Respiratory Deformation Estimation in X-ray Guided IMRT Using a Bilinear Model

Tobias Geimer[1,2,3], Stefan B. Ploner[1], Paul Keall[4], Christoph Bert[2,3], and Andreas Maier[1,2]

[1]Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg
[2]Erlangen Graduate School of Advanced Optical Technologies, FAU ER-N
[3]Department of Radiation Oncology, Universitätklinikum Erlangen, FAU ER-N
[4]ACRF Image X Institute, The University of Sydney, Australia
tobias.geimer@fau.de

Abstract. Driving a respiratory motion model in X-ray guided radiotherapy can be challenging in treatments with continuous rotation such as VMAT, as data-driven respiratory signal extraction suffers from angular effects overlapping with respiratory changes in the projection images. Compared to a linear model trained on static acquisition angles, the bilinear model gains flexibility in terms of handling multiple viewpoints at the cost of accuracy. In this paper, we evaluate both models in the context of serving as the surrogate input to a motion model. Evaluation is performed on the 20 patient 4D CTs in a leave-one-phase-out approach yielding a median accuracy drop of only 0.14 mm in the 3D error of estimated vector fields of the bilinear model compared to the linear one.

1 Introduction

In intensity modulated radiotherapy (IMRT) the radiation beams are shaped to closely envelope the tumor region. With the linear accelerator (LINAC) rotating around the patient, a therapeutic dose is accumulated within the malignent cells while healthy tissue is spared. Treatment is either performed from predefined angles (Step&Shoot) or in a continuous rotation (VMAT). Requiring a complex dose optimization planned on pre-treatment (4D) CT, intra-fractional respiratory motion can hinder accurate dose delivery. To compensate, a gimbaled beam following the tumor motion [1] coupled with respiratory motion models [2] have found success in image-guided radiotherapy (IGRT).

A motion model comprises a motion representation trained from 4D CT, e.g. via principal component analysis (PCA). Correspondence is established to a highly-correlated surrogate available in the treatment room, from which the internal motion is inferred. Among surrogate sources, many LINACs come equipped with an on-board imager to provide kV fluoroscopy. In this context, X-ray guided RT shares similarities with cone-beam CT reconstruction where respiratory signal extraction from X-ray projections is prominent. Most data-driven approaches [3] provide only a 1D signal insufficient for a low reconstruction error of the PCA-based motion representation [2]. While unsupervised learning on

X-ray fluoroscopy [4] yields multiple respiratory features, they are restricted to static acquisition and only applicable to Step&Shoot but not to continuously rotating VMAT. Recently, Geimer et al. [5] presented a bilinear decoupling of angular and respiratory variation in rotational X-ray scans. Based on digitally reconstructed radiographs (DRRs), a rotational and respiratory feature representation is learned. While it was shown that the decoupled respiratory features can drive a respiratory motion model, no quantitative evaluation was performed against comparable PCA-based models. With the bilinear model being an extension to the static case, this paper aims to provide insight into how much accuracy is potentially sacrificed to gain independence from the trajectory angle.

In the following, we provide a brief overview of the structure of motion models and illustrate how both the linear (Sec. 2.1) and bilinear model (Sec. 2.2) serve as surrogate input. Sec. 2.3 explains the performance comparison between static and bilinear surrogate as the main contribution of this paper. Results are presented in Sec. 3 and wrapped up with a concise discussion and conclusion in Sec. 4.

2 Material & Methods

2.1 Respiratory Motion Models

McClelland et al. [2] identify four components, (1) the representation of the motion to be described, (2) the choice of surrogate signal and processing thereof, (3) a correlation model linking surrogate to motion, and (4) a fitting method to determine model parameters from training data. In the following, we will give a possible choice of these components for X-ray guided RT and demonstrate how the bilinear fluoroscopic model can slot in as the surrogate component.

Motion Representation. Given F binned volumes $\boldsymbol{v}_j \in \mathbb{R}^{N^3}$, $j \in \{1, \dots, F\}$ in a 4D CT, B-spline based deformable image registration (DIR) over the entire lung w.r.t. the end-exhale phase (0_{In}) yields displacement vector fields (VF) $\boldsymbol{d}_j \in \mathbb{R}^{3N^3}$ (see Fig. 1), where N is the arbitrary dimension of the volume. In order to suppress noise and prevent overfitting, PCA is often applied resulting in a low-dimensional representation $\{\tilde{\boldsymbol{\Theta}}, \bar{\boldsymbol{d}}\}$, where $\tilde{\boldsymbol{\Theta}} = \left(\tilde{\boldsymbol{\theta}}_1, \dots, \tilde{\boldsymbol{\theta}}_f\right) \in \mathbb{R}^{3N^3 \times f}$, $f \ll F$, are the first f eigenvectors and $\bar{\boldsymbol{d}}$ is the mean VF. As a result, every $\boldsymbol{d}_j$ can be expressed as a linear combination by the respiratory PC scores $\tilde{\boldsymbol{a}}_j \in \mathbb{R}^f$ up to a residual error $\boldsymbol{\epsilon} \in \mathbb{R}^{N^3}$

$$\boldsymbol{d}_j = \tilde{\boldsymbol{\Theta}}\, \tilde{\boldsymbol{a}}_j + \bar{\boldsymbol{d}} + \boldsymbol{\epsilon}. \tag{1}$$

Static X-ray Surrogate. An X-ray projection $\boldsymbol{p}_{i,j} \in \mathbb{R}^{N^2}$ of volume $\boldsymbol{v}_j$ under the acquisition angle ϕ_i is given by the X-ray transform $\boldsymbol{R}_i \in \mathbb{R}^{N^2 \times N^3}$, such that $\boldsymbol{p}_{i,j} = \boldsymbol{R}_i\, \boldsymbol{v}_j$, where N again denotes arbitrary dimension of projection images and/or volume. An analogous decomposition to Eqn. 1 for the volume $\boldsymbol{v}_j$ yields

$$\boldsymbol{p}_{i,j} = \boldsymbol{R}_i \left(\boldsymbol{\Theta} \boldsymbol{a}_j + \bar{\boldsymbol{v}}\right) = \boldsymbol{\Theta}_i^R \boldsymbol{a}_j + \bar{\boldsymbol{p}}_i. \tag{2}$$

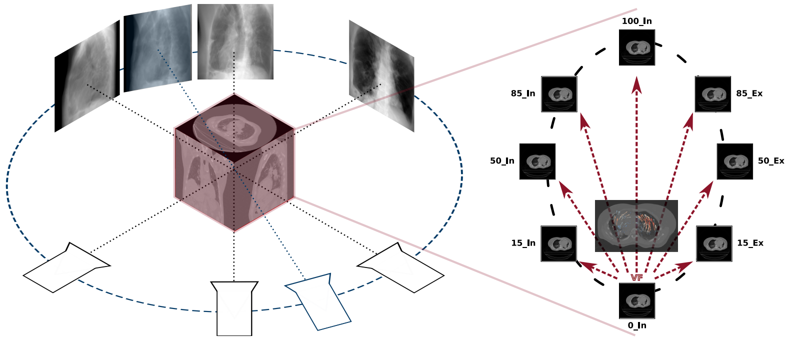


Fig. 1. Illustration of the bilinear decoupling idea. Respiratory changes are inherently 3-dimensional, but can only be observed in 2D when projected according to the X-ray transform, where they overlap with angular changes due to rotation.

Here, $\boldsymbol{\Theta}_i^R \in \mathbb{R}^{N^2 \times f}$ describes respiration-induced variation observable in the 2D projections under the specific angle ϕ_i. Consequently, such a PCA decomposition can be trained on the 4D CT forward projected under said acquisition angle.

Correlation Model. A popular model in literature for the correlation between internal and surrogate scores $\tilde{\boldsymbol{A}}, \boldsymbol{A} \in \mathbb{R}^{F \times f}$ is multi-linear regression (MLR) [6]

$$\boldsymbol{W} = \underset{\hat{\boldsymbol{W}}}{\operatorname{argmin}} \left(\frac{1}{2} ||\hat{\boldsymbol{W}} \boldsymbol{A} - \tilde{\boldsymbol{A}}||_2^2 + \alpha \frac{1}{2} ||\hat{\boldsymbol{W}}||_2^2 \right), \tag{3}$$

with the Moore-Penrose pseudoinverse as the closed-form solution.

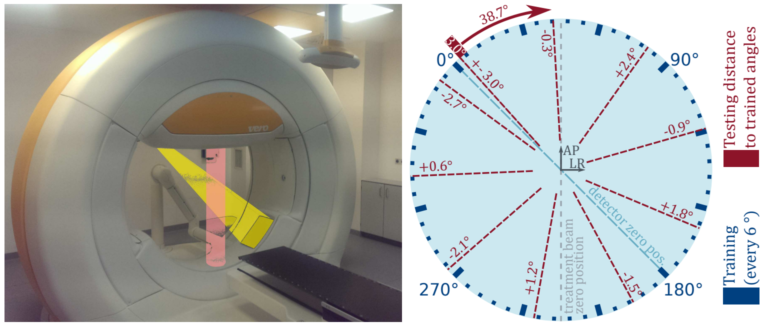
2.2 Bilinear Model for Rotational X-ray

A static angle model as described in Sec. 2.1 is unable to explain variation caused by rotation. As an extension to the linear case, a bilinear model can decouple angular and respiratory variation into distinct feature spaces [5], such that a projection $\boldsymbol{p}_{i,j}$ at angle ϕ_i and respiratory phase t_j can be written as

$$\boldsymbol{p}_{i,j} = \mathcal{M} \times_1 \boldsymbol{a}_j \times_2 \boldsymbol{b}_i, \tag{4}$$

where $\boldsymbol{a}_j \in \mathbb{R}^f$, $\boldsymbol{b}_i \in \mathbb{R}^g$ are respiratory and rotational feature weights, and $\mathcal{M} \in \mathbb{R}^{N^2 \times f \times g}$ is a model tensor trained from DRRs of a prior 4D CT. $\times_k$ denotes the kth mode product [7]. Bilinear training and application will be outlined in the following. For a detailed derivation we refer the reader to [5].

Model Training. Simulating a circular scan that mimics the VMAT arc the F volumes $\boldsymbol{v}_j$ are projected at G angles ϕ_i, $i \in \{1, \ldots, G\}$. Higher-order SVD [7] is applied to the data tensor $\mathcal{P} \in \mathbb{R}^{N^2 \times F \times G}$ such that $\mathcal{P} = \mathcal{M} \times_1 \boldsymbol{A} \times_2 \boldsymbol{B}$, where $\boldsymbol{A} \in \mathbb{R}^{F \times f}$ and $\boldsymbol{B} \in \mathbb{R}^{G \times g}$ contain the model weights of the training set.



(a) Vero LINAC (b) Experimental Setup

Fig. 2. (a) Vero with kV imager mounted on the same ring gantry, offset by 45° to the MV treatment beam. (b) Bilinear training at every 6° (blue) with test angles (red) every 38.7° starting at 3°, resulting in varying distances to neighboring training angles.

Weight Estimation. Both rotational and respiratory bilinear weights need to be determined for a new projection image $\boldsymbol{p}(t, \phi)$ at unknown respiratory state t. While the acquisition angle is known, the corresponding rotational weights $\boldsymbol{b}(\phi)$ are not. However, given similarity between neighboring views, we adopt the B-spline interpolation of [5] interpolate $\boldsymbol{b}(\phi)$ from the training weights $\boldsymbol{B}$. Mode-multiplying $\boldsymbol{b}\,(\phi)$ into $\mathcal{M}$ removes the angular variation and yields

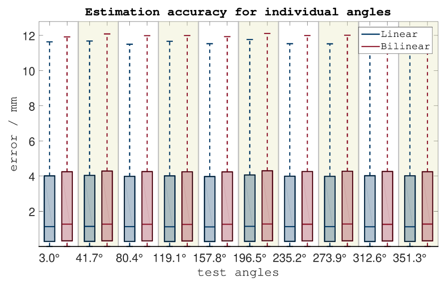
$$\mathcal{M}_\phi^R = \mathcal{M} \times_2 \boldsymbol{b}\,(\phi) \in \mathbb{R}^{N^2 \times f \times 1} \quad \rightarrow \quad \boldsymbol{M}_\phi^R \in \mathbb{R}^{N^2 \times f} \tag{5}$$

an angle-dependent model matrix such that $p(\phi, t) = \boldsymbol{M}_\phi^R\, \boldsymbol{a}_t$ as in Eqn. 2.

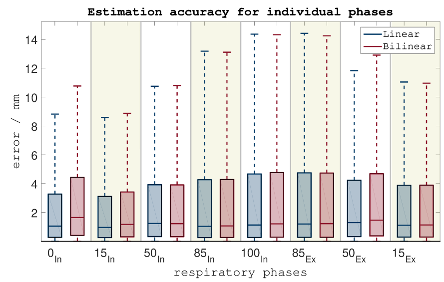
2.3 Performance Comparison

Data. We evaluate the performance of the linear and bilinear surrogate model on the 4D CTs of 20 patients being treated for lung carcinoma or metastasis at the University Hospital Erlangen. Each 4D CT consists of $F = 8$ volumes reconstructed at respiratory states 0 %, 15 %, 50 %, 85 %, 100 % inhale, and 85 %, 50 %, 15 % exhale. DIR of these volumes provided the VF for training the patient-specific PCA-based internal motion representation as described in Sec. 2.1. For training the bilinear surrogate model, each phase was forward projected along a circular trajectory according to the Vero SBRT (Brainlab AG, Munich, Germany) geometry at $G = 60$ angles in steps of 6° from 0° to 354°. For testing, additional DRRs were created at ten angles ϕ_k every 38.7° starting at 3°. The interval was chosen to ensure varying distances to neighboring training angles. Fig. 2 illustrates the choice of training and testing angles, including °-distances.

Experiments. A leave-one-phase-out evaluation was performed for each patient. For the motion representation (feature dimensionality $f = 4$) each respiratory phase was subsequently removed prior to PCA. The linear surrogate model



(a) Individual angles.



(b) Individual phases.

Fig. 3. Estimation error of the linear and bilinear model averaged for angles and phases.

with $f = 4$ was trained from the DRRs of each test angle directly minus the left-out phase. In contrast, the bilinear model was trained on the $G = 60$ training angles without the left-out phase and $g = 40$ as rotational dimensionality ensuring flexibility towards rotation. Here, $f = 5$ as the 1st bilinear component is near constant and represents the shift towards the mean due to missing mean normalization [5]. Finally, the regression matrices were computed between internal scores and the two different surrogate features of the training set.

The projection $\boldsymbol{p}_{k,j}$ for left-out phase t_j and test angle ϕ_k is then fed to both surrogate models, with the bilinear model also requiring the angle ϕ_k. The extracted weights are regressed to an internal representation a_j and the VF is reconstructed according to Eqn. 1. Estimation accuracy is reported based on the voxelwise euclidean error to the ground truth VF from DIR.

3 Results

Average accuracy over all 20 patients, 10 angles, and 8 phases was 1.13 ± 0.58 mm (median: 0.78 mm) for the static model and 1.27 ± 0.67 mm (median: 0.89 mm) for the bilinear surrogate. Given 1600 observation per model, a Levene test was performed on the mean errors indicating significance at a p-value of $1.2e^{-6}$. Fig. 3 shows the average error over all patients displayed for individual angles and phases. Neither surrogate model is sensitive to the viewing angle. Estimation error outliers increased with the inhale state, which seems reasonable given that the 100_{In} state corresponds to largest VF magnitude. As seen in Fig. 4 displaying the mean error for individual patients average over all phases and angles, the estimation error mostly relies on the actual patient and the quality of the prior 4D CT. Overall, the bilinear surrogate model was only 0.14 mm worse on average than the linear model specifically trained for that particular test angle.

4 Discussion

The bilinear surrogate model performed only slightly worse than the one trained on static acquisition angles. This is unsurprising, as the static model has seen

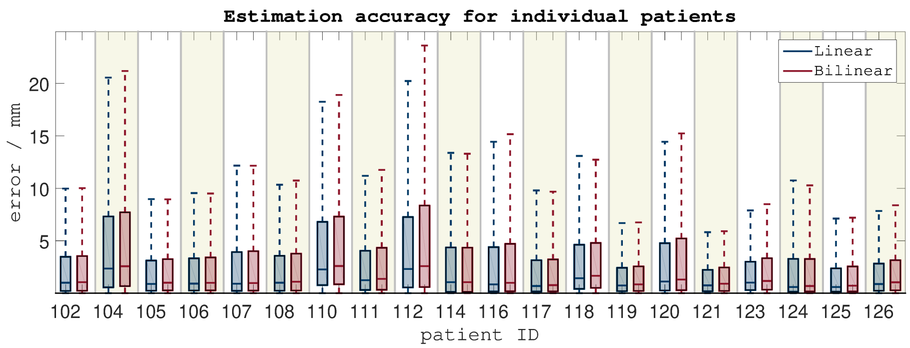


Fig. 4. Average estimation error of the linear and bilinear model for individual patients.

the test angle except for the test phase while the bilinear one is flexible over the entire trajectory. As such, the major gain in flexibility comes at the cost of only a small drop in accuracy. Similar to [5], the leave-one-out evaluation suffers from two shortcomings in assuming a perfect baseline registration between diagnostic 4D CT and the in-room patient, and no inter-fractional changes. However, the linear model also benefits from this simplification and, thus, the comparison is still valid. Relying on two distinct 4D CT per patient for training and testing can help model these conditions.

In conclusion, we showed in a retrospective patient study that a bilinear model operating on a circular X-ray sequence can be used to drive a respiratory motion model during continuously rotating VMAT treatments.

References

1. Akimoto M, Nakamura M, Miyabe Y, Mukumoto N, Yokota K, Mizowaki T, et al. Long-term stability assessment of a 4D tumor tracking system integrated into a gimbaled linear accelerator. J Appl Clin Med Phys. 2015;16(5):373–380.
2. McClelland JR, Hawkes DJ, Schaeffter T, King AP. Respiratory motion models: A review. Med Image Anal. 2013;17(1):19–42.
3. Yan H, Wang X, Yin W, Pan T, Ahmad M, Mou X, et al. Extracting respiratory signals from thoracic cone beam CT projections. Phys Med Biol. 2013;58(5):1447–64.
4. Fischer P, Pohl T, Faranesh A, Maier A, Hornegger J. Unsupervised Learning for Robust Respiratory Signal Estimation from X-Ray Fluoroscopy. IEEE Trans Med Imaging. 2017;36(4):865–877.
5. Geimer T, Keall P, Breininger L, Caillet V, Dunbar M, Bert C, et al. Decoupling Respiratory and Angular Variation in Rotational X-ray Scans Using a Prior Bilinear Model. In: Proc. GCPR; 2018. p. 1–12.
6. Wilms M, Werner R, Ehrhardt J, Schmidt-Richberg A, Schlemmer HP, Handels H. Multivariate regression approaches for surrogate-based diffeomorphic estimation of respiratory motion in radiation therapy. Phys Med Biol. 2014;59(5):1147–1164.
7. De Lathauwer L, De Moor B, Vandewalle J. A Multilinear Singular Value Decomposition. SIAM J Matrix Anal Appl. 2000 Jan;21(4):1253–1278.