CrossMark

Ensemble framework based real-time respiratory motion prediction for adaptive radiotherapy applications

Sivanagaraja Tatinati[a], Kianoush Nazarpour[b,c], Wei Tech Ang[d], Kalyana C. Veluvolu[a,d,*]

[a] School of Electronics Engineering, College of IT Engineering, Kyungpook National University, Daegu, South Korea
[b] School of Electrical and Electronic Engineering, Newcastle University, Newcastle-upon-Tyne NE1-7RU, UK
[c] Institute of Neuroscience, Newcastle University, Newcastle-upon-Tyne NE2 4HH, UK
[d] School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore

ARTICLE INFO

ABSTRACT

Successful treatment of tumors with motion-adaptive radiotherapy requires accurate prediction of respiratory motion, ideally with a prediction horizon larger than the latency in radiotherapy system. Accurate prediction of respiratory motion is however a non-trivial task due to the presence of irregularities and intra-trace variabilities, such as baseline drift and temporal changes in fundamental frequency pattern. In this paper, to enhance the accuracy of the respiratory motion prediction, we propose a stacked regression ensemble framework that integrates heterogeneous respiratory motion prediction algorithms. We further address two crucial issues for developing a successful ensemble framework: (1) selection of appropriate prediction methods to ensemble (level-0 methods) among the best existing prediction methods; and (2) finding a suitable generalization approach that can successfully exploit the relative advantages of the chosen level-0 methods. The efficacy of the developed ensemble framework is assessed with real respiratory motion traces acquired from 31 patients undergoing treatment. Results show that the developed ensemble framework improves the prediction performance significantly compared to the best existing methods.

1. Introduction

Radiotherapy is a therapeutic alternative for patients whom are diagnosed with inoperable tumor or who cannot go through surgery due to various complications [1,2]. Tumor movement due to respiration is one of the most challenging problems in robotic radiotherapy especially when tumors are in the lungs. In Wilbert et al. [1], with correlation analysis, it was shown that movement of external body can be a good predictor of the tumor movement that was calculated using computed tomography. In the current commercial systems, such as CyberKnife and VERO, there is however an inevitable time delay of 70–400 ms between the actual movement of tumor and the movement obtained from the correlation model. Furthermore, this delay is device-dependent and is mainly due to image data acquisition and processing and mechanical limitations of the radiotherapy systems [2,9]. To overcome this time delay and hence the positioning error, *prediction* of tumor movement with a prediction horizon equivalent to

the radiotherapy system latency was proposed [1,13]. Accurate prediction of tumor motion is however challenging because the respiratory motion traces are often subject to irregularities and intra-trace variabilities such as baseline, frequency, and temporal changes in their fundamental pattern [14,17].

Over the last two decades, several signal processing techniques have been developed to model the physiological signals [1,2,7–9,13,29,35–38]. For respiratory motion prediction, few of these methods are customized to learn its characteristics [1,2,7–9,13,29]. Univariate prediction approaches include methods based on state-space modeling with adaptive algorithms such as least mean squares (LMS) [13] or extended Kalman filtering (EKF) [4], or wavelet-based multi-scale regression (wLMS) [5] and methods based on machine learning techniques such as support vector machines (SVM) [27], accurate-online SVM (SVRPred) [6], artificial neural networks (ANN) [18], and ensemble learning [29] have been developed. Recently, to further enhance the prediction performance, multi-variate prediction approaches based on Bayesian inference [10,11] and Gaussian processes [12] have been developed. A hybrid method was formulated by cascading EKF and SVM focused on the prediction horizons between 200 ms and 600 ms which are relevant to image-guided radiotherapy [15].

* Corresponding author. Tel.: +82 539507232.
E-mail address: veluvolu@ee.knu.ac, veluvolukc@gmail.com (K.C. Veluvolu).

A comparative analysis of these algorithms with latencies of the VERO and the CyberKnife systems was provided in [9,10]. Results showed that at large prediction horizons, machine learning techniques provide better prediction performance for most of the traces compared to the state-space modeling methods.

Each respiratory motion prediction method has a signal model formulated to represent certain degree of the underlying true respiratory phenomenon. The signal model is solely based on the accumulated prior knowledge which is gained by either stochastic state-space modeling [3,4] or heuristic machine learning techniques [10,18]. By the virtue of its subject-dependent and non-stationary nature, accumulated prior knowledge does not represent the whole evolution of the respiratory motion because irregularities and intra-trace variabilities lead to temporal variations. Consequently, no prediction method can be superior to other methods for all subjects [19]. A comparison analysis conducted among all the prediction methods on the data acquired from 31 subjects supports this hypothesis [9,15].

Motivated by these observations and to enhance the respiratory motion prediction performance, we propose an ensemble learning framework based on stacked regression [25]. Ensemble learning is one of the innovative ideas from the machine learning community and has been successful used in classification and time series forecasting [20,21,23–25]. Stacked regression is a way of integrating multiple prediction models to enhance the prediction performance [21,23]. It has been shown that stacked regression significantly reduces the modeling errors and its variance when compared to the methods choose to ensemble [22,24,25]. Based on the bias-variance trade-off analysis for ensemble methods [20,28], we hypothesized that ensemble of the best existing respiratory motion prediction methods can enhance the prediction performance.

A necessary and sufficient condition to design a successful ensemble learning framework for respiratory motion prediction is "the prediction methods chosen to ensemble (referred as level-0 methods in Breiman et al. [21]) should be accurate and diverse" [23–25]. Based on the techniques proposed in Breiman et al. [21] we identified level-0 methods from a pool of the best existing respiratory motion prediction methods. To exploit the relative advantages of these level-0 methods and to ensemble, we employed machine learning techniques. We validated the proposed ensemble learning approaches with a comprehensive analysis conducted for four prediction horizon with publicly available respiratory motion database recorded from 31 patients [9,26]. The chosen prediction horizons were in-line with the latencies of commercially available systems.

2. Methods and materials

2.1. Ensemble learning: stacked regression

The proposed ensemble learning framework for respiratory motion prediction comprises of a training stage and a testing stage, as shown in Fig. 1.

(a) *Training stage*: It has three phases namely ensemble pruning, ensemble learning, and integration. In the ensemble pruning stage, the prediction methods that are employed as the level-0 methods, will be selected from the pool of existing respiratory motion prediction methods. To this end, diversity measures, such as correlation and mutual information, are employed to identify the diverse and accurate prediction methods (the pruning procedure is detailed in the following sub-section). In the ensemble learning phase, the chosen level-0 methods are trained independently on a dataset $\mathcal{L}$ to perform multi-step prediction: $\mathcal{L} = \{\mathbf{s}_k, s_{k+h}\}_{k=1}^{N}$ with $\mathbf{s}_k = [s_k, s_{k-1}, \ldots, s_{k-n}]$ as the input vector (the value of n
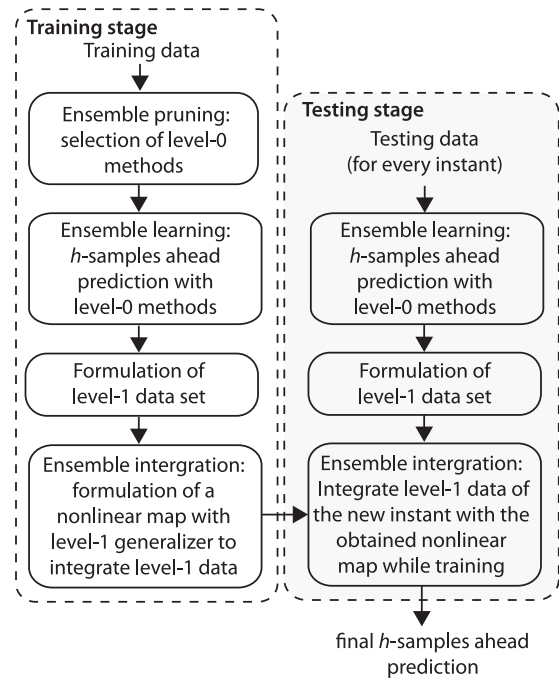

Fig. 1. Schematic diagram: framework of proposed ensemble learning approach for respiratory motion prediction.

depends on the prediction algorithm), s_{k+h} as the output vector, and N as the number of training samples. Finally, in the ensemble integration phase, for each sample $\mathbf{s}_k$ in $\mathcal{L}$, $\hat{s}_{k+h}^1, \hat{s}_{k+h}^2, \ldots, \hat{s}_{k+h}^{\mathcal{P}}$ denote h-samples ahead predicted values computed with the level-0 prediction methods ($\mathcal{P}$ represents number of level-0 methods chosen). Consequently for $\mathcal{L}$, the data set formulated by assembling the predicted values of level-0 prediction methods is named as level-1 data, $\mathcal{T} = \{s_{k+h}, [\hat{s}_{k+h}^1 \ \hat{s}_{k+h}^2 \ \cdots \ \hat{s}_{k+h}^{\mathcal{P}}]\}_{k=1}^{N}$. To derive a nonlinear map ($\mathcal{F}$) which can effectively integrate the level-1 data, machine learning techniques are employed as a level-1 generalizer.

(b) *Testing stage*: In this stage, for every new s_i, level-1 data is formulated by assembling the h-samples ahead predicted values obtained from level-0 methods, that is ($\hat{s}_{i+h}^1, \hat{s}_{i+h}^2, \ldots, \hat{s}_{i+h}^{\mathcal{P}}$). With the formulated level-1 data and the identified nonlinear map $\mathcal{F}$, the final h-samples ahead prediction for the corresponding sample is computed as $\hat{s}_{i+h} = \mathcal{F}(\hat{s}_{i+h}^1, \hat{s}_{i+h}^2, \ldots, \hat{s}_{i+h}^{\mathcal{P}})$.

2.1.1. Ensemble pruning

The success of an ensemble of prediction method relies highly upon the *diversity* of the individual prediction methods [24]. According to the bias-variance trade-off analysis, an ensemble method that comprises level-0 methods with more disagreement is more likely to attain a better generalization performance [24,28]. Thus, in this work, to identify appropriate level-0 methods for respiratory motion prediction, we formulated a pool of best respiratory motion prediction methods. According to the reports by Hong et al. [4] and Ernst et al. [9], LCM-EKF and wLMS methods provide better performance for most of the traces compared to other existing respiratory motion prediction methods such as kernel density estimation (KDE) [16], ANN [18] and SVM [9,27]. Accordingly, we formulated a pool of methods with LCM-EKF, autoregressive moving average model with fading memory Kalman filter (ARMA-FMKF), wLMS, normalized least mean squares (nLMS) and multi-step linear method (MULIN).

To assess the diversity between two prediction algorithms, we employed the correlation coefficient (ρ) and mutual information