Real-time prediction and gating of respiratory motion in 3D space using extended Kalman filters and Gaussian process regression network

To cite this article: W Bukhari and S-M Hong 2016 Phys. Med. Biol. 61 1947

View the article online for updates and enhancements.

Related content
- Real-time prediction and gating of respiratory motion using an extended Kalman filter and Gaussian process regression
  W Bukhari and S-M Hong
- Real-time prediction of respiratory motion using a cascade structure of an extended Kalman filter and support vector regression
  S-M Hong and W Bukhari
- Real-time prediction of respiratory motion based on a local dynamic model
  S-M Hong, B-H Jung and D Ruan

Recent citations
- A condition monitoring approach for real-time monitoring of degrading systems using Kalman filter and logistic regression
  E. Skordilis and R. Moghaddass
Real-time prediction and gating of respiratory motion in 3D space using extended Kalman filters and Gaussian process regression network

W Bukhari and S-M Hong

School of Electronics Engineering, Kyungpook National University, Daegu 41566, Korea

E-mail: waqas@knu.ac.kr and smhong@ee.knu.ac.kr

Received 11 December 2015
Accepted for publication 25 January 2016
Published 15 February 2016

Abstract
The prediction as well as the gating of respiratory motion have received much attention over the last two decades for reducing the targeting error of the radiation treatment beam due to respiratory motion. In this article, we present a real-time algorithm for predicting respiratory motion in 3D space and realizing a gating function without pre-specifying a particular phase of the patient’s breathing cycle. The algorithm, named EKF-GPRN+, first employs an extended Kalman filter (EKF) independently along each coordinate to predict the respiratory motion and then uses a Gaussian process regression network (GPRN) to correct the prediction error of the EKF in 3D space. The GPRN is a nonparametric Bayesian algorithm for modeling input-dependent correlations between the output variables in multi-output regression. Inference in GPRN is intractable and we employ variational inference with mean field approximation to compute an approximate predictive mean and predictive covariance matrix. The approximate predictive mean is used to correct the prediction error of the EKF. The trace of the approximate predictive covariance matrix is utilized to capture the uncertainty in EKF-GPRN+ prediction error and systematically identify breathing points with a higher probability of large prediction error in advance. This identification enables us to pause the treatment beam over such instances. EKF-GPRN+ implements a gating function by using simple calculations based on the trace of the predictive covariance matrix. Extensive numerical experiments are performed based on a large database of 304 respiratory motion traces to evaluate EKF-
The experimental results show that the EKF-GPRN\(^+\) algorithm reduces the patient-wise prediction error to 38\%, 40\% and 40\% in root-mean-square, compared to no prediction, at lookahead lengths of 192 ms, 384 ms and 576 ms, respectively. The EKF-GPRN\(^+\) algorithm can further reduce the prediction error by employing the gating function, albeit at the cost of reduced duty cycle. The error reduction allows the clinical target volume to planning target volume (CTV-PTV) margin to be reduced, leading to decreased normal-tissue toxicity and possible dose escalation. The CTV-PTV margin is also evaluated to quantify clinical benefits of EKF-GPRN\(^+\) prediction.

Keywords: respiratory motion, motion prediction, gating, CTV-PTV margin, Gaussian process regression network, Kalman filter

(Some figures may appear in colour only in the online journal)

1. Introduction

The motion of thoracic and abdominal tumours induced by respiratory motion can significantly compromise the conformal dose delivery in radiotherapy (Ozhasoglu and Murphy 2002, Keall et al 2006, Poulsen et al 2010). Motion-adaptive radiotherapy aims to deliver a conformal dose to the target tumour by compensating for its motion in real-time. Prediction of respiratory motion is necessary in real-time for the motion compensation and it has received much attention for over a decade (Sharp et al 2004, Ruan and Keall 2010, Krauss et al 2011, Lee et al 2012, Purdy et al 2012, Li and Sharp 2013, Dürichen et al 2014). Motion prediction approaches can be classified into two groups: model-based and model-free (Ernst et al 2013). The former predicts respiratory motion based on a mathematical motion model. The Kalman filter based on a state-space model is an example and it is computationally simple and responsive to irregular changes in respiratory motion. Model-free approaches do not assume an explicit mathematical model and predict the future position by learning a mapping from previous observations with, for example, support vector regression (SVR) and artificial neural networks (ANNs) (Riaz et al 2009, Murphy and Pokhrel 2009). While the model-free algorithms are effective for learning and adapting to slowly time-varying dynamics of respiratory motion, their prediction can largely deviate from the true position over a period of irregular respiratory motion. Recently, a real-time prediction algorithm has been proposed that utilizes both model-based and model-free approaches by combining them in a cascade form (Bukhari and Hong 2015). The algorithm exploits the advantages of both of the approaches in predicting respiratory motion.

In this article, we present a real-time algorithm for predicting respiratory motion in 3D space and realizing a gating function without pre-specifying a particular phase of the patient’s breathing cycle. The algorithm, named EKF-GPRN\(^+\), first employs an extended Kalman filter (named LCM-EKF) independently along each coordinate to predict the respiratory motion and then uses Gaussian process regression network (GPRN) to correct the error of LCM-EKF prediction in 3D space. The GPRN is a nonparametric Bayesian algorithm for modeling input-dependent correlations between output variables in multi-output regression (Wilson et al 2012, Wilson 2014). We employ variational inference with mean field approximation for GPRN to compute an approximate predictive mean and covariance matrix of the prediction errors of LCM-EKF (Nguyen and Bonilla 2013). The approximate predictive mean is used to correct the prediction error of LCM-EKF. The trace of the approximate
predictive covariance matrix is utilized to capture the uncertainty in EKF-GPRN+ prediction error and systematically identify breathing points with a higher probability of large prediction error in advance. This identification enables us to pause the treatment beam over such instances. EKF-GPRN+ implements a gating function by using simple calculations based on the trace of the predictive covariance matrix. It is demonstrated through extensive numerical experiments that EKF-GPRN+ reduces the prediction error to 38%, 40% and 40% in root-mean-square (RMS), compared to no prediction, at lookahead lengths of 192 ms, 384 ms and 576 ms, respectively. It is observed from the experiments that the use of the input-dependent correlations between the prediction error of EKF in 3D space plays a key role in improving the prediction accuracy. The error reduction allows the clinical target volume to planning target volume (CTV-PTV) margin to be reduced, leading to decreased normal-tissue toxicity and possible dose escalation (Vedam et al 2003, Keall et al 2006). We evaluate the CTV-PTV margin as well to demonstrate clinical benefits of EKF-GPRN+ prediction. It is also demonstrated through numerical experiments that EKF-GPRN+ can improve prediction accuracy and reduce the CTV-PTV margin further by employing the gating function, albeit at the cost of a reduced duty cycle.

2. Methods

In this section, we present EKF-GPR and EKF-GPRN+ prediction algorithms, starting with a brief introduction to LCM-EKF, EKF-GPR and EKF-GPR+ prediction algorithms and GPRN model.

2.1. LCM model and LCM-EKF prediction algorithm

The LCM model characterizes one-dimensional respiratory motion locally as a projection of a circular motion to capture the temporal evolution of the respiratory motion (Hong et al 2011). Let \( x(k), \dot{x}(k), y(k), \dot{y}(k), \Omega(k) \) denote the position, the velocity along the \( x \)-axis, the velocity along the \( y \)-axis, and the angular velocity, respectively. Defining the discrete-time state vector as \( \mathbf{x}(k) = [x(k) \; \dot{x}(k) \; y(k) \; \Omega(k)]^T \), we can represent the local circular motion as a discrete-time state equation

\[
\mathbf{x}(k+1) = \mathbf{f}(\mathbf{x}(k)) + \mathbf{v}(k),
\]

with a nonlinear mapping \( \mathbf{f} \) that characterizes one-step evolution of the state vector in the discrete time. The evolution of position \( x(k) \) is the projection of the planar circular motion onto the \( x \)-axis. The \( y \)-axis is an auxiliary axis augmented to define the circular motion. The process noise vector \( \mathbf{v}(k) \) is a zero-mean white noise sequence. The measurement equation of position \( x(k) \) is

\[
z(k) = x(k) + w(k),
\]

where \( z(k) \) is the position measurement at time \( k \), and \( w(k) \) denotes the additive measurement noise. Based on the LCM model, the LCM-EKF algorithm employs the first-order EKF to recursively update the state estimate \( \hat{x}(k|k) \) with measurement \( z(k) \), and evaluates the \( l \)-step lookahead prediction \( \hat{x}_{LCM}(k+1|k) \). We use the LCM-EKF algorithm to predict the respiratory motion in 3D space by applying it independently along each coordinate. More details on the LCM-EKF prediction algorithm can be found in Hong et al (2011) and Hong and Bukhari (2014).
2.2. EKF-GPR and EKF-GPR$^+$ prediction algorithms

The EKF-GPR algorithm employs LCM-EKF to predict the respiratory motion and then uses the Gaussian process regression (GPR) algorithm to estimate and correct the error of the LCM-EKF prediction (Bukhari and Hong 2015). The structure of the EKF-GPR algorithm is illustrated in figure 1. First, let us denote the LCM-EKF prediction error by $\xi_{\text{LCM}}(k + l | k)$, that is,

$$\xi_{\text{LCM}}(k + l | k) = z(k + l) - \hat{x}_{\text{LCM}}(k + l | k), \quad (3)$$

and let $\hat{x}(k | k)$ denote the state estimate of LCM-EKF at time $k$. The GPR algorithm predicts the conditional distribution of $\xi_{\text{LCM}}(k + l | k)$ given the state estimate $\hat{x}(k | k)$ under Gaussian assumptions. It evaluates the predictive mean (\(\bar{\xi}_{\text{LCM}}(k + l | k)\)) and the predictive variance (\(\text{var}(\xi_{\text{LCM}}(k + l | k))\)) of $\xi_{\text{LCM}}(k + l | k)$ as a function of $\hat{x}(k | k)$. We use $\bar{\xi}_{\text{LCM}}(\hat{x}(k | k))$ as the estimate of $\xi_{\text{LCM}}(k + l | k)$ and utilize it to correct the LCM-EKF prediction $\hat{x}_{\text{LCM}}(k + l | k)$ to obtain the prediction of EKF-GPR, denoted by $\hat{\xi}(k + l | k)$, as

$$\hat{\xi}(k + l | k) = \hat{x}_{\text{LCM}}(k + l | k) + \bar{\xi}_{\text{LCM}}(k + l | k), \quad (4)$$

where $\hat{x}_{\text{LCM}}(k + l | k)$ denotes the estimate of $\xi_{\text{LCM}}(k + l | k)$. The GPR algorithm is a supervised learning algorithm. It evaluates the predictive mean $\xi_{\text{LCM}}(\hat{x}(k | k))$ and the predictive variance $\text{var}(\xi_{\text{LCM}}(\hat{x}(k | k)))$ based on learning over a set of training samples $\{\hat{x}(i | i), \xi_{\text{LCM}}(i + l | i)\}_{i \in I}$ as illustrated in figure 1, where $I$ denotes the index set of training samples. The predictive mean is the best point-prediction of the prediction error $\xi_{\text{LCM}}(k + l | k)$. The predictive variance $\text{var}(\xi_{\text{LCM}}(\hat{x}(k | k)))$ is also a useful measure for estimating the strength of EKF-GPR prediction error after the error correction. The algorithm named EKF-GPR$^+$ utilizes the predictive variance to capture the strength in the EKF-GPR prediction error and systematically identify breathing points with a higher probability of large prediction variance in advance. This identification enables us to deactivate the treatment beam over such instances. EKF-GPR$^+$ implements a gating function by using simple calculations based on the predictive variance. The EKF-GPR$^+$ algorithm effectively reduces the prediction error in RMS by employing the gating function, albeit at the cost of a reduced duty cycle. We use EKF-GPR$^+$ ($D\%$) to denote the EKF-GPR$^+$ algorithm with a nominal duty cycle parameter set to $D$. Note that EKF-GPR$^+$ ($100\%$) is equivalent to EKF-GPR. More details on the EKF-GPR$^+$ algorithm can be found in Bukhari and Hong (2015).
We can use the EKF-GPR algorithm for predicting and gating respiratory motion in 3D space. First, we apply the EKF-GPR algorithm independently along each coordinate for prediction and error correction. Then we use the sum of all the predictive variances as an approximate measure of the strength of EKF-GPR prediction error in 3D space after the error correction. EKF-GPR utilize the sum to implement a gating function. The basic structure of EKF-GPR for 3D prediction is presented in figure 2.

2.3. Gaussian process regression networks (GPRN)

Modeling correlations between the output variables in the multi-output regression framework is known to improve regression performance (Breiman and Friedman 1997). Gaussian process (GP) models have been developed for the multi-output regression assuming fixed correlations between output variables (Bonilla et al. 2008, Alvarez and Lawrence 2011). Dürrichen et al. (2015) utilized a multi-task GP framework for motion compensation in radiation therapy. The fixed correlation assumption is, however, restrictive since correlations between output variables can be input-dependent.

GPRN is a new multi-output regression framework, which combines the structural properties of Bayesian neural networks with the nonparametric flexibility of Gaussian processes (Wilson et al. 2012, Wilson 2014). This GPRN model represents each output variable as an adaptive mixture of latent functions shared across all the output variables. This representation enables GPRN to model input-dependent correlations between multiple output variables and improve regression performance. GPRN also underlies a highly expressive nonstationary kernel that learns nonstationary covariance structures between input and output variables. Such a nonstationary kernel can express more complicated structures in the data than stationary kernels that GPR can utilize (Rasmussen and Williams 2006). The following is a brief summary of GPRN. More details can be found in Wilson et al. (2012) and Wilson (2014).

Suppose that we are given $N$ samples of input-output pairs, $\{((x(i), y(x(i))))\}_{i=1}^N$, where $x(i)$ denotes a $d$-dimensional input vector and $y(x(i))$ is the corresponding $p$-dimensional output. Our goal is to predict the mean and the covariance matrix of the predictive distribution $p(y(x)|D)$ at a new input $x$. The GPRN represents the output vector $y(x)$ as

$$y(x) = W(x)[f(x) + \sigma_f z_f] + \sigma_z z_r,$$

where $W(x)$ is a $p \times q$ matrix of mixing weights, each being an independent Gaussian process with $W(x) \sim GP(0, k_w)$, and $f(x)$ is a $q$-dimensional vector of latent functions, each being also an independent Gaussian process with $f(x) \sim GP(0, k_f)$, $z_f$ and $z_r$ are white noise processes with $\mathcal{N}(0, I_q)$ and $\mathcal{N}(0, I_p)$, respectively. Here, $I_q$ and $I_p$ are $q \times q$ and $p \times p$ identity matrices. This structure of GPRN is depicted in figure 3, where $\hat{f}(x)$ represents the latent function $f(x)$ with additive Gaussian noise $\sigma_f z_f$ such that $\hat{f}(x) = f(x) + \sigma_f z_f \sim GP(0, k_f)$. Here, $k_f$ is the covariance kernel of $\hat{f}(x)$ such that
where $\delta_{aw}$ denotes the Kronecker delta.

It can be seen from figure 3 that each output variable is an input-dependent mixture of latent Gaussian process basis functions with the mixing weights being also Gaussian processes. Each of the outputs $y_i(x)$, conditioned on the mixing weights $W(x)$, is a Gaussian process with kernel (Wilson 2014)

$$k_y(x_i, x_u) = \sum_{j=1}^{q} W_j(x_u) k_j(x_i, x_u) W_j(x_u) + \delta_{aw} \sigma_i^2,$$

Note that this underlying GPRN kernel is an input-dependent mixture of kernels used in latent functions. This underlying kernel is highly expressive that can extract a nonstationary covariance structure in the data.

Inference in GPRN is intractable. Variational inference methods have been usually utilized to approximate posterior distributions for GPRN inference (Wilson et al 2012, Nguyen and Bonilla 2013). The approximate distributions are employed in learning hyperparameters $\{\theta_i, \theta_u, \sigma_j, \sigma_i\}$, where $\theta_i$ and $\theta_u$ denote the hyperparameters for kernels $k_j$ and $k_u$, respectively. The approximations are also used in deriving the mean and covariance of the predictive distribution. The approximate predictive mean of an output variable $y_i(x)$ and the approximate predictive covariance between two output variables $y_i(x)$ and $y_j(x)$ over a new input $x$ are given, respectively, as

$$\hat{y}_i(x) = \sum_k E[W_k(x)]E[f_k(x)],$$

for all $i$, and

Figure 3. Structure of Gaussian process regression network (Wilson 2014).
\[
\text{cov}(y(x))_{ij} = \sum_k [E[W_i(x)]E[W_j(x)]\text{var}(f_k(x)) + \delta_i \text{var}(W_i(x))E[f_k^2(x)] + \delta_j E[\sigma^2_f]],
\]
for all \(i\) and \(j\). Note that the approximate predictive mean and covariance depend on the first two moments of latent functions and weight functions.

2.4. EKF-GPRN and EKF-GPRN+ prediction algorithms

The EKF-GPRN algorithm first employs LCM-EKF independently along each coordinate to predict the respiratory motion and then utilizes GPRN as a multi-output regression in 3D space to estimate and correct the errors of LCM-EKF prediction. The basic structure of EKF-GPRN for 3D prediction is presented in figure 4. In contrast to EKF-GPR in 3D (shown in figure 2), EKF-GPRN utilizes GPRN to jointly estimate and correct the error vector of LCM-EKF prediction in 3D space. Let \(x_i(k) = [x_i(k), \dot{x}_i(k), \ddot{x}_i(k)]^T\) denote the state vector of LCM-EKF in coordinate \(i\). Let \(\hat{x}(k) = [\hat{x}_1^T(k), \hat{x}_2^T(k), \hat{x}_3^T(k)]^T\) and use \(\hat{x}(k)\) to denote the estimate of \(x(k)\) at time \(k\). Similarly, we use \(z_i(k)\) to denote the position measurement in coordinate \(i\) and \(\hat{z}(k)\) to denote the vector \([z_1(k), z_2(k), z_3(k)]^T\). Also, we use \(\hat{z}_{\text{LCM}}(k + l|k)\) and \(\hat{z}(k + l|k)\) to denote the \(l\)-step lookahead prediction of LCM-EKF and EKF-GPRN in 3D space, respectively. In addition, we use \(\xi_{\text{LCM}}(k + l|k)\) to denote the LCM-EKF prediction error, that is, \(\hat{x}_{\text{LCM}}(k + l|k) = z(k + l) - \hat{z}_{\text{LCM}}(k + l|k)\), and use \(\hat{\xi}_{\text{LCM}}(k + l|k)\) to denote the estimate of \(\xi_{\text{LCM}}(k + l|k)\). Then, we can use figure 1 to describe the structure of EKF-GPRN, replacing \(z(k), \hat{z}_{\text{LCM}}(k + l|k), \hat{z}(k + l|k)\) and GPR in the figure with \(z(k), \hat{z}_{\text{LCM}}(k + l|k), \hat{z}(k + l|k)\) and GPRN, respectively. Note that the LCM-EKF block of EKF-GPRN in the figure consists of three LCM-EKFs implemented independently along each coordinate.

It has been reported that strong correlations exist between the state estimate \(\hat{x}(k|k)\) and the prediction error \(\xi_{\text{LCM}}(k + l|k)\) (Hong and Bukhari 2014). The EKF-GPR algorithm exploits this correlation in estimating \(\xi_{\text{LCM}}(k + l|k)\) based on \(\hat{x}(k|k)\), independently for each coordinate, and uses this estimate to correct the LCM-EKF prediction. We observe that strong correlations also exist between the components of the 3D error vector \(\xi_{\text{LCM}}(k + l|k)\) and the correlations are dependent on \(\hat{x}(k|k)\), as we will discuss later in this work. We adopt GPRN to learn and exploit these two different types of correlations in predicting the error \(\xi_{\text{LCM}}(k + l|k)\). Note that the GPRN framework can learn input-dependent correlations between multiple output variables as well as nonstationary covariance structures between input and output variables. While Bayesian models have been developed for multi-output regression to model fixed correlations between the output variables, GPRN is the only known Bayesian model that can represent the input-dependent correlations between output variables.

The GPRN algorithm in EKF-GPRN evaluates the predictive mean of \(\hat{x}_{\text{LCM}}(k + l|k)\) \(\langle \hat{\xi}_{\text{LCM}}(\hat{x}(k|k)) \rangle\) and the predictive covariance matrix \(\text{cov}(\hat{\xi}_{\text{LCM}}(\hat{x}(k|k)))\) as a function of \(\hat{x}(k|k)\)
using equations (8) and (9), respectively. We use \( \bar{\xi} \) \( \hat{()}/LCM(\hat{()}) \) as the estimate of \( \xi_LCM(k + l|k) \) as
\[
\bar{\xi}_{LCM}(k + l|k) = \bar{\xi}_{LCM}(\hat{x}(k|k)). \tag{10}
\]
Note that \( x \) and \( y(x) \) in equation (5) correspond to \( \hat{x}(k|k) \) and \( \xi_{LCM}(k + l|k) \), respectively. We use the error estimate \( \hat{\xi}_{LCM}(k + l|k) \) to correct the LCM-EKF prediction \( \hat{z}_{LCM}(k + l|k) \) and obtain the 1-step prediction of EKF-GPRN as
\[
\hat{z}(k + l|k) = \hat{z}_{LCM}(k + l|k) + \hat{\xi}_{LCM}(k + l|k). \tag{11}
\]

The GPRN algorithm evaluates the predictive mean \( \bar{\xi}_{LCM}(\hat{x}(k|k)) \) and the predictive covariance \( \text{cov}(\hat{z}_{LCM}(\hat{x}(k|k))) \) based on learning over a set of training samples \( \{\hat{x}(i|i), \xi_{LCM}(i + l|i)\}_{i \in I} \) as illustrated in figure 1.

We use \( \text{tr}[\text{cov}(\xi_{LCM}(\hat{x}(k|k)))] \) to denote the trace of the predictive covariance matrix and employ it to quantify the strength of EKF-GPRN prediction error (after the error correction). The trace approximately corresponds to the mean square error (MSE) of the EKF-GPRN prediction in 3D space. This relation can be made more precise. First, we can express the MSE of LCM-EKF prediction as
\[
E[\xi_i^T(k + l|k)\xi_{LCM}(k + l|k)] = |E[\xi_{LCM}(k + l|k)]|^2 + \text{tr}[\text{cov}(\xi_{LCM}(k + l|k))]. \tag{12}
\]

Suppose that GPRN learning is perfect and the predictive mean \( \bar{\xi}_{LCM}(\hat{x}(k|k)) \) and predictive covariance \( \text{cov}(\xi_{LCM}(\hat{x}(k|k))) \) are equal to \( E[\xi_{LCM}(k + l|k)] \) and \( \text{cov}(\xi_{LCM}(k + l|k)) \), respectively. Then, the EKF-GPRN prediction has zero mean and covariance equal to the predictive covariance \( \text{cov}(\xi_{LCM}(\hat{x}(k|k))) \) as we correct the LCM-EKF prediction error \( \hat{\xi}_{LCM}(k + l|k) \) with the predictive mean \( \bar{\xi}_{LCM}(\hat{x}(k|k)) \). This implies that the MSE value of EKF-GPRN prediction is approximately equal to \( \text{tr}[\text{cov}(\xi_{LCM}(\hat{x}(k|k)))] \) as long as the predictive mean and covariance are reasonably accurate. It supports that we can use the trace of the predictive covariance as an approximate measure for estimating the strength of a possible prediction error of EKF-GPRN.

We utilize the trace of the predictive covariance matrix from the GPRN to identify breathing points with a higher probability of large prediction error in advance. This identification enables us to deactivate the treatment beam over such instances. We maintain a set of traces of the matrices sampled over a training window to determine a threshold on the trace. The threshold for a given duty cycle \( D\% \) is determined as the sample point at which \((100 - D)\%\) of the traces in the set lie above it. The trace associated with each of the subsequent predictions is then compared to this threshold. If the trace is greater than this threshold, we deactivate the treatment beam for that prediction. We name the EKF-GPRN algorithm with this detection mechanism as the EKF-GPRN+ algorithm. The EKF-GPRN+ algorithm effectively reduces the prediction error in RMS by employing the gating function, albeit at the cost of a reduced duty cycle. We use EKF-GPRN+ \( (D\%) \) to denote the EKF-GPRN+ algorithm with a nominal duty cycle parameter set to \( D \). Note that EKF-GPRN+ \( (100\%) \) is equivalent to EKF-GPRN.

2.5. Implementation

A simple grid search is performed over the parameter space of each algorithm to find the optimum parameters that minimize the prediction error of the corresponding algorithm. For the grid search, we utilized the first 10 min of 31 respiratory motion traces, each from a different patient, in a large database of 304 respiratory motion traces that will be described in
We implemented the LCM-EKF algorithm as described in Hong and Bukhari (2014). The dynamics of motion along different coordinates, however, are different. We optimized the process noise parameters of LCM-EKF separately for each coordinate. The magnitude of respiratory motion along coordinate 2 is the largest, and the optimum values of process noise parameters for this coordinate are identical with those obtained over PCA-processed one-dimensional traces in Hong and Bukhari (2014): $q_1 = 0.1, q_2 = 0.1, q_3 = 10^{-4}$ and $q_4 = 10^{-5}$. The process noise parameters for coordinates 1 and 3 were identically set as $q_1 = 2.5 \times 10^{-2}, q_2 = 2.5 \times 10^{-2}, q_3 = 2.5 \times 10^{-5}$ and $q_4 = 10^{-5}$. The variance of the measurement noise $w(k)$ was set to $10^{-4}$ for all the coordinates.

Inference in GPRN is intractable and we use the mean-field approximation for the inference which is a variational Bayesian method. The method approximates the posterior distribution as a factorized distribution over the parameters of GPRN with each factor distribution being a Gaussian (Nguyen and Bonilla 2013). As input to GPR and GPRN, we use only the velocity components in the state estimate vector of LCM-EKF. To be more specific, we use the estimate of $[x_i(k) \dot{x}_i(k)]^T$ for GPR of coordinate $i$ and the estimate of $[\dot{x}_i(k) \ddot{x}_i(k) \dot{z}_i(k) \ddot{z}_i(k)]^T$ for GPRN. The number of latent functions in GPRN is an important factor which affects the prediction performance. We obtained the prediction performance of EKF-GPRN with the number of latent functions from 1 to 6 over the first 10 min of 31 traces for each lookahead length. The best prediction performance was achieved when the number of latent functions was equal to or greater than 4, and we set it to 4.

EKF-GPRN$^+$ has been implemented in Matlab and run on a PC (Intel(R) Core(TM) i7-3770 CPU @ 3.40 GHz). In our implementation, we use the first 1 min data from 31 traces, each from a different patient, to train and obtain an initial GPRN model off-line. The GPRN algorithm is initialized with this initial GPRN model for all traces and it is updated every minute. Note that the GPRN algorithm learns the mapping from LCM-EKF state estimates to its prediction error. The mapping does not change fast in time and the GPRN update rate was high enough. The maximum number of iterations in learning hyperparameters for the GPRN update was set to 5 and the computation time for the update was a maximum of 1.5 sec. The update is not time-sensitive and it can be performed in the background. We use the updated GPRN model over a following 1 min interval. The computation time for LCM-EKF prediction in 3D is approximately 150 $\mu$sec and the total computation time is 370 $\mu$sec for completing one cycle of EKF-GPRN$^+$ with a sample data.

The first principal component of a 3D trace has been widely used in studies on the prediction of respiratory motion in 1D space. For a practical assessment of prediction over the first principal component, we implement an online PCA processing and evaluate real-time prediction of 3D respiratory motion in the principal component direction. We obtain the first principal component of a given trace and its direction through the online PCA processing over a 1 min sliding window, and use LCM-EKF and EKF-GPR algorithms to predict the first principal component and project the prediction back to the original 3D space. The prediction error is the difference between the prediction and the corresponding measurement in 3D space. The algorithms are referred to as LCM-EKF(PCA) and EKF-GPR(PCA). The structure of EKF-GPR(PCA) is shown in figure 5, whereas LCM-EKF(PCA) does not include the GPR components.

![Figure 5. Basic structure of EKF-GPR(PCA) for 3D prediction.](image-url)
3. Numerical experiments

3.1. Materials

A large database of 304 respiratory motion traces is employed to evaluate the proposed algorithms. The 3D traces in the database were obtained during CyberKnife treatment at Georgetown University Hospital from a group of 31 patients (Ernst 2012). The database also includes PCA-processed 1D traces extracted over the longest interval in between the large couch movements. The PCA traces have been employed to evaluate 1D prediction algorithms (Ernst et al. 2013, Hong and Bukhari 2014). Likewise, we extracted 3D traces from the original traces over the intervals each corresponding to its PCA counterpart and we use them to evaluate our 3D prediction algorithms. The database lists the traces sorted in ascending alphanumeric order. For convenience in our presentation, we assigned the numeric IDs from 1 to 304 to these traces in the same order.

We implemented the proposed algorithms and performed experiments to evaluate their prediction performance over the 3D traces at lookahead lengths of 192 ms, 384 ms and 576 ms. In our experiments, we downsample the traces to 5.2 Hz for conformity to image-guided treatment practices including MRI-guided radiotherapy (Cervino et al. 2011, Cho et al. 2011, Tehrani et al. 2013, Lagendijk et al. 2014, Yip et al. 2014). The experimental results in this section are based on the data that we obtained over each trace after the first 1 min.

3.2. Performance measures

We use the RMS prediction error and the CTV-PTV margin to evaluate the performance of prediction algorithms. First, let $\mathbf{z}^{(i)}(k+l)$ denote the measurement at $k+l$ for a trace $i$, and let $\mathbf{Z}^{(i)}(k+l)$ denote its corresponding $l$-step lookahead prediction. Let $\mathbf{e}^{(i)}(k+l) = \mathbf{z}^{(i)}(k+l) - \mathbf{Z}^{(i)}(k+l)$ and let $||\mathbf{e}^{(i)}(k+l)||$ denote the magnitude of the 3D prediction error. We evaluate the RMS prediction error (RMSE$^{(i)}$) for 3D trace $i$ as

$$\text{RMSE}^{(i)} = \frac{1}{|\mathbf{K}^{(i)}|} \sum_{k \in \mathbf{K}^{(i)}} ||\mathbf{e}^{(i)}(k+l)||^2,$$

where $\mathbf{K}^{(i)}$ is an index set of samples of trace $i$ and $|\mathbf{K}^{(i)}|$ denotes the number of indices in the set. Let RMSE$(j)$ denote the patient-wise RMS prediction error of patient $j$, averaged across all traces of the patient. We obtain the patient-wise RMS error as

$$\text{RMSE}(j) = \frac{1}{|I(j)|} \sum_{i \in I(j)} (\text{RMSE}^{(i)})^2,$$

where $I(j)$ is the index set of traces of patient $j$ and $|I(j)|$ is the number of traces of patient $j$. We evaluate the mean, standard deviation, 90th percentile point and maximum value of the set $\{\text{RMSE}(j)\}_{j \in J}$, where $J$ is the index set of all patients. These statistics indicate the interpatient variability of RMS prediction error of an algorithm. The RMS prediction error averaged across all of the patients is obtained as

$$\text{RMSE}_p = \frac{1}{|J|} \sum_{j \in J} \text{RMSE}(j)^2,$$

where $|J|$ is the total number of patients.
The CTV-PTV margin is also evaluated to quantify clinical benefits of EKF-GPRN+. The margin is added around the CTV to ensure that the prescribed dose is actually delivered to the CTV, accounting for the effect of targeting error of the treatment beam (Vedam et al 2003, Keall et al 2006). The targeting error can be reduced by predicting the position of tumour and directing the treatment beam towards the predicted position. The error reduction allows the CTV-PTV margin to be reduced, leading to decreased normal-tissue toxicity and possible dose escalation.

In our experiment, we calculate the CTV-PTV margin that is required to maintain the CTV to lie within the PTV for 95% of each respiratory motion trace at a given lookahead length. Note that the CTV-PTV margin corresponds to the 95th percentile point of the set \( \{ \| e^{(k+l)} \| \}_{k \in \mathbb{K}} \), where \( \| e^{(k+l)} \| \) denotes the magnitude of the \( l \)-step lookahead prediction error of trace \( i \) as described above. Let \( m^{(i)} \) denote the CTV-PTV margin of trace \( i \) for a given lookahead length. We obtain the patient-wise CTV-PTV margin averaged across all the traces of each patient. Let \( m^{(j)} \) denote the patient-wise CTV-PTV margin of patient \( j \). For the set \( \{ m^{(j)} \}_{j \in J} \), we evaluate its mean, standard deviation, 90th percentile point and maximum value.

3.3. Prediction performance of EKF-GPR+ and EKF-GPRN+ algorithms

Firstly, we evaluated the RMS prediction error of LCM-EKF, EKF-GPR+ and EKF-GPRN+ for each patient. Table 1 presents a statistical summary of the RMS error. Note that ‘no prediction’ in the table corresponds to the case that we employ the current position measurement itself as the lookahead prediction. The values in the parentheses in the table are the percent ratios relative to no prediction. The table shows that LCM-EKF, EKF-GPR and EKF-GPRN algorithms reduce the RMS error of the baseline no prediction approximately by 28%, 45% and 60%, respectively, across all of the lookahead lengths. Similar observations hold for the other statistics. These results confirm that EKF-GPRN reduces the prediction error as well as its inter-patient variability significantly and more effectively than EKF-GPR.

The left column of figure 6 presents scatter plots of the patient-wise RMS error of EKF-GPR and EKF-GPRN against the RMS error of no prediction for lookahead lengths of 192 ms, 384 ms and 576 ms. The two straight lines in each plot, corresponding to the 40% RMS error and 50% RMS error of no prediction, serve as reference lines for comparison against no prediction. The vertical difference between a pair of markers corresponds to the reduction in prediction error achieved by EKF-GPRN over EKF-GPR, and we can see that it increases as the lookahead length increases. In all of the plots, EKF-GPRN markers are always located below the corresponding EKF-GPR markers, indicating that EKF-GPRN reduces the RMS error of EKF-GPR over all of the 31 patients across all of the lookahead lengths. The EKF-GPR markers are populated slightly above the 50% RMS error line, whereas most of the EKF-GPRN markers are located below that line around 40%. This advantage of EKF-GPRN over EKF-GPR is statistically significant. We performed the paired Student’s \( t \)-test to confirm the statistical significance of EKF-GPRN in prediction accuracy over EKF-GPR. The test rejected the null hypothesis with strong evidence for all the lookahead lengths with \( p \)-values less than \( 10^{-6} \).

Table 1 also presents a statistical summary of the patient-wise RMS prediction error of EKF-GPR+ and EKF-GPRN+ for the duty cycles of 70%, 80% and 90%. The table shows that EKF-GPRN+(70%), EKF-GPRN+(80%) and EKF-GPRN+(90%) reduce the patient-wise RMS error approximately to 32%, 33% and 35%, respectively, compared to no prediction.
Table 1. Statistical summary of patient-wise RMS prediction error (mm) (figures in parentheses are percent ratios relative to no prediction).

<table>
<thead>
<tr>
<th>Lookahead length</th>
<th>No prediction</th>
<th>LCM-EKF</th>
<th>EKF-GPR</th>
<th>EKF-GPR + Duty cycle D</th>
<th>EKF-GPRN + Duty cycle D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE&lt;sub&gt;P&lt;/sub&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>192 ms</td>
<td>0.79</td>
<td>0.57 (72)</td>
<td>0.44 (56)</td>
<td>0.34 (43)</td>
<td>0.36 (45)</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.75</td>
<td>0.54 (72)</td>
<td>0.42 (55)</td>
<td>0.32 (43)</td>
</tr>
<tr>
<td></td>
<td>std dev</td>
<td>0.25</td>
<td>0.19 (76)</td>
<td>0.15 (58)</td>
<td>0.11 (43)</td>
</tr>
<tr>
<td></td>
<td>90th percentile</td>
<td>1.01</td>
<td>0.82 (81)</td>
<td>0.67 (66)</td>
<td>0.46 (46)</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>1.40</td>
<td>0.96 (69)</td>
<td>0.76 (55)</td>
<td>0.65 (47)</td>
</tr>
<tr>
<td>384 ms</td>
<td>RMSE&lt;sub&gt;P&lt;/sub&gt;</td>
<td>1.43</td>
<td>1.03 (72)</td>
<td>0.77 (54)</td>
<td>0.59 (41)</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>1.36</td>
<td>0.98 (72)</td>
<td>0.73 (54)</td>
<td>0.56 (41)</td>
</tr>
<tr>
<td></td>
<td>std dev</td>
<td>0.45</td>
<td>0.33 (73)</td>
<td>0.25 (56)</td>
<td>0.19 (42)</td>
</tr>
<tr>
<td></td>
<td>90th percentile</td>
<td>1.87</td>
<td>1.37 (73)</td>
<td>1.03 (55)</td>
<td>0.84 (45)</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>2.58</td>
<td>1.77 (69)</td>
<td>1.29 (50)</td>
<td>0.95 (37)</td>
</tr>
<tr>
<td>576 ms</td>
<td>RMSE&lt;sub&gt;P&lt;/sub&gt;</td>
<td>1.96</td>
<td>1.44 (73)</td>
<td>1.08 (55)</td>
<td>0.85 (43)</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>1.87</td>
<td>1.37 (73)</td>
<td>1.02 (54)</td>
<td>0.80 (43)</td>
</tr>
<tr>
<td></td>
<td>std dev</td>
<td>0.60</td>
<td>0.45 (75)</td>
<td>0.35 (59)</td>
<td>0.27 (45)</td>
</tr>
<tr>
<td></td>
<td>90th percentile</td>
<td>2.43</td>
<td>1.77 (73)</td>
<td>1.29 (53)</td>
<td>1.11 (46)</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>3.52</td>
<td>2.43 (69)</td>
<td>1.87 (53)</td>
<td>1.44 (41)</td>
</tr>
</tbody>
</table>
Figure 6. Patient-wise RMS prediction error and CTV-PTV margin for comparison of EKF-GPR and EKF-GPRN. Left half: RMS prediction error; right half: CTV-PTV margin. (a), (b) 192 ms. (c), (d) 384 ms. (e), (f) 576 ms.
across all of the lookahead lengths. These results correspond to an additional error reduction by 10% to 15% over EKF-GPR in terms of the RMS prediction error of no prediction.

The threshold on the trace of the predictive covariance matrix of EKF-GPRN was determined by setting the duty cycle parameter $D$ to 70%, 80% and 90% in the above experiments. Recall that we use a set of traces of the approximate predictive covariance matrices from GPRN sampled over a 1 min training window to determine the threshold and employ this threshold over the following minute. As a result, the true duty cycle ($D^*$) is not exactly the same as set by the nominal parameter $D$. In our experiments, the true duty cycles have their values in an interval ($\{D-3\%\}, D\%)$ for a given duty cycle $D\%$ for all the patients across all the lookahead lengths. It indicates that EKF-GPRN controls the duty cycle with a high accuracy. The average of the true duty cycles were 68%, 79% and 89% for the nominal duty cycles of 70%, 80% and 90%, respectively, across all of the lookahead lengths.

### 3.4. CTV-PTV margin of EKF-GPR+ and EKF-GPRN+ algorithms

Table 2 presents a statistical summary of the patient-wise CTV-PTV margin. It shows that the average CTV-PTV margins required with no prediction are 1.30 mm, 2.35 mm and 3.23 mm, respectively, for lookahead lengths of 192 ms, 384 ms and 576 ms. EKF-GPR reduces the average margins to 0.49 mm, 0.98 mm and 1.38 mm, respectively, which correspond to 38%, 42%, and 43% of the baseline no prediction. The table also shows that LCM-EKF and EKF-GPR algorithms reduce the margin to approximately 73% and 57% of no prediction, respectively, across all of the lookahead lengths. Note that the average margins in table 2 are between 1.65 and 1.75 times greater than the corresponding RMS prediction errors ($\text{RMSE}_P$) presented in table 1. Similar observations hold for the other statistics in the table.

The right column of figure 6 presents scatter plots of the patient-wise CTV-PTV margins required with EKF-GPR and EKF-GPRN against that of no prediction for lookahead lengths of 192 ms, 384 ms and 576 ms. The two straight lines in each plot correspond to the 40% margin and 50% margin of no prediction. The distributions of the margin are not quite different from those of the RMS prediction error presented in the left column, except that their axis scales are different. As a result, the preceding discussion on the RMS prediction error similarly holds for the margin except that the vertical difference between a pair of markers (corresponding to the reduction in the CTV-PTV margin achieved by EKF-GPRN over EKF-GPR) increased approximately by a factor of 1.7. The advantage of EKF-GPRN over EKF-GPR in the margin is also statistically significant for all the lookahead lengths with $p$-values less than $10^{-6}$.

Table 2 also presents a statistical summary of the patient-wise CTV-PTV margin required with EKF-GPR+ and EKF-GPRN+. The table shows that EKF-GPRN+ (70%), EKF-GPRN+ (80%) and EKF-GPRN+ (90%) reduce the patient-wise margin to 29%, 31% and 33%, respectively, compared to that of no prediction at the lookahead length of 192 ms, 32%, 34% and 37%, respectively, at 384 ms, and 34%, 36% and 38%, respectively, at 576 ms. These results correspond to an additional reduction of the CTV-PTV margin by 12% to 16% over EKF-GPR+ relative to no prediction. Note that EKF-GPRN reduces the prediction error and the margin even more than EKF-GPR+ (70%).

Figure 7 presents scatter plots of the patient-wise CTV-PTV margin of EKF-GPRN, EKF-GPRN+ (70%) and EKF-GPRN+ (90%) against that of no prediction at the lookahead lengths of 192 ms, 384 ms and 576 ms. The two straight lines in each plot correspond to the 30% margin and 40% margin of no prediction. The scatter plots show that the CTV-PTV margin of EKF-GPRN+ decreases as the duty cycle decreases for all of the 31 patients across all of the lookahead lengths, confirming that the gating function satisfies its purpose.
Table 2. Statistical summary of patient-wise CTV-PTV margin (mm) (figures in parentheses are percent ratios relative to no prediction).

<table>
<thead>
<tr>
<th>Lookahead length</th>
<th>No prediction</th>
<th>LCM-EKF</th>
<th>EKF-GPR</th>
<th>70%</th>
<th>80%</th>
<th>90%</th>
<th>EKF-GPRN</th>
<th>70%</th>
<th>80%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>192 ms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>1.30</td>
<td>0.93 (72)</td>
<td>0.73 (56)</td>
<td>0.57 (44)</td>
<td>0.60 (46)</td>
<td>0.64 (49)</td>
<td>0.49 (38)</td>
<td>0.38 (29)</td>
<td>0.40 (31)</td>
<td>0.43 (33)</td>
</tr>
<tr>
<td>std dev</td>
<td>0.46</td>
<td>0.35 (76)</td>
<td>0.26 (57)</td>
<td>0.18 (39)</td>
<td>0.20 (43)</td>
<td>0.21 (46)</td>
<td>0.18 (39)</td>
<td>0.13 (28)</td>
<td>0.14 (30)</td>
<td>0.15 (33)</td>
</tr>
<tr>
<td>90th percentile</td>
<td>1.80</td>
<td>1.42 (79)</td>
<td>1.14 (63)</td>
<td>0.89 (49)</td>
<td>0.94 (52)</td>
<td>1.02 (57)</td>
<td>0.72 (40)</td>
<td>0.58 (32)</td>
<td>0.64 (36)</td>
<td>0.66 (37)</td>
</tr>
<tr>
<td>Max</td>
<td>2.44</td>
<td>1.74 (71)</td>
<td>1.26 (52)</td>
<td>0.99 (41)</td>
<td>1.01 (41)</td>
<td>1.08 (44)</td>
<td>0.97 (40)</td>
<td>0.71 (29)</td>
<td>0.75 (31)</td>
<td>0.81 (33)</td>
</tr>
<tr>
<td>384 ms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>2.35</td>
<td>1.72 (73)</td>
<td>1.34 (57)</td>
<td>1.03 (44)</td>
<td>1.09 (46)</td>
<td>1.17 (50)</td>
<td>0.98 (42)</td>
<td>0.75 (32)</td>
<td>0.79 (34)</td>
<td>0.86 (37)</td>
</tr>
<tr>
<td>std dev</td>
<td>0.83</td>
<td>0.63 (76)</td>
<td>0.48 (58)</td>
<td>0.36 (43)</td>
<td>0.38 (46)</td>
<td>0.41 (49)</td>
<td>0.38 (46)</td>
<td>0.29 (35)</td>
<td>0.30 (36)</td>
<td>0.33 (40)</td>
</tr>
<tr>
<td>90th percentile</td>
<td>3.16</td>
<td>2.46 (78)</td>
<td>1.85 (59)</td>
<td>1.47 (47)</td>
<td>1.54 (49)</td>
<td>1.63 (52)</td>
<td>1.31 (41)</td>
<td>0.96 (30)</td>
<td>1.02 (32)</td>
<td>1.14 (36)</td>
</tr>
<tr>
<td>Max</td>
<td>4.53</td>
<td>3.22 (71)</td>
<td>2.43 (54)</td>
<td>1.88 (42)</td>
<td>1.99 (44)</td>
<td>2.12 (47)</td>
<td>2.18 (48)</td>
<td>1.69 (37)</td>
<td>1.78 (39)</td>
<td>1.90 (42)</td>
</tr>
<tr>
<td>576 ms</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>3.23</td>
<td>2.37 (73)</td>
<td>1.90 (59)</td>
<td>1.51 (47)</td>
<td>1.60 (50)</td>
<td>1.70 (53)</td>
<td>1.38 (43)</td>
<td>1.09 (34)</td>
<td>1.16 (36)</td>
<td>1.23 (38)</td>
</tr>
<tr>
<td>std dev</td>
<td>1.11</td>
<td>0.81 (73)</td>
<td>0.68 (61)</td>
<td>0.53 (48)</td>
<td>0.56 (50)</td>
<td>0.59 (53)</td>
<td>0.56 (50)</td>
<td>0.43 (39)</td>
<td>0.46 (41)</td>
<td>0.49 (44)</td>
</tr>
<tr>
<td>90th percentile</td>
<td>4.51</td>
<td>3.20 (71)</td>
<td>2.62 (58)</td>
<td>1.99 (44)</td>
<td>2.06 (46)</td>
<td>2.22 (49)</td>
<td>1.99 (44)</td>
<td>1.46 (32)</td>
<td>1.54 (34)</td>
<td>1.70 (38)</td>
</tr>
<tr>
<td>Max</td>
<td>6.22</td>
<td>4.23 (68)</td>
<td>3.53 (57)</td>
<td>2.89 (46)</td>
<td>3.02 (49)</td>
<td>3.18 (51)</td>
<td>3.11 (50)</td>
<td>2.54 (41)</td>
<td>2.66 (43)</td>
<td>2.79 (45)</td>
</tr>
</tbody>
</table>
Figure 7. Patient-wise CTV-PTV margin comparison of EKF-GPRN and EKF-GPRN\(^+\) at 70\% and 90\% duty cycles. (a) 192 ms. (b) 384 ms. (c) 576 ms.
4. Discussion

EKF-GPR and EKF-GPRN algorithms estimate the error of LCM-EKF prediction and correct it by exploiting the correlations between the state estimate and the prediction error of LCM-EKF. Both of the algorithms significantly improve prediction accuracy over LCM-EKF. Prediction of respiratory motion in 3D space provides an additional opportunity for EKF-GPRN to exploit the correlations existing between LCM-EKF prediction errors in different coordinates, resulting in further improvement in prediction accuracy over EKF-GPR. The correlations between the prediction errors are dependent on the state vector estimate $\hat{x}(k|k)$ of LCM-EKF. Note that the latent functions in GPRN are input-dependent and sharing these latent functions across the output variables enables GPRN to represent the input-dependent correlations between the output variables. We utilized GPRN to model the input-dependent correlations between the prediction error of EKF, which is pivotal to improve the prediction accuracy in 3D space.

A showcase example is presented in figure 8 to highlight how effectively EKF-GPRN exploits the input-dependent correlations. Firstly, figure 8(a) illustrates a square grid of the correlation coefficients between the prediction errors of LCM-EKF in coordinates 1 and 2, where $x$ and $y$-axis correspond to the velocity estimate $\dot{x}(k|k)$ along coordinate 1 and the velocity estimate $\dot{y}(k|k)$ along coordinate 2, respectively. The data was obtained for all sample points over trace 13 at a lookahead length of 384 ms (this trace corresponds to the trace with ID p7_f22_m3 in the CyberKnife treatment database). The grid consists of 129 squares, each with $0.16 \text{ mm sec}^{-1} \times 1.6 \text{ mm sec}^{-1}$ in the input space and each square includes at least 10 samples. The figure shows that the LCM-EKF prediction errors have relatively strong correlations. The RMS prediction error of LCM-EKF for this trace is 1.93 mm at the lookahead length of 384 ms. EKF-GPR and EKF-GPRN reduce this error to 1.40 mm and 1.20 mm, respectively. Although EKF-GPR significantly reduces the LCM-EKF prediction errors, it cannot model the correlations between the prediction errors in 3D space. Figure 8(b) presents the distribution of input-dependent correlations between the EKF-GPR prediction errors after GPR error correction. We can see from the figure that the correlations become weaker but still maintain a certain level of strength even after the GPR error correction. The prediction errors of EKF-GPRN, in contrast, are almost decorrelated after the GPRN error correction; see figure 8(c). This example illustrates how effectively EKF-GPRN can exploit the input-dependent correlations between LCM-EKF prediction errors in improving its prediction accuracy.

Although not reported in this article, we implemented EKF-GPR and EKF-GPRN with different configurations and performed experiments to evaluate their performances. Through the experiments, we learned that the underlying covariance between LCM-EKF prediction error and its state estimates along each coordinate was not stationary and more complicated than can be reasonably expressed by simple GPR covariance functions. However, the advantage of the expressive underlying kernel of GPRN model contributed only in a small part to overall performance gain of EKF-GPRN over EKF-GPR. We learned in the experiments that the use of the input-dependent correlations between the output variables plays the key role in improving the prediction accuracy in 3D space.

The unscented Kalman filter (UKF) was developed as a nonlinear estimation method to overcome possible limitations arising from linearization that EKF relies on (Julier and Uhlmann 2004). We implemented an UKF based on the LCM model (named LCM-UKF) to compare with LCM-EKF. The comparison shows that LCM-UKF and LCM-EKF are comparable in prediction accuracy, but LCM-EKF is much faster in computation time than LCM-UKF. We also implemented LCM-EKF(PCA) and EKF-GPR(PCA) and evaluated their performances.
Figure 8. Input-dependent correlation coefficients between the prediction errors: (a) LCM-EKF (without error correction); (b) EKF-GPR (after GPR error correction); (c) EKF-GPRN (after GPRN error correction).
The results show that LCM-EKF(PCA) and EKF-GPR(PCA) yield significantly worse prediction performance compared to LCM-EKF and EKF-GPR, respectively. It suggests that lateral motion is significant in the plane normal to the first principal component direction and real-time PCA prediction is not a good choice for predicting 3D respiratory motion.

Irregularities in respiratory motion are difficult to predict and prediction accuracy deteriorates over such events, irrespective of prediction algorithms. Traces 223–231 are examples with irregular patterns persisting over the whole period of treatment. The traces are belong to the patient that corresponds to the EKF-GPRN marker singly lying above the 50% lines in figure 6. It can be seen from figure 7 that, despite of such irregularities, EKF-GPRN+ (70%) reduces the CTV-PTV margin from 60% to 40% that of no prediction. The marker at the far right (with a largest no prediction error) in figures 6 and 7 corresponds to traces 10–18. The traces represent respiratory motion with a large amplitude and high frequency, resulting in a largest prediction error and CTV-PTV margin.

Finally, we summarize the prediction and clinical advantages of EKF-GPRN+ as follows. The patient-wise RMS prediction errors of no prediction are 0.79 mm, 1.43 mm and 1.96 mm, respectively, for lookahead lengths of 192 ms, 384 ms and 576 ms. EKF-GPRN+ (70%) reduces the errors to 0.24 mm, 0.44 mm and 0.63 mm. The average CTV-PTV margins of no prediction are 1.30 mm, 2.35 mm and 3.23 mm, respectively, for lookahead lengths of 192 ms, 384 ms and 576 ms. EKF-GPRN+ (70%) reduces the margins to 0.38 mm, 0.75 mm and 1.09 mm.

5. Conclusion

A real-time algorithm for predicting and gating respiratory motion in 3D space has been presented. The algorithm, named EKF-GPRN+, first employs an extended Kalman filter (LCM-EKF) along each coordinate to predict the respiratory motion and then uses a Bayesian GPRN model for multi-output regression to jointly estimate and correct the errors of the LCM-EKF prediction in all the coordinates. The EKF-GPRN+ algorithm also implements an automated gating function without pre-specifying a particular region of the patient’s breathing cycle. Specifically, EKF-GPRN+ systematically identifies breathing points with a higher probability of large prediction error in advance so that we can deactivate the treatment beam over such instances. Extensive numerical experiments were performed based on a large database of 304 traces from 31 patients. The experimental results have shown that EKF-GPRN+ improve prediction accuracy significantly and more effectively than EKF-GPR+, allowing the CTV-PTV margin to be reduced further. The experiments also confirmed that EKF-GPRN+ controls the duty cycle with a high accuracy.

The extended Kalman filter seems more flexible and easier to implement information fusion from multiple imaging sources to guide radiation treatment compared to methods relying on learning. The EKF-GPRN+ algorithm has an attractive structure for this purpose with an extended Kalman filter at its first stage. We plan to investigate this direction in the future.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF-2015R1D1A3A01017331) and by the Institute for Information and Communications Technology Promotion (IITP) grant funded by the Korean Government (MSIP) (10041145, Self-Organizing Software Platform for Welfare Devices). The authors declare that they have no competing interests.
References


Ernst F 2012 Compensating for Quasi-Periodic Motion in Robotic Radiosurgery (Berlin: Springer)


Rasmussen C E and Williams C K I 2006 Gaussian Processes for Machine Learning (Cambridge, MA: MIT)


1966
Wilson A G 2014 Covariance kernels for fast automatic pattern discovery and extrapolation with Gaussian processes PhD Dissertation University of Cambridge