Marker-less Reconstruction of Dense 4-D Surface Motion Fields using Active Laser Triangulation for Respiratory Motion Management

Sebastian Bauer[1], Benjamin Berkels[3], Svenja Ettl[2], Oliver Arold[2], Joachim Hornegger[1], Martin Rumpf[4]

[1] Pattern Recognition Lab, Dept. of Computer Science
[2] Institute of Optics, Information and Photonics
Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
`sebastian.bauer@cs.fau.de`
[3] Interdisciplinary Mathematics Institute,
University of South Carolina, Columbia, SC, USA
[4] Institute for Numerical Simulation,
Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn, Germany

Abstract. To manage respiratory motion in image-guided interventions a novel sparse-to-dense registration approach is presented. We apply an emerging laser-based active triangulation (AT) sensor that delivers sparse but highly accurate 3-D measurements in real-time. These sparse position measurements are registered with a dense reference surface extracted from planning data. Thereby a dense displacement field is reconstructed which describes the 4-D deformation of the complete patient body surface and recovers a multi-dimensional respiratory signal for application in respiratory motion management. The method is validated on real data from an AT prototype and synthetic data sampled from dense surface scans acquired with a structured light scanner. In a study on 16 subjects, the proposed algorithm achieved a mean reconstruction accuracy of ± 0.22 mm w.r.t. ground truth data.

1 Introduction

Respiration-synchronized image-guided radiation therapy (IGRT) techniques aim at tracking the tumor location and reposition the beam dynamically. To reduce additional radiation exposure, recent hybrid tumor-tracking techniques combine episodic radiographic imaging with continuous monitoring of external breathing surrogates based on the premise that the internal tumor position can be accurately predicted from external motion. The underlying correlation model can be established from a series of simultaneously acquired external-internal position measurements [1] or 4-D CT planning data [2]. Clinically available solutions for hybrid tumor-tracking [1,3] measure external motion using a single or a few passive markers on the patient's chest as a low-dimensional surrogate. Thus, these techniques are incapable of depicting the full complexity of respiratory motion, they involve extensive patient preparation, and require reproducible

marker placement with a substantial impact on model accuracy. Modern IGRT solutions that allow to monitor the motion of the complete external patient surface help to reduce correlation model uncertainties. In particular, range imaging (RI) technologies can acquire a dense 3-D surface model of the patient [4–6]. Based on the estimation of a dense displacement field representing the deformation of the instantaneous torso surface w.r.t. a reference surface (either from RI or planning CT data), a highly accurate correlation model can be established [7, 8]. The deformation estimation from dense surface scans for application in RT has been investigated recently [8, 9]. Available RI-based IGRT solutions are capable of delivering dense surface information in a marker-less manner but focus on patient positioning, do not support dense sampling in real-time [4, 5] or at the cost of a limited field of view [6], often imply high costs in terms of hardware and are subject to measurement uncertainties due to the sampling principles e.g. active stereo [6] or swept lasers [4, 5]. The temporal resolution of these solutions may be insufficient to characterize respiratory motion. In this paper, we propose a marker-less system based on a non-moving active laser triangulation (AT) sensor that delivers sparse but accurate measurements in real-time (30 Hz). Using prior patient shape knowledge from planning data, a variational model is proposed to recover a dense and accurate displacement field and to reconstruct a complete and reliable patient surface model at the instantaneous respiration phase. Estimating the dense deformation is combined with recovering a sparse displacement field from AT measurements to planning data, thus the approach is closely related to the field of inverse-consistent registration [10, 11]. The variational model is discretized using Finite Elements, the optimization is guided by a step-size controlled gradient flow to guarantee fast and smooth relaxation.

2 Method

In this section, we derive the variational model for the reconstruction of a dense displacement field from sparse measurements. Given is a reference shape $\mathcal{G} \subset \mathbb{R}^3$ extracted from planning data and the instantaneous body surface $\mathcal{M}$ represented by a sparse sampling Y. For instance, let us assume that the AT sensor acquires a set of n measurements $Y = \{y_1, \ldots, y_n\}$, $y_i \in \mathbb{R}^3$, arranged in a grid-like structure (Fig. 1). We assume that $\mathcal{G}$ is given as a graph, i.e. there is a domain $\Omega \subset \mathbb{R}^2$ usually associated with the plane of the patient table and a function $g : \Omega \to \mathbb{R}$ such that $\mathcal{G} = \big\{(\zeta, g(\zeta)) \in \mathbb{R}^3 : \zeta \in \Omega\big\}$. Due to respiration, the intra-fractional sampling Y is not aligned with $\mathcal{G}$. Now, the goal is to estimate the unknown, non-rigid, dense deformation ϕ of $\mathcal{G}$ with $Y \subset \phi(\mathcal{G})$. For this purpose, in a joint manner, we estimate ϕ together with an inverse deformation ψ matching Y and $\mathcal{G}$ in the sense that $\psi(Y) \subset \mathcal{G}$. When registering Y onto $\mathcal{G}$ we solely deal with a sparse displacement field $(\psi(y_i))_{i=1,\ldots,n}$ on the n positions measured by the AT sensor. A geometric sketch of the registration configuration is depicted in Fig. 1. Estimating ψ allows us to establish a correspondence between the AT measurements and the reference patient surface, whereas the dense deformation ϕ enables the reconstruction of the complete instantaneous patient surface. We represent
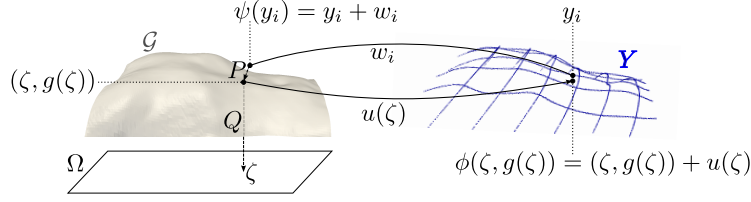
Fig. 1. Geometric configuration for reconstructing the dense deformation ϕ with $\phi(\zeta, g(\zeta)) = (\zeta, g(\zeta)) + u(\zeta)$ from sparse sampling data $Y = \{y_1, \ldots, y_n\}$ and the approximate sparse inverse ψ with $\psi(y_i) = y_i + w_i$ (for a better visibility $\mathcal{G}$ and Y have been pulled apart). Furthermore, the projection P onto $\mathcal{G}$ and the orthogonal projection Q from the graph $\mathcal{G}$ onto the parameter domain Ω are sketched.

ψ by a vector of displacements $W = \{w_1, \ldots, w_n\}$ with $\psi(y_i) = y_i + w_i$. Furthermore, the deformation ϕ is represented by a displacement $u : \Omega \to \mathbb{R}^3$ defined on the parameter domain Ω of the graph $\mathcal{G}$ with $\phi(\zeta, g(\zeta)) = (\zeta, g(\zeta)) + u(\zeta)$. To quantify the matching of $\psi(Y)$ onto $\mathcal{G}$ let us assume that the signed distance function d with respect to $\mathcal{G}$ is precomputed in a sufficiently large neighborhood in $\mathbb{R}^3$. We set $d(x) := \pm \mathrm{dist}(x, \mathcal{G})$, where the sign is positive outside the body, i.e. above the graph, and negative inside. Then $\nabla d(x)$ is the outward pointing normal on $\mathcal{G}$ and $|\nabla d(x)| = 1$. Based on this signed distance map d we can define the projection $P(x) := x - d(x) \nabla d(x)$ of a point $x \in \mathbb{R}^3$ in a neighborhood of $\mathcal{G}$ onto the closest point on $\mathcal{G}$ and compute the mismatch of $\psi(Y)$ and $\mathcal{G}$ pointwise via $|P(\psi(y_i)) - \psi(y_i)| = |d(y_i + w_i)|$. Let us emphasize that we do not expect ψ to be equal to the projection P. Indeed, the computational results discussed below underline that it is the prior in the deformation ϕ which leads to general matching correspondences for a minimizer of our variational approach.

2.1 Definition of the Registration Energy

Now, we define a functional $\mathcal{E}$ on dense displacement fields u and sparse vectors of displacements W such that a minimizer represents a suitable matching of the planning data and AT measurements:

$$\mathcal{E}[u, W] := \mathcal{E}_{\mathrm{match}}[W] + \kappa \, \mathcal{E}_{\mathrm{con}}[u, W] + \lambda \, \mathcal{E}_{\mathrm{reg}}[u] \tag{1}$$

where κ and λ are nonnegative constants controlling the contributions of the individual terms. $\mathcal{E}_{\mathrm{match}}$ denotes a term measuring closeness of $\psi(Y)$ to $\mathcal{G}$. The consistency functional $\mathcal{E}_{\mathrm{con}}$ is responsible for establishing the relation between both displacement fields. Finally, $\mathcal{E}_{\mathrm{reg}}$ ensures a regularization of the dense displacement u. The detailed definitions of these functionals are as follows.

Matching Energy. In order to measure closeness of $\psi(Y)$ to $\mathcal{G}$, we use the pointwise mismatch measure discussed above and define

$$\mathcal{E}_{\mathrm{match}}[W] := \frac{1}{2n} \sum_{i=1}^{n} |d(y_i + w_i)|^2 \,. \tag{2}$$

Consistency Energy. For a known instantaneous deformation ϕ of the patient surface $\mathcal{G}$ and an exact deformation correspondence $\psi(Y) \subset \mathcal{G}$ of the AT measurement Y the identity $\phi(\psi(Y)) = Y$ holds. But for an arbitrary deformation ψ described by some vector of displacements W in general $\psi(Y) \not\subset \mathcal{G}$. To relate ϕ and ψ in this case we have to incorporate the projection P because ϕ is only defined on $\mathcal{G}$. In fact, to ensure that $(\phi \circ P \circ \psi)(W) \approx W$ for a minimizer of the total energy we introduce the consistency energy

$$\mathcal{E}_{\mathrm{con}}[u, W] := \frac{1}{2n} \sum_{i=1}^{n} |P(y_i + w_i) + u(Q\,P(y_i + w_i)) - y_i|^2 , \qquad (3)$$

where $Q \in \mathbb{R}^{2 \times 3}$ denotes the orthographic projection matrix with $Q(\zeta, g(\zeta)) = \zeta$. Here, we have used that $\phi(P(\psi(y_i))) = P(y_i + w_i) + u(Q\,P(y_i + w_i))$. Indeed, this definition of the consistency energy allows us to compute a dense smooth displacement of the patient planning surface even though only a sparse set of measurements is available.

Prior for the Displacement. To ensure smoothness of the deformation ϕ on $\mathcal{G}$ we incorporate a thin plate spline type regularization of the corresponding displacement u [12] and define

$$\mathcal{E}_{\mathrm{reg}}[u] := \frac{1}{2} \int_{\Omega} |\triangle u|^2 \, \mathrm{d}x , \qquad (4)$$

where $\triangle u = (\triangle u_1, \triangle u_2, \triangle u_3)$ and thus $|\triangle u|^2 = \sum_{k=1}^{3} (\triangle u_k)^2$. Indeed, since our input data Y only implicitly provide information for ϕ on a sparse set, a first order regularizer is inadequate to ensure sufficient regularity for the deformation. Let us emphasize that (discrete) smoothness of the approximate inverse deformation ψ is implicitly controlled by the regularization of ϕ.

2.2 Numerical Optimization

To minimize the functional $\mathcal{E}$ (Eq. 1), we apply a Finite Element approximation and optimize the functional using a gradient descent scheme. In particular, after an appropriate scaling of $\mathcal{G}$ we choose $\Omega = [0,1]^2$ and consider a piecewise bilinear, continuous Finite Element approximation on a uniform rectangular mesh covering Ω. In the experiments we used a 129×129 grid. Furthermore, the signed distance function d is precomputed using a fast marching method on a uniform rectangular 3-D grid covering the unit cube $[0,1]^3$ and stored on the nodes of this grid. In the algorithm d and ∇d are evaluated using trilinear interpolation of nodal values. For the gradient descent, derivatives of the energy have to be computed numerically. The derivatives of $\mathcal{E}_{\mathrm{match}}$ and $\mathcal{E}_{\mathrm{con}}$ w.r.t. w_j are given as:

$$\partial_{w_j} \mathcal{E}_{\mathrm{match}}[W] = \frac{1}{n} d(y_j + w_j) \nabla d(y_j + w_j)$$

$$\partial_{w_j} \mathcal{E}_{\mathrm{con}}[u, W] = \frac{1}{n} \left(P(y_j + w_j) + u(Q\,P(y_j + w_j)) - y_j \right)^T$$
$$\left(DP(y_j + w_j) + \nabla u(Q\,P(y_j + w_j)) Q\,DP(y_j + w_j) \right)$$

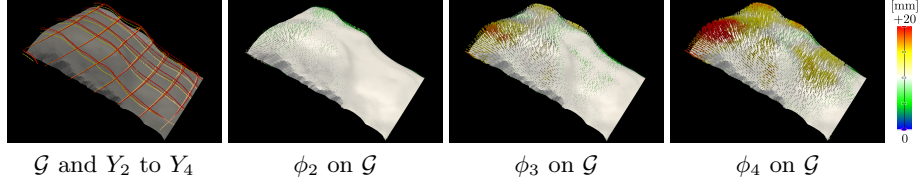| $\mathcal{G}$ and Y_2 to Y_4 | ϕ_2 on $\mathcal{G}$ | ϕ_3 on $\mathcal{G}$ | ϕ_4 on $\mathcal{G}$ |

Fig. 2. Validation on real AT data. Estimation of ϕ_p transforming $\mathcal{G}$ into $\mathcal{M}_p$, from AT sampling data Y_p. For the glyph visualization of ϕ_p on $\mathcal{G}$, $|u(\zeta)|$ is color coded [mm].

where DP denotes the Jacobian of the projection P. The variations of $\mathcal{E}_{\mathrm{con}}$ and $\mathcal{E}_{\mathrm{reg}}$ w.r.t. u in a direction $\vartheta : \Omega \to \mathbb{R}^3$ are given by:

$$\langle \partial_u \mathcal{E}_{\mathrm{con}}[u, W], \vartheta \rangle = \frac{1}{n} \sum_{i=1}^{n} \left(P(y_i + w_i) + u(Q\,P(y_i + w_i)) - y_i \right) \vartheta(Q\,P(y_i + w_i))$$

$$\langle \partial_u \mathcal{E}_{\mathrm{reg}}[u], \vartheta \rangle = \sum_{k=1}^{3} \int_{\Omega} \triangle u_k \triangle \vartheta_k \, \mathrm{d}x$$

The evaluation of DP involves the Hessian $D^2 d(x)$ of the distance function. One can either compute $D^2 d(x)$ based on second order finite differences or - as actually implemented here - replace the projection direction in P by the already computed direction from the last update. Furthermore, the Laplacian of a Finite Element function is evaluated by the discrete Finite Element Laplacian. In the gradient descent scheme we stop iterating as soon as the energy decay is smaller than a threshold value ϵ, $\epsilon = 10^{-4}$ proved to be sufficient to achieve the accuracy reported below. For the first frame of the respiratory motion we initialize $u = 0$ and $w_j = P(y_j) - y_j$ leading to approx. 60 gradient descent steps on average. For all subsequent frames we take u from the previous step and $w_j = P(y_j) - y_j$ as initial data resulting in approx. 45 descent steps on average.

3 Experiments and Results

Experimental Setup. For validation of the method, we have used an eye-safe AT prototype that acquires a sparse grid of 11×10 accurate 3-D sampling lines in real-time (30 Hz), using two perpendicular laser line pattern projection systems and a 1024×768 px resolution CCD chip [13]. Within the measurement volume, the mean AT measurement uncertainty is $\sigma = 0.39$ mm. The evaluation dataset is composed of 32 datasets from 16 subjects, each performing abdominal and thoracic breathing, respectively. Per subject, we synchronously acquired both real AT data and surface data using a moderately accurate but rather dense structured light (SL) system with a resolution of 320×240 px. Both sensors were mounted at a height of 1.2 m above the patient table, at a viewing angle of $30°$. AT and SL data were aligned using calibration. From each dataset, we extracted sparse AT measurements Y_p and dense SL meshes $\mathcal{M}_p$ for 8 phases within one
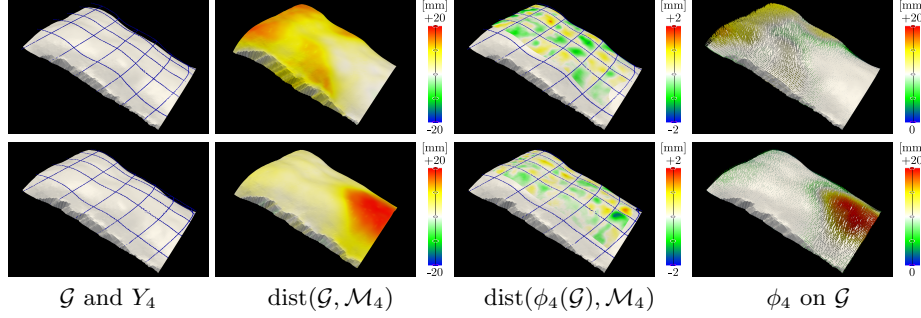
Fig. 3. Estimation of ϕ_p transforming $\mathcal{G}$ into $\mathcal{M}_p$ from realistic AT sampling data Y_p, for thoracic (top row) and abdominal respiration (bottom row). $p = 4$ represents the respiration state of fully inhale, roughly. For the glyph visualization of ϕ on $\mathcal{G}$, $|u(\zeta)|$ is color coded [mm]. Please note that the color coding differs by a factor of 10.
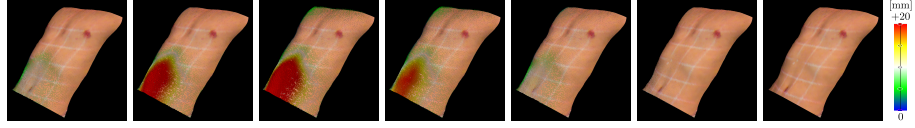

Fig. 4. Glyph visualization of ϕ_2 to ϕ_8 on $\mathcal{G}$ for an abdominal respiration cycle.

respiration cycle, the index p denotes the phase. In the experiments below, the subject's body surface at full expiration $\mathcal{M}_1$ is considered as the given planning data $\mathcal{G}$. The model parameters were empirically set to $\kappa = 8 \cdot 10^{-1}$, $\lambda = 4 \cdot 10^{-8}$.

Validation on Real AT Data. Results for the reconstruction of ϕ_p for phase p on real AT data are given in Fig. 2. A quantitative evaluation on real AT and aligned SL data was unfeasible, as the SL camera exhibited local sampling artifacts due to the underlying measurement principle and interferences between the laser grid (AT) and speckle pattern projections (SL) of the synchronously used modalities, which cause local deviations in the scale of several millimeters.

Quantitative Evaluation on Realistic AT Data. For quantitative evaluation, we developed a simulator for the generation of realistic AT sampling data from dense SL surfaces. For this purpose, the noise characteristics of our AT sensor prototype were measured in an optics lab and used to augment the synthetic sampling, providing realistic AT data. We considered the reconstruction of the displacement field ϕ_p from realistic AT data Y_p, $p = \{2, \ldots, 8\}$, sampled from $\mathcal{M}_p$. An evaluation is given in Fig. 3 and the displacements for a full respiration cycle are shown in Fig. 4. The accuracy of the deformation estimation is assessed by the absolute error $|\text{dist}(\phi_p(\mathcal{G}), \mathcal{M}_p)|$ in Fig. 5 representing the mismatch between the transformed reference surface and the ground truth surface. To discard boundary effects, the evaluation is performed within the central surface of interest covering the torso. Over all subjects and phases, the mean reconstruction error was 0.22 mm w.r.t. ground truth dense SL data. This indicates that the method can reliably recover the dense displacement field from a
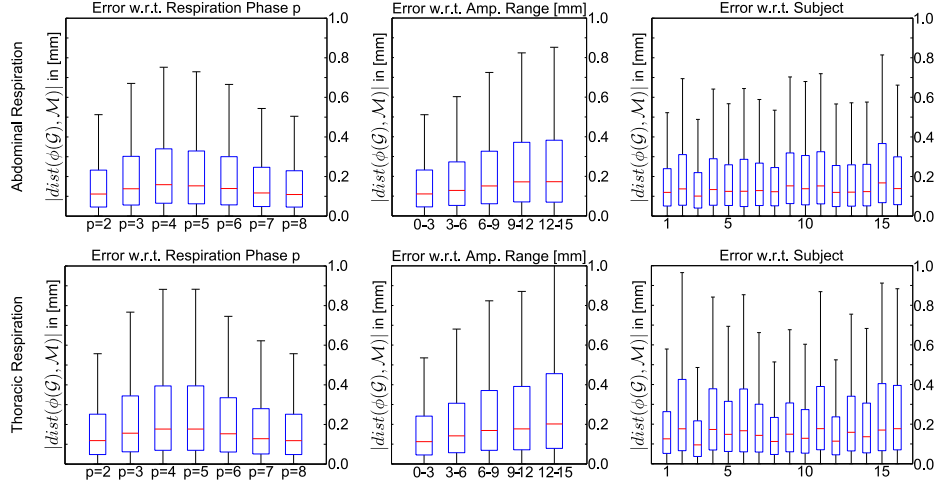
Fig. 5. Box plots of $|\text{dist}(\phi_p(\mathcal{G}), \mathcal{M}_p)|$ for realistic AT sampling data from 16 subjects, for abdominal (top row) and thoracic (bottom row) respiration. Given are plots for different phases of the respiration cycle (left), w.r.t. the respiration amplitude (center), and for the individual subjects (right). The reconstruction error scales approximately linearly with the respiration amplitude observing a peak at the respiration state of fully inhale (phase 4/5). The whiskers indicate that >99% of the residual error is <1 mm.

sparse sampling of the instantaneous patient state using prior shape knowledge.
Performance. With our proof of concept implementation, a single gradient descent step on a single core of a Xeon X5550 2.67GHz CPU takes ≈ 60 ms. Over all subjects, we achieved total runtimes of 2.6 ± 0.7 s, thus significantly outperforming related work on dense-to-dense surface registration [8] with runtimes in the scale of minutes (25 iterations, 11.9 s per iteration on comparable CPU and for a surface mesh with a comparable number of vertices).

4 Conclusions and Outlook

In this paper, a variational approach to marker-less reconstruction of dense non-rigid 4-D surface motion fields from sparse but accurate AT sampling data has been introduced. The algorithm can precisely reconstruct the dense respiratory displacement field using prior shape knowledge from planning data. The implications for RT motion management are manifold. The motion fields can be used as multi-dimensional respiration surrogates, as input for accurate external-internal motion correlation models, and to reconstruct the body shape for patient positioning. Beyond its application in RT, the approach holds potential for motion-compensated tomographic reconstruction and image-guided interventions.

References

1. Hoogeman, M., Prvost, J.B., Nuyttens, J., Pll, J., Levendag, P., Heijmen, B.: Clinical accuracy of the respiratory tumor tracking system of the Cyberknife: assessment by analysis of log files. Int J Radiat Oncol Biol Phys **74**(1) (2009) 297–303
2. Verellen, D., Depuydt, T., Gevaert, T., Linthout, N., Tournel, K., Duchateau, M., Reynders, T., Storme, G., Ridder, M.D.: Gating and tracking, 4D in thoracic tumours. Cancer Radiother **14**(67) (2010) 446–454
3. Willoughby, T.R., Forbes, A.R., Buchholz, D., Langen, K.M., Wagner, T.H., Zeidan, O.A., Kupelian, P.A., Meeks, S.L.: Evaluation of an infrared camera and X-ray system using implanted fiducials in patients with lung tumors for gated radiation therapy. Int J Radiat Oncol Biol Phys **66**(2) (2006) 568–575
4. Brahme, A., Nyman, P., Skatt, B.: 4D laser camera for accurate patient positioning, collision avoidance, image fusion and adaptive approaches during diagnostic and therapeutic procedures. Med Phys **35**(5) (2008) 1670–1681
5. Moser, T., Fleischhacker, S., Schubert, K., Sroka-Perez, G., Karger, C.P.: Technical performance of a commercial laser surface scanning system for patient setup correction in radiotherapy. Phys Med **27**(4) (2011) 224–232
6. Peng, J.L., Kahler, D., Li, J.G., Samant, S., Yan, G., Amdur, R., Liu, C.: Characterization of a real-time surface image-guided stereotactic positioning system. Med Phys **37**(10) (2010) 5421–5433
7. Fayad, H., Pan, T., Clement, J.F., Visvikis, D.: Correlation of respiratory motion between external patient surface and internal anatomical landmarks. Med Phys **38**(6) (2011) 3157–3164
8. 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
9. 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 (2012) 98–109
10. Cachier, P., Rey, D.: Symmetrization of the non-rigid registration problem using inversion-invariant energies: Application to multiple sclerosis. In Delp, S., DiGoia, A., Jaramaz, B., eds.: MICCAI 2000. Volume 1935 of LNCS. Springer Berlin/Heidelberg (2000) 697–708
11. Christensen, G., Johnson, H.: Consistent image registration. IEEE Trans Med Imaging **20**(7) (2001) 568–582
12. Modersitzki, J., Fischer, B.: Curvature based image registration. Journal of Mathematical Imaging and Vision **18**(1) (2003) 81–85
13. Ettl, S., Arold, O., Yang, Z., Häusler, G.: Flying triangulation—an optical 3D sensor for the motion-robust acquisition of complex objects. Appl Opt **51**(2) (2012) 281–289