complex relations between the internal tumor motion and external surrogates accurately. Due to the local nature of weighting functions, the method is inherently robust to outliers in the training data. Moreover, both training and adapting to new data is performed almost instantaneously with memory-based learning, making it suitable for dynamically following variable internal/external relations. We evaluated the method using respiratory motion data from 11 patients. The data set consists of simultaneous measurement of 3D tumor motion and 1D abdominal surface (used as the surrogate signal in this study). There are a total of 171 respiratory traces, with an average peak-to-peak amplitude of ~15 mm and average duration of ~115 seconds per trace. Given only 5-second (roughly one breath) pretreatment training data, the method achieved an average 3D error of 1.5 mm and 95th percentile error of 3.4 mm on unseen test data. The average 3D error was further reduced to 1.4 mm when the model was tuned to its optimal setting for each respiratory trace. In one trace where a few outliers are present in the training data, the proposed method achieved an error reduction of as much as ~50% compared with the best linear model (1.0 mm versus 2.1 mm). The memory-based learning technique is able to accurately capture the highly complex and nonlinear relations between tumor and surrogate motion in an efficient manner (a few ms per estimate). Furthermore, the algorithm is particularly suitable to handle situations where the training data are contaminated by large errors or outliers. These desirable properties make it an ideal candidate for accurate and robust tumor gating/tracking using respiratory surrogates.

* Author to whom correspondence should be addressed: Ruijiang Li, Ph.D., Department of Radiation Oncology, Stanford University School of Medicine, 875 Blake Wilbur Drive, Stanford, CA 94305-5847, Telephone: (650) 725-6711, Fax: (650) 498-5008, rli2@stanford.edu.

Keywords

respiratory tumor motion; surrogate; memory-based learning; locally weighted regression; real-time

1. Introduction

Respiratory tumor motion is a major challenge in cancer radiotherapy. Because tumors can move significantly (~2–3 cm) during respiration, it is very difficult to deliver sufficient radiation dose without damaging the surrounding healthy tissue (Jiang, 2006a). A promising solution to this problem is respiratory gating or tumor tracking (Keall, et al., 2006). For either technique to work effectively, the location of the tumor must be known, with high precision and in real time.

The current state of the art for real-time tumor localization is based on fiducial markers implanted inside or near the tumor. These can be either radiopaque metal markers tracked fluoroscopically (Shirato, et al., 2003; Murphy, 2004; Tang, et al., 2007) or small wireless transponders tracked using non-ionizing electromagnetic fields (Balter, et al., 2005). While the accuracy of marker tracking is clinically sufficient, the implantation procedure is invasive and may cause serious side effects such as pneumothorax (Arslan, et al., 2002), etc if performed percutaneously. Another issue with fluoroscopic tracking is the associated imaging dose, which can lead to an increased risk of secondary cancer or genetic defects (Sachs, et al., 2005; Cardis, et al., 2007). Due to these and other risks and limitations, tumor tracking with implanted fiducial markers is not widely accepted in the clinic.

An attractive approach to localizing tumors indirectly is through external respiratory surrogates, such as patient skin surface or tidal volume (Bert, et al., 2005; Jiang, 2006b). For instance, the real-time position management (RPM) system (Varian Medical Systems, Inc., Palo Alto, CA, USA) has been extensively implemented and widely adopted around the world (Mageras, et al., 2005). These external respiratory surrogates are particularly appealing because: 1) they are noninvasive to the patient; 2) they do not involve ionizing radiation; 3) they are inexpensive and convenient to use. Due to the highly complex and nonlinear relations between tumor and surrogate motion, its ultimate success hinges on the ability to accurately infer the tumor motion from respiratory surrogates. Given their widespread use in the clinic, such a method is critically needed.

Most early studies have focused on investigating the internal/external correlations (Hoisak, et al., 2004; Chi, et al., 2006; Yan, et al., 2006; Ionascu, et al., 2007; Wu, et al., 2008). More recently, some work has been done in actively seeking to decode the functional relations between the internal organ motion and external surrogates. The approaches include physical respiratory models (Li, et al., 2009), linear or polynomial regression models (Kanoulas, et al., 2007; Seppenwoolde, et al., 2007; Zhang, et al., 2007; Ruan, et al., 2008; K. T. Malinowski, et al., 2010; K. Malinowski, et al., 2012), or more sophisticated methods such as support vector machines (D'Souza, et al., 2009), etc. However, they either fall short in providing sufficiently accurate or robust estimates of the tumor motion, or due to their high complexities, are difficult and time consuming to adapt or re-optimize whenever new measurement data become available and the internal/external relation needs to be rebuilt on the fly.

In this paper, we present a memory-based learning approach to the respiratory tumor motion estimation problem. The method first stores the training data in memory and then finds relevant data to answer a particular query. Nearby data points are assigned high relevance

(or weights) and conversely distant data are assigned low relevance. By fitting relatively simple models to local patches instead of fitting one single global model, it is able to capture highly nonlinear and complex relations between the internal tumor motion and external surrogates accurately. Due to the local nature of weighting functions, it is inherently robust to outliers in the training data. Furthermore, both training and adapting to new data is performed almost instantaneously with memory-based learning. These properties make it a suitable candidate for accurate and robust tumor gating/tracking approaches using external respiratory surrogates. Similar approaches have been adopted to solve a related but different problem of predicting future respiratory motion (Ruan, et al., 2007; Ruan, 2010). The difference from this previous work is that instead of decoding the temporal dynamics in a single time series, we aim to decode the functional relations among different time series (e.g., tumor motion and surrogate signal).

Another major difficulty in modeling respiratory tumor motion is hysteresis. The idea of “state augmentation” to model hysteretic respiratory motion has been proposed by Zhang et al. (2007) and Ruan et al. (2008), where the input is augmented by a single lagged surrogate sample with a certain time lag. In this study, we take a slightly extended approach, where the input is augmented by several time-lagged surrogate samples. This gives a more complete picture of the recent respiratory dynamics (Ruan, 2010).

This paper is organized as follows. In Section 2, we present the general theoretical framework and specific learning algorithms. Some practical aspects of the method will also be discussed. We then describe the patient respiratory motion data and evaluation metrics used in this study. Sections 3 and 4 show the quantitative results and furnish some discussions. Finally, we conclude our paper in Section 5.

2. Methods and materials

2.1 General description and formulation of the problem

Given the concurrent measurement of tumor motion and surrogate signal over a short period of time (e.g., with pre-treatment imaging), we wish to establish the relationship between the two and apply it for an extended period when the continuous measurement of tumor motion is not available (e.g., for treatment guidance or intervention). The concurrent measurement of tumor motion and surrogate signal is called training data. For simplicity, we focus on 1D surrogate signal in this study. If there are multiple external surrogates available, e.g., in the CyberKnife SynchronyTM system, our formulation naturally extends to such cases.

While it is conceptually easier to use the current surrogate signal to predict the corresponding tumor motion, here we consider a more general setting where the input is expanded by combining the current surrogate signal (a scalar) and its previous samples. This can potentially lead to more accurate estimation if the input and output relation is highly complex or the tumor exhibits hysteric type of motion. Denote the 1D surrogate signal as s_t , where t is a time index. The input now becomes a vector: $\mathbf{x}_t = [s_{t-D\tau}, s_{t-(D-1)\tau}, \dots, s_t]^T$, with D being the dimension of input vector. The parameter τ indicates the time lag between successive components when augmenting the inputs. While it is possible to use the original high-frequency data as the input, this creates a difficulty for regression methods: the higher the input dimension is, the more the fitting parameters there will be. This may lead to overfitting of the model parameters and thus increase the overall errors on unseen test data. In this work, it was fixed at 0.4 s to achieve a balance between respiratory dynamics and measurement noise (Ruan, 2010).

Because the tumor motion is in 3D space, a different model can be built independently for each direction, namely, anterior-posterior (AP), left-right (LR), and superior-inferior (SI).

This can be either a simple linear instantaneous model, or a nonlinear model incorporating respiratory dynamics. In this work, we utilize a powerful memory-based learning method to find the internal/external relations, which will be described in details in the next section.

2.2 Memory-based learning

Memory-based learning defers processing of the training data until a query needs to be answered. It involves storing of the training data in memory, and finding relevant data to answer a particular query. Relevance is usually measured using a distance function, with nearby points (surrogate input vectors $\mathbf{x}_t$) having high relevance. The simplest example is the nearest neighbor method, which chooses the closest point and uses its output value. This, however, leads to discontinuous functional mapping between the input and output and gives poor results in general. We propose to use some nontrivial memory-based learning methods, which use locally weighted training to average, interpolate between, extrapolate from, or otherwise combine training data (Vapnik, 1992).

2.2.1 Locally weighted training—Locally weighted training aims to fit relatively simple models to local patches instead of one single global model. It requires the local model to fit geometrically nearby data points well, with less concern for distant points. Given a training data set denoted by $\mathcal{D} = \{(\mathbf{x}_t, y_t) \mid t = 1, \dots, T\}$, a general locally weighted training criterion is:

$$C(\mathbf{x}) = \sum_{t=1}^T [(f(\mathbf{x}_t, \beta) - y_t)^2 \cdot K_H(\mathbf{x} - \mathbf{x}_t)] \quad (1)$$

where, $K_H(\cdot) = \frac{1}{\det(\mathbf{H})} K(\mathbf{H}^{-1} \cdot)$ is called the kernel function or weighting function, and $\mathbf{H}$ is a bandwidth matrix (its determination will be explained in details later). In this work, we have used the least square criterion for the cost function. For the distance weighting function, we use the Gaussian kernel:

$$K_H(\mathbf{x}) = \frac{1}{(2\pi)^{D/2} \cdot \det(\mathbf{H})} \exp(-\mathbf{x}^T \mathbf{H}^{-2} \mathbf{x} / 2) \quad (2)$$

The function $f(\mathbf{x}_t, \beta)$ is the local model, with β being the model parameters, whose optimal values will be determined by the locally weighted learning technique described later. Depending on the exact functional form, different methods can arise. Two techniques will be investigated here: 1) *kernel regression* or *weighted average*, which averages the outputs of nearby points, inversely weighted by their distance to the query point; 2) *locally weighted regression* (LWR), which fits a surface to nearby points using a distance weighted regression.

2.2.2 Kernel regression and locally weighted regression—In kernel regression, we try to find the best estimate for the output using a local model that is a constant, $\hat{y}$. The best estimate $\hat{y}(\mathbf{x})$ that minimizes the cost function in Eq. (1) has a closed form solution and is given by a weighted average of the training outputs (Atkeson, et al., 1997):

$$\hat{y}(\mathbf{x}) = \frac{\sum_{t=1}^T [K_H(\mathbf{x} - \mathbf{x}_t) \cdot y_t]}{\sum_{t=1}^T K_H(\mathbf{x} - \mathbf{x}_t)} \quad (3)$$

In locally weighted regression, the local model is a linear function of the input: $f(\mathbf{x}_b, \beta) = \beta^T \mathbf{x}_b$, where each input vector in the training data has been appended with a constant term 1. In this case, the best estimate $\hat{y}(\mathbf{x})$ that minimizes the cost function in Eq. (1) is given by (Atkeson, et al., 1997):

$$\hat{y}(\mathbf{x}) = \mathbf{x}^T (\mathbf{X}^T \mathbf{W}^T \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}^T \mathbf{W} \mathbf{y} \quad (4)$$

where, $\mathbf{X}$ is a matrix whose t -th row is $\mathbf{x}_t$, $\mathbf{y}$ is a vector whose t -th element is y_t , and $\mathbf{W}$ is a diagonal matrix with diagonal elements given by $w_t = \sqrt{K_H(\mathbf{x} - \mathbf{x}_t)}$. Note that both the training criterion $C(\mathbf{x})$ and the output $\hat{y}(\mathbf{x})$ depend on the location of the query point $\mathbf{x}$ and is a nonlinear function of the training input vectors.

The key concept of memory-based learning is to use relevant data to infer the output value. In this work, relevance is measured using the Gaussian kernel function defined in Eq. (2). As can be seen, the distance function is embedded into the optimal solutions in Eq. (3–4) as weighting factors, which are applied to each data point (surrogate input vector) in the training data set. As can be seen from Eq. (3–4), in order to calculate the output value, every data point in the training data set is used. Since the distance function $K_H(\mathbf{x} - \mathbf{x}_t)$ depends on the new input $\mathbf{x}$, every training data point needs to be stored in the memory in order to generate the appropriate output value. This is in contrast to conventional linear regression, where once the optimal parameters are computed, the original input data can be discarded.

2.2.3 Determining the bandwidth matrix—At this point, it remains to determine the bandwidth matrix $\mathbf{H}$ for its practical applications. In order to exploit the inherent correlation among different components of the input vector, a common choice for the bandwidth matrix is one that is proportional to the square root of the sample covariance matrix of the data:

$$\mathbf{H} = h \cdot [(\mathbf{X} - \bar{\mathbf{X}})^T (\mathbf{X} - \bar{\mathbf{X}}) / (T - 1)]^{1/2} \quad (5)$$

where $\bar{\mathbf{X}}$ is a matrix formed by the sample mean of the input. In this way, the data are transformed to have an identity covariance matrix and the kernel function is represented by a single scalar h . By judiciously selecting a kernel bandwidth h , one can well approximate the true probability density function (PDF) with a finite number of samples. For this purpose, one can use cross validation to select the optimal bandwidth. However, when the data size is large, this technique can be computationally expensive. An alternative is to consider the statistical property of the estimator under certain assumptions for the underlying PDF and choose one that minimizes the expected mean square error, i.e., achieves the optimal balance between bias and variance. To this end, we use the Silverman's rule (Silverman, 1986) for bandwidth selection:

$$\hat{h} = \left(\frac{4}{D+2} \cdot \frac{1}{T} \right)^{1/(D+4)} \quad (6)$$

where D is the dimension of the input vector, and T is the number of input vectors in the training data set.

2.3 Useful properties of the proposed method

The proposed method has several properties that are important in the context of finding internal/external relations. In contrast to global models, where a single function is used to fit all of the training data, locally weighted learning adopts a divide-and-conquer approach and aims to fit simpler models to local patches instead of the entire region. As such, it can

potentially reduce the bias and lead to more accurate estimation if the internal/external relations are highly complex and nonlinear.

Another important property is the robustness of the proposed method. By focusing on nearby data points and diminishing the effects of distant points with the help of an appropriate distance function, locally weighted learning has a built-in mechanism to effectively reject outliers in the training data, in contrast to other add-on and often heuristic approaches for outlier rejection. This point will be illustrated later in the Results section. Finally, both training and adapting to new data is performed almost instantaneously, because this process is deferred until a query needs to be answered in memory-based learning. It turns out that for both kernel regression and locally weighted regression, the solution has a closed form as shown in Eqs. (3–4). This feature becomes important when the overall internal/external relations are changing over time and the regression model needs to be rebuilt on the fly.

2.4 Evaluations

The proposed method was evaluated using respiratory motion data from 11 patients treated with the Mitsubishi real-time radiation therapy (RTTR) system at the Nippon Telegraph and Telephone Corporation (NTT) Hospital in Sapporo, Japan. Patients with abdominal and thoracic tumors, typically have two to four 1.5 mm diameter gold ball bearings implanted in or near the tumor, which are tracked in real time with diagnostic x-ray fluoroscopy systems consisting of two pairs of x-ray tubes and imagers (Shirato, et al., 2003). The 1D movement of the patient's abdominal surface is monitored by the AZ-733V external respiratory gating system (Anzai Medical, Tokyo, Japan). The data set consists of simultaneous measurement of 3D tumor motion and 1D abdominal surface sampled at 30 frame/s. There are a total of 171 respiratory traces, with an average duration of ~115 s per trace. The average peak-to-peak amplitude of the tumor motion is ~15 mm in this group of patients. A more detailed description of the data set can be found in Berbeco et al (2005).

In this study, the 1D abdominal surface motion is used as the surrogate signal to find the internal tumor motion. A different regression model is built for each of the three directions in each respiratory trace. The number of input dimensions ranges from 1, 2, 3, 4, 5; the length of training data ranges from 5 s, 10 s, 15 s, 20 s, 25 s, 30 s, with the rest of the trace being the test data (average: ~115 s, range: ~21 and ~294 s). The minimum amount of training data (5 s) roughly corresponds to one breath of an average patient. To evaluate the performance of the algorithms with noisy measurements, we added Gaussian random noise to the training data under different noise levels, i.e., 10%, 20%, 50% of the standard deviation of the training data.

For comparison, we also trained global linear regression models for each of the respiratory motion traces, whose output is simply given by:

$$\hat{y}(x) = x^T (X^T X)^{-1} X^T y \quad (7)$$

The algorithm was evaluated in terms of 3D localization error. For each time point, the 3D localization error is defined as:

$$\sqrt{(y(1) - \hat{y}(1))^2 + (y(2) - \hat{y}(2))^2 + (y(3) - \hat{y}(3))^2} \quad (8)$$

where, $y(1)$, $y(2)$, $y(3)$ are the ground truth tumor location in LR, AP, SI directions, respectively. The different methods were compared in terms of: 1) average error; 2) root-mean-square (RMS) error; and 3) 95th percentile error.

3. Results

Figure 1 shows examples of the complex relations between surrogate and tumor motion (in the AP direction) in two patients. In Fig. 1 a) and c), only the current surrogate signal is shown along with the AP motion; it is obvious that no instantaneous model (linear or nonlinear) can accurately capture such a complex internal/external relation, due to the hysteric respiratory motion. This emphasizes the importance of incorporating the history of the respiratory traces. In Fig. 1 b), the input is augmented by a lagged sample of the surrogate signal. It appears that using a linear model would be fine in this case, although not ideal. However, using a global linear model for the same augmented input shown in Fig. 1 d) will inevitably lead to a bias in the estimation. The use of locally weighted learning proves to be most advantageous in such complex cases.

Table 1 summarizes the 3D localization error for all the 171 traces using three different methods, where the dimension of the input vector is 2, and the training data is the first 5 s of each of the respiratory trace and the rest is used as the test data. In terms of all three error metrics, locally weighted regression outperforms linear regression, which in turn outperforms kernel regression. To assess the statistical significance of the results, we performed a paired *t*-test on the average 3D error obtained with locally weighted regression and linear regression, with $p < 0.001$. This demonstrates that the error reduction is statistically significant.

Table 2 lists the average 3D localization error for all 171 traces under different noise levels. The degradation of performance for both algorithms is negligible until the noise level reaches 50% of the standard deviation of the training data. To evaluate the performance of the algorithm over an extended period of time, Table 3 shows the 3D localization error for 35 traces that are longer than 3 min. It turns out that the same trend in Table 1 holds here, i.e., locally weighted regression outperforms linear regression, which in turn outperforms kernel regression, although the errors are larger compared with those for all 171 traces with the same method.

Figure 2 shows the average 3D error versus the input dimension and training data length using locally weighted regression for three typical patients. In Fig. 2 a), the average 3D error is strongly dependent on the input dimension and is smallest with 2 inputs; it is minimally dependent on the training data length. In Fig. 2 b), the average 3D error is strongly dependent on the training data length and the error levels off with more than 10 s training data; it is insensitive to the input dimension with 2 inputs or more. In Fig. 2 c), the landscape is more complicated: the average 3D error is moderately dependent on the training data length with 1–3 inputs, and decreases monotonically as a function of training data length with more than 4 inputs; the error is smallest with 3 to 5 inputs, although the reduction is minimal with more than 4 inputs.

To optimize the algorithm performance, we selected trace-specific input dimensions with the training data length fixed at 5 s. The reason why 5 s was selected is because in order for training data to be acquired, the patient has to be exposed to radiation imaging dose, and 5 s is about the smallest duration in which a typical breathing cycle is completed. This selection thus ensures reasonable localization accuracy while minimizing imaging dose to the patient. The average 3D error for all the 171 traces was further reduced to 1.4 mm from 1.5 mm, using locally weighted regression. Again, the error reduction was found to be statistically

significant, with $p < 0.001$. Although the training data is preferred to be as short as possible in order to reduce imaging dose, choosing patient-specific or even trace-specific input dimensions is recommended to fully utilize the flexibilities afforded by the proposed method and maximize its localization accuracy. In this work, the optimal input dimensions were selected retrospectively. In practice, this can be done by using a separate (and small) validation data set in addition to the training data. The validation data set can be acquired in the same way the training data set is acquired; it can be done just after the acquisition of the training data but before the treatment starts. The optimal algorithm parameters (e.g., input dimension) can then be automatically selected based on their respective performance on the validation data set.

Figure 3 shows the surrogate signal in one of the respiratory traces of patient 9, where the training data was contaminated by a few “bad” samples toward the end of the training data. Although the bad training samples appear to moderately differ from the respiratory dynamics and there are only a few of them, they can have some profound effects on the results if a regular regression model is used. Figures 4–6 show the tumor motion estimation results using linear regression, kernel regression, and locally weighted regression, with an average 3D error of 2.1, 1.5, and 1.0 mm, respectively. Figure 7 shows the histogram of the average 3D error for the three methods. It is evident that linear regression was unduly and adversely affected by a few bad training samples: a bias in the output is apparent in all three directions. The results were greatly improved with kernel regression, and in the case of locally weighted regression, the results do not seem to be affected by the bad training samples at all. This clearly demonstrates the robustness of the proposed method.

To demonstrate the performance of the algorithm under irregular breathing, Figure 8 shows the tumor motion estimation results using locally weighted regression for one of the irregular respiratory traces. Tracking is generally accurate except for a few breathing cycles in the lateral direction, where the motion is small compared with the other two directions. In another patient who shows irregular breathing, the algorithm is able to accurately track the tumor motion even during deep breathing, as shown in Fig. 9.

4. Discussions

In this paper, we have presented a memory-based learning method to infer the relations between tumor motion and respiratory surrogates. Two memory-based learning methods were investigated: kernel regression and locally weighted regression. Both use locally weighted training to fit relatively simple models to local patches and are able to capture highly nonlinear and complex internal/external relations in an accurate way. Due to the local nature of the kernel functions, they are inherently robust to outliers in the training data. One distinct advantage of memory-based learning over traditional global nonlinear models (such as neural networks and support vector machines) is that both training and adapting to new data (whenever it becomes available by acquiring new images after training) is performed almost instantaneously, making it suitable for dynamically following variable internal/external relations.

When evaluated on 171 respiratory traces from 11 patients, locally weighted regression performed favorably compared with linear regression; kernel regression was found to be generally less accurate than linear regression, although it gave more robust estimation. One reason is that kernel regression uses a very conservative function approximator, i.e., averaging, as shown in Eq. (3). Because all weights (or kernel functions) are positive, the output is thus bounded by those in the training set and cannot “extrapolate” beyond the output values in the training data. Kernel regression typically displays a bias near the peak inhale and exhale phases in the respiratory trace. On the other hand, locally weighted

regression with a planar local model exactly reproduces a line or plane and thus can both interpolate and extrapolate from training data appropriately.

The local model $f(\mathbf{x}, \beta)$ is not limited to a particular functional form. In addition to being a constant or a linear function, it can be easily extended to a polynomial function (quadratic, cubic, etc), or even functional mapping such as neural networks. These complex local models might lead to more accurate estimation. On the other hand, with more model complexity, come more parameters to optimize, which may become problematic (e.g., overfitting) with limited training data. We have implemented regression techniques with quadratic functions as local models. Our preliminary results show that they typically lead to an increase of localization error in the amount of >1 mm compared with their linear counterparts. Their ultimate utility needs to be evaluated on independent test data and should be considered on a case-by-case basis.

The overall internal/external relations may change over time. In this case, it is important to acquire new respiratory motion data and adapt to such changes. The results shown in Fig. 2 b) indicate that acquiring 5 s additional training data (from 5 s to 10 s) can lead to a reduction of average 3D error by as much as 40%, or ~ 1 mm. Alternatively, one can take into consideration the temporal correlation between the original and new training data and update the model whenever new training data becomes available. In order to effectively incorporate the new data into the training data, an exponential discounting factor determined by the time difference between current time and each training sample (Ruan, et al., 2008) can be inserted into each term in the training criterion in Eq. (1). With memory-based learning, such a procedure can be done instantly, i.e., the model is optimized every time a new training data sample is acquired. Deciding when and how to acquire new training data involves analysis of respiratory motion and change detection, which will be a topic of our future work.

The “bad” training samples in Fig. 3 can either be due to measurement errors, or due to true variations in patient breathing. In the former case, we may be able to just reject them. On the other hand, if they are due to true variations in patient breathing, then it is better to keep them in the training data so that the system ‘memorizes’ these breathing dynamics and will recognize them and give the correct output if they occur in the future. If outliers occur during treatment delivery, the weighting function will become equally small. In that case, kernel regression will simply give the mean value determined by the training data as the output, while locally weighted regression will default to standard linear regression.

The proposed method has been implemented on the MATLAB platform running on a PC with a 2.80 GHz CPU and 4GB RAM. Each 3D position estimate takes ~ 0.5 ms using locally weighted regression with 2 inputs and 5 s training data. The computational time is proportional to the amount of training data. Even with 30 s of training data, each estimate takes ~ 3 ms. Thus, the proposed technique can be readily applied in a real time setting.

5. Conclusion

We have presented a memory-based learning method to infer the relations between tumor motion and respiratory surrogates. The method is able to accurately capture the highly complex and nonlinear relations between tumor and surrogate motion in an efficient manner (a few ms per estimate). It has been demonstrated that compared with standard linear regression, the algorithm is particularly suitable to handle situations where the training data are contaminated by large errors or outliers. These desirable properties make it an ideal candidate for accurate and robust tumor gating/tracking using respiratory surrogates.

Acknowledgments

The authors would like to thank Dr. Seiko Nishioka of the Department of Radiology, NTT Hospital, Sapporo, Japan and Dr. Hiroki Shirato of the Department of Radiation Medicine, Hokkaido University School of Medicine, Sapporo, Japan for sharing the respiratory motion data with them. This work is partially supported by the National Cancer Institute (1R21 CA153587 and 1R01 CA133474) and National Science Foundation (0854492).

Reference

- Arslan S, Yilmaz A, Bayramgurler B, et al. CT- guided transthoracic fine needle aspiration of pulmonary lesions: accuracy and complications in 294 patients. *Med Sci Monit.* 2002; 8(7):CR493–CR497. [PubMed: 12118196]
- Atkeson CG, Moore AW, Schaal S. Locally weighted learning. *Artificial Intelligence Review.* 1997; 11(1–5):11–73.
- Balter JM, Wright JN, Newell LJ, et al. Accuracy of a wireless localization system for radiotherapy. *Int J Radiat Oncol Biol Phys.* 2005; 61(3):933–937. [PubMed: 15708277]
- Berbeco RI, Nishioka S, Shirato H, et al. Residual motion of lung tumours in gated radiotherapy with external respiratory surrogates. *Phys Med Biol.* 2005; 50(16):3655–3667. [PubMed: 16077219]
- Bert C, Metheany KG, Doppke K, et al. A phantom evaluation of a stereo-vision surface imaging system for radiotherapy patient setup. *Med Phys.* 2005; 32(9):2753–2762. [PubMed: 16266088]
- Cardis E, Vrijheid M, Blettner M, et al. The 15-country collaborative study of cancer risk among radiation workers in the nuclear industry: Estimates of radiation-related cancer risks. *Radiation Research.* 2007; 167(4):396–416. [PubMed: 17388693]
- Chi PC, Balter P, Luo D, et al. Relation of external surface to internal tumor motion studied with cine CT. *Med Phys.* 2006; 33(9):3116–3123. [PubMed: 17022203]
- D'Souza, W.; Malinowski, K.; Zhang, H. Machine learning for intrafraction tumor motion modeling with respiratory surrogates; Paper presented at the International Conference on Machine Learning and Applications; 2009 Dec 13–15.
- Hoisak JD, Sixel KE, Tirona R, et al. Correlation of lung tumor motion with external surrogate indicators of respiration. *Int J Radiat Oncol Biol Phys.* 2004; 60(4):1298–1306. [PubMed: 15519803]
- Ionascu D, Jiang SB, Nishioka S, et al. Internal-external correlation investigations of respiratory induced motion of lung tumors. *Med Phys.* 2007; 34(10):3893–3903. [PubMed: 17985635]
- Jiang SB. Radiotherapy of mobile tumors. *Semin Radiat Oncol.* 2006a; 16(4):239–248. [PubMed: 17010907]
- Jiang SB. Technical aspects of image-guided respiration-gated radiation therapy. *Med Dosim.* 2006b; 31(2):141–151. [PubMed: 16690455]
- Kanoulas E, Aslam JA, Sharp GC, et al. Derivation of the tumor position from external respiratory surrogates with periodical updating of the internal/external correlation. *Phys Med Biol.* 2007; 52(17):5443–5456. [PubMed: 17762097]
- Keall PJ, Mageras GS, Balter JM, et al. The management of respiratory motion in radiation oncology report of AAPM Task Group 76. *Med Phys.* 2006; 33(10):3874–3900. [PubMed: 17089851]
- Li G, Xie H, Ning H, et al. A novel analytical approach to the prediction of respiratory diaphragm motion based on external torso volume change. *Phys Med Biol.* 2009; 54(13):4113–4130. [PubMed: 19521009]
- Mageras, GS.; Yorke, E.; Jiang, SB. “4D” IMRT Delivery. In: Bortfeld, T.; Schmidt-Ullrich, R.; De Neve, W.; Wazer, DE., editors. *Image-Guided IMRT*. Heidelberg, Germany: Springer-Verlag; 2005. p. 269-288.
- Malinowski K, McAvoy TJ, George R, et al. Incidence of changes in respiration-induced tumor motion and its relationship with respiratory surrogates during individual treatment fractions. *Int J Radiat Oncol Biol Phys.* 2012; 82(5):1665–1673. [PubMed: 21498009]
- Malinowski KT, Pantarotto JR, Senan S, et al. Inferring positions of tumor and nodes in Stage III lung cancer from multiple anatomical surrogates using fourdimensional computed tomography. *Int J Radiat Oncol Biol Phys.* 2010; 77(5):1553–1560. [PubMed: 20605343]

- Murphy MJ. Tracking moving organs in real time. *Semin Radiat Oncol.* 2004; 14(1):91–100. [PubMed: 14752737]
- Ruan D. Kernel density estimation-based real-time prediction for respiratory motion. *Phys Med Biol.* 2010; 55(5):1311–1326. [PubMed: 20134084]
- Ruan D, Fessler JA, Balter JM. Real-time prediction of respiratory motion based on local regression methods. *Phys Med Biol.* 2007; 52(23):7137–7152. [PubMed: 18029998]
- Ruan D, Fessler JA, Balter JM, et al. Inference of hysteretic respiratory tumor motion from external surrogates: a state augmentation approach. *Phys Med Biol.* 2008; 53(11):2923–2936. [PubMed: 18460744]
- Sachs RK, Brenner DJ. Solid tumor risks after high doses of ionizing radiation. *Proceedings of the National Academy of Sciences of the United States of America.* 2005; 102(37):13040–13045. [PubMed: 16150705]
- Seppenwoolde Y, Berbeco RI, Nishioka S, et al. Accuracy of tumor motion compensation algorithm from a robotic respiratory tracking system: a simulation study. *Med Phys.* 2007; 34(7):2774–2784. [PubMed: 17821984]
- Shirato H, Harada T, Harabayashi T, et al. 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.* 2003; 56(1):240–247. [PubMed: 12694845]
- Silverman, B. *Density Estimation for Statistics and Data Analysis.* Chapman and Hall; 1986.
- Tang X, Sharp GC, Jiang SB. Fluoroscopic tracking of multiple implanted fiducial markers using multiple object tracking. *Phys Med Biol.* 2007; 52(14):4081–4098. [PubMed: 17664596]
- Vapnik, V. *Principles of riskminimization for learning theory.* Vol. (Vol. 4). San Mateo, CA: Morgan Kaufman; 1992.
- Wu H, Zhao Q, Berbeco RI, et al. Gating based on internal/external signals with dynamic correlation updates. *Phys Med Biol.* 2008; 53(24):7137–7150. [PubMed: 19033643]
- Yan H, Yin FF, Zhu GP, et al. The correlation evaluation of a tumor tracking system using multiple external markers. *Med Phys.* 2006; 33(11):4073–4084. [PubMed: 17153387]
- Zhang Q, Pevsner A, Hertanto A, et al. A patient-specific respiratory model of anatomical motion for radiation treatment planning. *Med Phys.* 2007; 34(12):4772–4781. [PubMed: 18196805]

Figure 1.
Examples of the complex relations between surrogate and tumor motion: a-b), patient 4; c-d), patient 10.

Figure 2. Average 3D error versus the input dimension and training data length using locally weighted regression for one of the traces in: a), patient 4; b), patient 11; c), patient 10.

Figure 3.

Surrogate signal in one of the respiratory traces of patient 9 for training and testing (only the first 30 s is shown), which is separated by the vertical dotted line. The short period of bad training data is indicated by an arrow. Notice that due to the input vector expansion, a total of 7 s of surrogate data is needed in the case of the 5 s training period.

Figure 4. Tumor motion estimation results with 2 inputs and 5 s training using linear regression, for the respiratory traces corresponding to Fig. 3 (solid line: ground truth tumor positions; dashed line: estimated positions).

Figure 5. Tumor motion estimation results with 2 inputs and 5 s training using kernel regression, for the respiratory traces corresponding to Fig. 3 (solid line: ground truth tumor positions; dashed line: estimated positions).

Figure 6.

Tumor motion estimation results with 2 inputs and 5 s training using locally weighted regression, for the respiratory traces corresponding to Fig. 3 (solid line: ground truth tumor positions; dashed line: estimated positions).

Figure 7. Histogram of the average 3D error for the respiratory traces corresponding to the patient data shown in Figs. 4–6 using: a), linear regression; b), kernel regression; c), locally weighted regression.

Figure 8.

Tumor motion estimation results using locally weighted regression for one of the irregular respiratory traces (solid line: ground truth tumor positions; dashed line: estimated positions). Tracking is generally accurate except for a few breathing cycles in the lateral direction, where the motion is small compared with the other two directions. The average 3D localization error is 1.7 mm in this case.

Figure 9.

Tumor motion estimation results using locally weighted regression for one of the irregular respiratory traces (solid line: ground truth tumor positions; dashed line: estimated positions). The algorithm is able to accurately track the tumor motion during deep breathing. The average 3D localization error is 0.9 mm in this case.

Table 1

3D localization error (mean $\pm$ std) for all 171 traces using three different methods.

Method/Metric	Average (mm)	RMS (mm)	95 th percentile (mm)
Linear Regression	1.6 $\pm$ 1.0	1.9 $\pm$ 1.3	3.5 $\pm$ 2.6
Kernel Regression	1.7 $\pm$ 1.1	2.2 $\pm$ 1.6	4.2 $\pm$ 3.6
Locally Weighted Regression	1.5 $\pm$ 1.0	1.9 $\pm$ 1.3	3.4 $\pm$ 2.7

Table 2

Average 3D localization error (mean $\pm$ std) for 171 traces under different noise levels.

Method/Noise Std (%)	0	10	20	50
Linear Regression	1.6 $\pm$ 1.0	1.6 $\pm$ 1.0	1.6 $\pm$ 1.0	1.9 $\pm$ 1.0
Kernel Regression	1.7 $\pm$ 1.1	1.7 $\pm$ 1.1	1.8 $\pm$ 1.1	2.1 $\pm$ 1.1
Locally Weighted Regression	1.5 $\pm$ 1.0	1.5 $\pm$ 1.0	1.5 $\pm$ 1.0	1.8 $\pm$ 1.0

Table 3

3D localization error (mean $\pm$ std) for 35 traces that are longer than 3 min.

Method/Metric	Average (mm)	RMS (mm)	95 th percentile (mm)
Linear Regression	2.0 $\pm$ 1.2	2.5 $\pm$ 1.5	4.6 $\pm$ 2.8
Kernel Regression	2.3 $\pm$ 1.4	3.1 $\pm$ 2.2	6.2 $\pm$ 5.0
Locally Weighted Regression	2.0 $\pm$ 1.1	2.4 $\pm$ 1.4	4.6 $\pm$ 2.8