Joint Surface Reconstruction and 4-D Deformation Estimation from Sparse Data and Prior Knowledge for Marker-Less Respiratory Motion Tracking

Benjamin Berkels

Institute for Numerical Simulation,

Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn, Germany[*]

Sebastian Bauer

Pattern Recognition Lab, Department of Computer Science,

Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany[†]

Svenja Ettl and Oliver Arold

Institute of Optics, Information and Photonics,

Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

Joachim Hornegger

Pattern Recognition Lab, Department of Computer Science,

Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany[‡]

Martin Rumpf

Institute for Numerical Simulation,

Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn, Germany

Abstract

Purpose: The intra-procedural tracking of respiratory motion has the potential to substantially improve image-guided diagnosis and interventions. We have developed a sparse-to-dense registration approach that is capable of recovering the patient's external 3-D body surface and estimating a 4-D (3-D+time) surface motion field from sparse sampling data and patient-specific prior shape knowledge.

Methods: The system utilizes an emerging marker-less and 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 recovered, which describes the spatio-temporal 4-D deformation of the complete patient body surface, depending on the type and state of respiration. It yields both a reconstruction of the instantaneous patient shape and a high-dimensional respiratory surrogate for respiratory motion tracking. The method is validated on a 4-D CT respiration phantom and evaluated on both real data from an AT prototype and synthetic data sampled from dense surface scans acquired with a structured-light scanner.

Results: In the experiments, we estimated surface motion fields with the proposed algorithm on 256 datasets from 16 subjects and in different respiration states, achieving a mean surface reconstruction accuracy of ± 0.23 mm w.r.t. ground truth data – down from a mean initial surface mismatch of 5.66 mm. The 95^{th} percentile of the local residual mesh-to-mesh distance after registration did not exceed 1.17 mm for any subject. On average, the total runtime of our proof of concept CPU implementation is 2.3 s per frame, outperforming related work substantially.

Conclusions: In external beam radiation therapy (RT), the approach holds potential for patient monitoring during treatment using the reconstructed surface, and for motion-compensated dose delivery using the estimated 4-D surface motion field in combination with external-internal correlation models.

[*] Interdisciplinary Mathematics Institute, University of South Carolina, Columbia, SC, USA; benjamin.berkels@ins.uni-bonn.de

[†] Graduate School of Information Science in Health (GSISH), Technische Universität München, München, Germany

[‡] Erlangen Graduate School of Advanced Optical Technologies (SAOT), Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

I. INTRODUCTION

Respiration-synchronized image-guided radiation therapy (IGRT) techniques aim at continuously tracking the moving target over its trajectory and re-position the treatment table [1] or radiation beam [2, 3] dynamically to follow the tumor's changing position [3–9]. This allows to reduce the tumor-motion margin in dose distribution and to increase the accelerator's duty cycle compared to gated RT [10]. Recent hybrid solutions 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 the deformation of the external body surface in the time interval between image acquisitions – using a correlation model trained from a series of simultaneously acquired external-internal position measurements [7, 11], 4-D CT [3, 12, 13] or 4-D MRI planning data [14]. The key issue with external-internal motion correlation models is the actual level of correlation accounting for the accuracy of dose delivery. Clinically available solutions that are in use or potentially suitable for hybrid tumor-tracking [7, 15, 16] typically measure external motion using a single or a few passive markers on the patient's chest as a low-dimensional (in most cases 1-D) surrogate. However, in practice, these low-dimensional techniques are incapable of depicting the full complexity of respiratory motion. Experimental studies by Fayad et al. [17] and Yan et al. [18] confirmed that using multiple external surrogates at different anatomical locations is superior to the conventional approach with a single 1-D respiratory signal for external-internal correlation modeling. Further drawbacks of marker-based external surrogates are the fact that they involve extensive patient preparation and require reproducible equipment and marker placement with a considerable impact on model accuracy.

Modern IGRT systems that allow to monitor the motion of the complete external patient surface have the potential to help reducing correlation model uncertainties. In particular, marker-less optical range imaging (RI) technologies can acquire a dense 3-D surface model of the patient [19–25] over time. Based on the estimation of a dense displacement field representing the deformation of the instantaneous torso shape with respect to a reference surface, a more reliable and accurate correlation model can be established [17, 26]. However, available RI-based IGRT solutions that are capable of delivering dense surface information in a marker-less manner [19–21, 23, 24] focus on patient positioning [25] and do not support dense sampling in real-time [20, 21] or at the cost of a limited field of view [19, 23, 24]. Furthermore, none of the commercially available systems feature non-rigid respiratory motion tracking.

In this paper, we propose a marker-less system based on a non-moving active laser triangulation (AT) sensor that delivers sparse but highly accurate measurements in real-time. Using prior patient shape knowledge from tomographic planning data, a variational model is introduced to recover a dense and accurate 4-D (3-D+time) displacement field that provides a high-dimensional breathing surrogate, and to reconstruct a reliable and complete patient surface model at the instantaneous respiration phase. Different from our first approach in [27] we propose here a scheme which much better reflects the underlying projective geometry and is substantially more efficient.

II. METHODS AND MATERIALS

Given is a pre-fractional reference shape $\mathcal{G} \subset \mathbb{R}^3$ that can be (1) extracted from planning tomographic data (CT/MRI), (2) captured with a dense range imaging sensor of low temporal resolution, or (3) acquired by an AT sensor in combination with a steerable treatment table. During dose delivery, the instantaneous patient body surface denoted by $\mathcal{M} \subset \mathbb{R}^3$ is represented by sparse AT sampling data $Y \subset \mathbb{R}^3$. In particular, the AT sensor acquires a finite set of n measurements $Y = \{y_1, \ldots, y_n\}$, $y_i \in \mathbb{R}^3$, arranged in a grid-like structure (Fig. 1). Note that the intra-fractional grid-like sampling Y is not aligned with $\mathcal{G}$ and depends on the respiration state.

Now, the goal is to estimate the unknown, dense and non-rigid deformation $\phi : \mathcal{G} \to \mathbb{R}^3$ that matches the reference shape $\mathcal{G}$ to the instantaneous patient body surface $\mathcal{M}$. Ideally, ϕ should be such that $\mathcal{M} = \phi(\mathcal{G})$, but since our data Y only contains information about a sparse subset of $\mathcal{M}$, the condition on ϕ appropriate for our problem setting is $Y \subset \phi(\mathcal{G})$. Along the lines of inverse-consistent registration [28, 29], in a joint manner, we estimate ϕ together with its inverse ψ. Again, due to the sparse nature of our input data, we do not try to estimate the inverse everywhere on $\mathcal{M}$ but only on the known sparse subset Y. In other words, instead of trying to find $\psi : \mathcal{M} \to \mathcal{G}$ with $\psi(\mathcal{M}) = \mathcal{G}$, we estimate a sparse deformation $\Psi : Y \to \mathbb{R}^3$ such that $\Psi(Y) \subset \mathcal{G}$. Here, dense and sparse deformations are distinguished by using lower and upper case letters respectively. Let us underline that Ψ is fully represented by the discrete set $\{\Psi(y_1), \ldots, \Psi(y_n)\}$ containing the deformed positions of

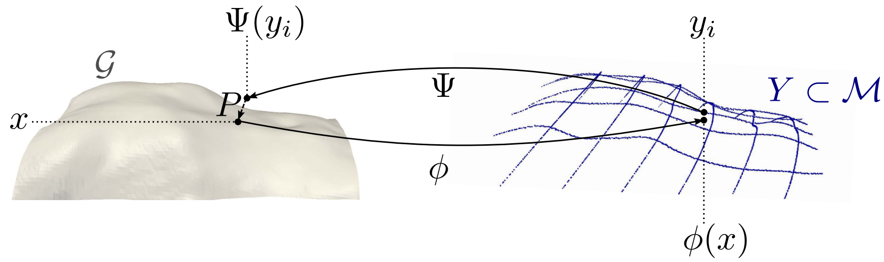


FIG. 1. Geometric configuration for the reconstruction of the dense deformation ϕ and the approximate sparse inverse Ψ from sparse sampling data $Y = \{y_1, \ldots, y_n\}$ and reference shape data $\mathcal{G} \subset \mathbb{R}^3$. For a better visibility $\mathcal{G}$ and Y have been pulled apart. Furthermore, the projection P onto $\mathcal{G}$ is sketched.

the n points acquired by the AT sensor. A geometric sketch that illustrates the deformations ϕ and Ψ 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 can be used as a high-dimensional breathing surrogate and enables the reconstruction of the complete instantaneous patient surface for intra-fractional monitoring.

To quantify the matching of $\Psi(Y)$ onto $\mathcal{G}$ let us assume that the signed distance function (SDF) $d_\mathcal{G}$ with respect to $\mathcal{G}$ is precomputed in a sufficiently large neighborhood in $\mathbb{R}^3$. Using this signed distance function $d_\mathcal{G}(x) := \pm\mathrm{dist}(x, \mathcal{G})$, we can construct the projection P of a point $x \in \mathbb{R}^3$ in a neighborhood of $\mathcal{G}$ onto the closest point on $\mathcal{G}$. Let us emphasize that, even though $P(Y) \subset \mathcal{G}$ holds by construction, we do not expect any biologically reasonable Ψ to be equal to the projection P.

A. Definition of the Registration Energy

Now, we define a functional $\mathcal{E}$ on on a dense deformation ϕ and a sparse deformation Ψ whose minimizer represents a suitable matching of the planning data $\mathcal{G}$ and AT measurements Y:

$$\begin{aligned}\mathcal{E}[\phi, \Psi] &:= \mathcal{E}_{\text{match}}[\Psi] + \kappa\, \mathcal{E}_{\text{con}}[\phi, \Psi] + \lambda\, \mathcal{E}_{\text{reg}}[\phi] \\ &= \frac{1}{2n} \sum_{i=1}^{n} |d_\mathcal{G}(\Psi(y_i))|^2 + \frac{\kappa}{2n} \sum_{i=1}^{n} |\phi(P(\Psi(y_i))) - y_i|^2 \\ &\quad + \frac{\lambda}{2} \int |\Delta(\phi - \mathrm{id})|^2 \,\mathrm{d}x\end{aligned} \tag{1}$$

Here, κ and λ are nonnegative constants controlling the contributions of the individual terms and id denotes the identity mapping. $\mathcal{E}_{\text{match}}$ is a matching energy that encodes the condition $\Psi(Y) \subset \mathcal{G}$ by measuring closeness of $\Psi(Y)$ to $\mathcal{G}$. The consistency functional $\mathcal{E}_{\text{con}}$ is responsible for establishing the relation between both displacement fields, constraining Ψ and ϕ to be approximately inverse to each other on the sparse set of positions Y where Ψ is defined. Thereby, it implicitly encodes the condition $Y \subset \phi(\mathcal{G})$. Finally, $\mathcal{E}_{\text{reg}}$ ensures a regularization of the dense deformation ϕ. Since $\mathcal{E}_{\text{con}}$ only controls ϕ on a sparse set, it is necessary to use a higher order regularization here. For a detailed construction and discussion of these functionals we refer to the appendix ([38], Sect. I).

B. Discretization and Minimization

To minimize the highly non-convex objective functional $\mathcal{E}$ w.r.t. the unknowns ϕ and Ψ (Eq. 1), we apply a multi-linear FE discretization in space and use a regularized gradient descent [30] to guarantee a fast and smooth relaxation. The gradient descent is discretized explicitly in time, the step size is controlled with the Armijo rule [31]. We stop the gradient descent iteration as soon as the energy decay is smaller than a specified threshold value ϵ, for practical values we refer to Sect. IV C. By default, both deformations ϕ and Ψ are initialized with the identity mapping. In the experiments (Sect. IV C), we further study the benefit of initializing ϕ with the estimates from the previous step and initializing Ψ with $\Psi(y_j) = P(y_j)$ for $j = 1, \ldots, n$.

Details on the numerical optimization, including the derivatives of the energy, are given in the appendix ([38], Sect. II). Note that the numerical evaluation of the projection onto $\mathcal{G}$ is based on the signed distance function $d_{\mathcal{G}}$ and the expression $P(x) = x - d_{\mathcal{G}}(x)\nabla d_{\mathcal{G}}(x)$. Thus, the variation of $\mathcal{E}_{\text{con}}$ w.r.t. Ψ involves the derivative of P which in turn involves second derivatives of $d_{\mathcal{G}}$. To avoid these second derivatives we use a projection approximation scheme. Compared to the scheme used in [27], the scheme used here treats the distance $d_{\mathcal{G}}$ implicitly and the direction $\nabla d_{\mathcal{G}}$ explicitly, while [27] treated both $d_{\mathcal{G}}$ and $\nabla d_{\mathcal{G}}$ explicitly. Thus, the improved projection approximation scheme reflects the underlying projective geometry much better and is substantially more efficient. The appendix ([38], Sect. III) describes the projection scheme in detail. For a quantitative analysis of the impact of this modification on reconstruction accuracy and the convergence speed of the entire algorithm,

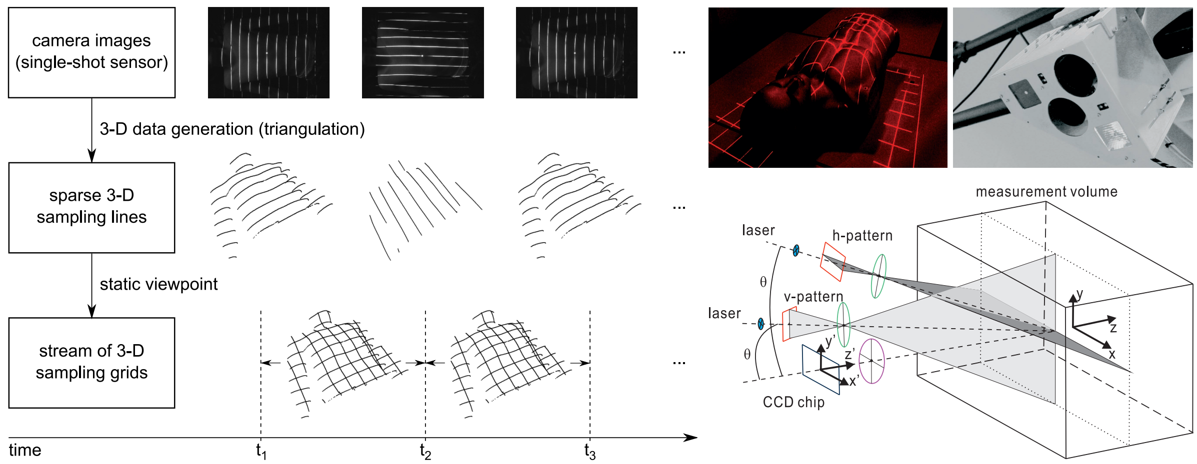


FIG. 2. Left: AT measurement principle for respiratory motion tracking. Top right: AT sensor and projected laser grid on a phantom. Here, for illustration purposes, both projection units were activated simultaneously. In practice, projecting a grid would hinder a precise 3-D data generation at the observed intersection points. Hence, the horizontal and vertical projection units are activated alternately. Bottom right: Schematic setup of the AT sensor consisting of two orthogonal projection units and a CCD camera.

respectively, we refer to Sect. IV C.

C. Active Triangulation Prototype

The active triangulation sensor used in this work was recently introduced for interactive reconstruction of dense and accurate 3-D models [32]: A hand-guided sensor is moved around an object while continuously capturing camera images of a projected line pattern. Each camera image delivers sparse 3-D measurements, which are aligned to precedingly acquired data in real-time. The sensors work at their physical limits, hence having minimal measurement uncertainty for the corresponding measurement volume.

In this work, the AT sensor is rigidly mounted and acquires sparse 3-D data of the respiring patient from one static viewpoint [27]. Alternately, sets of 11 horizontal and 10 vertical lines are projected onto the scene observed by a synchronized CCD camera with a resolution of 1024×768 px and a framerate of 30 Hz [32] (see Fig. 2). Due to the alternating projection, a fully updated set of horizontal and vertical measurements is available at an effective framerate of 15 Hz. Within the employed measurement volume of 80×80×35 cm^3, the mean measurement uncertainty is $\sigma = 0.39$ mm.

III. EXPERIMENTAL SETUP

The experimental evaluation divides into two parts. In Sect. III A, the validation setup of the proposed model on surface data from a synthetic 4-D CT respiration phantom is described. In the second part (Sect. III B), we present a comprehensive study on data from 16 healthy subjects. In particular, we have quantified the accuracy in 4-D deformation estimation and surface reconstruction, respectively, and analyzed the performance of the proposed framework w.r.t. relevant system parameters.

All experiments below (Sect. III A, III B) were performed with a parameter setting of $\kappa = 0.8$, $\lambda = 4 \cdot 10^{-8}$. These weighting factors were determined empirically. The convergence threshold was empirically set to $\epsilon = 10^{-4}$ (Sect. IV C). To generate AT sampling data from synthetic datasets, we have developed a virtual simulator that mimics the sampling principle of the AT sensor by intersecting a given triangulated surface with a set of sampling rays. These rays are arranged in a grid-like structure and the default grid and sampling density of the simulator are set in accordance to the specifications of the actual AT prototype used in the experiments on real data, see Sect. III B. Due to occlusion constraints in a clinical RT environment, the simulator's sampling plane and viewing angle, respectively, is set 30° off from an orthogonal camera position w.r.t. the treatment table.

A. Validation on 4-D CT Respiration Phantom Data

For model validation, we have investigated the reconstruction of respiratory deformation fields from surface data of an established 4-D CT respiration phantom (NCAT [33]). For the experiments, we generated dense surface meshes $\mathcal{M}_p$ for 8 phases within one respiration cycle, for male and female phantom data. The index $p \in \{1, \ldots, 8\}$ denotes the phase. We considered both scenarios of arms-up and arms-down patient posture. The NCAT parameters were set to default values [33]. The phantom surface at the state of full expiration $\mathcal{M}_1$ ($p = 1$) was considered as the planning geometry $\mathcal{G}$. The remaining set of surfaces was used to generate synthetic sampling data $Y_2, \ldots, Y_8$ using our AT simulator. The accuracy of the deformation estimation is assessed by the absolute distance of the points in $\phi_p(\mathcal{G})$ to $\mathcal{M}_p$, representing the residual mismatch in terms of mesh-to-mesh distance between the transformed reference surface $\phi_p(\mathcal{G})$ and the ground truth surface $\mathcal{M}_p$. Here, we exploit the SDF w.r.t. $\mathcal{M}_p$ to establish a correspondence between $\phi_p(\mathcal{G})$ and $\mathcal{M}_p$ by computing the distance of a point in the transformed reference surface to the closest point on the ground truth surface, i. e. computing $|d_{\mathcal{M}_p}|$ on $\phi_p(\mathcal{G})$. To discard boundary effects at the body-table transition, the evaluation is performed in the central volume of interest that covers the trunk of the phantom.

B. Prototype Study on Healthy Subjects

In order to demonstrate the clinical feasibility of the presented system and to evaluate it under realistic conditions, we have conducted a study on 16 healthy subjects, male and female. In particular, we have investigated the performance of our modified projection approximation compared to [27], the impact of initializing the displacement fields with estimates of the preceding respiration phase, and the influence of the convergence threshold ϵ. Using the AT simulator and the measured noise characteristics of our prototype, we performed realistic simulations to study the influence of the AT laser grid density. Along with quantitative and qualitative results in terms of reconstruction accuracy, we have analyzed the performance of our implementation in terms of convergence and runtime.

Using an eye-safe prototype as described in Sect. II C, we acquired 32 datasets from 16 subjects, each performing (i) abdominal and (ii) thoracic breathing. In addition to AT data, we synchronously acquired surface data with a moderately accurate but rather dense structured-light (SL) system as dense ground truth data for quantitative evaluation of our approach, as detailed in [27]. SL data were pre-processed using edge-preserving bilateral filtering. From each dataset, we extracted sparse AT measurements Y_p and dense SL meshes $\mathcal{M}_p$ for 8 phases within one respiration cycle. This results in a total number of $32 \times 8 = 256$ datasets. For the experiments, we considered the reconstruction of the displacement field ϕ_p from a given planning surface $\mathcal{G}$ and intra-fractional AT data Y_p, $p \in \{2, \ldots, 8\}$. The subject's body surface at full expiration $\mathcal{M}_1$ was considered as the planning data surface $\mathcal{G}$. As with the 4-D phantom study (Sect. III A), the accuracy of the deformation estimation is assessed by the residual mismatch $|d_{\mathcal{M}_p}|$ on $\phi_p(\mathcal{G})$ within the central volume of interest.

In practice, a quantitative evaluation on synchronously acquired real AT and 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 caused local deviations in the scale of millimeters. Hence, the evaluation on real AT data is restricted to qualitative results. For quantitative evaluation, we employed our simulator for the generation of realistic AT sampling data from dense SL surfaces $\mathcal{M}_p$. In order to generate realistic AT data, the noise characteristics of our AT sensor prototype were used to augment the synthetic sampling of dense SL data.

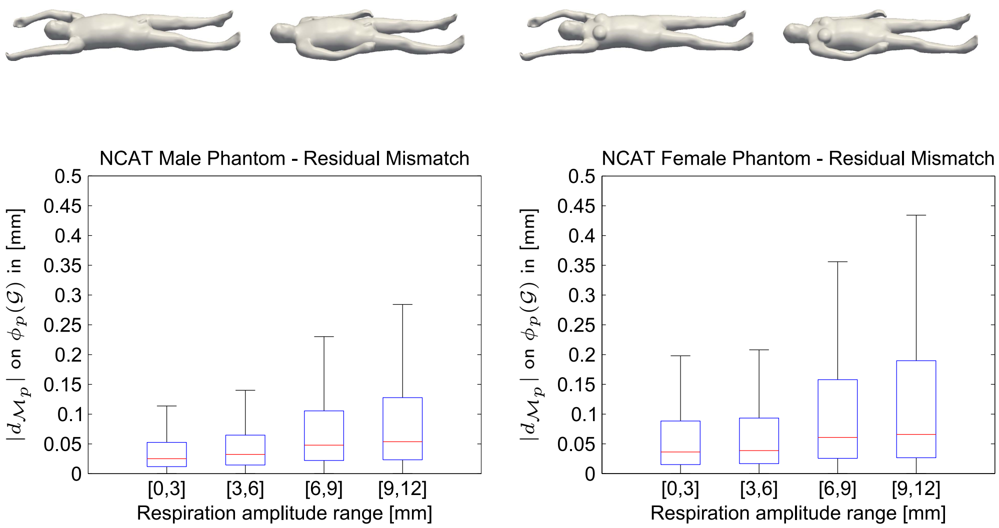
Let us stress here that the aforementioned interferences do not hinder the practical application of the proposed method – which only requires an AT sensor – but the generation of ground truth data necessary for evaluation.

C. Parameter Study

First, we investigated the benefits of initializing ϕ with the estimates from the previous step and Ψ with $\Psi(y_i) = P(y_j)$ to reduce the number of iterations needed for the optimization scheme to convergence. Note that for the first frame we initialize $\phi = \mathrm{id}$. Furthermore, to determine a suitable value for the convergence threshold ϵ, we studied the impact of ϵ on both reconstruction accuracy and runtime.

Compared to the first formulation in [27], we have improved the approximation scheme w.r.t. a proper representation of the underlying projective geometry as discussed in the appendix ([38], Sect. III). In order to study the computational impact of this modification, we compared the results with the improved projection approximation to previous results [27] for $\epsilon = 10^{-7}$.

With regard to upcoming generations of AT sensors that are expected to feature more dense laser grids, we further investigated the influence of the grid density on the registration error. The evaluation was performed on realistic AT simulator data with grids of 11×11, 22×22, 33×33 and 44×44 lines.



a *b*

FIG. 3. Validation of the model on a male (a) and female (b) NCAT phantom. Given are boxplots of the absolute registration error in [mm] in terms of $|d_{\mathcal{M}_p}|$ on $\phi_p(\mathcal{G})$ w.r.t. discrete ranges of respiration amplitude.

IV. RESULTS

A. Validation on 4-D CT Respiration Phantom Data

Quantitative results on NCAT phantom data are given in Fig. 3. The boxplots illustrate the absolute registration error w.r.t. discrete ranges of respiration amplitude for the male and female phantom. The results for the arms-up and arms-down datasets are combined per gender. Even for instances with a large initial surface mismatch in the scale of 9-12 mm, the median residual error in terms of $|d_{\mathcal{M}_p}|$ on $\phi_p(\mathcal{G})$ is substantially smaller than 0.1 mm. The error scales directly proportional to the respiration amplitude (cf. Sect. IV B). Qualitative results are depicted in Fig. 4. With the female phantom, the AT coverage of the breast is limited. This becomes evident with an increased local error around the outer part of the female breast. However, the impact is moderate due to the incorporation of prior shape knowledge and the higher order regularization of ϕ ([38], Sect. I C). These model priors are also beneficial in cases of (self-)occlusion. For instance, due to the viewing angle of 30° w.r.t. the treatment table plane, the upper part of the female breast in Fig. 4 is self-occluded,

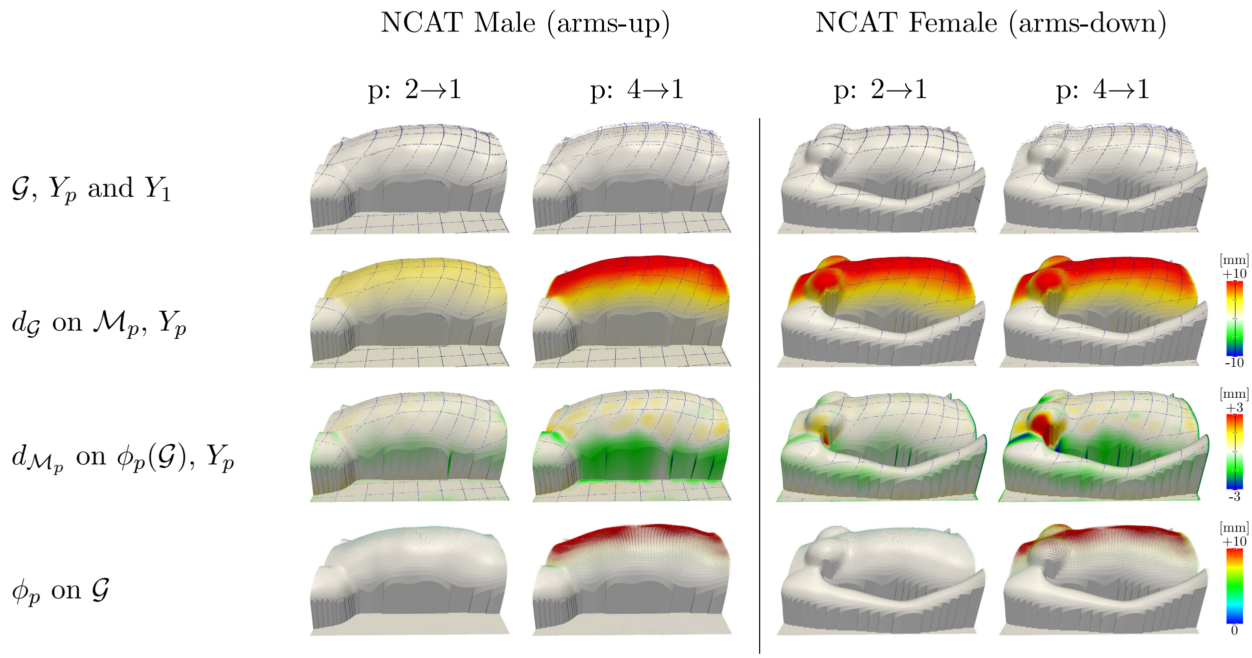


FIG. 4. Qualitative NCAT results for reconstruction of the deformation field for phases $p = 2, 4$ w.r.t. $p = 1$ as reference (full expiration), for male arms-up (left) and female arms-down (right) data. First row: $\mathcal{G}$, Y_p (outer contour) and $Y_1 \subset \mathcal{G}$. Second row: Initial mismatch in terms of $d_{\mathcal{G}}$ on $\mathcal{M}_p$, and Y_p. Third row: Residual mismatch after application of the proposed method in terms of $d_{\mathcal{M}_p}$ on $\phi_p(\mathcal{G})$, and Y_p. Fourth row: Glyph visualization of the displacement field ϕ_p on $\mathcal{G}$, $|\phi_p - \mathrm{id}|$ is color coded.

but can be reconstructed in a robust manner.

B. Prototype Study on Healthy Subjects

Qualitative results of the prototype study on healthy subjects are depicted in Fig. 5. To facilitate an anatomic interpretation of the deformation, we overlaid the color texture on $\mathcal{G}$ that was acquired with our SL device. Note that an analysis of the deformation ϕ_p allows for a distinct differentiation between abdominal and thoracic respiratory deformation patterns and inter-subject variations in the respiration amplitude. For instance, in the case of thoracic respiration, subject S_1 and subject S_2 exhibit a similar motion pattern in the thorax region but substantial differences in the abdominal region.

Quantitative results over all subjects on realistic AT data are given in Fig. 6. Fig. 6a

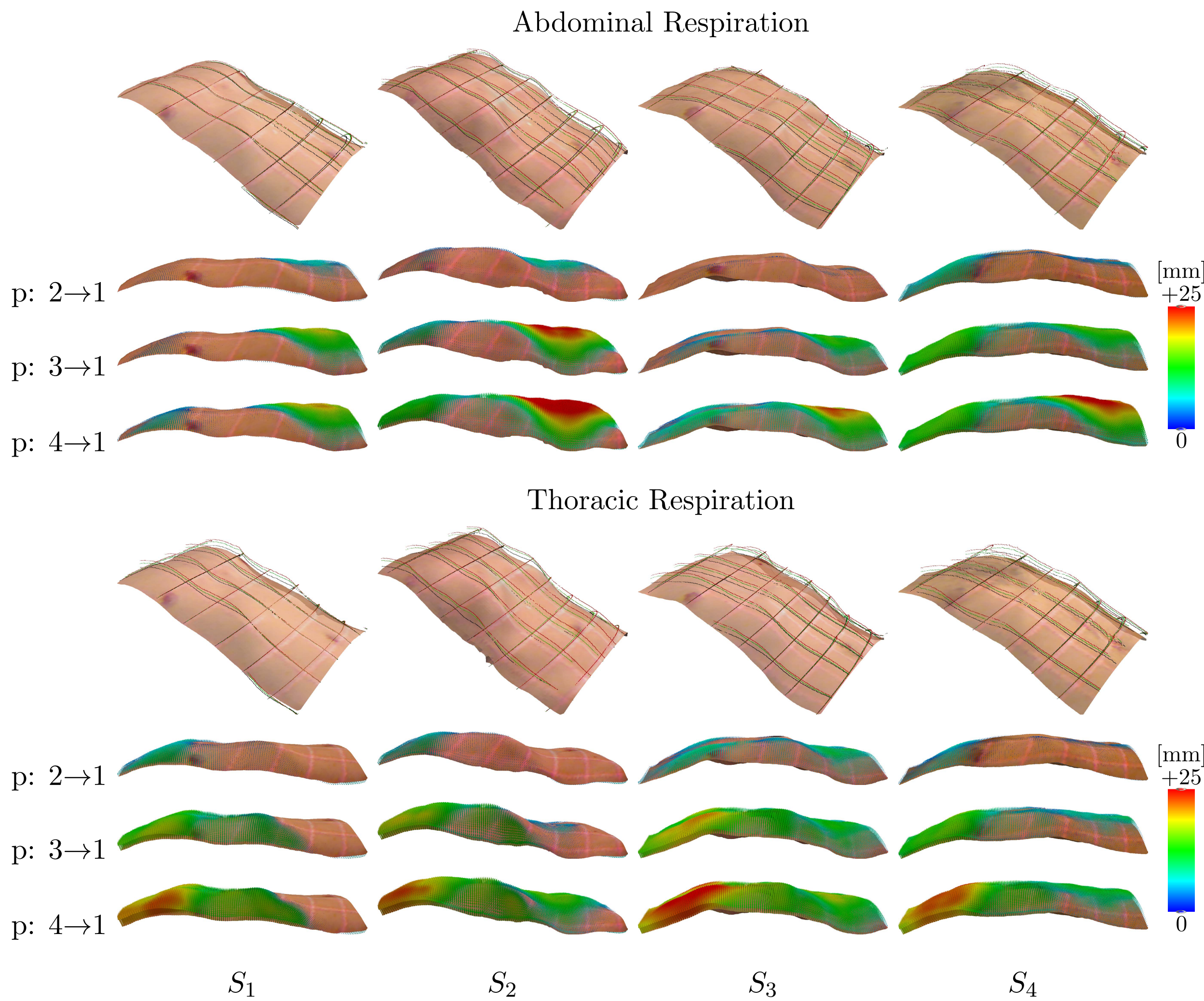


FIG. 5. Results on real AT data from four subjects (left to right), for abdominal (top) and thoracic (bottom) respiration, for phases $p = 2, 3, 4$. For each subject, the reference surface $\mathcal{G} = \mathcal{M}_1$ and the AT sampling data Y_p are shown in the first row (Y_2 most inner contour, Y_4 most outer contour, Y_3 in between). The following three rows illustrate the estimated displacement fields ϕ_2, ϕ_3, ϕ_4 on $\mathcal{G}$. For the glyph visualization of ϕ_p on $\mathcal{G}$, $|\phi_p - \mathrm{id}|$ is color coded in [mm].

depicts boxplots of the initial mismatch $|d_{\mathcal{G}}|$ on $\mathcal{M}_p$ and residual mismatch $|d_{\mathcal{M}_p}|$ on $\phi_p(\mathcal{G})$ over all 16 subjects. Here, the results for abdominal and thoracic respiration are evaluated in a common plot. While Fig. 6a gives an impression about the overall performance, Fig. 6b shows the residual mismatch in a more detailed scale. Figs. 6c,d depict the residual error for discrete respiration phases over all subjects, for abdominal (Fig. 6c) and thoracic respiration (Fig. 6d). The reconstruction error scales approximately linearly with the respiration

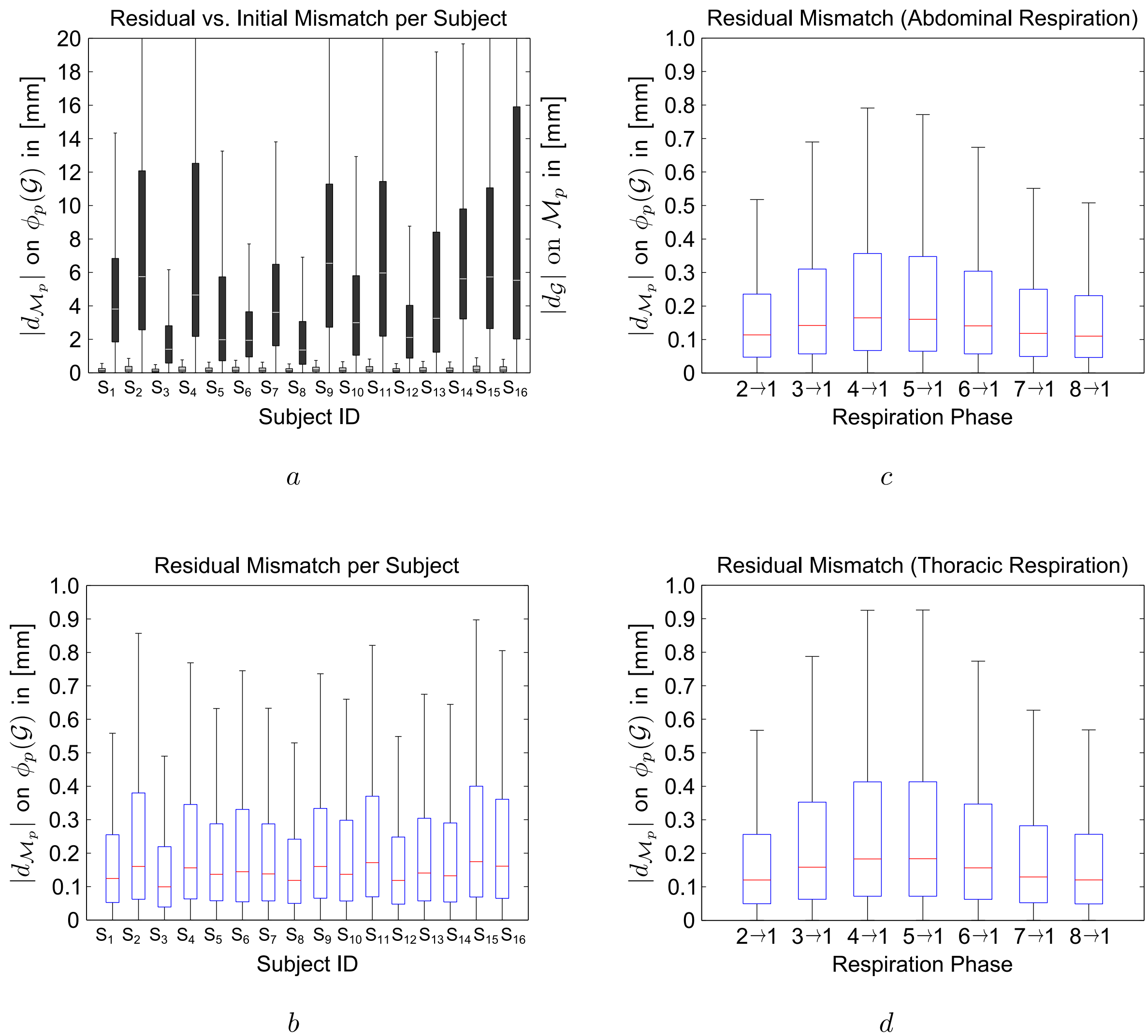


FIG. 6. Quantitative results of the prototype study.(a) Results per individual subject, comparing the initial mismatch(dark gray bars) vs. residual mismatch(light gray bars) as boxplots over both respiration types (abdominal and thoracic) and all phases. (b) Residual mismatch per subject. (c,d) Boxplots of the residual mismatch for discrete phases of the respiration cycle, for abdominal (c) and thoracic (d) respiration, over all subjects.

amplitude observing a peak at the respiration state of fully inhale (phase 4/5). The boxplot whiskers indicate that >99% of the residual error is < 1 mm.

Over all subjects, respiration types and respiration phases, the mean reconstruction error in terms of residual mismatch $|d_{\mathcal{M}_p}|$ on $\phi_p(\mathcal{G})$ was $\pm$ 0.21 mm and $\pm$ 0.25 mm for abdominal and thoracic respiration, respectively, see Table I. The 95$^{\text{th}}$ percentile did not exceed 0.93 mm for abdominal respiration and 1.17 mm for thoracic respiration, for any subject. For

TABLE I. Results over all subjects, respiration types and phases. Given are the mean, median and 95$^{\text{th}}$ percentile of the initial and residual mismatch in [mm], for abdominal respiration (A), thoracic respiration (T) and the entire dataset covering both respiration types (A/T). The last row states numbers in terms of residual mesh-to-mesh mismatch from related work by Schaerer et al. [26].

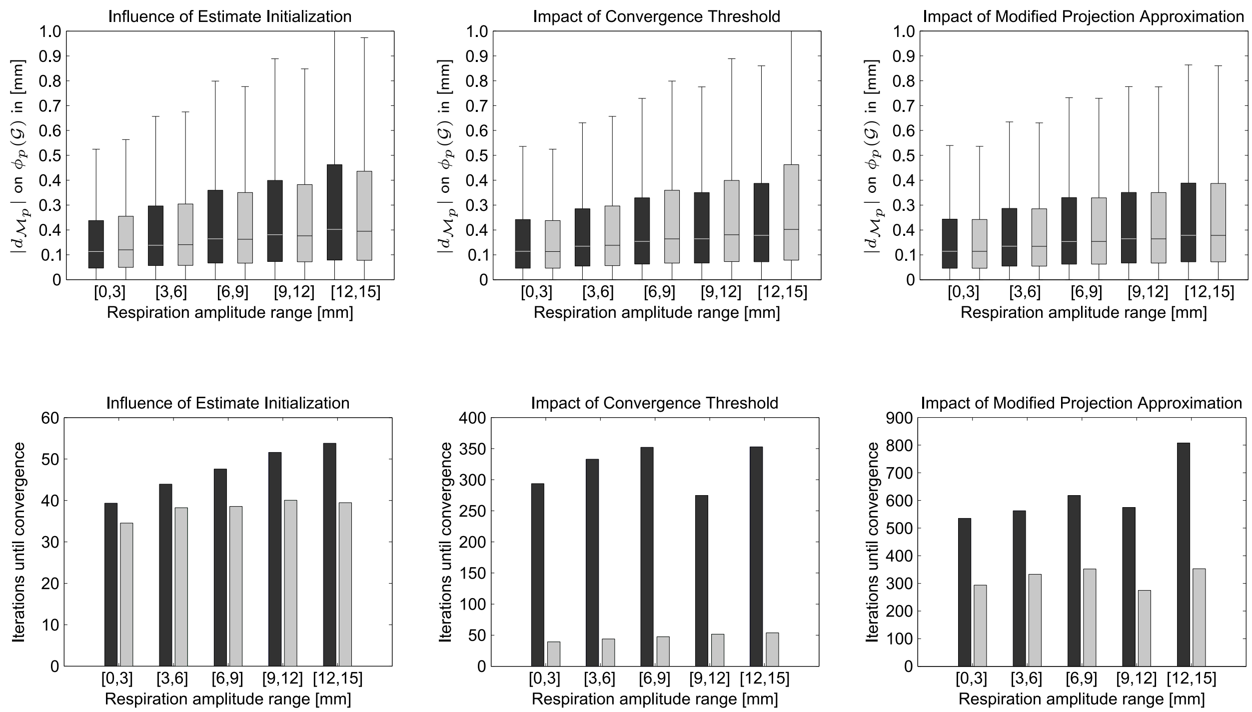
	Initial Mismatch [mm]			Residual Mismatch [mm]		
	A	T	A/T	A	T	A/T
Mean	5.09	6.24	5.66	0.21	0.25	0.23
Median	3.95	4.66	4.23	0.13	0.15	0.14
95$^{\text{th}}$ Percentile	14.0	17.1	15.2	0.69	0.82	0.76
[26], 95$^{\text{th}}$ Percentile	-	-	6.1	-	-	1.08

a detailed overview of the initial and residual mismatch (95$^{\text{th}}$ percentile) for the individual subjects, separated for abdominal and thoracic respiration, we refer to the appendix ([38], Sect. IV). We assume the moderately higher reconstruction error for thoracic respiration to result from the higher initial mismatch of thoracic respiration data (mean: 6.24 mm) compared to abdominal data (mean: 5.09 mm).

C. Parameter Study

Reconstruction accuracy and runtime with and without an appropriate initialization of ϕ and Ψ are compared in Fig. 7a. The experiments illustrate that initializing ϕ and Ψ reduces both the registration error and the number of iterations needed for the optimization scheme to convergence. Over all datasets, estimate initialization reduced runtime by 19.2%. The results for different convergence thresholds ($\epsilon = 10^{-4}$ and $\epsilon = 10^{-7}$) are depicted in Fig. 7b. The boxplots indicate that a reduction of ϵ by a factor of 10^3 results in a small improvement in reconstruction accuracy at the cost of a substantial increase in solver iterations. In an empirical study, we found that the convergence threshold of $\epsilon = 10^{-4}$ used in the experiments gave the best tradeoff between accuracy and runtime.

Fig. 7c illustrates that both projection approximations result in a comparable reconstruction accuracy while the new improved approximation reduces runtime substantially (48.2%



a *b* *c*

FIG. 7. Parameter Study. Given are the residual mismatch (top row) and the number of iterations until convergence (bottom row), respectively. (a) Results without (dark gray) and with (light gray) initialization of ϕ and Ψ. (b) Impact of the convergence threshold, results for $\epsilon = 10^{-7}$ are depicted in dark gray, results for $\epsilon = 10^{-4}$ in light gray. (c) Impact of the improved projection approximation (light gray) compared to our previous work [27] (dark gray). To investigate the impact of different convergence thresholds and projection approximations independent from the effect of initializing ϕ and Ψ, the results in (b) and (c) were generated without initialization.

over all subjects). Note that we used $\epsilon = 10^{-7}$ instead of $\epsilon = 10^{-4}$ as convergence threshold to separate the effect of the improved approximation from the influence of the convergence threshold.

Quantitative results w.r.t. the influence of the AT grid density are depicted in Fig. 8. Indeed, a denser grid of 22×22 lines compared to the currently used 11×11 lines leads to a reduction of the reconstruction error by 49.0%. Further refinement does not noticeably improve the results - probably due to the low-frequent surface geometry of the human torso

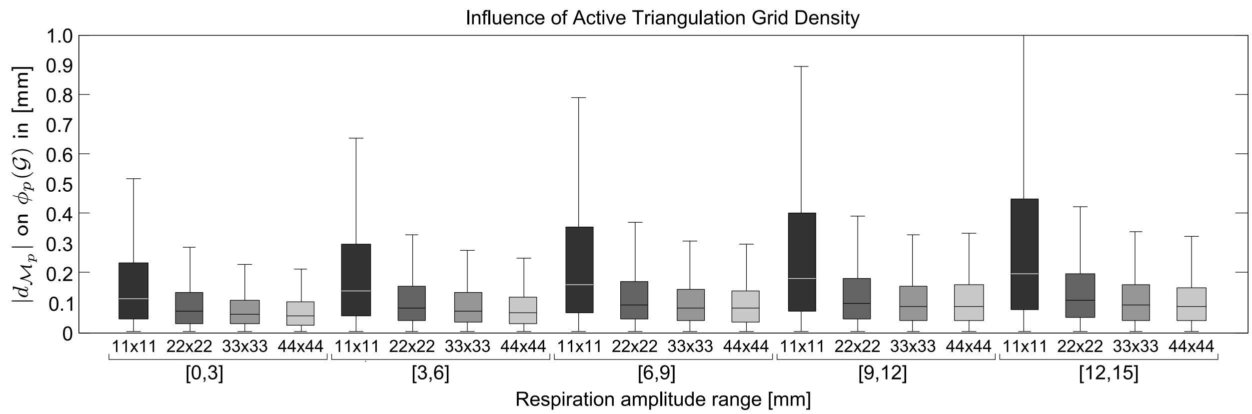


FIG. 8. Influence of AT grid density on registration accuracy. Given are boxplots for laser grid resolutions of 11×11, 22×22, 33×33, 44×44 sampling lines (grouped as four adjacent entries colored from dark to light gray), for increasing ranges of respiration amplitude (from left to right).

D. Runtime Performance

Let us comment on the runtime performance in our volunteer study. The total runtime per frame was 2.3 s, measured as mean over all datasets of the volunteer study. In detail, when initializing ϕ with the estimates from the previous phase and Ψ with $\Psi(y_j) = P(y_j)$, the optimization process took 38.2±2.1 iterations to converge for one subject, in average over all subjects, respiration types and respiration phases. 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. The resulting per-frame runtime of 2.3 s substantially outperforms related work on dense-to-dense surface registration [26] 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).

V. DISCUSSION

We have investigated the performance of the proposed method on synthetic, realistic and real data. On 256 datasets from 16 subjects with an average initial mismatch of 5.66 mm, the mean reconstruction error was ±0.23 mm and the 95^{th} percentile did not exceed 1.17 mm for any subject. In the experiments (Sect. IV C), it was further shown that a proper initialization of the displacements ϕ and Ψ and the improved approximation of the projection compared to our first approach [27] reduces the runtime by 19.2% and 48.2%, respectively. Higher framerates are possible for both the line pattern projection systems and the observing camera. Moreover, denser laser lines can be realized by adapting the setup to the required measurement volume.

With regard to the state-of-the-art in non-rigid surface deformation estimation, in particular in RT, let us compare our results to recent work by Schaerer et al. [26] on motion tracking with dense surfaces. In their study on five male subjects and dense surface acquisitions from three respiration phases, the authors achieved a residual mismatch of 1.08 mm (95^{th} percentile) in terms of mesh-to-mesh surface distance using non-rigid ICP surface registration [34]. Note that our result of 0.76 mm for the 95^{th} percentile residual mismatch over all subjects and both respiration types slightly outperform these numbers, see Table I. In addition, let us remark that compared to to the volunteers from Ref. [26], many of the subjects in our study exhibited a considerably higher respiration amplitude and initial mismatch (15.2 mm vs. 6.1 mm), cf. Table I. Hence, the low residual mismatch indicates that our method can reliably recover the dense displacement field from a sparse sampling of the instantaneous patient state using prior shape knowledge, even in the presence of strong respiration.

Let us further distinguish our approach from commercial RI-based IGRT solutions. First and foremost, available solutions do not support dense sampling in real-time [20, 21] or at the cost of a limited field of view [19, 23, 24]. For instance, note that the Sentinel system (C-RAD AB, Uppsala, Sweden) [20] and the Galaxy system (LAP GmbH, Lüneburg, Germany) [21] take several seconds for a complete scan of the torso, and the real-time mode of the VisionRT stereo system (VisionRT Ltd., London, UK) [23] is limited to interactive framerates of 1.5-7.5 Hz, depending on the size of the surface of interest. The temporal resolution of these solutions may be insufficient to characterize respiratory motion [35]. We expect the low framerates to result from the underlying measurement technologies. Beside its limitations in terms of sampling density and speed, respectively, commercially available solutions often imply high costs in terms of hardware and are subject to measurement uncertainties due to the underlying sampling principles e.g. active stereo photogrammetry [19, 23, 24] or consecutive light-sectioning using mechanically swept lasers [20, 21]. Third and last, the general focus of these systems is on patient positioning [25] and none of them features dense and non-rigid respiratory motion tracking.

In addition to the aforementioned hardware-related benefits of the AT sensor for respiratory motion tracking, the CPU implementation of our surface registration approach outperforms the non-rigid ICP used by Schaerer et al. substantially in terms of runtime performance (two orders of magnitude). As our approach exhibits an inherently high degree of data parallelism we will consider a GPU implementation [36] in future work to achieve real-time operation required for clinical applications.

IGRT solutions are typically expected to serve both tasks of patient positioning prior to treatment and respiratory motion management during dose delivery. RI-based systems perform patient positioning by acquiring a 3-D sampling of the patient shape and registering it to a reference surface extracted from planning data. This requires a dense sampling of the patient's surface in the treatment position and a statically mounted AT sensor would only recover sparse 3-D information along the projection grid. However, note that using the AT sensor in combination with the steerable treatment couch enables the reconstruction of dense 3-D sampling data.

VI. CONCLUSIONS AND OUTLOOK

We have introduced a variational approach to marker-less reconstruction of dense non-rigid 4-D surface motion fields from sparse but accurate AT sampling data. In a comprehensive study on 16 subjects, we demonstrated the capability of the algorithm to precisely reconstruct the dense respiratory displacement field using prior shape knowledge from planning data. In the field of RT, the 4-D motion fields can be used as biologically reasonable high-dimensional respiration surrogates for gated RT, as input for accurate external-internal motion correlation models in respiration-synchronized RT, for motion compensated patient positioning [37], and to reconstruct the intra-fractional body shape for patient setup monitoring during dose delivery. Beyond its application in RT, the approach holds potential for diagnostic and therapeutic applications.

ACKNOWLEDGMENTS

B. Berkels and M. Rumpf acknowledge support by the Hausdorff Center for Mathematics. Furthermore, B. Berkels acknowledges support by the DFG Priority Program 1335.

S. Bauer acknowledges support by the European Regional Development Fund (ERDF) and the Bayerisches Staatsministerium für Wirtschaft, Infrastruktur, Verkehr und Technologie under grant IUK338/001, and by the Graduate School of Information Science in Health (GSISH) and the Technische Universität München (TUM) Graduate School. J. Hornegger acknowledges funding of the Erlangen Graduate School in Advanced Optical Technologies (SAOT) by the German National Science Foundation (DFG) in the framework of the excellence initiative.

---

[1] J. Wilbert, J. Meyer, K. Baier, M. Guckenberger, C. Herrmann, R. Hess, C. Janka, L. Ma, T. Mersebach, A. Richter, M. Roth, K. Schilling, and M. Flentje, "Tumor tracking and motion compensation with an adaptive tumor tracking system (ATTS): System description and prototype testing," *Med. Phys.*, vol. 35, pp. 3911–3921, 2008.

[2] W. Kilby, J. R. Dooley, G. Kuduvalli, S. Sayeh, and C. R. Maurer, "The CyberKnife robotic radiosurgery system in 2010," *Technol. Cancer Res. Treat.*, vol. 9, pp. 433–452, Oct 2010.

[3] D. Verellen, T. Depuydt, T. Gevaert, N. Linthout, K. Tournel, M. Duchateau, T. Reynders, G. Storme, and M. D. Ridder, "Gating and tracking, 4D in thoracic tumours," *Cancer Radiother.*, vol. 14, no. 6–7, pp. 446–454, 2010.

[4] T. Depuydt, D. Verellen, O. Haas, T. Gevaert, N. Linthout, M. Duchateau, K. Tournel, T. Reynders, K. Leysen, M. Hoogeman, G. Storme, and M. D. Ridder, "Geometric accuracy of a novel gimbals based radiation therapy tumor tracking system," *Radiother. Oncol.*, vol. 98, no. 3, pp. 365–372, 2011.

[5] S. Dieterich, C. Cavedon, C. F. Chuang, A. B. Cohen, J. A. Garrett, C. L. Lee, J. R. Lowenstein, M. F. d'Souza, D. D. Taylor, X. Wu, and C. Yu, "Report of AAPM TG 135: quality assurance for robotic radiosurgery.," *Med. Phys.*, vol. 38, pp. 2914–2936, Jun 2011.

[6] H. Fayad, T. Pan, C. Roux, C. Le Rest, O. Pradier, J. Clement, and D. Visvikis, "A patient specific respiratory model based on 4D CT data and a time of flight camera (TOF)," in *IEEE NSS MIC*, pp. 2594–2598, 2009.

[7] M. Hoogeman, J.-B. Prévost, J. Nuyttens, *et al.*, "Clinical accuracy of the respiratory tumor tracking system of the CyberKnife: assessment by analysis of log files.," *Int. J. Radiat. Oncol., Biol., Phys.*, vol. 74, no. 1, pp. 297–303, 2009.

[8] J. McClelland, J. Blackall, S. Tarte, A. Chandler, S. Hughes, S. Ahmad, D. Landau, and D. Hawkes, "A continuous 4D motion model from multiple respiratory cycles for use in lung radiotherapy," *Med. Phys.*, vol. 33, no. 9, pp. 3348–3358, 2006.

[9] M. J. Murphy, "Tracking moving organs in real time.," *Semin. Radiat. Oncol.*, vol. 14, pp. 91–100, Jan 2004.

[10] P. J. Keall, G. S. Mageras, J. M. Balter, R. S. Emery, K. M. Forster, S. B. Jiang, J. M. Kapatoes, D. A. Low, M. J. Murphy, B. R. Murray, C. R. Ramsey, M. B. V. Herk, S. S. Vedam, J. W. Wong, and E. Yorke, "The management of respiratory motion in radiation oncology report of AAPM task group 76," *Med. Phys.*, vol. 33, pp. 3874–3900, Oct 2006.

[11] A. Muacevic, C. Drexler, A. Wowra, A. Schweikard, A. Schlaefer, R. T. Hoffmann, R. Wilkowski, and H. Winter, "Technical description, phantom accuracy, and clinical feasibility for single-session lung radiosurgery using robotic image-guided real-time respiratory tumor tracking," *Technol. Cancer Res. Treat.*, vol. 6, pp. 321–328, Aug 2007.

[12] J. Eom, X. G. Xu, S. De, and C. Shi, "Predictive modeling of lung motion over the entire respiratory cycle using measured pressure-volume data, 4DCT images, and finite-element analysis.," *Med. Phys.*, vol. 37, pp. 4389–4400, Aug 2010.

[13] J. Vandemeulebroucke, S. Rit, J. Kybic, P. Clarysse, and D. Sarrut, "Spatiotemporal motion estimation for respiratory-correlated imaging of the lungs.," *Med. Phys.*, vol. 38, pp. 166–178, Jan 2011.

[14] M. Miquel, J. Blackall, S. Uribe, D. Hawkes, and T. Schaeffter, "Patient-specific respiratory models using dynamic 3d mri: Preliminary volunteer results," *Phys. Med.*, 2012.

[15] E. C. Ford, G. S. Mageras, E. Yorke, K. E. Rosenzweig, R. Wagman, and C. C. Ling, "Evaluation of respiratory movement during gated radiotherapy using film and electronic portal imaging.," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 52, pp. 522–531, Feb 2002.

[16] T. R. Willoughby, A. R. Forbes, D. Buchholz, K. M. Langen, T. H. Wagner, O. A. Zeidan, P. A. Kupelian, and S. L. Meeks, "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.*, vol. 66, no. 2, pp. 568–575, 2006.

[17] H. Fayad, T. Pan, J. F. Clement, and D. Visvikis, "Correlation of respiratory motion between external patient surface and internal anatomical landmarks.," *Med. Phys.*, vol. 38, no. 6, pp. 3157–3164, 2011.

[18] H. Yan, F.-F. Yin, G.-P. Zhu, M. Ajlouni, and J. H. Kim, "The correlation evaluation of a tumor tracking system using multiple external markers," *Medical Physics*, vol. 33, no. 11, pp. 4073–4084, 2006.

[19] C. Bert, K. G. Metheany, K. Doppke, and G. T. Y. Chen, "A phantom evaluation of a stereo-vision surface imaging system for radiotherapy patient setup.," *Med. Phys.*, vol. 32, pp. 2753–2762, Sep 2005.

[20] A. Brahme, P. Nyman, and B. Skatt, "4D laser camera for accurate patient positioning, collision avoidance, image fusion and adaptive approaches during diagnostic and therapeutic procedures.," *Med. Phys.*, vol. 35, no. 5, pp. 1670–1681, 2008.

[21] T. Moser, S. Fleischhacker, K. Schubert, G. Sroka-Perez, and C. P. Karger, "Technical performance of a commercial laser surface scanning system for patient setup correction in radiotherapy.," *Phys. Med.*, vol. 27, no. 4, pp. 224–232, 2011.

[22] D. P. Gierga, M. Riboldi, J. C. Turcotte, G. C. Sharp, S. B. Jiang, A. G. Taghian, and G. T. Chen, "Comparison of target registration errors for multiple image-guided techniques in accelerated partial breast irradiation," *Int. J. Radiat. Oncol., Biol., Phys.*, vol. 70, no. 4, pp. 1239–1246, 2008.

[23] J. L. Peng, D. Kahler, J. G. Li, S. Samant, G. Yan, R. Amdur, and C. Liu, "Characterization of a real-time surface image-guided stereotactic positioning system.," *Med. Phys.*, vol. 37, no. 10, pp. 5421–5433, 2010.

[24] P. J. Schöffel, W. Harms, G. Sroka-Perez, W. Schlegel, and C. P. Karger, "Accuracy of a commercial optical 3D surface imaging system for realignment of patients for radiotherapy of the thorax.," *Phys. Med. Biol.*, vol. 52, pp. 3949–3963, Jul 2007.

[25] T. Willoughby, J. Lehmann, J. A. Bencomo, S. K. Jani, L. Santanam, A. Sethi, T. D. Solberg, W. A. Tome, and T. J. Waldron, "Quality assurance for nonradiographic radiotherapy localization and positioning systems: report of task group 147.," *Med. Phys.*, vol. 39, pp. 1728–1747, Apr 2012.

[26] J. Schaerer, A. Fassi, M. Riboldi, P. Cerveri, G. Baroni, and D. Sarrut, "Multi-dimensional respiratory motion tracking from markerless optical surface imaging based on deformable mesh registration.," *Phys. Med. Biol.*, vol. 57, no. 2, pp. 357–373, 2012.

[27] S. Bauer, B. Berkels, S. Ettl, O. Arold, J. Hornegger, and M. Rumpf, "Marker-less reconstruction of dense 4-D surface motion fields using active laser triangulation for respiratory motion

management," in *MICCAI, Part I* (N. Ayache, H. Delingette, P. Golland, and K. Mori, eds.), vol. 7510 of *LNCS*, pp. 414–421, 2012.

[28] P. Cachier and D. Rey, "Symmetrization of the non-rigid registration problem using inversion-invariant energies: Application to multiple sclerosis," in *MICCAI 2000* (S. Delp, A. DiGoia, and B. Jaramaz, eds.), vol. 1935 of *LNCS*, pp. 697–708, Springer, 2000.

[29] G. Christensen and H. Johnson, "Consistent image registration," *IEEE Trans. Med. Imaging.*, vol. 20, no. 7, pp. 568–582, 2001.

[30] G. Sundaramoorthi, A. Yezzi, and A. Mennucci, "Sobolev active contours," *Int. J. Comput. Vision*, vol. 73, no. 3, pp. 345–366, 2007.

[31] L. Armijo, "Minimization of functions having Lipschitz continuous first partial derivatives," *Pacific J. Math.*, vol. 16, no. 1, pp. 1–3, 1966.

[32] S. Ettl, O. Arold, Z. Yang, and G. Häusler, "Flying triangulation—an optical 3D sensor for the motion-robust acquisition of complex objects," *Appl. Opt.*, vol. 51, no. 2, pp. 281–289, 2012.

[33] W. Segars, S. Mori, G. Chen, and B. Tsui, "Modeling respiratory motion variations in the 4D NCAT phantom," in *IEEE NSS/MIC*, vol. 4, pp. 2677–2679, 2007.

[34] B. Amberg, S. Romdhani, and T. Vetter, "Optimal step nonrigid ICP algorithms for surface registration," in *IEEE CVPR*, pp. 1–8, Jun 2007.

[35] T. Waldron, "External surrogate measurement and internal target motion: Photogrammetry as a tool in IGRT," in *ACMP Annual Meeting*, p. 10059, 2009.

[36] A. Bruhn, T. Pock, and X.-C. Tai, "Efficient algorithms for global optimisation methods in computer vision," *Dagstuhl Reports*, vol. 1, no. 11, pp. 66–90, 2012.

[37] J. Wasza, S. Bauer, and J. Hornegger, "Real-time motion compensated patient positioning and non-rigid deformation estimation using 4-d shape priors," in *MICCAI, Part II* (N. Ayache, H. Delingette, P. Golland, and K. Mori, eds.), vol. 7511 of *LNCS*, pp. 576–583, 2012.

[38] See supplementary material at http://dx.doi.org/[tbd] for the appendix. [EPAPS]