Spatio-Temporal Motion Estimation for Respiratory-Correlated Imaging of the Lungs

Jef Vandemeulebroucke[1,2,3] Simon Rit[1,2] Jan Kybic[3,4]
Patrick Clarysse[1] David Sarrut[1,2]

[1]CREATIS-Laboratory; CNRS UMR5220; INSA-Lyon; University of Lyon, France
[2]Léon Bérard Cancer Center, University of Lyon, F-69373, Lyon, France
[3]Center for Machine Perception, Czech Technical University in Prague, Czech Republic
[4]Computer Vision Lab, Ecole Polytechnique Fédérale de Lausanne, Switzerland

September 21, 2010

Abstract

Purpose: Four-dimensional computed tomography (4D CT) can provide patient-specific motion information for radiotherapy planning and delivery. Motion estimation in 4D CT is challenging due to the reduced image quality and the presence of artifacts. We aim to improve the robustness of deformable registration applied to respiratory-correlated imaging of the lungs, by using a global problem formulation and pursuing a restrictive parametrization for the spatio-temporal deformation model.
Methods: A spatial transformation based on free-form deformations was extended to the temporal domain, by explicitly modelling the trajectory using a cyclic temporal model based on B-splines. A global registration criterion allowed to consider the entire image sequence simultaneously, and enforce the temporal coherence of the deformation throughout the respiratory cycle. To ensure a parametrization capable of capturing the dynamics of respiratory motion, a prestudy was performed on the temporal dimension separately. The temporal parameters were tuned by fitting them to diaphragm motion data acquired for a large patient group. Suitable properties were retained and applied to spatio-temporal registration of 4D CT data. Registration results were validated using large sets of landmarks and compared to consecutive spatial registrations. To illustrate the benefit of the spatio-temporal approach, we also assessed the performance in the presence of motion-induced artifacts.
Results: Cubic B-splines gave better or similar fitting results as lower orders, and were selected because of their inherently stronger regularization. The fitting and registration errors increased gradually with the temporal control point spacing, representing a trade-off between achievable accuracy and sensitivity to noise and artifacts. A piecewise smooth trajectory model, allowing for a discontinuous change of speed at end-inhale, was found most suitable to account for the sudden changes of motion at this breathing phase. The spatio-temporal modelling allowed a reduction of the number of parameters of 45%, while maintaining registration accuracy within 0.1 mm. The approach reduced the sensitivity to artifacts.
Conclusions: Spatio-temporal registration can provide accurate motion

estimation for 4D CT and improves the robustness to artifacts.
Keywords: deformable registration, respiratory motion, 4D CT

1 Introduction

The advent of four-dimensional (4D) computed tomography (CT) has allowed patient-specific respiratory motion information to be incorporated into radiation therapy planning and delivery. 4D CT provides multiple three-dimensional (3D) CT volumes, representing the patient at different stages of the breathing cycle[42,6,15,25]. The additional patient data, implies an order of magnitude increase in the workload required to obtain a 4D treatment plan. Deformable registration is the tool that can facilitate partial automation of the 4D planning process[11]. It can provide the motion fields which are required for automating tasks such as re-contouring of anatomic structures[16], patient-specific margin definition[50] or 4D treatment plan evaluation[8]. Deformable image registration is also an enabling tool for alternative applications of respiratory-correlated imaging such as ventilation imaging[9], motion compensation[45,28] or motion modelling[19,49]. Although extensive validation is required before extending the clinical use of deformable image registration, it is expected to become a standard methodology in radiotherapy[31,12].

Deformable image registration can be described as the task of finding a suitable geometric transformation between corresponding image data, such that a transformed image becomes similar to another one[21]. While the concept of image registration is easily described, the underlying numerical problem is difficult to solve. Mainly because the registration problem is *ill-posed*. Small changes of the input images may lead to very different registration results. Moreover, the solution might not be unique. Salient image information might be sparse or ambiguous, and the acquisition process might have introduced noise and artifacts. To facilitate the process, prior knowledge about the deformation should be incorporated in the registration framework in order to favour solutions with plausible physical characteristics. Explicit parametric restrictions can constrain the optimization to transformations that represent suitable properties. This approach can offer a reduction of the search space by making the description more problem-specific, and consequently improve the robustness of the optimization process.

An example are spatio-temporal registration schemes, which consist in a global formulation of the motion estimation problem for temporal image sequences. Rather than estimating frame-to-frame displacements individually, the entire sequence is considered simultaneously, allowing to enforce the temporal coherence of the deformation across the sequence. By making assumptions such as smoothness about the temporal variations of the transformation, these approaches often enable a more compact and restrictive description of the full motion estimation problem. Spatio-temporal deformable registration has received considerable attention in literature, mostly in cardiac image analysis[10,4,20,3,26,14,5,39,27,33], but more recently also for respiratory-correlated imaging of the thorax[43,35,2]. Usually, a 3D-4D formulation is utilized to find a smooth time-dependent deformation field that aligns all images from a given input sequence with a reference image, which can be a frame of that same sequence[4,14,5,39,43,2]. Sometimes, spatial as well as temporal alignment of multiple image sequences is desirable,

leading to a 4D-4D registration framework[3,26,35,27].

In comparison to conventional diagnostic CT, 4D CT images tend to be acquired at lower spatial resolution and are characterized by higher noise levels because of the low radiation dose per image. In addition, an alarmingly high number of acquisitions contains motion-induced artifacts[48], mainly due to irregular patient breathing during image acquisition. In the case of artifacts, the image information can be considered locally invalid, as it does not correspond to the patient anatomy. Clinical use of the estimated motion fields requires them to be as close to the unknown reality as possible. A problem-specific, spatio-temporal deformation model could contribute in reducing sensitivity to local image irregularities and render the motion estimate more plausible and potentially more representative of the patient's breathing motion under these challenging circumstances.

In this study, we develop a spatio-temporal registration scheme for lung motion quantification in respiratory-correlated sequences. Our primary objective is to obtain a low-dimensional representation of the 4D deformation model, capable of accurately representing the respiratory motion, while being more robust to artifacts and increased noise levels. The approach consists of a 3D-4D problem formulation in which temporal regularization is pursued by explicitly modelling the trajectory of moving structures. With respect to previous work on spatio-temporal registration, we specifically focus on respiratory-correlated image sequences, and develop and evaluate a cyclic trajectory model for representing the motion over an entire breathing cycle. In addition, the chosen parametrization reflects our aim to improve registration robustness by rendering the deformation model more problem-specific.

2 Method

The spatio-temporal transformation will be developed incrementally. We first describe a conventional spatial registration, of which the proposed method can be seen as an extension. Next, the temporal dimension is considered separately and the method for modelling the trajectory is detailed. The sought spatio-temporal deformation function is obtained by combining both.

2.1 Problem Description

Consider a 4D sequence, represented by an intensity function $f(\boldsymbol{i}, k) \in \mathbb{R}$ with $\boldsymbol{i} \in \mathbf{I} \subset \mathbb{Z}^3$ and $k \in \mathbf{K} \subset \mathbb{Z}$; $\mathbf{I}$ and $\mathbf{K}$ being the set of spatial and temporal sample indices, respectively. We wish to analyze the motion with respect to the 3D reference frame at time index $k_\mathrm{r} \in \mathbf{K}$. The task of motion estimation throughout the 4D sequence is formulated as the search for the unknown spatio-temporal transformation $\mathcal{T}_\mathrm{st}$, defined for $\mathbf{I} \times \mathbf{K} \mapsto \mathbb{R}^3$, where $\mathcal{T}_\mathrm{st}(\boldsymbol{i}, k)$ represents the location of a point at time k which was at position $\boldsymbol{i}$ at time k_r.

2.2 Spatial Registration

Consider the subproblem of retrieving the transformation $\mathcal{T}_\mathrm{s}$ (in which the s stands for spatial) between the reference volume and the frame at time k. A

continuous representation is employed for the spatial transformation using free-form deformations based on B-splines[30]

$$\mathcal{T}_{\mathrm{s}}(\boldsymbol{x}) = \boldsymbol{x} + \sum_{\boldsymbol{j}\in\mathbf{J}} \boldsymbol{a}_{\boldsymbol{j}} \phi_{\boldsymbol{j}}(\boldsymbol{x}) \tag{1}$$

where $\boldsymbol{x} \in \mathbf{X}$ is the continuous spatial coordinate associated with $\mathbf{I}$; $\mathbf{J} \subset \mathbb{Z}^3$ is the set of spatial parameter indices considered for basis functions $\phi_{\boldsymbol{j}}(\boldsymbol{x}) = \beta^n(\boldsymbol{x}/h - \boldsymbol{j})$ with $h \in \mathbb{R}$ the uniform spatial control point spacing, and β^n the tensor product of one-dimensional B-spline kernels of degree n. We used cubic B-splines for the spatial basis functions ($n = 3$). The parameters of $\mathcal{T}_{\mathrm{s}}$ are the B-spline coefficients $\boldsymbol{a}_{\boldsymbol{j}} \in \mathbb{R}^3$ (one for each component of the deformation), i.e. $\mathcal{T}_{\mathrm{s}}$ is fully characterized by specifying $\boldsymbol{a} = \{\boldsymbol{a}_{\boldsymbol{j}}\}_{\boldsymbol{j}\in\mathbf{J}}$.

We define a similarity criterion $\mathcal{J}_{\mathrm{s}}$, based on the mean squared intensity differences with respect to the samples of the reference volume

$$\mathcal{J}_{\mathrm{s}}(\mathcal{T}_{\mathrm{s}}, k) = \frac{1}{N_{\mathbf{I}}} \sum_{\boldsymbol{i}\in\mathbf{I}} \Big(f\big(\mathcal{T}_{\mathrm{s}}(\boldsymbol{i}), k\big) - f\big(\boldsymbol{i}, k_{\mathrm{r}}\big) \Big)^2 \tag{2}$$

with $N_{\mathbf{I}}$ the number of spatial samples considered. We chose this criterion because of its fast computation time and the smoothness of the resulting search space. For simplicity, no explicit regularization term was included in the criterion. For now, only the influence of the parametrization of the deformation function was explored. Evaluating the intensity function f at non-grid positions, requires a continuous representation for which we used cubic B-spline interpolation

$$f(\boldsymbol{x}, k) = \sum_{\boldsymbol{i}\in\mathbf{I}} \boldsymbol{d}_{\boldsymbol{i}} \beta^n(\boldsymbol{x} - \boldsymbol{i}) \,. \tag{3}$$

Coefficients $\boldsymbol{d}_{\boldsymbol{i}}$ are found quickly from the image intensities using recursive filtering[40].

Solving the spatial registration problem for frame k comes down to estimating the optimal parameters $\boldsymbol{a}^*$ in the sense of the criterion $\mathcal{J}_{\mathrm{s}}$

$$\boldsymbol{a}^* = \arg\min_{\boldsymbol{a}} \mathcal{J}_{\mathrm{s}}(\mathcal{T}_{\mathrm{s}}; k) \,. \tag{4}$$

By solving (4) consecutively for all $k \in \mathbf{K}$ except k_{r}, a solution to the 4D motion estimation problem can be composed. Solutions obtained for previous k values, can be used to initialize subsequent registrations.

2.3 Trajectory Modelling

Temporal sequences enable modelling the temporal variations of the estimated deformations. Tissue trajectories are expected to evolve smoothly and continuously over time, allowing to introduce constraints which enforce the temporal coherence of the deformation across the sequence. This is similar to the approach described for the spatial dimensions. Nonetheless, the temporal dimension is handled separately as it is inherently different. For instance, in the case of respiratory-correlated CT, the sequence is periodic and the number of temporal samples is low compared to spatial samples.

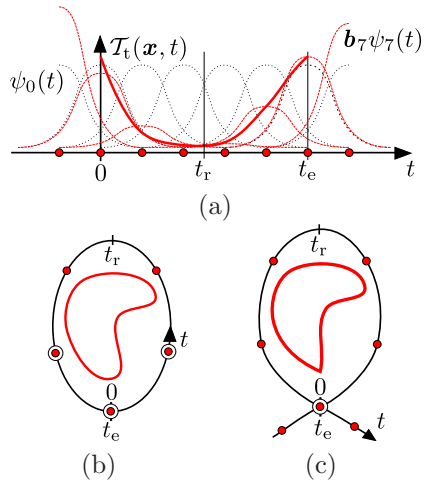


Figure 1: (a) Schematic representation of a trajectory model based on cubic B-splines, with eight control points (*red dots*) placed uniformly inside and just outside $[0, t_\mathrm{e})$. Each corresponds to a B-spline kernel ψ_l (*dotted line*) and $\mathcal{T}_\mathrm{t}(\boldsymbol{x}, t)$ (*red solid curve*) is found by combining the scaled kernels (*dashed red line*). (b, c) Alternative, representation with a cyclic temporal axis wrapped around the trajectory. Large control points indicate a constraint is applied. (*b*) The smooth trajectory model $\mathcal{T}_\mathrm{t}$ and, (*c*) the piecewise smooth trajectory model $\mathcal{T}_\mathrm{t}^*$.

Trajectory Model Let $t \in \mathbf{T}$ be the continuous coordinate associated with $\mathbf{K}$ and suppose for simplicity $\mathbf{T} = [0, t_\mathrm{e})$. Let $\mathcal{T}_\mathrm{t}(\boldsymbol{x}, t)$ denote the trajectory of a point at position $\boldsymbol{x}$ at the reference time t_r. The search for $\mathcal{T}_\mathrm{t}$ is limited to continuous and smooth functions of t, by expressing it using a suitable set of basis functions $\{\psi_l\}_{l \in \mathbf{L}}$

$$\mathcal{T}_\mathrm{t}(\boldsymbol{x}, t) = \boldsymbol{x} + \sum_{l \in \mathbf{L}} \boldsymbol{b}_l \psi_l(t) \,. \tag{5}$$

$\mathbf{L} \subset \mathbb{Z}$ is the set of temporal parameter indices and $\boldsymbol{b_l} \in \mathbb{R}^3$ the coefficients of the basis functions. We adopted temporal B-spline basis functions[10,3,14,43] of order $m \in \mathbb{N}$, $\psi_l(t) = \beta^m(t/s - l)$ with $s \in \mathbb{R}$ the temporal control point spacing, because of their good approximation properties, computational simplicity and implicit smoothness. In Ledesma-Carbayo et al. [14] temporal B-splines were found to work at least as well as harmonic functions[4,20,5]. Figure 1a shows a schematic, one-dimensional representation of a trajectory model based on cubic B-splines ($m = 3$), with five control points ($s = t_\mathrm{e}/5$) placed uniformly along the considered interval $[0, t_\mathrm{e})$. Evaluating $\mathcal{T}_\mathrm{t}$ near the borders of the interval, requires taking into account control points with non-zero weight just outside the interval. It can be seen that a total of eight degrees of freedom is considered, represented by the B-spline coefficients $\boldsymbol{b}_0$ to $\boldsymbol{b}_7$.

Smooth Trajectory Model The trajectory model can be further constrained by incorporating *a priori* knowledge of the motion, leading to a more restrictive parametrization. For instance, 4D CT data is inherently periodic. In addition, trajectories can be expected to be smooth functions of time. The trajectory can be made periodic and smooth throughout the entire cycle by imposing the same order of smoothness to the endpoints as the rest of the trajectory, thus obtaining $\mathcal{T}_\mathrm{t} \in C^{m-1}(\mathbf{T})$. This leads to the set of m conditions

$$\frac{\partial^z \mathcal{T}_\mathrm{t}(\boldsymbol{x}, 0)}{\partial t^z} = \frac{\partial^z \mathcal{T}_\mathrm{t}(\boldsymbol{x}, t_\mathrm{e})}{\partial t^z} \quad \text{for} \quad z = [0, \ldots, m-1\,]. \tag{6}$$

As will be shown, each condition results in a linear equation for the model parameters, allowing to express one of the parameters in function of the others.

A schematic representation of a trajectory satisfying (6) is shown in figure 1b. A cyclic time axis is shown to illustrate the placement and the influence of the control points. The banded control points indicate that a constraint is applied.

Piecewise Smooth Trajectory Model Due to the limited temporal resolution of 4D CT, and depending on the degrees of freedom considered in (5), the smoothness constraint might be too restrictive, leading to locally reduced representation accuracy in regions where the velocity is varying rapidly. This can be the case for the end-inhale phase where fast inversion of the motion takes place. Alternatively, we can locally relax the smoothness constraints and propose a piecewise smooth trajectory representation $\mathcal{T}_{\mathrm{t}}^{*}$. A similar expansion as (5) is utilized for $\mathcal{T}_{\mathrm{t}}^{*}$ but, assuming end-inhale corresponds to $t = 0$, a single constraint is applied at end-inhale

$$\mathcal{T}_{\mathrm{t}}^{*}(\boldsymbol{x}, 0) = \mathcal{T}_{\mathrm{t}}^{*}(\boldsymbol{x}, t_{\mathrm{e}}) \,. \tag{7}$$

This condition leads to periodic trajectories, but allows a discontinuity in the velocity at end-inhale (figure 1c). In this case, the sections near end-inhale are parametrized independently which implies a local increase of control points and degrees of freedom.

Temporal Constraints It is instructive to consider the effect of the temporal constraints on the trajectory model. For instance, as deformation is estimated with respect to a reference, by definition

$$\mathcal{T}_{\mathrm{t}}(\boldsymbol{x}, t_{\mathrm{r}}) = \boldsymbol{x} \,. \tag{8}$$

This condition allows to express one B-spline coefficient in terms of the others, effectively removing one degree of freedom from the system, i.e.

$$\boldsymbol{b}_{l_{\mathrm{r}}} = - \sum_{l \in \mathbf{L}, l \neq l_{\mathrm{r}}} \boldsymbol{b}_l \frac{\psi_l(t_{\mathrm{r}})}{\psi_{l_{\mathrm{r}}}(t_{\mathrm{r}})} \,, \tag{9}$$

in which we constrained the parameter $\boldsymbol{b}_{l_{\mathrm{r}}}$ associated to the basis function $\psi_{l_{\mathrm{r}}}$, which is non-zero at t_{r}. Following Ledesma-Carbayo et al. [14], introducing (9) into (5), and regrouping terms for each of the parameters allows the temporal model to be expressed using a smaller set of constrained basis functions

$$\psi_l^{\mathrm{c}}(t) = \psi_l(t) - \frac{\psi_l(t_{\mathrm{r}})\psi_{l_{\mathrm{r}}}(t)}{\psi_{l_{\mathrm{r}}}(t_{\mathrm{r}})} \,, \tag{10}$$

that only generates trajectories that satisfy (8).

In the following, we will denote $\mathbf{L}^{\mathrm{c}}$ the set of temporal indices of basis functions to which constraints (6) and (8) have been applied.

2.4 Spatio-Temporal Registration

Estimating the motion in a 4D CT sequence by performing consecutive spatial registrations using (1) fails to exploit the temporal relation between the frames.

This is remedied by modelling the trajectory as in (5). A global approach is found by coupling the temporal and the spatial deformation model

$$\mathcal{T}_{\mathrm{st}}(\boldsymbol{x}, t) = \boldsymbol{x} + \sum_{\boldsymbol{j} \in \mathbf{J}} \sum_{l \in \mathbf{L}^c} \boldsymbol{c}_{\boldsymbol{j},l}\, \phi_{\boldsymbol{j}}(\boldsymbol{x}) \psi_l^c(t)\,. \tag{11}$$

The result is a linear, spatio-temporal deformation function, separable in space and time. A straightforward extension to the temporal dimension of (2), leads to the criterion

$$\mathcal{J}_{\mathrm{st}}(\mathcal{T}_{\mathrm{st}}) = \frac{1}{N_{\mathbf{K}}} \sum_{k \in \mathbf{K}} \mathcal{J}_{\mathrm{s}}(\mathcal{T}_{\mathrm{st}}, k) \tag{12}$$

to be optimized with respect to the parameters $\boldsymbol{c} = \{\boldsymbol{c}_{\boldsymbol{j},l}\}_{\boldsymbol{j} \in \mathbf{J}, l \in \mathbf{L}^c}$. We will use $\mathcal{T}_{\mathrm{st}}$ and $\mathcal{T}_{\mathrm{st}}^*$ in reference to the spatio-temporal deformation models obtained when using the smooth and piecewise smooth temporal models $\mathcal{T}_{\mathrm{t}}$ and $\mathcal{T}_{\mathrm{t}}^*$, respectively.

2.5 Motion Mask Extraction

Breathing motion is characterized by sliding of the liver and lungs, resulting in a discontinuity of the motion at the pleural wall[47,34]. Accurate matching in these regions requires a complex spatial transformation, even though the remainder of the deformation can be considered smooth. We previously addressed this issue[41] by automatically extracting a motion mask, dividing the thorax into moving (lungs, mediastinum and abdomen) and less-moving regions (the remainder).

Motion masks were computed for all frames of f. The result is the division of the thorax into two subregions $\mathbf{I}_{\mathrm{in}}, \mathbf{I}_{\mathrm{out}} \subset \mathbf{I}$, roughly representing the inner and outer thoracic structures. For each of the subregions, a separate registration problem can be formulated following[47], with the advantage that the search can be limited to spatially smooth deformations. In the following, we focussed on the inner thoracic structures.

2.6 Optimization

The spatio-temporal approach allows a more restrictive parametrization of the transform and reduces the total number of degrees of freedom of the 4D motion estimation problem compared to consecutively applying $\mathcal{T}_{\mathrm{s}}$. However, directly minimizing (12) considers all degrees of freedom simultaneously, increasing the dimensionality of the optimization problem with respect to one 3D-3D registration. In response, a multiresolution approach was employed, allowing to gradually increase the complexity of the problem. The resolution of the spatial dimensions of both the image sequence and the B-spline control point grid of the transformation was doubled in each of three consecutive resolution levels. The final image resolution was set to 2 mm. We previously found that, in combination with a motion mask, a control point grid spacing $h = 32$ mm provided a good compromise between registration efficiency and accuracy. The temporal dimension, characterized by low resolution, remained unmodified throughout the optimization.

Each level was handled using a Quasi-Newton approach in the form of the limited memory BFGS method[23], because of its high precision and improved

rate of convergence with respect to simple gradient descent algorithms[13]. The procedure started from a zero deformation, and subsequent levels were initialized by upsampling the solution found at the previous level. The required partial derivatives of the similarity criterion can be calculated explicitly, for instance for $\mathcal{T}_{\mathrm{st}}$

$$\frac{\partial \mathcal{J}_{\mathrm{st}}(\mathcal{T}_{\mathrm{st}})}{\partial \boldsymbol{c}_{\boldsymbol{j},l}} = \frac{2}{N_{\mathbf{K}} N_{\mathbf{I}}} \sum_{k \in \mathbf{K}} \sum_{\boldsymbol{i} \in \mathbf{I}} \Big(f\big(\mathcal{T}_{\mathrm{st}}(\boldsymbol{i},k),k\big) - f\big(\boldsymbol{i},k_{\mathrm{r}}\big) \Big) \frac{\partial f\big(\mathcal{T}_{\mathrm{st}}(\boldsymbol{i},k),k\big)}{\partial \boldsymbol{x}} \frac{\partial \mathcal{T}_{\mathrm{st}}(\boldsymbol{i},k)}{\partial \boldsymbol{c}_{\boldsymbol{j},l}}. \tag{13}$$

$\partial f / \partial \boldsymbol{x}$ is found by deriving (3), while $\partial \mathcal{T}_{\mathrm{st}} / \partial \boldsymbol{c}_{\boldsymbol{j},l}$ can be calculated considering (11). For instance, for the q^{th} spatial component

$$\frac{\partial f(\boldsymbol{x},k)}{\partial x_q} = \sum_{\boldsymbol{i} \in \mathbf{I}} \boldsymbol{d}_{\mathbf{i}} \frac{\partial \beta^n(x_q)}{\partial x_q} \prod_{\zeta \neq q} \beta^n(x_\zeta - i_\zeta) \tag{14}$$

$$\frac{\partial \mathcal{T}_{\mathrm{st}}(\boldsymbol{i},k)}{\partial \boldsymbol{c}_{\boldsymbol{j},l,q}} = \phi_{\boldsymbol{j}}(\boldsymbol{x}) \psi_l^{\mathrm{c}}(t)\,. \tag{15}$$

2.7 Implementation

The registration algorithms were implemented in C++. Evaluating $\mathcal{T}_{\mathrm{st}}(\boldsymbol{i},k)$ was performed using B-LUTs[32]: a fast, low memory B-spline implementation based on a look-up table of B-spline tensor products $\beta^n(\boldsymbol{x})$. Registration algorithms were multi-threaded and executed on an eight-core system. The execution times depended on the specifics of the 4D CT data set. Consecutively registering all frames of f using $\mathcal{T}_{\mathrm{s}}$, required between 5 and 10 hours, whereas T_{st} and $\mathcal{T}_{\mathrm{st}}^*$ required about twice as much time.

In comparison, the most expensive step for the spatio-temporal approach is the calculation of $\partial \mathcal{T}_{\mathrm{st}} / \partial \boldsymbol{c}$ in (13). In the current implementation, this requires multiple table look-ups due to the presence of the modified basis functions ψ_l^{c} in (15), compared to only one for calculating $\partial \mathcal{T}_{\mathrm{s}} / \partial \boldsymbol{a}$.

3 Experiments

Three types of experiments were performed to validate the spatio-temporal deformation model. First, we conducted a prestudy on the temporal dimension of the model separately. Breathing patterns are patient-specific and strong interpatient and intercycle variability has been reported which can affect cycle duration, motion amplitude and speed of the movement[7,46,29]. By fitting the trajectory models to motion data covering many cycles and measured on a large set of patients, we ensured the temporal parametrization is flexible enough to capture the dynamics of respiratory motion.

The most suitable temporal parameter values were retained and used for the spatio-temporal deformation models. The latter were applied to the registration of 4D CT images of the thorax. Extensive spatial validation of the registrations was performed using large sets of landmarks. The registration accuracy was compared to the conventional frame-to-frame approach. In a final experiment, the benefit of the spatio-temporal approach was illustrated, by assessing the performance in the presence of artifacts.

3.1 Temporal Fit of Diaphragm Motion Data

Data Description We used projection sequences of cone-beam computed tomography (CBCT) acquired at the Netherlands Cancer Institute (Amsterdam, the Netherlands) for image-guidance of 33 lung cancer patients treated by radiotherapy with the protocol described in Sonke et al. [38]. Cone-beam projections consist of planar X-ray images, acquired from rotating views around the patient. They were acquired at 5.5 fps over 200° with a 50°/min gantry rotation speed for 4D CBCT imaging[37]. Their resolution was 512^2 pixels of 0.8^2 mm^2 (0.52^2 mm^2 at the isocenter). 257 sequences of cone-beam projection images (5 to 19 per patient) were analyzed.

The motion was analyzed by extracting the cranio-caudal position of a diaphragm dome using an adapted version[29] of the algorithm developed to extract a respiratory signal for 4D CBCT reconstruction[51,37]. The extraction resulted in a 2 min 1D+t signal per acquisition with 0.52 mm and 5.5 fps resolution, i.e. 2 of the 4 dimensions of the sought 4D model at fine resolution but for only one point of space. In addition to the largeness of the dataset, the projection images are advantageous because they have higher cranio-caudal and temporal resolutions than 4D CT images. As such, the diaphragm motion data provided a valuable benchmark for tuning the temporal parametrization of the deformation models.

Experiments Each signal was split in respiratory cycles by detecting the end-inhale peaks after smoothing out the local minima. Each cycle was analyzed separately by assuming periodicity, similar to 4D CT images. The temporal models described in section 2.3, were fitted to each cycle with the optimal solution in the least square sense. The influence of the trajectory model parameters was evaluated: we varied the B-spline order m and the control point spacing s, or equivalently the number of control points. In addition, we verified the suitability of a smoothness constraint at end-inhale by comparing $\mathcal{T}_\mathrm{t}$ and $\mathcal{T}_\mathrm{t}^*$.

The similarity between the measured and the fitted signals was evaluated using the root mean square (RMS) of their difference. Results for each patient were averaged and the group mean over all patients was computed. The results were analyzed both globally and per respiratory phase by dividing each cycle into ten equitemporal-temporal phase bins, as it is typically done in current 4D CT scanners.

3.2 Spatio-Temporal Registration of 4D CT

Data Description We used 4D CT data sets of six non-small cell lung cancer patients acquired at the Léon Bérard Cancer Center (Lyon, France) for the purpose of radiotherapy planning on a Philips 16-slice Brilliance Big Bore Oncology Configuration (Phillips Medical Systems, Cleveland, OH). Acquisitions were performed in helical mode using a table pitch of 0.1, 400 mAs effective exposure (80 mA tube current) at 120 kV.

Respiratory-correlated reconstruction was performed through simultaneous acquisition of a respiratory surrogate signal, provided by the Pneumo Chest pressure belt (Lafayette Instrument, Lafayette, IN). Reconstruction yielded ten 3D CT frames at approximately $1 \times 1 \times 2$ mm^3 resolution.

Experiments The spatio-temporal deformable registration approaches $\mathcal{T}_{\mathrm{st}}$ and $\mathcal{T}_{\mathrm{st}}^{*}$, as described in section 2.4, were applied to all 4D CT data sets. In addition to the normal set of constraints used for $\mathcal{T}_{\mathrm{t}}$ and $\mathcal{T}_{\mathrm{t}}^{*}$, constraint (8) was enforced for all registrations. Deformable registration was performed with respect to the middle frame ($k = 5$), roughly corresponding to end-exhale. The position of end-exhale has been reported to be more reproducible than end-inhale[36,38], making it a suitable reference to analyze breathing motion.

For validation purposes, anatomical landmarks were identified in the lung region using a semi-automatic software tool[22]. The system automatically provided a set of well-distributed, distinctive points with index $\boldsymbol{p}_{\mathrm{r}} \in \mathbf{I}$ in the lung region of the end-exhale frame. Observers identified the corresponding positions $\boldsymbol{p}_k \in \mathbf{I}$ of the points in frame k, using a custom designed interface and aided by initial estimates provided by the system. Points coinciding with artifacts were excluded. The system initially provided 130 distinctive points and the procedure was stopped after 100 points were successfully identified in the corresponding frames.

For all six patients, 100 point correspondences were provided between the end-exhale and the end-inhale frame, and the process was repeated by a second observer. The mean distance between the annotations was 0.5 mm (0.9 mm standard deviation). For Patients 1-3, a single observer provided 100 correspondences for each of the frames of the 4D CT, resulting in a total of 900 manually identified landmarks for each of the three data sets. The manual annotations were compared to the corresponding point positions estimated through registration by computing the target registration error (TRE)

$$\mathrm{TRE} = \left\| \mathcal{T}_{\mathrm{st}}(\boldsymbol{p}_{\mathrm{r}}, k) - \boldsymbol{p}_k \right\| . \tag{16}$$

The registration results were also compared to those obtained when performing consecutive 3D registrations using $\mathcal{T}_{\mathrm{s}}$, described in section 2.2. The same multiresolution and optimization scheme was applied as in the spatio-temporal case. Since no temporal regularization is applied in the case of $\mathcal{T}_{\mathrm{s}}$, the results are considered as a reference indicating the achievable registration accuracy when allowing all temporal degrees of freedom.

3.3 Registration of 4D CT with Artifacts

Data Description To dispose of a ground truth, we constructed 4D CT acquisition with artifacts f_{a} by introducing a simulated, motion-induced artifact in the 4D CT of patient 2, characterized by large motion. A mid-inhalation frame ($k = 8$) was altered by modifying a series of axial slices halfway the lungs. Ten slices starting from slice index $i_2 = i_{\mathrm{a}}$ in the end-exhale frame ($k = 5$), were copied to the same location in the target frame, i. e.

$$f_{\mathrm{a}}(\boldsymbol{i}, k) = \begin{cases} f(\boldsymbol{i}, 5) & \text{for } k = 8, i_2 \in [i_{\mathrm{a}}, i_{\mathrm{a}} + 10), \\ f(\boldsymbol{i}, k) & \text{otherwise.} \end{cases} \tag{17}$$

The procedure resulted in an axial slab of 20 mm along the cranio-caudal direction, containing an inconsistent view of the patient anatomy with respect to the surrounding slices. This resembles the situation of a frame locally influenced by erroneous tagging of the respiratory phase, or irregular breathing during image acquisition.

Spline	Internal Control Points						
Degree m	4	5	6	7	8	9	10
	Group mean RMS (mm) of $\mathcal{T}_{\mathrm{t}}$						
0	2.65	2.25	1.95	1.72	1.53	1.39	1.27
1	1.25	0.98	0.83	0.72	0.63	0.57	0.51
2	1.17	0.89	0.77	0.67	0.60	0.54	0.49
3	1.17	0.87	0.76	0.66	0.59	0.53	0.48
	Group mean RMS (mm) of $\mathcal{T}_{\mathrm{t}}^{*}$						
2	0.70	0.56	0.47	0.40	0.34	0.30	0.27
3	0.67	0.55	0.47	0.40	0.35	0.30	0.27

Table 1: Group mean of the RMS of the difference between the measured motion of the diaphragm dome in the cranio-caudal direction and fitted functions for the two temporal models with different number of control points and B-spline degrees. For degrees 0 and 1, both models are equivalent.

Experiments The simulated sequence f_{a} was registered in the same way as described in the previous experiment. The registration results were compared to those obtained using consecutive 3D registrations. By comparing also to the results obtained for the original sequence, the sensitivity of the method to locally introduced artifacts was evaluated.

The registration accuracy was assessed by using the landmarks identified in the original, unmodified 4D CT acquisition. The analysis is performed at two levels. First, we computed a global evaluation of the TRE, taking into account all landmarks. Second, a local analysis was performed by only considering landmarks located within five slices of the artifact.

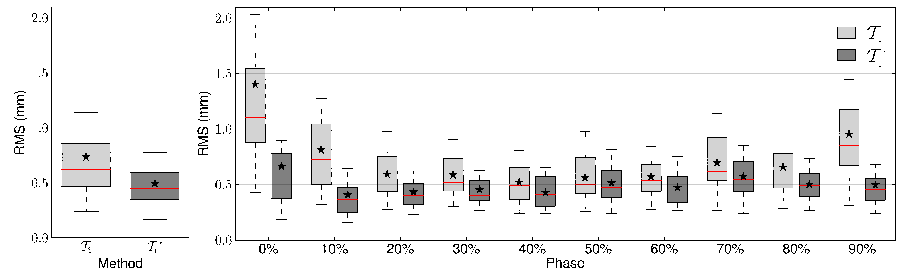
4 Results

4.1 Temporal Fit of Diaphragm Motion Data

The fitting results are summarized in table 1 for a variety of B-spline functions. Constant B-splines ($m = 0$) gave much poorer results than other degrees. This is not surprising, since they produce piecewise constant functions which can not describe the continuity of the respiratory motion. Linear B-splines ($m = 1$) gave residuals of the same order, but were found significantly worse than quadratic ($m = 2$) and cubic splines ($m = 3$) for all tested models ($p < 3 \times 10^{-4}$). Cubic B-splines consistently gave better results than quadratic splines, although the difference was not significant for all tested models in table 1. As they also inherently impose a stronger temporal regularization, which is our purpose, they were selected for the rest of the study.

The influence of the number of control points can also be seen from table 1. For both models, the residual of the fit was proportional to the spacing of the control points s: the Pearson's product-moment correlation coefficient was greater than 0.99. As expected, the number of control points is a trade-off between the achievable representation accuracy and the parameters of the fitted function.

We used box and whisker plots to further illustrate the distribution of fitting



(a) (b)

Figure 2: Box and whisker plots of the RMS errors per patient after fitting the models to the diaphragm motion data, using with five internal control points and $m = 3$. (a) The RMS over the entire cycle; (b) the RMS per phase bin.

errors. The box extends from the lower ($p_{25\%}$) to upper quartile ($p_{75\%}$) of the data, with a red horizontal line at the median and a $*$ symbol at the mean. The whiskers extend from the box to the most extreme value below $p_{25\%} + 0.75 \times (p_{75\%} - p_{25\%})$, the remaining points are considered outliers. Outliers were not plotted for clarity.

Figure 2 illustrates the global fitting errors and the errors per respiratory phase bin for $\mathcal{T}_\mathrm{t}$ and $\mathcal{T}_\mathrm{t}^*$ (using five internal control points and $m = 3$). The overall performance of $\mathcal{T}_\mathrm{t}^*$ was much better than for $\mathcal{T}_\mathrm{t}$. The largest discrepancies can be observed near end-inhale (0%). At this phase, the change of speed was too sudden to be described by the smooth trajectory model $\mathcal{T}_\mathrm{t}$ and the residual was found to be significantly higher ($p < 2 \times 10^{-3}$) than at all other phases, for all tested values of m and s. This was not the case for the piecewise smooth $\mathcal{T}_\mathrm{t}^*$ model, which resulted in more homogeneous residuals per phase because the smoothness constraint is relaxed at end-inhale.

4.2 Spatio-Temporal Registration of 4D CT

We retained the temporal representations with four and five internal control points for the spatio-temporal model, which corresponds to $s = 2.5$ and 2 frames, respectively. Table 2 summarizes the temporal characteristics of the registration methods.

Trajectories obtained for landmarks with large displacements are plotted in figure 3. The landmark positions identified manually throughout the 4D CT are also shown and were linearly interpolated for clarity. The estimates obtained using $\mathcal{T}_\mathrm{s}$ were interpolated using cubic splines. The trajectories of the spatio-temporal methods were directly obtained from the continuous 4D transforms. All trajectories were projected on the sagittal plane, where motion predominantly occurs. Overall, the obtained trajectories appear very similar. The spatio-temporal trajectories tend to be smoother than $\mathcal{T}_\mathrm{s}$. The main difference between $\mathcal{T}_\mathrm{st}$ and $\mathcal{T}_\mathrm{st}^*$ is visible at end-inhale (bottom of the plot). At this point, $\mathcal{T}_\mathrm{st}^*$ tends to be pointier and in some cases visibly closer to the corresponding landmark. Note that deviations between the estimated trajectories and the measured landmark trajectories are partially due to the landmark identification process, which was performed in voxel index space, while the trajectories evolve in the continuous space. This effect will also contribute to the registration errors

Properties	Representation				
	$\mathcal{T}_{\mathrm{s}}$	$\mathcal{T}_{\mathrm{st}}$		$\mathcal{T}^*_{\mathrm{st}}$	
	$s=1$	$s=2$	$s=2.5$	$s=2$	$s=2.5$
Temporal CP	9	8	7	8	7
Continuity at $t=0$	$\diagup$	C^2	C^2	C^0	C^0
Constraints	0	4	4	2	2
Temporal DOF	9	4	3	6	5
Parameters	63 882	28 392	21 294	42 588	35 490

Table 2: Summary of the temporal properties for the registration methods when using cubic splines for the spatio-temporal methods, and a control point spacing of either 2 or 2.5 frames. The amount of temporal control points (CP) reflect the internal CP as well as the ones required at the border. The number of degrees of freedom (DOF) are the number of CP, reduced by the number of constraints. As an example, we list the resulting number of parameters required to register the inner thoracic region ($\mathbf{I}_{\mathrm{in}}$) for Patient 1.

Patient	TRE for 4D CT (mm)			
	Original	$\mathcal{T}_{\mathrm{s}}$	$\mathcal{T}_{\mathrm{st}}$	$\mathcal{T}^*_{\mathrm{st}}$
1	3.47 ± 2.14	0.96 ± 0.66	1.02 ± 0.71	1.00 ± 0.69
2	6.41 ± 3.99	1.20 ± 0.96	1.37 ± 1.13	1.27 ± 1.09
3	3.65 ± 3.04	1.11 ± 1.14	1.17 ± 1.08	1.16 ± 1.15
GM	4.51 ± 3.15	1.09 ± 0.94	1.19 ± 0.99	1.14 ± 1.00

Table 3: The mean TRE obtained over the nine frames for Patients 1-3 based on 900 landmarks each, and its group mean (GM). The registration error (± 1 SD) of the 3D registration is compared to the accuracy obtained for the spatio-temporal algorithms with $m = 3$ and $s = 2$ frames. The original landmark distance (*Original*) is given to illustrate the magnitude of the motion.

evaluated using the landmarks.

For Patients 1-3, landmarks were available for all frames of the 4D CT. The global registration accuracy is summarized in table 3 in terms of the mean TRE based on 900 landmarks each. For a spacing $s = 2$ frames, the group mean TRE of both spatio-temporal methods was within 0.1 mm of $\mathcal{T}_{\mathrm{s}}$. When increasing the temporal control point spacing from 2 to 2.5, the mean TRE increased gradually (1.27 ± 1.17 mm for $\mathcal{T}^*_{\mathrm{st}}$ and 1.18 ± 1.03 mm for $\mathcal{T}^*_{\mathrm{st}}$), but remained comparable to $\mathcal{T}_{\mathrm{s}}$. For clarity, only results using a spacing of 2 frames will be shown in the following.

For Patients 1-3, the registration errors were also analyzed for each frame separately. Figure 4a corresponds to the group mean TRE of the entire 4D CT and figure 4b shows the group mean TRE per frame. The mean TRE over the entire 4D CT was comparable for all methods, though $\mathcal{T}_{\mathrm{st}}$ performed slightly worse. The analysis per phase revealed that most discrepancies in TRE are located near end-inhale (0%, 10% and 90 %). $\mathcal{T}^*_{\mathrm{st}}$ generally obtained an accuracy closer to $\mathcal{T}_{\mathrm{s}}$ for these phases.

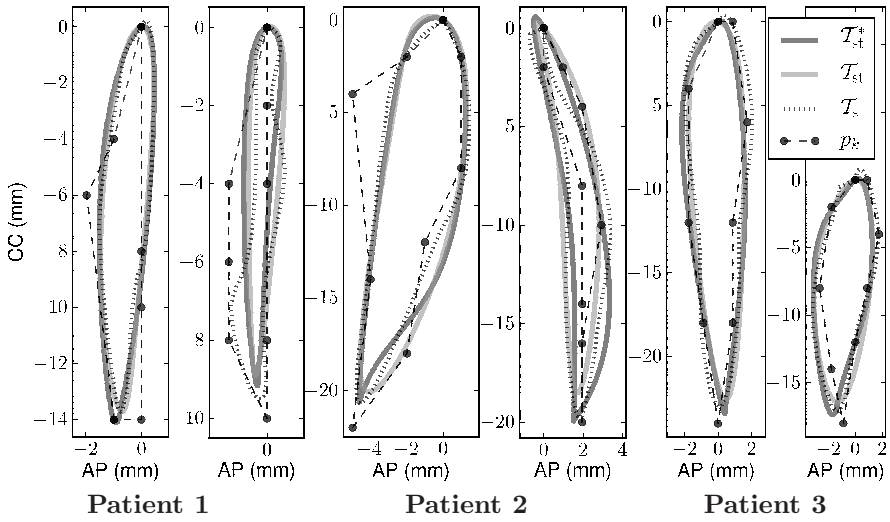
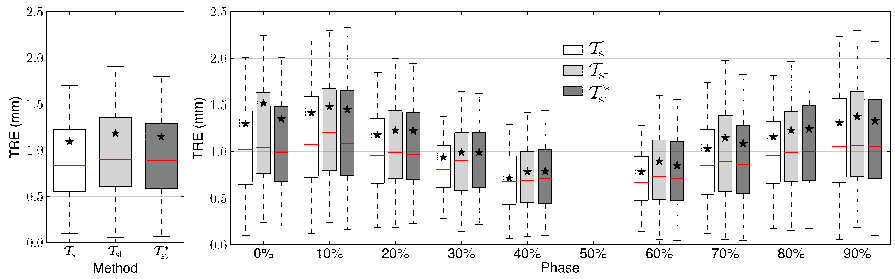


Figure 3: Trajectories projected on the sagittal plane, for some landmarks with large displacements of Patients 1-3. The trajectories obtained using the respective registration methods are plotted with the manually identified landmark positions throughout the 4D CT (p_k). For all trajectories shown, at least two landmarks positions coincided at end-exhale, corresponding to position $(0, 0)$.



(a) (b)

Figure 4: Box and whisker plots of the group mean TRE for Patients 1-3 for which landmarks were available in all frames, using $m = 3$ and $s = 2$ frames. (a) The combined registration errors for the entire 4D registration. Each box is drawn based on 2700 landmarks. (b) TRE per phase bin. For each frame, the registration error is estimated from 300 landmarks. The 50% phase bin corresponds to the reference frame.

Patient	TRE for End-Inhale (mm)			
	Original	$\mathcal{T}_{\mathrm{s}}$	$\mathcal{T}_{\mathrm{st}}$	$\mathcal{T}_{\mathrm{st}}^{*}$
1	6.34 ± 2.94	0.94 ± 0.51	0.98 ± 0.56	0.96 ± 0.57
2	14.00 ± 7.17	1.44 ± 1.04	1.95 ± 1.88	1.56 ± 1.34
3	7.67 ± 5.03	1.51 ± 1.66	1.63 ± 1.66	1.53 ± 1.70
4	7.33 ± 4.86	1.79 ± 2.71	1.97 ± 3.00	1.96 ± 2.92
5	7.09 ± 5.08	1.43 ± 1.39	1.54 ± 1.49	1.48 ± 1.39
6	6.68 ± 3.67	1.18 ± 0.80	1.32 ± 1.13	1.25 ± 0.95
GM	8.19 ± 4.97	1.38 ± 1.53	1.57 ± 1.78	1.46 ± 1.65

Table 4: The mean TRE (± 1 SD) obtained by evaluating the registration only at end-inhale for Patients 1-6 based on 100 landmarks each, and its group mean (GM). The registration error of the 3D registration ($\mathcal{T}_{\mathrm{s}}$) is compared to the accuracy obtained for the spatio-temporal algorithms ($\mathcal{T}_{\mathrm{st}}$ and $\mathcal{T}_{\mathrm{st}}^{*}$) using $m = 3$ and $s = 2$ frames. The original landmark distance (*Original*) is given to illustrate the magnitude of the motion.

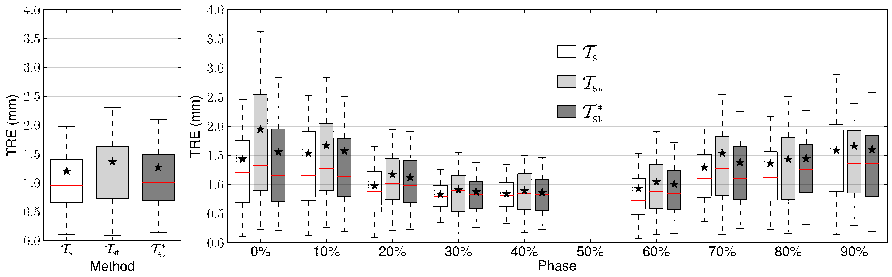
A separate, more extensive evaluation of the accuracy of the end-exhale to end-inhale registration is listed in table 4 for Patients 1-6. $\mathcal{T}_{\mathrm{st}}^{*}$ consistently outperforms $\mathcal{T}_{\mathrm{st}}$ in terms of mean TRE. The difference in group mean TRE between $\mathcal{T}_{\mathrm{st}}^{*}$ and $\mathcal{T}_{\mathrm{s}}$ was below 0.1 mm. In contrast, the difference in mean TRE between $\mathcal{T}_{\mathrm{s}}$ and $\mathcal{T}_{\mathrm{st}}$ was above 0.1 mm for five out of six patients, and the group mean TRE was almost 0.2 mm higher. This confirms the results reported when fitting the diaphragm motion data in section 4.1, where it was found that the smooth temporal model resulted in larger errors near end-inhale.

Table 3 shows relatively small differences in group mean TRE over the entire 4D CT, suggesting comparable performance for all registration methods. This measure was found misleading, as it tends to average out the differences due to the large numbers of measurements (2700 landmarks for each method). Further analysis showed that the performance of $\mathcal{T}_{\mathrm{st}}$ varied considerably from patient to patient. While for Patient 1, all methods obtained very similar results, differences in TRE of the order of 0.5 mm were found at certain breathing phases for Patient 2 (figure 5, note the change in scale with respect to figure 4).

4.3 Spatio-Temporal Registration of 4D CT with Artifacts

The registration accuracy obtained for the sequence f_{a} is summarized in table 5. We only report results using the piecewise smooth spatio-temporal model $\mathcal{T}_{\mathrm{st}}^{*}$. We also list the TRE obtained for the original 4D CT, corresponding to Patient 2. With respect to the original 4D CT, the local and global TRE of $\mathcal{T}_{\mathrm{s}}$ and $\mathcal{T}_{\mathrm{st}}^{*}$ are within 0.1 mm . After inserting the artifact, the global TRE more than doubles for $\mathcal{T}_{\mathrm{s}}$, while the TRE of the spatio-temporal method, increases only marginally. Locally, the influence of the artifact is even more noticeable for $\mathcal{T}_{\mathrm{s}}$. For the spatio-temporal approach however, the local TRE remains below 2 mm for $\mathcal{T}_{\mathrm{st}}^{*}$.

Figure 6 shows the motion fields obtained using $\mathcal{T}_{\mathrm{s}}$ and $\mathcal{T}_{\mathrm{st}}^{*}$ for the inner thoracic region. The top row corresponds to the original 4D CT acquisition of Patient 2. Both methods produce very similar motion fields. The main



(a) (b)

Figure 5: Box and whisker plots of TRE for Patients 2 for which landmarks were available in all frames, using $m = 3$ and $s = 2$ frames. (a) The combined registration errors for the entire 4D registration. Each box is drawn based on 900 landmarks. (b) TRE per phase bin. For each frame, the registration error is estimated from 100 landmarks. The 50% phase bin corresponds to the reference frame.

Measure	**Data**	**TRE (mm)**		
		Original	$\mathcal{T}_{\mathrm{s}}$	$\mathcal{T}^{*}_{\mathrm{st}}$
global	f	9.00 ± 3.93	1.42 ± 1.30	1.44 ± 1.16
	f_{a}		3.17 ± 3.47	1.57 ± 1.20
local	f	11.40 ± 3.74	1.38 ± 1.44	1.46 ± 1.05
	f_{a}		6.82 ± 4.38	1.90 ± 1.22

Table 5: The mean TRE for the 4D CT sequence with simulated artifacts (f_{a}) and for the original, unmodified 4D CT (f) corresponding to Patient 2. The evaluation is limited to the frame where the artifact is introduced. The global TRE is based on 100 landmarks. The local TRE is based on 24 landmarks, all within 5 slices of the inserted artifact. The registration error of the 3D registration ($\mathcal{T}_{\mathrm{s}}$) is compared to the accuracy obtained for the spatio-temporal algorithm ($\mathcal{T}^{*}_{\mathrm{st}}$ using $m = 3$ and $s = 2$ frames). The original landmark distance (*Original*) is given to illustrate the magnitude of the motion.

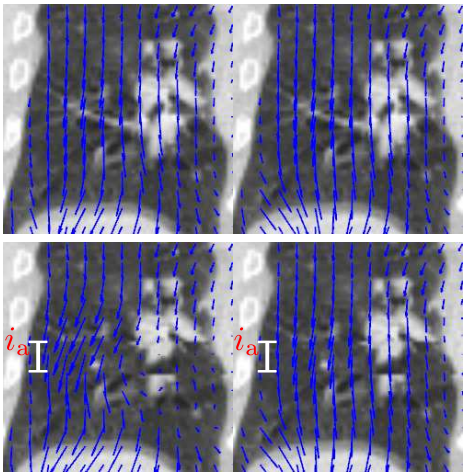


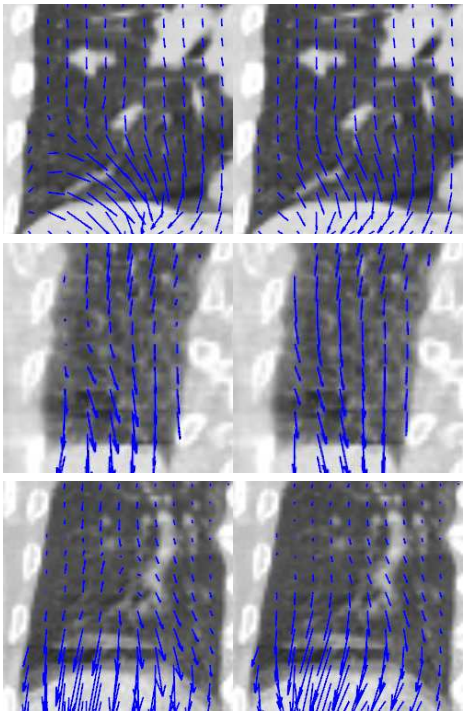
Figure 6: Motion fields in the presence of simulated artifacts: the top row corresponds to the original 4D CT acquisition of Patient 2, the bottom row corresponds to the modified sequence f_a in which an artifact was inserted at position i_a. (a) Coronal view of the motion field obtained for $\mathcal{T}_s$, (b) corresponding view of the motion field obtained for $\mathcal{T}_{st}^*$ using $m = 3$ and $s = 2$ frames.

differences can be observed near the diaphragm. The bottom row corresponds to the sequence with artifacts f_a. In this case, strong perturbations can be observed in the motion field obtained using $\mathcal{T}_s$. The influence of the introduced artifact is also noticeable for $\mathcal{T}_{st}^*$, but the changes in the motion field are less dramatic.

In figure 7, three examples are given of motion fields in the presence of real artifacts. The artifacts are shown in the first column and are generally most clearly visible near the diaphragm. The second column shows the motion fields obtained using $\mathcal{T}_s$. Strong perturbations can be seen, mainly in the part of the motion field that maps to region of the artifact; i.e., slightly above the location of the artifacts. The spatio-temporal approach tends to be less influenced by the artifacts. The resulting motion fields are noticeably smoother making them more plausible from a physiological point of view.

5 Discussion

Temporal Constraints In section 4.1, the smooth and piecewise smooth temporal models were compared at equal control point spacing (see figure 2). The difference in temporal constraints between $\mathcal{T}_t$ and $\mathcal{T}_t^*$, results in a different number of degrees of freedom at equal control point spacing. We therefore also performed a comparison between both models at equal degrees of freedom. The global performance of $\mathcal{T}_t$ was still significantly worse ($p < 4 \times 10^{-2}$) compared to the corresponding $\mathcal{T}_t^*$ models (see table 1 for $\mathcal{T}_t^*$ with two control points less than $\mathcal{T}_t$). In addition, despite the *global* increase in degrees of freedom, the high fitting residual at end-inhale remained for $\mathcal{T}_t$.



(a) (b)

Figure 7: Three examples of motion fields in the presence of real artifacts (a) Motion fields obtained for (b) $\mathcal{T}_{\mathrm{s}}$ and (b) $\mathcal{T}_{\mathrm{st}}^{*}$ using $m = 3$ and $s = 2$ frames.

This confirms that a *local* increase in control points (as is the case for the piecewise smooth model $\mathcal{T}_{\mathrm{t}}^{*}$) is more efficient in terms of number of parameters, to obtain an accurate representation throughout the respiratory cycle. It should not be excluded that other piecewise models can be found, requiring less degrees of freedom, while obtaining a similar accuracy. These could consist in making sensible assumptions about the trajectory near end-inhale that can be translated into constraints, eliminating one or both of the additional degrees of freedom.

The end-exhale phase, though also characterized with inversion of the motion, did not require further investigation of the constraints. It has been reported that respiratory motion tends to be asymmetrical[18,36], spending more time near end-exhale than end-inhale. Phase bins near end-exhale will represent relatively small deformation with respect to each other. Uniformly spaced control points with respect to these bins will lead to a spatially higher control point density near end-exhale, allowing a more accurate representation, even in the presence of smoothness constraints. This is confirmed by the trajectories depicted in figure 3, where two to three landmarks coincided with the end-exhale position.

From a temporal constraints point of view, $\mathcal{T}_{\mathrm{t}}^{*}$ is related to the trajectory model developed in Castillo et al. [2], where a compressible flow algorithm is extended with local trajectory modelling to perform 4D motion estimation for 4D CT. In this case however, one-way (and not cyclic) trajectories were sought between end-inhale and end-exhale, thus not requiring further attention at end-inhale. Cubic polynomials (equivalent to four degrees of freedom) were found to provide sufficient flexibility to parametrize the sought trajectories spanning six frames of the 4D CT. This corresponds well to the six degrees of freedom ($\mathcal{T}_{\mathrm{st}}^{*}$ with a temporal control point spacing of 2) describing the trajectory over 10 frames.

Spatio-Temporal Registration The principal aim of this study was to develop a low-dimensional spatio-temporal deformation model to improve robustness of the subsequent registration. We pursued a restrictive parametrization and strong temporal regularization, as these were expected to reduce sensitivity to noise and artifacts. The parametrization was thoroughly investigated, both spatially and temporally, to ensure an accurate representation of breathing motion was maintained.

Based on the fitting experiments of the diaphragm motion data, cubic temporal B-splines were found to perform best and were selected for the temporal parametrization. The value of the temporal control point spacing s was found to represent a trade-off between achievable accuracy on one hand and an increase of parameters on the other, the latter likely to increase sensitivity to noise and artifacts. In practice, its value should reflect the needs of the application and the quality of the images. Using $s = 2$, the spatio-temporal models obtained results comparable to the reference $\mathcal{T}_{\mathrm{s}}$ method, and was considered a suitable compromise for the 4D CT images dealt with in this study.

From a parametrization point of view, the $\mathcal{T}_{\mathrm{st}}$ model represents interesting characteristics. Minimal curvature is enforced throughout the entire cycle, and about a third less parameters are required with respect to $\mathcal{T}_{\mathrm{st}}^{*}$. Unfortunately, detailed analysis revealed larger TRE near end-inhale for $\mathcal{T}_{\mathrm{st}}$, indicating the smooth model fails to capture the full extent of the motion. Even though trajectories are expected to be smooth functions of time, a temporally smooth

parametrization was found to provide a less accurate representation, due to the low temporal resolution of respiratory-correlated imaging and the control point grid. Using $\mathcal{T}_{\text{st}}^*$ a uniform performance over the breathing cycle was obtained for all patients, and the group mean TRE was within 0.1 mm of the reference $\mathcal{T}_{\text{s}}$, for both $s = 2$ and 2.5 frames.

The improved matching of $\mathcal{T}_{\text{st}}^*$ at end-inhale comes at the price of two additional temporal degrees of freedom with respect to the smooth model. Nonetheless, with respect to $\mathcal{T}_{\text{s}}$, this model reduces the number of parameters to be estimated during registration by 33% and 45% for $s = 2$ and 2.5 frames, respectively (table 2). The impact of this compact, spatio-temporal parametrization was illustrated in the experiment in which a simulated motion-induced artifact was introduced in a 4D CT sequence. The motion field obtained using spatio-temporal registration was found to be considerably less influenced by the artifact, in comparison to the result obtained using $\mathcal{T}_{\text{s}}$.

Influence of tagging and binning In section 2, we made the assumption that the fourth image dimension was time. This allowed us to interpret $\mathcal{T}_{\text{t}}$ as a trajectory in function of time, and its derivatives as velocity and acceleration. In the case of 4D CT imaging, each frame is composed of data acquired at different times and different table positions. The interpretation of the fourth image dimension is closely related to the binning of acquired data, which is usually based on a surrogate signal.

The 4D CT data presented here was obtained from phase-based binning, which is by far the most common procedure. For each acquired cycle, the reconstruction of the different frames is performed by selecting projections equally spaced in time. Ignoring the non-periodic nature of breathing motion, the images obtained in this fashion can be considered equivalent to a temporal sequence. Alternative binning criteria[17,24] have been proposed. In particular, amplitude-based binning is expected to provide frames uniformly spaced with respect to organ displacement. For such images, other smoothness constraints in end-inhale and even end-exhale might be more suitable.

For all data presented, individual cycles were detected by tagging at end-inhale. The diaphragm motion data was artificially made periodic at end-inhale, to allow fitting the cyclic trajectory models. This procedure can be held partly responsible for the rapid changes at end-inhale, and contributes to the fitting residuals. To quantify the effect, we repeated the experiments when tagging at end-exhale. Comparatively larger residuals were observed at end-exhale, indicating the influence of the tagging position. The highest residuals were however still observed at end-inhale, confirming they are indeed caused by the sudden change in motion.

Robustness to Artifacts Thoroughly evaluating the performance in the presence of artifacts is difficult due to the absence of a ground truth for the underlying image. The improved robustness of the spatio-temporal approach was therefore illustrated using a simple experiment based on simulated data, and through visual inspection of motion fields for real artifacts. Further analysis of the behaviour of the spatio-temporal model in the presence of artifacts or noise is required. In particular, the influence of the temporal size and location of artifacts merits attention.

In the 3D (2) and 4D criterion (12), regularization terms penalizing undesirable properties of the deformation field were not included. The present study was limited to investigating the impact of explicit parametric restrictions. Regularization penalties will provide additional robustness and are expected to be complementary to the parametric contributions. In addition, the spatio-temporal framework allows regularization schemes to be extended to the temporal dimension, as in [1].

Applications of Spatio-temporal Motion Analysis The spatio-temporal deformation model (11) was applied to 4D CT of the thorax in a 3D-4D registration framework. The model can also be applied to spatio-temporal motion analysis between sequences. By relaxing the condition (8) which constrains the deformation at the reference frame, the model can be applied to a 4D-4D, frame-to-frame registration framework. By replacing criterion (12) with a suitable similarity measure, other modalities and even multi-modal problems can be studied.

This is similar to the 4D-4D registration method for respiratory-correlated images described by Schreibmann et al. [35] or the spatio-temporal alignment of cardiac sequences presented by Perperidis et al. [26]. In comparison, our method assumes only frame-to-frame spatial deformations without temporal shifts, and constrains deformations to a cyclic trajectory. These assumptions limit the degrees of freedom, and should be well suited to analyse motion patterns between, for instance, respiratory-correlated 4D CT, 4D cone-beam CT [37] or 4D magnetic resonance images [44].

6 Conclusion

We developed a spatio-temporal deformation model for deformable registration of respiratory-correlated images of the thorax. The model was obtained by extending spatial free-form deformations to the temporal domain, using a cyclic trajectory model based on cubic B-splines. A piecewise smooth temporal parametrization was found most suitable to account for the rapid changes in velocity at end-inhale. The spatio-temporal modelling resulted in a considerably more compact description of the deformation model. Spatio-temporal registration leads to comparable registration results while improving the robustness to artifacts.

Acknowledgement

The authors would like to thank Jan-Jakob Sonke of the Netherlands Cancer Institute-Antoni van Leeuwenhoek Hospital (Amsterdam, the Netherlands) for providing the cone-beam projection data, and Keelin Murphy of the Image Sciences Institute at the University of Utrecht (Utrecht, the Netherlands) for providing the sofware tool for landmark identification.

Jef Vandemeulebroucke was funded by the EC Marie Curie grant WARTHE, Jan Kybic was sponsored by the Czech Ministry of Education, Project MSM6840770012.

References

[1] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert. High accuracy optical flow estimation based on a theory for warping. In *Computer Vision - ECCV 2004*, pages 25–36, 2004.

[2] E. Castillo, R. Castillo, J. Martinez, M. Shenoy, and T. Guerrero. Four-dimensional deformable image registration using trajectory modeling. *Phys Med Biol*, 55(1):305–327, 2010.

[3] R. Chandrashekara, R. H. Mohiaddin, and D. Rueckert. Analysis of 3-D myocardial motion in tagged MR images using nonrigid image registration. *IEEE Trans Med Imaging*, 23(10):1245–1250, Oct 2004.

[4] P. Clarysse, C. Basset, L. Khouas, P. Croisille, D. Friboulet, C. Odet, and I. E. Magnin. Two-dimensional spatial and temporal displacement and deformation field fitting from cardiac magnetic resonance tagging. *Med Image Anal*, 4(3):253–268, Sep 2000.

[5] B. Delhay, P. Clarysse, C. Pera, and I. Magnin. A spatio-temporal deformation model for dense motion estimation in periodic cardiac image sequences. In *From Statistical Atlases to Personalized Models: Understanding Complex Diseases in Populations and Individuals, Satellite Workshop MICCAI 2006*, pages in–press, Copenhagen, Denmark, October 2006.

[6] E. C. Ford, G. S. Mageras, E. Yorke, and C. C. Ling. Respiration-correlated spiral CT: a method of measuring respiratory-induced anatomic motion for radiation treatment planning. *Med Phys*, 30(1):88–97, Jan 2003.

[7] R. George, S. S. Vedam, T. D. Chung, V. Ramakrishnan, and P. J. Keall. The application of the sinusoidal model to lung cancer patient respiratory motion. *Med Phys*, 32(9):2850–2861, Sep 2005.

[8] T. Guerrero, G. Zhang, W. Segars, T. Huang, S. Bilton, G. Ibbott, L. Dong, K. Forster, and K. Ping Lin. Elastic image mapping for 4-D dose estimation in thoracic radiotherapy. *Radiat Prot Dosimetry*, 115(1-4):497–502, 2005.

[9] T. Guerrero, K. Sanders, E. Castillo, Y. Zhang, L. Bidaut, T. Pan, and R. Komaki. Dynamic ventilation imaging from four-dimensional computed tomography. *Phys Med Biol*, 51(4):777–791, Feb 2006.

[10] J. Huang, D. Abendschein, V. G. Dvila-Romn, and A. A. Amini. Spatio-temporal tracking of myocardial deformations with a 4-D B-spline model from tagged MRI. *IEEE Trans Med Imaging*, 18(10):957–972, Oct 1999.

[11] P. Keall. 4-dimensional computed tomography imaging and treatment planning. *Semin Radiat Oncol*, 14(1):81–90, Jan 2004.

[12] M. L. Kessler. Image registration and data fusion in radiation therapy. *Br J Radiol*, 79 Spec No 1:S99–108, Sep 2006.

[13] S. Klein, M. Staring, and J. P. W. Pluim. Evaluation of optimization methods for nonrigid medical image registration using mutual information and B-splines. *IEEE Trans Image Process*, 16(12):2879–2890, Dec 2007.

[14] M. J. Ledesma-Carbayo, J. Kybic, M. Desco, A. Santos, M. Sühling, P. Hunziker, and M. Unser. Spatio-temporal nonrigid registration for ultrasound cardiac motion estimation. *IEEE Trans Med Imaging*, 24(9):1113–1126, Sep 2005.

[15] D. A. Low, M. Nystrom, E. Kalinin, P. Parikh, J. F. Dempsey, J. D. Bradley, S. Mutic, S. H. Wahab, T. Islam, G. Christensen, D. G. Politte, and B. R. Whiting. A method for the reconstruction of four-dimensional synchronized CT scans acquired during free breathing. *Med Phys*, 30(6):1254–1263, Jun 2003.

[16] W. Lu, G. H. Olivera, Q. Chen, M.-L. Chen, and K. J. Ruchala. Automatic re-contouring in 4D radiotherapy. *Phys Med Biol*, 51(5):1077–1099, Mar 2006.

[17] W. Lu, P. J. Parikh, J. P. Hubenschmidt, J. D. Bradley, and D. A. Low. A comparison between amplitude sorting and phase-angle sorting using external respiratory measurement for 4d ct. *Med Phys*, 33(8):2964–2974, Aug 2006.

[18] A. E. Lujan, E. W. Larsen, J. M. Balter, and R. K. T. Haken. A method for incorporating organ motion due to breathing into 3D dose calculations. *Med Phys*, 26(5):715–720, May 1999.

[19] J. R. McClelland, J. M. Blackall, S. Tarte, A. C. Chandler, S. Hughes, S. Ahmad, D. B. Landau, and D. J. Hawkes. A continuous 4D motion model from multiple respiratory cycles for use in lung radiotherapy. *Med Phys*, 33(9):3348–3358, Sep 2006.

[20] J. C. McEachen, A. Nehorai, and J. S. Duncan. Multiframe temporal estimation of cardiac nonrigid motion. *IEEE Trans Image Process*, 9(4):651–665, 2000.

[21] J. Modersitzki. *Numerical Methods for image Registration*. Oxford University Press, 2004.

[22] K. Murphy, B. van Ginneken, J. P. W. Pluim, S. Klein, and M. Staring. Semi-automatic reference standard construction for quantitative evaluation of lung CT registration. *Med Image Comput Comput Assist Interv*, 11(Pt 2):1006–1013, 2008.

[23] J. Nocedal. Updating quasi-newton matrices with limited storage. *Mathematics of computation*, 35(151):773–782, July 1980.

[24] J. R. Olsen, W. Lu, J. P. Hubenschmidt, M. M. Nystrom, P. Klahr, J. D. Bradley, D. A. Low, and P. J. Parikh. Effect of novel amplitude/phase binning algorithm on commercial four-dimensional computed tomography quality. *Int J Radiat Oncol Biol Phys*, 70(1): 243–252, Jan 2008.

[25] T. Pan, T.-Y. Lee, E. Rietzel, and G. T. Y. Chen. 4D-CT imaging of a volume influenced by respiratory motion on multi-slice CT. *Med Phys*, 31(2):333–340, Feb 2004.

[26] D. Perperidis, R. H. Mohiaddin, and D. Rueckert. Spatio-temporal free-form registration of cardiac MR image sequences. *Med Image Anal*, 9(5):441–456, Oct 2005.

[27] J. Peyrat, H. Delingette, M. Sermesant, C. Xu, and N. Ayache. Registration of 4D cardiac CT sequences under trajectory constraints with multichannel diffeomorphic demons. *IEEE Trans Med Imaging*, Mar 2010.

[28] S. Rit, D. Sarrut, and L. Desbat. Comparison of analytic and algebraic methods for motion-compensated cone-beam CT reconstruction of the thorax. *IEEE Trans Med Imaging*, Feb 2009.

[29] S. Rit, M. van Herk, L. Zijp, and J.-J. Sonke. Intrafraction and interfraction variability of the respiratory motion amplitude. In *Proceedings of the 28th Annual ESTRO Meeting*, Maastricht, The Netherlands, 2009.

[30] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans Med Imaging*, 18(8):712–721, Aug 1999.

[31] D. Sarrut. Deformable registration for image-guided radiation therapy. *Zeitschrift fr Medizinische Physik*, 16:285–297, 2006.

[32] D. Sarrut and J. Vandemeulebroucke. B-LUT: Fast and low memory B-spline image interpolation. *Comput Methods Programs Biomed*, 99:172–178, 2010.

[33] J. Schaerer, C. Casta, J. Pousin, and P. Clarysse. A dynamic elastic model for heart segmentation and motion estimation. *Submitted to Med Image Anal*, 2010.

[34] A. Schmidt-Richberg, J. Ehrhardt, R. Werner, and H. Handels. Slipping objects in image registration: improved motion field estimation with direction-dependent regularization. *Med Image Comput Comput Assist Interv*, 12(Pt 1):755–762, 2009.

[35] E. Schreibmann, B. Thorndyke, T. Li, J. Wang, and L. Xing. Four-dimensional image registration for image-guided radiotherapy. *Int J Radiat Oncol Biol Phys*, 71(2):578–586, Jun 2008.

[36] Y. Seppenwoolde, H. Shirato, K. Kitamura, S. Shimizu, M. van Herk, J. V. Lebesque, and K. Miyasaka. Precise and real-time measurement of 3D tumor motion in lung due to breathing and heartbeat, measured during radiotherapy. *Int J Radiat Oncol Biol Phys*, 53(4):822–834, Jul 2002.

[37] J. Sonke, L. Zijp, P. Remeijer, and M. van Herk. Respiratory correlated cone beam CT. *Med Phys*, 32(4):1176–1186, Apr 2005.

[38] J.-J. Sonke, J. Lebesque, and M. van Herk. Variability of four-dimensional computed tomography patient models. *Int J Radiat Oncol Biol Phys*, 70(2):590–598, Feb 2008.

[39] H. Sundar, H. Litt, and D. Shen. Estimating myocardial motion by 4D image warping. *Pattern Recognition*, 42(11):2514 – 2526, 2009. ISSN 0031-3203.

[40] M. Unser. Splines: A perfect fit for signal and image processing. *IEEE Signal Processing Magazine*, 16(6):22–38, November 1999.

[41] J. Vandemeulebroucke, O. Bernard, J. Kybic, P. Clarysse, and D. Sarrut. Automatic motion mask extraction for deformable registration of the lungs. In *XVIth International Conference on the Use of Computers in Radiation Therapy*, Amsterdam, June 2010.

[42] S. S. Vedam, P. J. Keall, V. R. Kini, H. Mostafavi, H. P. Shukla, and R. Mohan. Acquiring a four-dimensional computed tomography dataset using an external respiratory signal. *Phys Med Biol*, 48(1):45–62, Jan 2003.

[43] D. Visvikis, M. Ledesma-Carbayo, F. Lamare, O. Mawlawi, P. Jarritt, P. Bruyant, A. Santos, G. Kontaxakis, N. Boussion, and C. Cheze Le Rest. A spatiotemporal image registration algorithm for respiratory motion correction in PET/CT. *J Nucl Med, Meeting Abstracts*, 47(suppl 1):234P–a–, 2006.

[44] M. von Siebenthal, G. Székely, U. Gamper, P. Boesiger, A. Lomax, and P. Cattin. 4D MR imaging of respiratory organ motion and its variability. *Phys Med Biol*, 52(6):1547–1564, Mar 2007.

[45] J. W. H. Wolthaus, J. J. Sonke, M. van Herk, and E. M. F. Damen. Reconstruction of a time-averaged midposition CT scan for radiotherapy planning of lung cancer patients using deformable registration. *Med Phys*, 35(9):3998–4011, Sep 2008.

[46] H. Wu, G. C. Sharp, Q. Zhao, H. Shirato, and S. B. Jiang. Statistical analysis and correlation discovery of tumor respiratory motion. *Phys Med Biol*, 52(16):4761–4774, Aug 2007.

[47] Z. Wu, E. Rietzel, V. Boldea, D. Sarrut, and G. C. Sharp. Evaluation of deformable registration of patient lung 4DCT with subanatomical region segmentations. *Med Phys*, 35(2):775–781, Feb 2008.

[48] T. Yamamoto, U. Langner, B. W. Loo, J. Shen, and P. J. Keall. Retrospective analysis of artifacts in four-dimensional CT images of 50 abdominal and thoracic radiotherapy patients. *Int J Radiat Oncol Biol Phys*, 72(4):1250–1258, Nov 2008.

[49] Q. Zhang, A. Pevsner, A. Hertanto, Y.-C. Hu, K. E. Rosenzweig, C. C. Ling, and G. S. Mageras. A patient-specific respiratory model of anatomical motion for radiation treatment planning. *Med Phys*, 34(12):4772–4781, Dec 2007.

[50] T. Zhang, N. P. Orton, and W. A. Tom. On the automated definition of mobile target volumes from 4D-CT images for stereotactic body radiotherapy. *Med Phys*, 32(11):3493–3502, Nov 2005.

[51] L. Zijp, J. Sonke, and M. van Herk. Extraction of the respiratory signal from sequential thorax cone-beam X-ray images. In *14th International Conference on the Use of Computers in Radiation Therapy, Seoul, Korea*, pages 507–509, May 2004.