

Real-time Motion Compensated Patient Positioning and Non-rigid Deformation Estimation using 4-D Shape Priors

Jakob Wasza¹, Sebastian Bauer¹, Joachim Hornegger^{1,2}

¹ Pattern Recognition Lab, Department of Computer Science

² Erlangen Graduate School in Advanced Optical Technologies (SAOT)

Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

{jakob.wasza,sebastian.bauer,jochim.hornegger}@cs.fau.de

Abstract. Over the last years, range imaging (RI) techniques have been proposed for patient positioning and respiration analysis in motion compensation. Yet, current RI based approaches for patient positioning employ rigid-body transformations, thus neglecting free-form deformations induced by respiratory motion. Furthermore, RI based respiration analysis relies on non-rigid registration techniques with run-times of several seconds. In this paper we propose a real-time framework based on RI to perform respiratory motion compensated positioning and non-rigid surface deformation estimation in a joint manner. The core of our method are pre-procedurally obtained 4-D shape priors that drive the intra-procedural alignment of the patient to the reference state, simultaneously yielding a rigid-body table transformation and a free-form deformation accounting for respiratory motion. We show that our method outperforms conventional alignment strategies by a factor of 3.0 and 2.3 in the rotation and translation accuracy, respectively. Using a GPU based implementation, we achieve run-times of 40 ms.

1 Introduction

Accurate patient positioning and respiratory motion analysis are key issues for the success of medical procedures. For example, in fractionated radiotherapy, the patient must be aligned with respect to planning data and continuously monitored to account for respiratory motion and spontaneous movements [1]. Recently, techniques based on real-time range imaging (RI) have been proposed for positioning and respiration analysis [2–6]. Compared to commonly employed imaging techniques, RI sensors are marker-less, non-intrusive and do not imply radiation exposure. However, current RI based positioning systems employ rigid registration techniques neglecting free-form deformations induced by respiratory motion. In this context, recent work reported an error scale of non-rigid motion up to 25 mm [5, 6]. Another motivation for non-rigid registration of respiration-induced surface deformations arises from observations that an analysis of multiple surface regions allows for an improved prediction of internal organ movement

compared to a single external surrogate [7]. Apart from applications in radio-therapy, the prediction of internal movement from external surrogates also holds potential for statistical lung motion modelling [8] or 4-D CT reconstruction and sorting [9]. However, current surrogates based on RI techniques rely on heuristic surface partitioning strategies based on points and patches [7] or manually selected bounding boxes of anatomical regions [4]. Though dense non-rigid surface registration techniques to generate respiration surrogates were recently proposed [6, 10], run-times of several seconds are not acceptable in clinical practice.

In this paper we propose a framework that by design allows to perform respiratory motion compensated patient positioning and non-rigid surface deformation estimation in a joint manner. Our method is based on pre-procedurally obtained patient-specific 4-D shape priors derived from statistical analysis of non-rigidly registered 3-D RI patient surface data from different respiration states. The intra-procedural alignment of the patient to the reference state is then driven by the shape priors, simultaneously yielding a rigid-body table transform and a free-form deformation accounting for respiratory motion. With real-time constraints in mind, our framework outsources the computationally expensive task of non-rigid surface registration to a pre-procedural training phase. Intra-procedurally, this allows to employ real-time algorithms taking advantage of the pre-procedurally obtained priors. In particular, we designed our method to support parallel computation on many-core architectures such as GPUs.

2 Method

Our method relies on RI devices that deliver dense and metric surface information of the captured body in real-time. Clinically available systems are for example AlignRT (Vision RT, London, UK) or Catalyst (C-RAD AB, Uppsala, Sweden). We denote $\mathbf{x}(\xi_i) \in \mathbb{R}^3$ the world coordinate associated with ξ_i in a 2-D sampling domain $\Omega \subset \mathbb{R}^2$ discretized with $N \times M$ pixels. We denote a point cloud or surface as $\mathcal{S} = \{\mathbf{x}(\xi_1), \dots, \mathbf{x}(\xi_{N \cdot M})\}$ which can be linearized as:

$$\mathcal{S} \equiv \mathbf{s} = \left[\mathbf{x}(\xi_1)^\top, \dots, \mathbf{x}(\xi_{N \cdot M})^\top \right]^\top \in \mathbb{R}^{3NM}. \quad (1)$$

We now briefly explain the general steps of our method. Pre-procedurally, we non-rigidly register RI surfaces $\{\mathcal{S}_1, \dots, \mathcal{S}_T\}$ acquired at different respiration states $t = 1 \dots T$ to a reference surface \mathcal{S}_{Ref} . The reference can be acquired using RI sensors or imported from volumetric planning data by (i) segmenting the body from the background, (ii) isosurface extraction, and (iii) rendering the surface mesh to a z-buffer representation and sample it from Ω . We obtain T displacement fields $\mathcal{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_T\}$, $\mathbf{u}_t \in \mathbb{R}^{3NM}$ that are used to build a deformable model $\mathcal{M}(\mathbf{b})$ representing the priors of our method. Here, \mathbf{b} denotes a parameter vector to control the model. The intra-procedural respiratory motion compensated alignment of the patient to the reference state is then computed by finding a rigid-body transformation $(\hat{\mathbf{R}}, \hat{\mathbf{t}})$ with rotation $\hat{\mathbf{R}}$ and translation $\hat{\mathbf{t}}$ and the parameters $\hat{\mathbf{b}}$ such that the corresponding model instance $\mathcal{M}(\hat{\mathbf{b}})$ fits the patients instantaneous state. An illustration is given in Fig. 1.

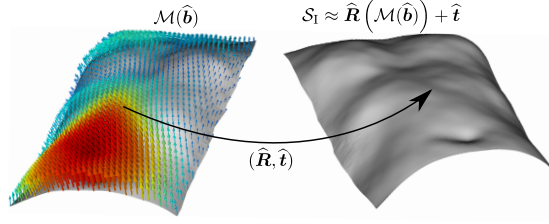


Fig. 1. Respiratory motion compensated patient alignment. The reference state (left gray shape) is deformed according to $\mathcal{M}(\hat{\mathbf{b}})$ and transformed by $(\hat{\mathbf{R}}, \hat{\mathbf{t}})$ to fit the instantaneous state \mathcal{S}_t (right shape). The displacement field corresponding to $\mathcal{M}(\hat{\mathbf{b}})$ is color coded with red tones denoting large and blue tones denoting small displacements.

2.1 Non-rigid Surface Registration

Training a deformable surface model requires a set of displacement fields \mathcal{U} that describe the elastic deformation ϕ_t matching a patient reference surface \mathcal{S}_{Ref} and patient surface data \mathcal{S}_t captured at different respiration states t . We represent the deformation ϕ_t by a displacement field \mathbf{u}_t defined on Ω with

$$\phi_t(\mathbf{x}(\boldsymbol{\xi})) = \mathbf{x}(\boldsymbol{\xi}) + \mathbf{u}_t(\boldsymbol{\xi}), \quad \mathbf{x}(\boldsymbol{\xi}) \in \mathcal{S}_{\text{Ref}}. \quad (2)$$

Now, the goal is to estimate ϕ_t in a sense that $\phi_t(\mathcal{S}_{\text{Ref}}) \approx \mathcal{S}_t$. For this purpose, we represent \mathcal{S}_t at time t by its corresponding signed distance function $d_t(\mathbf{x}) := \pm \text{dist}(\mathbf{x}, \mathcal{S}_t)$, where the sign is positive outside the body and negative inside. Further, $\nabla d_t(\mathbf{x})$ is the outward pointing normal on \mathcal{S}_t and $|\nabla d_t(\mathbf{x})| = 1$. Based on $d_t(\mathbf{x})$, we can define the projection $P(\mathbf{x}) := \mathbf{x} - d_t(\mathbf{x})\nabla d_t(\mathbf{x})$ of a point $\mathbf{x} \in \mathbb{R}^3$ in a neighborhood of \mathcal{S}_t onto the closest point on \mathcal{S}_t . Thus, we quantify the closeness of a displaced reference surface point $\phi_t(\mathbf{x}), \mathbf{x} \in \mathcal{S}_{\text{Ref}}$ to \mathcal{S}_t using $|P(\phi_t(\mathbf{x})) - \phi_t(\mathbf{x})| = |d_t(\phi_t(\mathbf{x}))|$ as a point-wise measure. Based on this closeness measure, we use a variational formulation for estimating \mathbf{u}_t as a minimizer of the functional

$$\mathcal{E}[\mathbf{u}_t] = \mathcal{E}_{\text{match}}[\mathbf{u}_t] + \alpha \mathcal{E}_{\text{reg}}[\mathbf{u}_t] = \int_{\Omega} \left(d_t(\mathbf{x}(\boldsymbol{\xi}) + \mathbf{u}_t(\boldsymbol{\xi}))^2 + \alpha \|D\mathbf{u}_t(\boldsymbol{\xi})\|_F^2 \right) d\boldsymbol{\xi} \quad (3)$$

where $D\mathbf{u}(\boldsymbol{\xi})$ denotes the Jacobian of $\mathbf{u}(\boldsymbol{\xi})$ and α the regularization weight. The matching term $\mathcal{E}_{\text{match}}$ ensures that $\phi_t(\mathcal{S}_{\text{Ref}}) \approx \mathcal{S}_t$. As a smoothness prior, we took a quadratic regularization term on the Jacobian of the displacement into account. For numerical minimization of Eq. 3, we considered a Finite Element approximation on a uniform rectangular $N \times M$ grid covering Ω for spatial discretization and applied a multi-scale gradient descent scheme, c.f. [10].

2.2 Deformable Model Generation

Based on the patient-specific set of non-rigidly registered surfaces \mathcal{S}_t we apply a principal component analysis (PCA) to the training set $\mathcal{V} = \{\mathbf{s}_{\text{Ref}}, \mathbf{v}_1, \dots, \mathbf{v}_T\}$

with $\mathbf{v}_t = \mathbf{s}_{\text{Ref}} + \mathbf{u}_t$. The $P \leq T$ eigenvectors \mathbf{e}_i with the largest non-zero eigenvalues of the covariance matrix of the centered version of \mathcal{V} define the model's modes of variation as

$$\Phi = [\mathbf{e}_1, \dots, \mathbf{e}_P] \in \mathbb{R}^{3NM \times P}. \quad (4)$$

A linear combination of the P principal modes of variation spans a subset of RI surfaces \mathcal{S} composed of the given modes of variation as

$$\mathcal{S} \equiv \mathcal{M}(\mathbf{b}) = \bar{\mathbf{v}} + \Phi \mathbf{b}, \quad \bar{\mathbf{v}} = \frac{1}{T+1} \left(\mathbf{s}_{\text{Ref}} + \sum_{t=1}^T \mathbf{v}_t \right). \quad (5)$$

Here, $\mathbf{b} \in \mathbb{R}^P$ is a parameter vector holding the coefficients for the modes of variation, thus accounting for the model's inherent free-form deformations.

2.3 Respiratory Motion Compensated Patient Alignment

The intra-procedural alignment of the patient is performed by fitting the deformable model \mathcal{M} to the patient's instantaneous respiration state \mathcal{S}_I :

$$\hat{\mathbf{R}}, \hat{\mathbf{t}}, \hat{\mathbf{b}} = \underset{\mathbf{R}, \mathbf{t}, \mathbf{b}}{\operatorname{argmin}} \operatorname{dist}(\mathcal{S}_I, \mathbf{R}(\bar{\mathbf{v}} + \Phi \mathbf{b}) + \mathbf{t}). \quad (6)$$

The rotation matrix $\hat{\mathbf{R}}$ and translation vector $\hat{\mathbf{t}}$ define the rigid-body table transformation whereas the model parameter vector $\hat{\mathbf{b}}$ accounts for non-rigid deformations induced by respiratory motion. Furthermore, $\operatorname{dist}(\mathcal{S}_i, \mathcal{S}_j)$ quantifies the distance between two surfaces and may for example denote a point-to-point or point-to-plane measure. In this work we employ a model fitting strategy derived from the *iterative closest point* (ICP) algorithm [11]. In each iteration k , the current model instance \mathcal{S}_k is aligned to the instantaneous surface \mathcal{S}_I via a rigid-body transformation $(\mathbf{R}_k, \mathbf{t}_k)$ estimated by a closest-point relationship. Based on the estimated transformation, the model's closest points \mathbf{Y}_k on the instantaneous surface \mathcal{S}_I are then projected onto the model basis Φ to update the parameters \mathbf{b}_k . See Algorithm 1 for details.

A benefit of this model fitting scheme is the inherent high degree of parallelism in each iteration of Algorithm 1, thus allowing for a real-time capable implementation on many-core systems such as GPUs. In particular, we employ the recently proposed *random ball cover* for efficient closest point search [12].

2.4 Non-rigid Deformation Estimation using 4-D Shape Priors

By design, our method allows for efficient estimation of respiration induced deformations. For two surfaces $(\mathcal{S}_i, \mathcal{S}_j)$ and its associated estimated model coefficients $(\mathbf{b}_i, \mathbf{b}_j)$, the estimation of a dense displacement field $\mathbf{u}_{i,j}$ between these surfaces breaks down to a linear mapping as $\mathbf{u}_{i,j} = \Phi(\mathbf{b}_i - \mathbf{b}_j)$. This follows directly from Eq. (5). We further note that the computational complexity of non-rigid deformation estimation is therefore decoupled from the actual non-rigid registration technique creating the priors contained in Φ . Inherently, this allows to employ computational expensive non-rigid registration techniques as proposed by Schaerer et al. [6] or Bauer et al. [10] in the model generation stage.

Algorithm 1 Algorithm for Model Fitting

Input: Instantaneous surface \mathcal{S}_1 and deformable model $\bar{\mathbf{v}}, \Phi$
Input: Initial guess for table transformation $\mathbf{R}_0, \mathbf{t}_0$ and model parameters \mathbf{b}_0
for $k = 1, \dots, N_{\text{iter}}$ **do**
 $\mathcal{S}_k = \mathbf{R}_{k-1}(\bar{\mathbf{v}} + \Phi \mathbf{b}_{k-1}) + \mathbf{t}_{k-1}$
 $\mathbf{Y}_k \leftarrow \text{closestPoints}(\mathcal{S}_k, \mathcal{S}_1)$
 $(\mathbf{R}_k, \mathbf{t}_k) \leftarrow \text{estimateTransformation}(\mathcal{S}_k, \mathbf{Y}_k, \mathbf{R}_{k-1}, \mathbf{t}_{k-1})$
 $\mathbf{b}_k = \Phi^\top (\mathbf{R}_k^{-1}(\mathbf{Y}_k - \mathbf{t}_k) - \bar{\mathbf{v}})$
end for
Output: Estimated table transformation $(\hat{\mathbf{R}}, \hat{\mathbf{t}})$ and model coefficients $\hat{\mathbf{b}}$

3 Evaluation and Results

For our experiments we captured RI data from six subjects using a structured light sensor (30 Hz, 640×480 px) in a clinical environment. RI data were enhanced using edge-preserving filtering, cropped to a region of interest covering the subject and re-sampled in the sampling domain Ω to a resolution of 128×128 pixels. For model generation, the subjects were instructed to perform thoracic and abdominal breathing subsequently. For each breathing mode we then extracted surface data \mathcal{S}_t from $T = 8$ phases covering one respiration cycle. We found that more cycles did not improve accuracy. For all subjects, the body surface at full abdominal expiration was chosen as the reference \mathcal{S}_{Ref} for non-rigid registration of the remaining 15 shapes to form the training set \mathcal{V} . PCA was then performed on the training set \mathcal{V} with the number of modes P chosen such that 99% of the input variance is explained. This resulted in four variation modes for all subjects. To validate the extrapolation ability of our model to unseen data, the subjects were asked to perform regular breathing over several respiration cycles. Surfaces for validation were generated from 32 RI frames sampled every 10^{th} frame, starting at an arbitrary phase.

Patient Positioning The accuracy of our method for respiratory motion compensated patient positioning is assessed by evaluating the difference between the estimated table transforms $(\mathbf{R}_i, \mathbf{t}_i)$ for validation frame i and the corresponding ground truth (GT) transform $\mathbf{R}_{\text{GT}} \equiv \{r_{\text{GT}}^1, r_{\text{GT}}^2, r_{\text{GT}}^3\}$ and $\mathbf{t}_{\text{GT}} \equiv \{t_{\text{GT}}^1, t_{\text{GT}}^2, t_{\text{GT}}^3\}$. The superscripts 1, 2, 3 denote rotation around or translation along the x, y and z axis, respectively. The GT transform was derived from the identity transform as the subjects did not change position between the training and testing phase. For each subject, the average rotation error Δr_i and translation error Δt_i for validation frame i is computed according to:

$$\Delta r_i = \frac{1}{3} \sum_{j=1}^3 |r_{\text{GT}}^j - r_i^j|, \quad \Delta t_i = \frac{1}{3} \sum_{j=1}^3 |t_{\text{GT}}^j - t_i^j|. \quad (7)$$

For comparison, we oppose our method to a conventional ICP-based rigid alignment strategy as used in [5]. For both methods, the initial estimate for the table

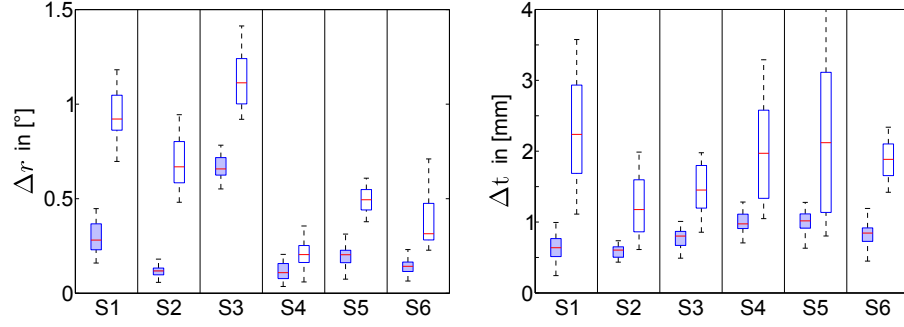


Fig. 2. Table transform error of our method (shaded) compared to conventional ICP-based alignment (not shaded) for subjects S1-S6. The left figure depicts the mean rotation error Δr in $^\circ$, the right figure the mean translation error Δt in [mm] compared to the ground truth transform.

transform was set to 10 mm and 5° off from the GT position. Fig. 2 shows quantitative results for the average rotation and translation error over all 32 validation frames. Note that our motion compensated positioning method significantly reduces that table transform error. Over all subjects, the rotation and translation error decreases by a factor of 3.0 and 2.3, respectively.

Deformation Estimation The capability of our method for non-rigid deformation estimation is assessed by computing the absolute error in terms of surface mismatch between the transformed estimated model instance and the instantaneous surface as $|dist(\mathbf{R}(\mathcal{M}(\mathbf{b})) + \mathbf{t}, \mathcal{S}_I)|$ using a point-to-plane measure. In order to account for boundary effects such as edges, we restricted the evaluation to a center region on the subjects surface. For comparison, we again oppose our method to a conventional ICP-based rigid alignment strategy. Fig. 3 exemplarily illustrates the surface mismatch for subject S1. In Fig. 4 left, the surface mismatch is shown for subject S4 and a subset of validation frames. The right plot in Fig. 4 depicts the surface mismatch for the individual subjects over all validation frames. We note the large residual displacements with ICP-based alignment that can be reduced significantly with our model-based deformation estimation. Over all subjects, our method reduces the surface mismatch by a factor of 1.9.

Performance Evaluation We implemented the model fitting routine as described in Algorithm 1 on an NVIDIA GTX570 GPU using the CUDA architecture. For performance reasons, we employed a sparse fitting scheme, e.g. a subset of uniformly sampled points in the model \mathcal{M} as well as the instantaneous surface \mathcal{S}_I where used. For $4 \cdot 10^3$ model points, $1.6 \cdot 10^4$ surface points and $N = 100$ iterations that proved to be sufficient to achieve the accuracy reported above, our method runs at 40 ms.

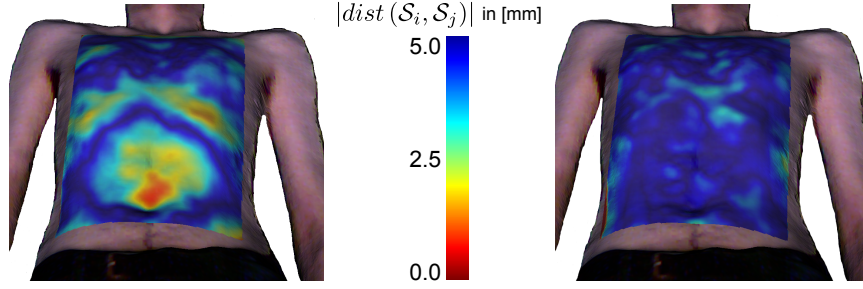


Fig. 3. Color-coded surface mismatch for ICP-based alignment (left) and our method (right). Average misalignment error is 1.3 mm with ICP and 0.5 mm with our method.

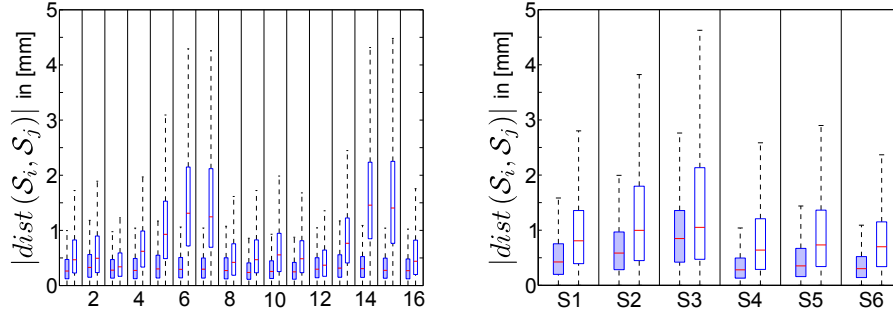


Fig. 4. Surface mismatch for our method (shaded) and ICP-based alignment (not shaded). The left figure shows the surface mismatch for an individual subject over several frames. The right figure depicts the surface mismatch for the individual subjects over all frames.

4 Conclusion

In this paper we have proposed a real-time capable range imaging based framework for joint respiratory motion compensated patient positioning and non-rigid surface deformation estimation. Both tasks are achieved by the main contribution of this paper, namely to employ 4-D patient-specific shape priors obtained from statistical analysis of non-rigidly registered surfaces. In experiments we showed that, in comparison to conventional positioning strategies, our method reduces the average alignment error by a factor of 3.0 and 2.3 for the rotational and translational components, respectively. We further demonstrated that a GPU-based implementation of our method allows to estimate dense surface deformations at a run-time of 40 ms.

Acknowledgments We gratefully acknowledge the support by the European Regional Development Fund (ERDF) and the Bayerisches Staatsministerium für Wirtschaft, Infrastruktur, Verkehr und Technologie (StMWIVT), in the context of the R&D program IuK Bayern under Grant No. IUK338.

References

1. Bert, C., Durante, M.: Motion in radiotherapy: particle therapy. *Phys Med Biol* **56**(16) (2011) R113–R144
2. Schaller, C., Rohkohl, C., Penne, J., Stürmer, M., Hornegger, J.: Inverse C-arm Positioning for Interventional Procedures Using Real-Time Body Part Detection. In Yang, G.Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C., eds.: *MICCAI 2009, Part I*. Volume 5761 of LNCS., Springer Berlin / Heidelberg (2009) 549–556
3. Fayad, H., Pan, T., Roux, C., Le Rest, C., Pradier, O., Clement, J., Visvikis, D.: A patient specific respiratory model based on 4D CT data and a time of flight camera (TOF). In: *IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*. (2009) 2594–2598
4. Hughes, S., McClelland, J., Tarte, S., Lawrence, D., Ahmad, S., Hawkes, D., Landau, D.: Assessment of two novel ventilatory surrogates for use in the delivery of gated/tracked radiotherapy for non-small cell lung cancer. *Radiother Oncol* **91**(3) (2009) 336 – 341
5. Placht, S., Stancanella, J., Schaller, C., Balda, M., Angelopoulou, E.: Fast time-of-flight camera based surface registration for radiotherapy patient positioning. *Med Phys* **39**(1) (2012) 4–17
6. Schaerer, J., Fassi, A., Riboldi, M., Cerveri, P., Baroni, G., Sarrut, D.: Multi-dimensional respiratory motion tracking from markerless optical surface imaging based on deformable mesh registration. *Phys Med Biol* **57**(2) (2012) 357–373
7. Fayad, H., Pan, T., Clement, J.F., Visvikis, D.: Technical note: Correlation of respiratory motion between external patient surface and internal anatomical landmarks. *Med Phys* **38**(6) (2011) 3157–3164
8. Klinder, T., Lorenz, C., Ostermann, J.: Prediction framework for statistical respiratory motion modeling. In Jiang, T., Navab, N., Pluim, J., Viergever, M., eds.: *MICCAI 2010, Part III*. Volume 6363 of LNCS., Springer Berlin / Heidelberg (2010) 327–334
9. Gianoli, C., Riboldi, M., Spadea, M.F., Travaini, L.L., Ferrari, M., Mei, R., Orecchia, R., Baroni, G.: A multiple points method for 4D CT image sorting. *Med Phys* **38**(2) (2011) 656–667
10. Bauer, S., Berkels, B., Hornegger, J., Rumpf, M.: Joint ToF image denoising and registration with a CT surface in radiation therapy. In Bruckstein, A., ter Haar Romeny, B., Bronstein, A., Bronstein, M., eds.: *Scale Space and Variational Methods in Computer Vision*. Volume 6667 of LNCS., Springer Berlin / Heidelberg (2011) 98–109
11. Besl, P., McKay, N.: A method for registration of 3-D shapes. *IEEE Trans Pattern Anal Mach Intell* **14**(2) (1992) 239–256
12. Cayton, L.: Accelerating nearest neighbor search on manycore systems. *Computing Research Repository* **abs/1103.2635** (2011)