

Precisely Forecasting the Location and Motion of Lung Tumors

Dan Cervone

PQ Talk

May 6, 2013

Introduction

External beam radiotherapy:

- Lung tumor patients are given an implant (fiducial) at the location of their tumor.
- X-ray tomography can reveal location of the fiducial, thus the tumor.
- Radiotherapy is applied to the tumor location in a narrow beam, minimizing exposure within healthy tissue.

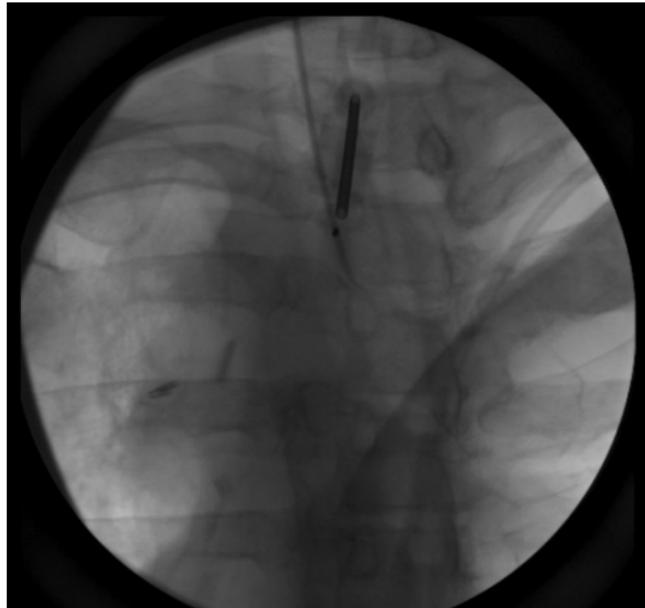
Introduction

Our problem:

- Patient's respiration means the tumor is in constant motion.
- Tracking the fiducial lags 0.1-2s behind, depending on specific machinery used.
- **We need to forecast the location of the tumor to overcome this latency and ensure concentrated, accurate radiotherapy.**

Illustration of the process

Fiducial



External beam radiotherapy



What we observe

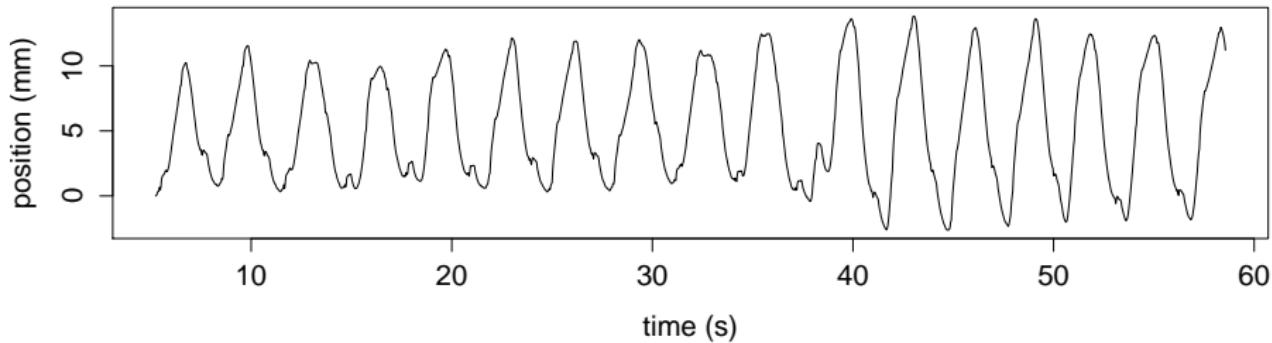
- Observed tumor position in 3D at 30 Hz.
- 1-5 min of observations per series.

What we predict

- Current observation predicts **future observation** 0.4s ahead.

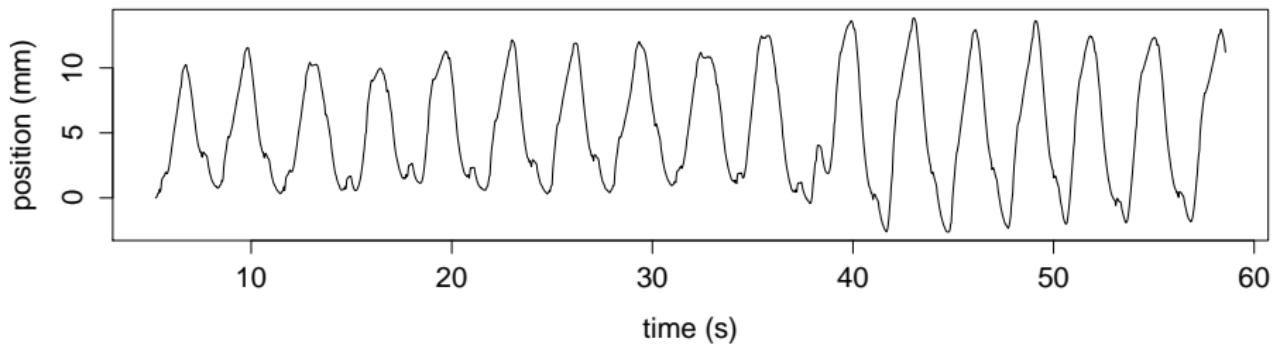
Modeling approaches

First principal component usually has over 99% of the variance, so we work with this alone:



Modeling approaches

First principal component usually has over 99% of the variance, so we work with this alone:



- Statistical methods: State space model [2], Kalman filtering [7], ridge regression [3], wavelets [6].
- Machine learning methods: Support Vector Regression [5, 1], Neural Networks [3, 4].
- Neural Networks are considered most effective. [3]

Philosophical musings

When predicting in a complex system, we must consider bias/variance tradeoff.

- Especially in prediction, more structural assumptions tend to lower variance but increase bias.
- Structure is often articulated in the form of distributional assumptions and dependence. Can also find structure by assuming certain types of invariance.
- When structure involves parameters, these are usually explored in the space of statistical models, not always in the space of the data.

Model overview

For a specific patient, we consider:

- Dictionary of d distinct motifs, $\Gamma = \{\gamma^{(1)}, \dots, \gamma^{(d)}\}$, characterize different breath cycle shapes.
- Location of tumor during breath cycle up to time t , μ_t , is a sequence of r transformed dictionary elements

$$\mu_t = (g_{\phi_1}(\gamma_1), g_{\phi_2}(\gamma_2), \dots, g_{\phi_r}(\gamma_r^*))$$

where ϕ_i parameterizes the transformation of γ_i .

Model overview

For a specific patient, we consider:

- Dictionary of d distinct motifs, $\Gamma = \{\gamma^{(1)}, \dots, \gamma^{(d)}\}$, characterize different breath cycle shapes.
- Location of tumor during breath cycle up to time t , μ_t , is a sequence of r transformed dictionary elements

$$\mu_t = (g_{\phi_1}(\gamma_1), g_{\phi_2}(\gamma_2), \dots, g_{\phi_r}(\gamma_r^*))$$

where ϕ_i parameterizes the transformation of γ_i .

- γ_r^* denotes a partial observation of γ_r ; the time series ends in the middle of a motif.
- We measure the tumor location with correlated noise:

$$\mathbf{Y}_t \sim N(\mu_t, K_\theta(\mathbf{t}))$$

Transformation parameters

For each $i = 1, \dots, r$, ϕ_i is four-dimensional.

- ϕ_i consists of four parameters: length scaling, position scaling, location drift, and skewness.
- System constrains μ_t to be continuous.
- We accumulate information for ϕ_i very quickly as we observe the i^{th} motif—this gives us precise predictions.
- Over time our prior for ϕ becomes more informative, increasing precision of future predictions both within and across motifs.

Illustration of (conditional) data-generating process



Illustration of (conditional) data-generating process

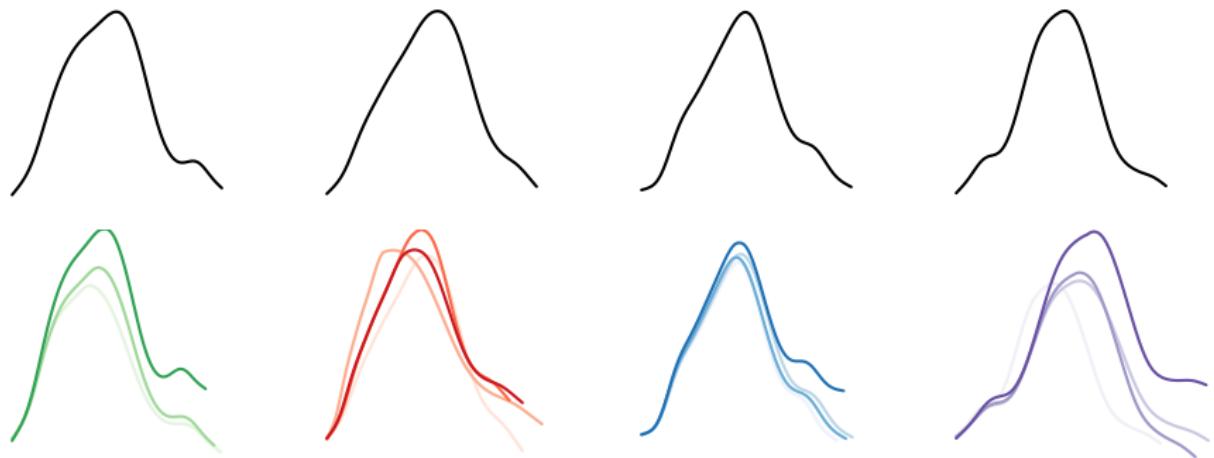


Illustration of (conditional) data-generating process

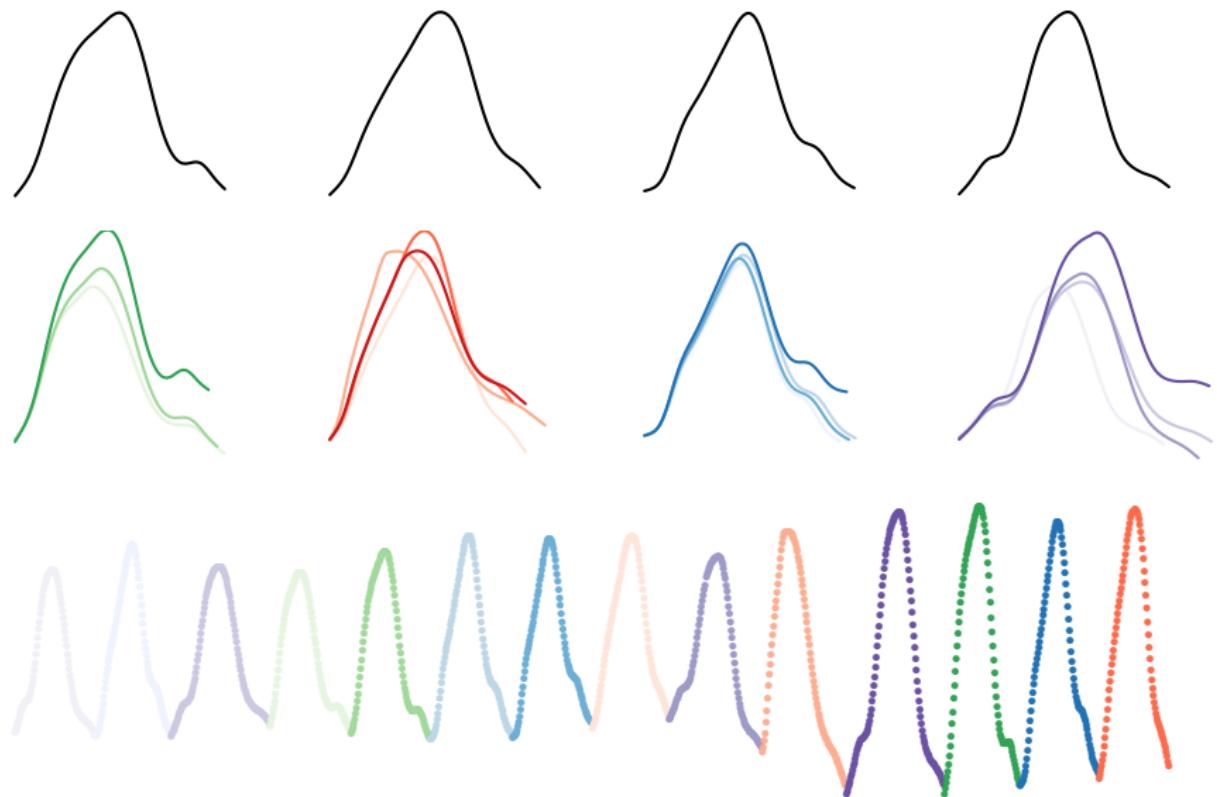
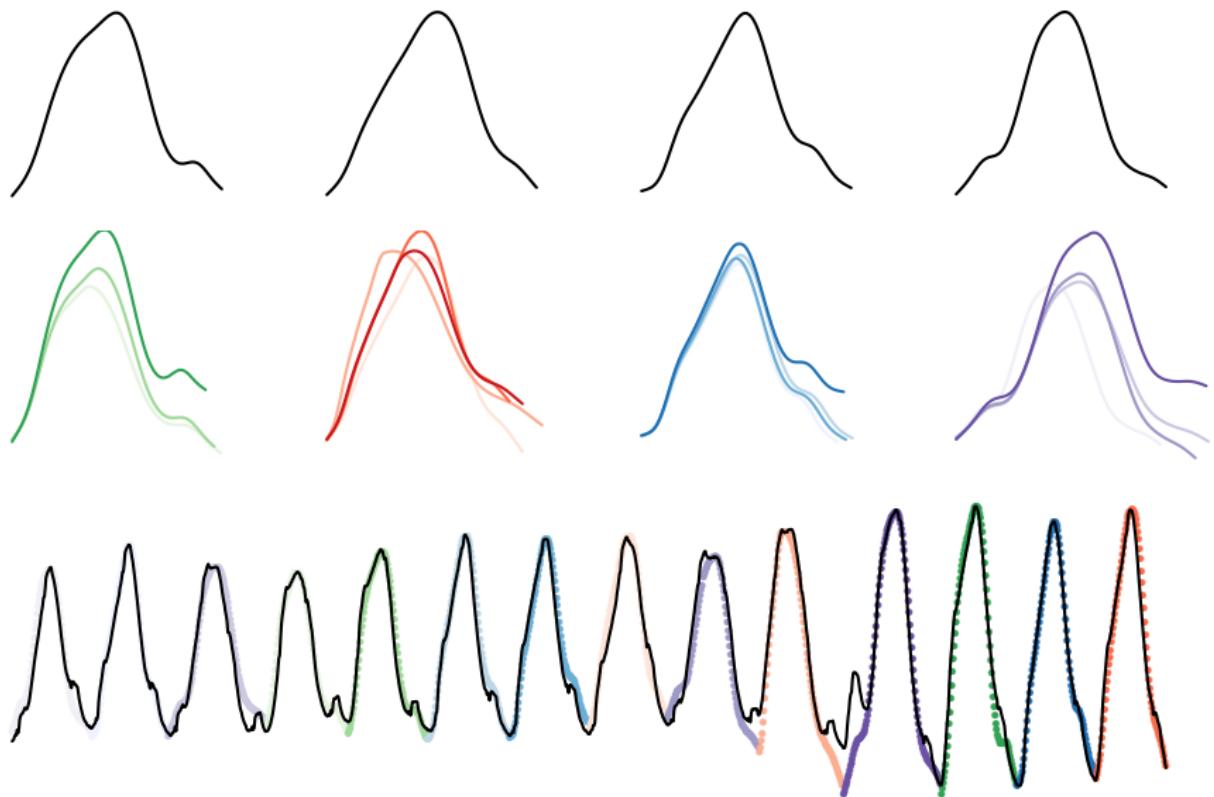


Illustration of (conditional) data-generating process



A few more modeling details

During the first 20-30 seconds:

- Learn $\Gamma = \{\gamma^{(1)}, \dots, \gamma^{(d)}\}$ nonparametrically. d can be determined by the data, but I fixed $d = 4$.
- Update priors for transformation parameters ϕ .
- Estimate and fix nuisance parameters θ .

A few more modeling details

During the first 20-30 seconds:

- Learn $\Gamma = \{\gamma^{(1)}, \dots, \gamma^{(d)}\}$ nonparametrically. d can be determined by the data, but I fixed $d = 4$.
- Update priors for transformation parameters ϕ .
- Estimate and fix nuisance parameters θ .

Then predictions for time τ in the future are made from the posterior predictive distribution:

$$\begin{aligned} P(y_{t+\tau} | \mathbf{y}_t) &\propto \int P(y_{t+\tau} | \mathbf{y}_t, \gamma_1, \gamma_2, \dots, \gamma_r, \phi_1, \dots, \phi_r) \\ &\quad \times P(\gamma_1, \gamma_2, \dots, \gamma_r, \phi_1, \dots, \phi_r | \mathbf{y}_t) d\gamma_1 \dots d\phi_r \end{aligned}$$

Notes:

- Γ and θ are assumed fixed and known!
- Do we know r ?

Implementation

Posterior distribution a mixture of discrete γ_i and continuous parameters (ϕ_i) .

- Very reasonable to fix parameters for motifs $r - 2$ and earlier due to approximate independence of y_t on the distant past.
- Thus posterior involves sampling from the space of $(\gamma_{r-1}, \gamma_r, \phi_{r-1}, \phi_r)$, which lives in $\Gamma^2 \times \mathbb{R}^8$.
- If d is small, we can work with marginal posterior $P(\phi_{r-1}, \phi_r | \mathbf{y}_t)$ using parallel tempering (marginalizing γ_{r-1}, γ_r yields multimodality).
- If d is large, we can use population MCMC methods.

Results

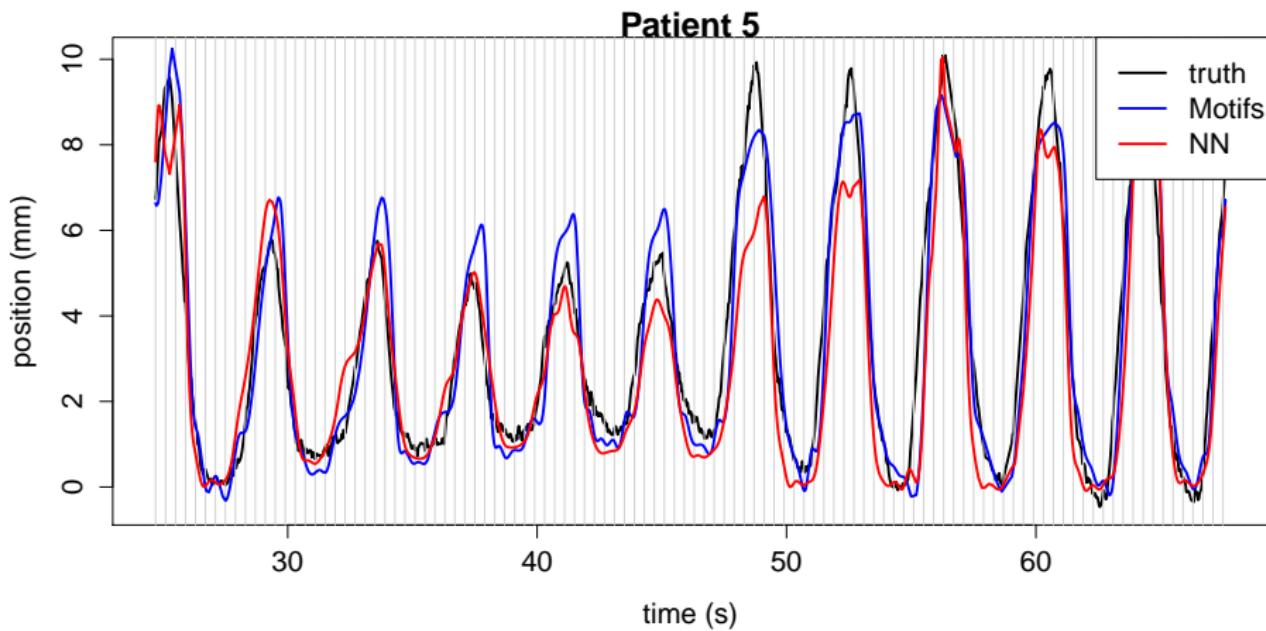


Figure: Comparison of Neural Network and Motif Sequencing predictions for patient 2 at a lag of 0.4s (gray lines). Root mean squared errors were 1.13 and 0.93 respectively.

Results

Comparison with NN

Table: Comparison of Motif Sequence method with Neural Networks across 11 patients

	RMSE	90% CI coverage
Motif Sequence	1.63	0.88
Neural Networks	1.81	0.77

- NN hyperparameters were tuned for each patient curve after observing 20s of training data.
- These estimates are not very precise: considerable within-curve and between-curve variance.

Conclusion

Benefits:

- Strong predictive performance—seemingly at least as strong as current methods.
- Better evaluation of prediction uncertainty.
- Scales linearly for fixed d .

Drawbacks:

- Computation is not real-time, thanks to difficulty of posterior sampling.
- Predictions somewhat sensitive to nuisance parameters θ .

Acknowledgements

Thanks to

- Natesh Pillai
- Ross Berbeco (HMS)
- John Henry Lewis (HMS)

References I

-  Floris Ernst and Achim Schweikard.
Forecasting respiratory motion with accurate online support vector regression (SVRpred).
International Journal of Computer Assisted Radiology and Surgery, 4(5):439–447, September 2009.
-  Alan Kalet, George Sandison, Huanmei Wu, and Ruth Schmitz.
A state-based probabilistic model for tumor respiratory motion prediction.
Physics in Medicine and Biology, 55(24):7615, December 2010.
-  A. Krauss, S. Nill, and U. Oelfke.
The comparative performance of four respiratory motion predictors for real-time tumour tracking.
Physics in medicine and biology, 56(16):5303, 2011.

References II

-  Martin J. Murphy and Sonja Dieterich.
Comparative performance of linear and nonlinear neural networks to predict irregular breathing.
Physics in Medicine and Biology, 51(22):5903, November 2006.
-  Nadeem Riaz, Piyush Shanker, Rodney Wiersma, Olafur Gudmundsson, Weihua Mao, Bernard Widrow, and Lei Xing.
Predicting respiratory tumor motion with multi-dimensional adaptive filters and support vector regression.
Physics in Medicine and Biology, 54(19):5735, October 2009.
-  Dan Ruan and Paul Keall.
Online prediction of respiratory motion: multidimensional processing with low-dimensional feature learning.
Physics in Medicine and Biology, 55(11):3011, June 2010.

References III

-  Gregory C. Sharp, Steve B. Jiang, Shinichi Shimizu, and Hiroki Shirato.
Prediction of respiratory tumour motion for real-time image-guided radiotherapy.
Physics in medicine and biology, 49(3):425, 2004.