

This document is downloaded from DR-NTU, Nanyang Technological University Library, Singapore.

Title	Predictive Local Receptive Fields Based Respiratory Motion Tracking For Motion-Adaptive Radiotherapy
Author(s)	Wang, Yubo; Tatinati, Sivanagaraja; Huang, Liyu; Hong, Kim Jeong; Shafiq, Ghufan; Veluvolu, Kalyana C.; Hoong, Andy Khong Wai
Citation	Wang, Y., Tatinati, S., Huang, L., Hong, K. J., Shafiq, G., Veluvolu, K. C., et al. (2017). Predictive Local Receptive Fields Based Respiratory Motion Tracking For Motion-Adaptive Radiotherapy. 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.
Date	2017
URL	http://hdl.handle.net/10220/42910
Rights	© 2017 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. The published version is available at: [http://dx.doi.org/10.1109/EMBC.2017.8037453].

Predictive Local Receptive Fields Based Respiratory Motion Tracking For Motion-Adaptive Radiotherapy

Yubo Wang¹, Sivanagaraja Tatinati^{2*}, Liyu Huang¹, Kim Jeong Hong³,
Ghufran Shafiq⁴, Kalyana C. Veluvolu⁴, and Andy W. H. Khong²

Abstract—Extracranial robotic radiotherapy employs external markers and a correlation model to trace the tumor motion caused by the respiration. The real-time tracking of tumor motion however requires a prediction model to compensate the latencies induced by the software (image data acquisition and processing) and hardware (mechanical and kinematic) limitations of the treatment system. A new prediction algorithm based on local receptive fields extreme learning machines (pLRF-ELM) is proposed for respiratory motion prediction. All the existing respiratory motion prediction methods model the non-stationary respiratory motion traces directly to predict the future values. Unlike these existing methods, the pLRF-ELM performs prediction by modeling the higher-level features obtained by mapping the raw respiratory motion into the random feature space of ELM instead of directly modeling the raw respiratory motion. The developed method is evaluated using the dataset acquired from 31 patients for two horizons in-line with the latencies of treatment systems like CyberKnife. Results showed that pLRF-ELM is superior to that of existing prediction methods. Results further highlight that the abstracted higher-level features are suitable to approximate the nonlinear and non-stationary characteristics of respiratory motion for accurate prediction.

I. INTRODUCTION

The goal of extracranial robotic radiotherapy (SBRT) is to deliver conformal radiation dose to the tumor site and minimize the damage of the surrounding healthy tissues by tracking the tumor motion in real-time [1]. Tumor site generally moves along with the intrinsic movements of internal organs caused mainly by the respiratory system. The magnitude of the tumor motion typically ranges within few millimeters and it largely depends on the location of tumor (up to 3 cm in cranio-cudal direction due to normal breathing [2]). These intrinsic movements introduce uncertainty in localizing the targeted tumor site with minimal position error. SBRT treatment systems employ external markers and a correlation model to obtain the tumor motion by tracking the external surface motion mainly caused by the respiration [3]. The performance of SBRT systems is however limited by the latency induced while localizing the targeted tumor site (image acquisition and processing) and delivering the conformal radiation dose to the tumor site (mechanical and kinematic limitations) [3], [4]. To this end, predicting the respiratory motion over the latency of the treatment systems

is proposed to localize the tumor site with minimal positional error.

Accurate prediction of respiratory motion is a non-trivial task and has received a great deal of attention over the last decade. Many signal processing algorithms have been customized for respiratory motion prediction in real-time [5]. Novel approaches including neural networks [6], least mean squares [5], and ensemble learning [7] techniques have been developed. Comparative analysis among six prediction techniques for 304 respiratory motion traces provided in [5] highlighted that there is no optimal algorithm that can outperform other algorithms across all datasets. This is primarily due to the non-stationary characteristics of respiratory motion traces. The results further underpin the fact that the performance of the algorithms is highly dependent on the characteristics of signal. Such a study has been done to identify the prediction algorithm based on the past observations of respiratory motion [8].

All the existing prediction methods aim to model the non-stationary and non-linear respiratory motion traces directly. These characteristics of respiratory motion make it challenging for the existing conventional prediction methods to learn the appropriate features for estimating the underlying generalized model that can yield accurate prediction of future values. The main contribution of this paper is to introduce the local receptive field (LRF) learning technique for predicting the respiratory motion. Till date, LRF based extreme learning machines (LRF-ELM) technique has been used for classification applications have shown to achieve good performance [9], [10]. Inspired by these results, we propose to customize LRF-ELM for prediction applications and named our algorithm as pLRF-ELM. The developed pLRF-ELM inherits the merits of LRF-ELM: a) modeling the higher-level features obtained by mapping the raw respiratory motion into the random feature space formulated by the feature maps and b) learn patterns that are informative and appropriate for accurate prediction automatically. The developed pLRF-ELM is evaluated with motion traces obtained from 31 patients (304 traces) for two prediction horizons well in-line with the latencies of treatment systems such as CyberKnife (115ms) and VERO system (308ms) [4], [5]. Results show that the proposed method exhibits higher prediction performance compared to existing methods.

The paper is organized as follows: In Section II, a brief description about ELM and the proposed pLRF-ELM are provided. Section III describes the implementation procedure, performance comparison with existing methods, and

¹ School of Life Science and Technology, Xidian University, Xi'an, Shaanxi, 710071, China. ³ School of Computer Science engineering, Kyungpook National University, Daegu, South Korea. ⁴ School of Electronics engineering, Kyungpook National University, Daegu, South Korea. ²School of Electrical Engineering, Nanyang Technological University, Singapore *tatinati@ntu.edu.sg

presents a brief discussion on the obtained results. Section IV concludes the paper.

II. METHODS

Throughout this article, the variable s denotes the respiratory motion and \hat{s} is the predicted trace. The prediction horizon is represented as k . The variable \hat{s}_{t+k} denotes the k samples ahead of the predicted value from sample t .

A. Extreme Learning Machines (ELM)

ELM is a type of single-layer feed-forward neural network. In this technique, input weights and hidden layer bias are randomly initialized and then solve for the output weights by using a simple generalized inverse operation. Given a set of N distinct samples $\mathcal{S} = \{(s_i, \mathbf{t}_i) | s_i \in \mathbb{R}^m, \mathbf{t}_i \in \mathbb{R}^n; i = 1, \dots, \tilde{N}\}$ with $\mathbf{s}_i = [s_{i,1}, \dots, s_{i,m}]^T$ being the input vector and $\mathbf{t}_i = [t_{i,1}, \dots, t_{i,n}]^T$ as its corresponding target vector, ELM finds the mapping between the input and its corresponding target using:

$$\mathbf{o}_j = f_L(\mathcal{S}) = \sum_{i=1}^L \beta_i g_i(\mathbf{w}_i \mathbf{s}_j + b_i); j = 1, \dots, \tilde{N}, \quad (1)$$

where \mathbf{w}_i is the input weights connecting the i -th hidden unit to the input vector \mathbf{s}_i , b_i is the hidden layer bias and β_i denotes output layer weights. The universal approximation of ELM requires the activation function $g_i(\bullet) : \mathbb{R} \rightarrow \mathbb{R}$ to be a bounded non-constant piecewise continuous function [10]. During the training phase, the input weights $\mathbf{w}_i \in \mathbb{R}^m$ and the hidden unit bias $b_i \in \mathbb{R}$ are randomly assigned according to any continuous probability distribution. Combining all \tilde{N} equations in (1), we have the following linear system

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{T}, \quad (2)$$

where \mathbf{H} is a $(L \times \tilde{N})$ matrix with each row representing the value of an input being activated through L hidden units and \mathbf{T} contains all \tilde{N} target vector for each input. The output weight matrix $\boldsymbol{\beta}$ can therefore be obtained as

$$\boldsymbol{\beta} = \begin{cases} \mathbf{H}^T(\frac{1}{c} + \mathbf{H}\mathbf{H}^T)^{-1}\mathbf{T} & \text{if } \tilde{N} < L; \\ (\frac{1}{c} + \mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{T} & \text{if } \tilde{N} > L, \end{cases} \quad (3)$$

where the superscript T denotes matrix transpose and c denotes the regularization constant. In case of time-series data such as respiratory motion signal, a new training sample can be formed immediately after receiving a new measurement. Taking advantage of such streaming data, an on-line sequential learning methods is also available for solving the system given in (2). For more details about ELM and its on-line sequential training methods, refer to [9].

B. Predictive Local Receptive Field ELM (pLRF-ELM)

The ELM structure described in the previous section assumes full connections between the inputs and the hidden layer units. When such structure is employed for predicting respiratory motion, the feature is usually the p recent history samples. The optimal value for p is usually less than a full respiratory cycle [2], [5]. However, a typical respiratory

motion exhibits strong local pattern which includes two distinct phases: inhale and exhale phases. Therefore, a better model for respiratory motion could benefit from modeling the local structure by including recent respiratory cycles.

The local receptive field based ELM is employed in this work to model the interdependence that is observed in the respiratory motion signal. The pLRF-ELM models the local correlation in the input space by using feature maps. We assume that the weights between the input and the hidden layer units are in the form of $\mathbf{A} \in \mathbb{R}^{r \times K}$, where r is the size of a feature map and K is the total number of feature maps. Thus, each feature map would have $r + p - 1$ dimensions. As proven by ELM theories [9], the value of \mathbf{A} can be randomly drawn from any probability distribution. To obtain better generalization performance, the input weights in \mathbf{A} are then orthogonalized by using singular value decomposition (SVD). Let $\hat{\mathbf{a}}_k = [a_{1,k}, \dots, a_{r,k}] \in \mathbb{R}^r$ be the k -th column of SVD transformed input weights in $\hat{\mathbf{A}}$. The value of i -th element at k -th feature map can then be computed using

$$c_{i,k,m} = \sum_{j=1}^r s_{i+j-1,m} a_{j,k}, \quad (4)$$

where $i = 1, \dots, r + p - 1$ and $k = 1, \dots, K$, such that $s_{j,m}$ represents j -th component of m -th input vector. Hence, the hidden layer would have the following form after feature map transformation

$$\mathbf{C} = \begin{bmatrix} g(c_{1,1,1}) & \cdots & g(c_{r+p-1,1,1}) & \cdots & g(c_{r+p-1,K,1}) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ g(c_{1,1,m}) & \cdots & g(c_{r+p-1,1,m}) & \cdots & g(c_{r+p-1,K,m}) \end{bmatrix}$$

where \mathbf{C} is a matrix of size $(\tilde{N} \times (r + p - 1)K)$ and $g(\bullet)$ is the activation function. Throughout this work, we choose sigmoid function as activation function for both ELM and pLRF-ELM.

Finally, the pLRF-ELM for modeling the respiratory motion signal is to solve the following linear system

$$\mathbf{C}\boldsymbol{\beta}_{LRF} = \mathbf{T}. \quad (5)$$

The solution for (5) can be obtained in the similar way that is provided in (3).

III. RESULTS

A. CyberKnife Respiratory motion database

The dataset comprises of 304 respiratory motion traces recorded from 31 patients during the radioactive therapy with CyberKnife at Georgetown university hospital. The traces were recorded using an optical tracking system known as the Synchrony respiratory motion tracking system by Accuracy, Inc. The sampling frequency was 26 Hz. The principal component obtained from the motion traces of the three markers placed on the patient is used for the analysis. For more information on the recording procedure and the rationale for pre-processing of the motion traces, refer [5].

B. Performance Indices

The prediction performance of a method is quantified by using the root mean square value of the prediction error (RMSE) of that particular method. The method that yield small RMSE provides better prediction performance. The *RMSE* can be defined as:

$$RMSE = \sqrt{\frac{\sum_{k=1}^m (s(k) - \hat{s}(k))^2}{m}}. \quad (6)$$

C. Parameter Selection & Implementation

Parameter selection for ELM: The hyper-parameters that requires optimal selection for ELM are the regularization constant (c), the number of hidden nodes (N). These parameters are chosen based on the grid search, as shown in Fig. 1(a). The combination that provided the smallest RMSE is considered as the optimal pair for initializing the ELM.

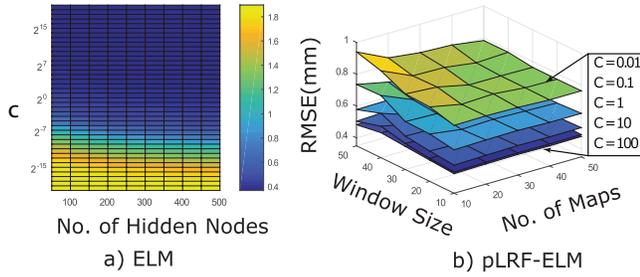


Fig. 1. Hyper-parameter selection: a) ELM and b) pLRF-ELM.

Implementation of ELM: The implementation of ELM for respiratory motion prediction is straight forward as described in [9]. In this work, the optimal number of previous samples (input vector length) is selected based on the auto correlation function. The number of samples p that provide the auto-correlation of more than 0.8 are considered as the input vector for ELM and for its online sequential variant. Two re-training scenarios a) after every 1 second and b) after every 10 seconds are considered to adapt to the non-stationary characteristics of respiratory motion.

Parameter selection for pLRF-ELM: The number of feature maps K , window size W_s , and regularization constant C are the hyper-parameters of pLRF-ELM that require optimal initialization. Grid search has been conducted over a wide range of values for each parameter and obtained RMSE for each triplet (K, W_s, C). The triplet that provided the smallest RMSE was chosen as the optimal parameter set for initialization. From Fig. 1(b), it can infer that the effect of window size has less significant effect on performance compared to the window size and regularization constant. Based on the grid search analysis, in this work, we chose the optimal values as $K = 10$, $W_s = 10$, and $C = 100$.

Implementation of pLRF-ELM: The main difference between LRF-ELM and pLRF-ELM is the input vector, where the former one uses images (two dimensional) whereas the later uses a time series (one dimensional) as the input vector. For respiratory motion prediction, we chose prior $p = 200$ samples as the input vector. According to the dimension

of input ($1 \times p$), feature maps ($K = 10$) considered for pLRF-ELM are one-dimensional with window size given as $W_s = 10$ (dimension of each feature map is $1 \times W_{s+p-1}$). As per the operations of ELM, the weights in each feature map were initialized randomly. The outputs at the k^{th} feature map are then computed in the following steps:

- The first W_s samples ($1, \dots, W_s$) in the input vector are multiplied with the randomly initialized weights of k^{th} feature map to obtain $F(k, W_1)$.
- As a second step, the next W_s samples ($2, \dots, W_s + 1$) in the input vector are considered to compute $F(k, W_2)$. The procedure is repeated until $F(k, W_{(p-W_s)})$.

The same procedure is applied for all feature maps to compute the feature vector outputs. The obtained feature vector outputs from all maps are concatenated as a single column vector and the output weights as provided in (5) is subsequently computed. Similar to ELM, the training of pLRF-ELM update after every one second and 10 seconds.

D. Comparison Analysis

The prediction performance of pLRF-ELM is compared with four existing methods: 1) ELM, 2) Online sequential ELM (OS-ELM), 3) wLMS, and 4) SVRPred. The wLMS and SVRPred methods are chosen as baseline due to their exemplary performance in respiratory motion prediction at the sampling rate of 26 Hz [5]. The comparison analysis was conducted for two prediction horizons, which are well-in-line with commercially available systems such as CyberKnife (115ms) and multi-leaf collimators (308ms) [4]. Performance comparison for all methods is performed on the complete database for both prediction horizons.

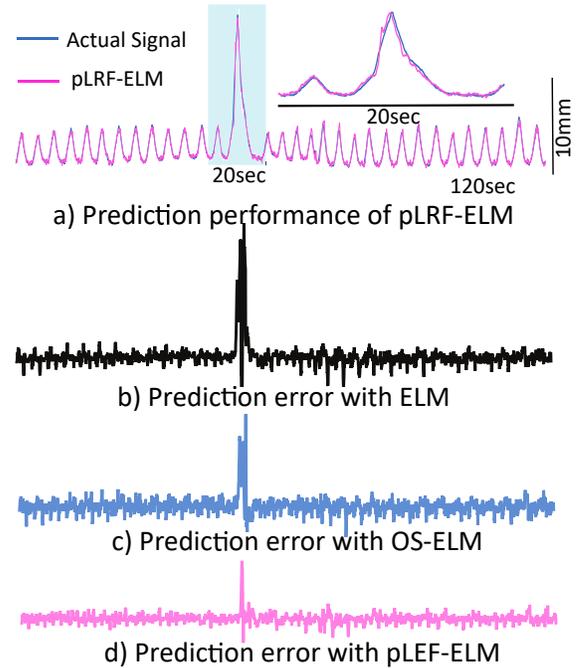


Fig. 2. Prediction performance with ELM variants.

Comparison among ELM variants: The prediction performance of pLRF-ELM along with the other variants of ELM at prediction horizon of 115 ms and re-training at every 10 s is shown in Fig. 2. It can be inferred that the tracking of respiratory motion with pLRF-ELM is accurate even in the presence of irregularities. For better illustration, a zoomed-in portion of the irregularity and the trace achieved with pLRF-ELM is shown in Fig. 2(a). The prediction error obtained for the same trace with ELM and OS-ELM are shown along with the prediction error of pLRF-ELM Fig. 2(b)-(d). The tracking performance show that pLRF-ELM yields less prediction error at all time instances. Furthermore, the reduction in prediction error at the irregularities is significant when compared to the other variants of ELM. The RMSE for prediction obtained with all variants of ELM is computed and plotted the mean with their respective standard deviation for both prediction horizon and both re-training periods in Fig. 3. Results show that pLRF-ELM reduces the prediction error considerably and improves the prediction accuracy by nearly 10% in the considered scenarios.

Comparison with existing methods: The optimal initialization provided in [5] is considered for wLMS and SVRPred methods in this analysis. In addition, both these methods update their parameters with every available new sample and as such the update rate for wLMS and SVRPred is 38 ms (one sample). The RMSE of prediction error obtained with these two methods along with the variants of ELM is shown in Fig. 3. Results show that among all the methods pLRF-ELM achieves the least prediction error in the considered scenarios, followed by OS-ELM, ELM, wLMS, and SVRPred. Compared to wLMS, pLRF-ELM improves the prediction performance by 25% at the update rate of one second and 20% at the update rate of 10 s at horizon of 115 ms. At the prediction horizon of 308 ms the improvement is 15% and 10% at the update rate of 1 s and 10 s, respectively.

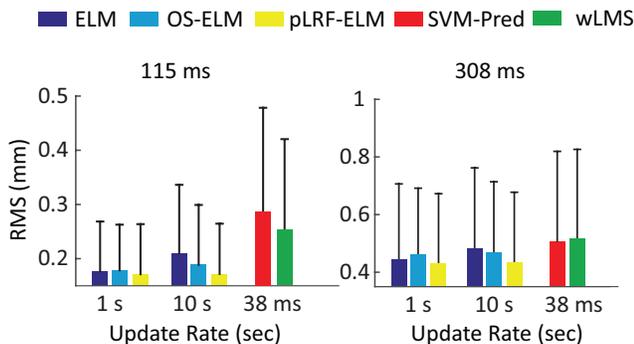


Fig. 3. Comparison analysis

The improvement in prediction performance with the proposed pLRF-ELM highlights the fact that the transformation of raw respiratory motion data with the randomly initialized feature maps aids in obtaining more generalized and stationary features to learn than the usual non-stationary features. These automatically learnt features yielded improvement in the respiratory motion prediction performance.

IV. CONCLUSION

pLRF-ELM, a regression variant for LRF-ELM is proposed in this work to enhance the prediction performance of respiratory motion traces. The prediction performance of pLRF-ELM was evaluated by the database collected from 31 subjects with CyberKnife. Performance comparison conducted with two horizons (115 ms and 308 ms) showed that pLRF-ELM achieves higher prediction accuracy than existing prediction methods. The improvement in prediction performance with the proposed pLRF-ELM underpins that instead of learning directly from the raw respiratory motion traces, the features abstracted by the randomly initialized feature maps provide more robust and reliable prediction model for respiratory motion prediction.

REFERENCES

- [1] S. Bhide and C. Nutting, "Recent advances in radiotherapy," *BMC Medicine*, vol. 8, no. 25, pp. 1–5, 2010.
- [2] S. C. Erridge, Y. Seppenwoolde, S. H. Muller, M. Van Herk, K. De Jaeger, J. Belderbos, L. J. Boersma, and J. V. Lebesque, "Portal imaging to assess set-up errors, tumor motion, and tumor shrinkage during conformal radiotherapy of non-small cell lung cancer," *Radiotherapy and Oncology*, vol. 66, no. 1, pp. 75–85, 2003.
- [3] A. Schweikard, G. Glosser, M. Bodduluri, M. J. Murphy, and J. R. Adler, "Robotic motion compensation for respiratory movement during radiosurgery," *Computer Aided Surgery*, vol. 5, no. 4, pp. 263–277, 2000.
- [4] P. R. Poulsen, B. C. Cho, A. Sawant, D. Ruan, and P. J. Keall, "Deformed analysis of latencies in image-based dynamic MLC tracking," *Medical Physics*, vol. 37, pp. 4998–5005, 2010.
- [5] F. Ernst, R. Durichen, A. Schaefer, and A. Schweikard, "Evaluating and comparing algorithms for respiratory motion prediction," *Physics in Medicine and Biology*, vol. 58, pp. 3911–3929, 2013.
- [6] A. Krauss, S. Nill, and U. Oelfke, "The comparative performance of four respiratory motion predictors for real-time tumor tracking," *Physics in Medicine and Biology*, vol. 56, pp. 5303–5317, 2011.
- [7] S. Tatinati, K. Nazarpour, W. T. Ang, and K. C. Veluvolu, "Ensemble framework based real-time respiratory motion prediction for adaptive radiotherapy applications," *Medical Engineering & Physics*, vol. 38, no. 8, pp. 749–757, 2016.
- [8] F. Ernst, A. Schaefer, and A. Schweikard, "Predicting the outcome of respiratory motion prediction," *Medical Physics*, vol. 38, no. 10, pp. 5569–5581, 2011.
- [9] G. B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Transactions on Systems Man and Cybernetics Part B-Cybernetics*, vol. 42, pp. 513–529, 2012.
- [10] G. B. Huang, Z. Bai, L. Lekamalage, C. Kasun, and C. M. Vong, "Local receptive fields based extreme learning machines," *IEEE Computational Intelligence Magazine*, vol. 10, no. 2, pp. 18–29, 2015.