A KERNEL-BASED FRAMEWORK FOR INTRA-FRACTIONAL RESPIRATORY MOTION ESTIMATION IN RADIATION THERAPY

T. Geimer[1,2,3], *M. Unberath*[1,2], *A. Birlutiu*[4], *O. Taubmann*[1,2], *J. Wölfelschneider*[3], *C. Bert*[2,3], *A. Maier*[1,2]

[1]Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg
[2]Graduate School in Advanced Optical Technologies (SAOT), Erlangen
[3]Department of Radiation Oncology, Universitätsklinikum Erlangen,
Friedrich-Alexander-Universität Erlangen-Nürnberg
[4]Computer Science Department, "1 December 1918" University of Alba Iulia, Romania

ABSTRACT

In radiation therapy, tumor tracking allows to adjust the beam such that it follows the respiration-induced tumor motion. However, most clinical approaches rely on implanted fiducial markers to locate the tumor and, thus, only provide sparse information. Motion models have been investigated to estimate dense internal displacement fields from an external surrogate signal, such as range imaging. With increasing surrogate complexity, we propose a respiratory motion estimation framework based on kernel ridge regression to cope with high-dimensional domains. This approach was validated on five patient datasets, consisting of a planning 4DCT and a follow-up 4DCT for each patient. Mean residual error was at best 2.73 ± 0.25 mm, but varied greatly between patients.

Index Terms— Motion Models, Range Imaging, Radiation Therapy, Kernel Trick, Active Shape Models

1. INTRODUCTION

In radiation therapy, malignant tumor cells are irradiated following an optimized fluence pattern that is the result of treatment planning based on CT imaging. A clinical target volume (CTV) and safety margins are defined to deliver the necessary therapeutic dose to the tumor while at the same time sparing healthy tissue. In this context, respiratory motion poses a problem as it is responsible for a continuous movement of the target volume that may result in clinically relevant over or under dosages in the CTV [1]. One advanced method to deal with free breathing during irradiation is tumor tracking.

Modern linear accelerators come equipped with imaging modalities that support the localisation of the tumor during treatment. Kilovoltage (kV) X-ray imaging [2] has been successfully applied, often coupled with one or more implanted radio-opaque fiducial markers to outline the tumor. In order to reduce the amount of additional non-therapeutic dose, correlation models [3] deduce internal information from an external surrogate signal. Clinical systems use low-frequent stereoscopic kV imaging as the ground truth to train and update the correlation model. Then, intra-procedurally only the external surrogate signal needs to be acquired. Unfortunately, these marker-based methods fall short with regards to non-rigid deformations. For real-time estimation of dense internal deformation fields, patient-specific motion models have been investigated [3, 4, 5]. The ground-truth deformation field is obtained from 4D imaging by registration to a reference phase. Recently, range imaging has emerged as a non-intrusive, markerless, high-dimensional surrogate suitable to drive said motion models [6]. However, with increasing surrogate complexity, the challenge is to extract meaningful information from high-dimensional domains. Wilms et al. [7] investigated multi-variate regression approaches based on range imaging among other surrogates, but only sparsely sampled the surface. Others added a dimensionality reduction step using active shape models [8] to describe physiological variation of a patient's breathing cycle [4, 9]. Alternatively, Li and Xing [10] proposed kernel-based respiratory motion estimation but only relied on a mono-dimensional surrogate.

In this work, we present a respiratory motion estimation pipeline using kernel ridge regression to model the internal-external correlation. Two approaches to reconstructing a newly observed phase were investigated: A weighted sum of samples (WSS) in the training and a linear combination of principal components by employing an active shape model (ASM). Evaluation was conducted in a study on five lung cancer patients. Each patient had at least two 4DCT acquisitions within eight weeks, allowing for strict separation of training on the planning CT and testing on a follow-up CTs. Thus, evaluation is closer to the actual application case than in previous studies where only one 4DCT was available.

2. MATERIAL & METHODS

The pipeline of our respiratory motion estimation framework is illustrated in Fig. 1. The planning CT (PCT) provides 4D images as the source for training a patient-specific motion

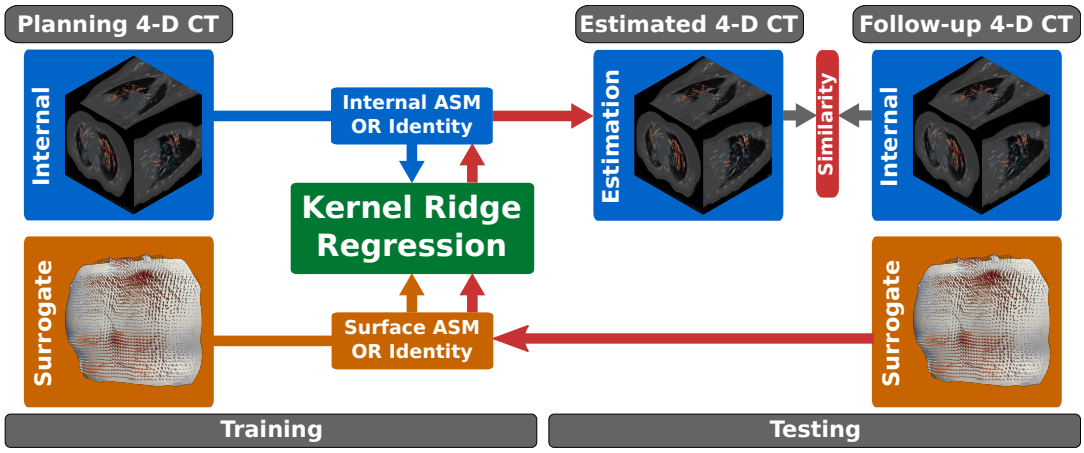

Fig. 1. Workflow of our motion estimation pipeline and evaluation. Both PCT and FCT provide internal deformation fields and the surface surrogate. Information from the PCT is used to train the motion model. Subsequently, surrogate information of the FCT is used to estimate a new deformation field which is then compared to the FCT internal ground truth.

model. Using surface deformation fields as the driving surrogate, a kernel-based correlation model is trained between the internal and external data. Using surrogate data extracted from a phase of the follow-up CT (FCT), a new deformation field is estimated for that particular phase and subsequently compared to the FCT serving as the ground truth. The following paragraphs will cover the individual parts more closely.

2.1. Patient Data

The datasets used for this simulation study originate from five patients with bronchial carcinoma or metastasis treated at the University Hospital Erlangen. Each patient received a time-resolved PCT consisting of eight equidistant phases. They were used as the source to train our patient-specific 4D motion model. Additionally, a FCT was acquired for each patient in order to monitor tumor shrinkage over the course of the treatment and adjust the CTV if needed. All CT volumes were resampled to a spacing of $1 \times 1 \times 2\,\text{mm}$.

2.2. Internal & Surrogate Data

Internal deformation fields were acquired via non-rigid registration w.r.t. the end-exhale phase, resulting in n vector fields $\{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_n\}, \boldsymbol{t}_i \in \mathbb{R}^{d_t}$ stored as columns of the data matrix $\boldsymbol{T} \in \mathbb{R}^{d_t \times n}$. Corresponding to that, $\boldsymbol{S} \in \mathbb{R}^{d_s \times n}$ carries the n surrogate observations $\boldsymbol{s}_i \in \mathbb{R}^{d_s}$, i.e. the dense surface deformation fields. For our simulation, the surface of the reference phase was extracted from the CT volume, followed by an interpolation of $\boldsymbol{t}_i$ at the mesh vertices to form the surrogate.

2.3. Kernel-based Regression

Correlation between target and surrogate domain can be considered a multi-variate, multi-linear regression problem [4] with the corresponding objective function

$$\underset{\boldsymbol{W}}{\operatorname{argmin}} \left(\frac{1}{2} ||\boldsymbol{W}^* \boldsymbol{S} - \boldsymbol{T}||_{\mathcal{F}}^2 + \alpha \frac{1}{2} ||\boldsymbol{W}^*||_{\mathcal{F}}^2 \right), \tag{1}$$

where $||\cdot||_{\mathcal{F}}$ denotes the Frobenius norm. A closed-form solution exists in

$$\boldsymbol{W} = \underbrace{\boldsymbol{T}\boldsymbol{S}^\top}_{d_t \times d_s} (\underbrace{\boldsymbol{S}\boldsymbol{S}^\top}_{d_s \times d_s} + \alpha I_{d_s})^{-1} \in \mathbb{R}^{d_t \times d_s} \tag{2}$$

representing the Moore-Penrose pseudo-inverse with Tikhonov regularization. This is also known as ridge regression. For high-dimensional data $\boldsymbol{T}$ and $\boldsymbol{S}$, computing $\boldsymbol{W}$ explicitly is computationally very expensive. Following Albert [11], Eqn. 2 is equivalent to

$$\boldsymbol{W} = \underbrace{\boldsymbol{T}}_{d_t \times n} (\underbrace{\boldsymbol{S}^\top \boldsymbol{S}}_{n \times n} + \alpha I_n)^{-1} \underbrace{\boldsymbol{S}^\top}_{n \times d_s}. \tag{3}$$

Further, we can now apply the kernel trick. The observations $\boldsymbol{s}_i$ are implicitly mapped to high-dimensional Hilbert space [12]: $\boldsymbol{\Phi} = \left[\boldsymbol{\phi}(\boldsymbol{s}_1), \ldots, \boldsymbol{\phi}(\boldsymbol{s}_n)\right]$. For the prediction of a target $\boldsymbol{t}_{\text{pred}}$ from a new observation $\boldsymbol{s}_{\text{new}}$, direct access to $\boldsymbol{\Phi}$ is never explicitly needed:

$$\begin{aligned} \boldsymbol{t}_{\text{pred}} &= \boldsymbol{T} \left(\boldsymbol{\Phi}^\top \boldsymbol{\Phi} + \alpha\, I_n \right)^{-1} \boldsymbol{\Phi}^\top \boldsymbol{\phi}(\boldsymbol{s}_{\text{new}}) \\ &= \boldsymbol{T} \left(\boldsymbol{K} + \alpha\, I_n \right)^{-1} \boldsymbol{\kappa}\left(\boldsymbol{s}_{\text{new}}\right). \end{aligned} \tag{4}$$

Instead, the estimation is performed entirely in terms of inner products in the higher-dimensional space, where $\boldsymbol{K}_{ij} = \boldsymbol{\phi}\left(\boldsymbol{s}_i\right)^\top \boldsymbol{\phi}\left(\boldsymbol{s}_j\right)$ and $\boldsymbol{\kappa}\left(\boldsymbol{s}_{\text{new}}\right)_i = \boldsymbol{\phi}\left(\boldsymbol{s}_i\right)^\top \boldsymbol{\phi}\left(\boldsymbol{s}_{\text{new}}\right)$. Besides enabling the framework to operate on high-dimensional data matrices, the approach also supports non-linear mappings. Common kernels used are the linear kernel $\kappa(\boldsymbol{a}, \boldsymbol{b}) = \boldsymbol{a}^\top \boldsymbol{b}$,

Gaussian kernel $\kappa_\sigma(\boldsymbol{a},\boldsymbol{b}) = e^{-\frac{||\boldsymbol{a}-\boldsymbol{b}||_2^2}{\sigma^2}}$, and polynomial kernel $\kappa_{d,o}(\boldsymbol{a},\boldsymbol{b}) = (\boldsymbol{a}^\top\boldsymbol{b}+o)^d$, where σ, o, d are parameters that require optimization via grid search.

2.4. Active Shape Model

With Eqn. 4, the estimated deformation field is the result of a weighted sum of training samples $\boldsymbol{T}$ (WSS). To omit redundant information, an additional generalisation step in the form of an ASM can be introduced [8] . The training sample $\boldsymbol{T}$ is decomposed into mutually orthogonal modes of variation. More precisely, an eigendecomposition of the covariance matrix $\boldsymbol{C}$ of the zero-mean data is performed. The first few components are often sufficient to represent more than 90% of the variance present in the dataset [13]. A data point $\boldsymbol{t}$ can be described in terms of the model as a linear combination:

$$\boldsymbol{t} = \bar{\boldsymbol{t}} + \boldsymbol{\mathcal{M}}_t\,\boldsymbol{f}_t + \boldsymbol{\epsilon}, \tag{5}$$

where $\boldsymbol{\mathcal{M}} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{p_t}]^\top \in \mathbb{R}^{d_t \times p_t}$ is the new basis formed by the first $p_t \leq n$ eigenvectors $\boldsymbol{v}_i$ of $\boldsymbol{C}$, $\bar{\boldsymbol{t}}$ is the consensus, $\boldsymbol{f} \in \mathbb{R}^{p_t}$ is the feature vector carrying the weights, and $\boldsymbol{\epsilon}$ is variance not explained by the model. Thus, instead of the data matrices $\boldsymbol{T}$ and $\boldsymbol{S}$, Eqn. 4 can also be computed using the set of features $\boldsymbol{F}_T \in \mathbb{R}^{p_t \times n}$ and $\boldsymbol{F}_S \in \mathbb{R}^{p_s \times n}$.

For the estimation process, the pipeline is extended by two steps. From the FCT, the observed surrogate $\boldsymbol{s}_{\text{new}}$ needs to be expressed in terms of the surface ASM:

$$\boldsymbol{f}_s = \boldsymbol{\mathcal{M}}_s^\top(\boldsymbol{s}_{\text{new}} - \bar{\boldsymbol{s}}). \tag{6}$$

Second, after regression the estimated internal features are used to reconstruct the internal deformation field up to an error $\boldsymbol{\epsilon}$ according to Eqn. 5.

2.5. Evaluation

Assessment of the estimation accuracy was performed for each patient individually. Linear, Gaussian and polynomial kernels were tested, both with and without additional generalisation by the ASM. The 4D motion model was trained on the PCT and tested on each phase of the FCT. Estimation accuracy was assessed using the L2-norm of the residual vectors between estimated field and FCT ground truth. A grid search was conducted to determine the regularization parameter α, the kernel parameters σ, o, d, and ASM dimensionality p_t, p_s with regards to the mean estimation error over all patients.

Both ASM and WSS are subject to a limited model accuracy, even if the internal weights were estimated perfectly. For the ASM, the lower bound is given by its generalisation ability. For WSS, we compared each FCT deformation field $\boldsymbol{t}_i$ with an optimal guess in a least-squares sense: $\boldsymbol{t}_{i,\text{optim}} = \boldsymbol{T}\boldsymbol{a}_i$, where $\boldsymbol{a}_i = \text{argmin}_{\tilde{\boldsymbol{a}}_i}\left(||\boldsymbol{T}\tilde{\boldsymbol{a}}_i - \boldsymbol{t}_i||_2^2\right)$. The result is a lower bound on the estimation error.

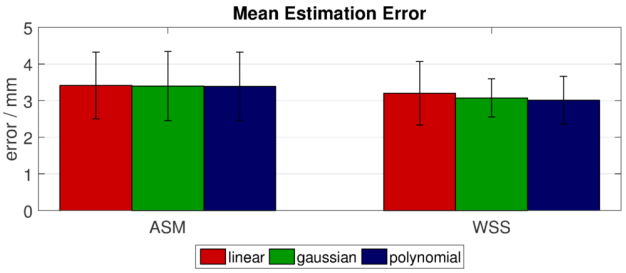

Fig. 2. Mean estimation error over all patients using ASM and WSS with a linear, Gaussian, and polynomial kernel. The mean breathing magnitude was 3.46 ± 0.50 mm.

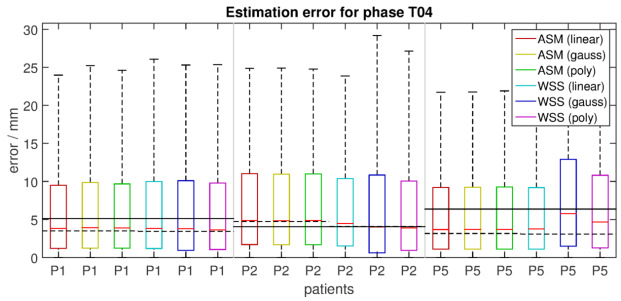

Fig. 3. Estimation error of end-inhale phase T04 for three FCTs. Black bars show the mean magnitude of the ground truth deformation field, while the dashed black lines indicate the lower bound achievable with the respective model.

3. RESULTS & DISCUSSION

Fig. 2 shows the overall estimation error for all patients and phases. All results are below the reference mean breathing magnitude of 3.40 ± 0.59 mm. The best result was achieved for the polynomial kernel operating directly on the data matrices with a mean residual error of 3.01 ± 0.65 mm.

Fig. 3 shows the end-inhale phase T04 for patients P1, P2, and P5, including the breathing magnitude and the model error. For patient P2, the follow-up CT was suffering from severe image artifacts (see Fig. 4). Therefore, the ASM, and to a minor degree the WSS, were unable to describe the corrupted motion fields. In total, two patients failed in this manner. Thus, we excluded them and repeated the experiment. According to Fig. 5, the mean residual error dropped below 3 mm, even though the reference magnitude increased. With the affected patients excluded, the best result was achieved by linear WSS at 2.73 ± 0.25 mm. However, a Student's t-test on the patient population showed statistical significance at a 5% significance level only for polynomial WSS compared to the three ASM approaches. In conclusion, estimation accuracy highly depends on the image quality of the planning and follow-up CTs. While all methods were within typical deformable image registration uncertainties [14], more data is necessary for statistically meaningful evaluation.

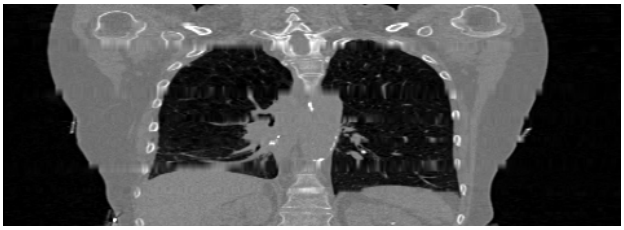

Fig. 4. FCT coronal slice of patient P2 at end-inhale. Artifacts are visible near the diaphragm and the top of the lung.

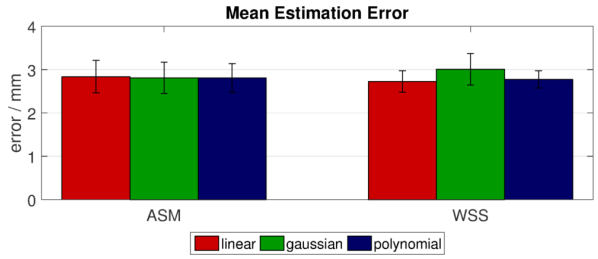

Fig. 5. Mean estimation error analogous to Fig. 2 with two corrupted patients excluded. The mean breathing magnitude was 3.57 ± 0.43 mm.

4. CONCLUSION

We proposed a respiratory motion estimation framework using kernel ridge regression. The method supports high-dimensional data domains as well as an additional generalisation step in the form of active shape models. Through the use of kernels, non-linear mappings are introduced into the regression problem. This approach was evaluated in a study on five lung cancer patients, each with a planning 4DCT for training and a follow-up 4DCT for testing. We obtained good results for three out of five testcases. Of the other two at least one CT was corrupted with severe image artifacts. For future work, we will acquire additional data and further investigate the failed cases as well as the inherent model error.

5. REFERENCES

[1] P J Keall, G S Mageras, J M Balter, R S Emery, K M Forster, S B Jiang, J M Kapatoes, D A Low, M J Murphy, B R Murray, C R Ramsey, M B Van Herk, S S Vedam, J W Wong, and E Yorke, "The management of respiratory motion in radiation oncology report of AAPM Task Group 76," *Med Phys*, vol. 33, no. 10, pp. 3874–3900, 2006.

[2] C-Y Huang, J N Tehrani, J Aun Ng, J Booth, and P J Keall, "Six Degrees-of-Freedom Prostate and Lung Tumor Motion Measurements Using Kilovoltage Intrafraction Monitoring," *Int J Radiat Oncol Biol Phys*, vol. 91, no. 2, pp. 368 – 375, 2015.

[3] J R McClelland, D J Hawkes, T Schaeffter, and A P King, "Respiratory motion models: A review," *Med Image Anal*, vol. 17, no. 1, pp. 19 – 42, 2013.

[4] O Taubmann, J Wasza, C Forman, P Fischer, J Wetzl, A Maier, and J Hornegger, "Prediction of Respiration-Induced Internal 3-D Deformation Fields From Dense External 3-D Surface Motion," in *28th Trans Comput Assist Radiol Surg (CARS)*, 2014, pp. 33–34.

[5] A Fassi, M Seregni, M Riboldi, P Cerveri, D Sarrut, G B Ivaldi, P Tabarelli de Fatis, M Liotta, and G Baroni, "Surrogate-driven deformable motion model for organ motion tracking in particle radiation therapy," *Phys Med Biol*, vol. 60, no. 4, pp. 1565, 2015.

[6] J Wasza, P Fischer, H Leutheuser, T Oefner, C Bert, A Maier, and J Hornegger, "Real-Time Respiratory Motion Analysis Using 4-D Shape Priors," *IEEE Trans Biomed Eng*, vol. 63, no. 3, pp. 485–495, 2016.

[7] M Wilms, R Werner, J Ehrhardt, A Schmidt-Richberg, H-P Schlemmer, and H Handels, "Multivariate Regression Approaches For Surrogate-based Diffeomorphic Estimation of Respiratory Motion in Radiation Therapy," *Phys Med Biol*, vol. 59, no. 5, pp. 1147, 2014.

[8] Timothy F Cootes, Christopher J Taylor, David H Cooper, and Jim Graham, "Active shape models - their training and application," *Comput Vis Image Underst*, vol. 61, no. 1, pp. 38–59, 1995.

[9] T Geimer, M Unberath, O Taubmann, C Bert, and A Maier, "Combination of Markerless Surrogates for Motion Estimation in Radiation Therapy," in *30th Trans Comput Assist Radiol Surg (CARS)*, 2016, pp. 59–60.

[10] R Li and L Xing, "A Kernel Method for Real-Time Respiratory Tumor Motion Estimation Using External Surrogates," in *10th Proc Int Conf Mach Learn Appl (ICMLA)*, Dec 2011, vol. 2, pp. 206–209.

[11] A Albert, "Regression and the Moore-Penrose pseudoinverse," Academic Press, New York, 1972.

[12] A Berlinet and C Thomas-Agnan, *Reproducing Kernel Hilbert Spaces in Probability and Statistics*, Springer Science & Business Media, 2011.

[13] M Unberath, A Maier, D Fleischmann, J Hornegger, and R Fahrig, "Open-source 4D Statistical Shape Model of the Heart for X-ray Projection Imaging," in *Proc IEEE Int Symp Biomed Imaging (ISBI)*. IEEE, 2015.

[14] K K Brock, "Results of a Multi-Institution Deformable Registration Accuracy Study (MIDRAS)," *Int J Radiat Oncol Biol Phys*, vol. 76, no. 2, pp. 583 – 596, 2010.