Fast time-of-flight camera based surface registration for radiotherapy patient positioning

Simon Placht[a)]
Metrilus GmbH, Henkestr. 91, 91052 Erlangen, Germany and Pattern Recognition Lab, Friedrich Alexander University Erlangen-Nuremberg, Martensstr. 3, 91058 Erlangen, Germany

Joseph Stancanello
German Research Cancer Centre, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany

Christian Schaller and Michael Balda
Metrilus GmbH, Henkestr. 91, 91052 Erlangen, Germany

Elli Angelopoulou
Pattern Recognition Lab, Friedrich Alexander University Erlangen-Nuremberg, Martensstr. 3, 91058 Erlangen, Germany

(Received 18 March 2011; revised 7 November 2011; accepted for publication 7 November 2011; published 9 December 2011)

Purpose: This work introduces a rigid registration framework for patient positioning in radiotherapy, based on real-time surface acquisition by a time-of-flight (ToF) camera. Dynamic properties of the system are also investigated for future gating/tracking strategies.
Methods: A novel preregistration algorithm, based on translation and rotation-invariant features representing surface structures, was developed. Using these features, corresponding three-dimensional points were computed in order to determine initial registration parameters. These parameters became a robust input to an accelerated version of the iterative closest point (ICP) algorithm for the fine-tuning of the registration result. Distance calibration and Kalman filtering were used to compensate for ToF-camera dependent noise. Additionally, the advantage of using the feature based preregistration over an "ICP only" strategy was evaluated, as well as the robustness of the rigid-transformation-based method to deformation.
Results: The proposed surface registration method was validated using phantom data. A mean target registration error (TRE) for translations and rotations of 1.62 ± 1.08 mm and $0.07° \pm 0.05°$, respectively, was achieved. There was a temporal delay of about 65 ms in the registration output, which can be seen as negligible considering the dynamics of biological systems. Feature based preregistration allowed for accurate and robust registrations even at very large initial displacements. Deformations affected the accuracy of the results, necessitating particular care in cases of deformed surfaces.
Conclusions: The proposed solution is able to solve surface registration problems with an accuracy suitable for radiotherapy cases where external surfaces offer primary or complementary information to patient positioning. The system shows promising dynamic properties for its use in gating/tracking applications. The overall system is competitive with commonly-used surface registration technologies. Its main benefit is the usage of a cost-effective off-the-shelf technology for surface acquisition. Further strategies to improve the registration accuracy are under development.

[DOI: 10.1118/1.3664006]

Key words: radiotherapy, registration, surface features, Kalman filtering, time-of-flight cameras

I. INTRODUCTION

Recent advances in computer technology and medical imaging devices have created new opportunities for medical applications such as image-guided diagnostics and interventional procedures. In order to benefit from the complementary information from these new imaging modalities, a registration of either mono-modal or multimodal image data are often required. Several image registration solutions are available, e.g., image-based or surface-based registration.[1]

Modern radiotherapy calls for effective and efficient, as well as safe treatments. Within this framework, image registration plays a key role in the treatment planning as well as in the delivery phase. The introduction of image guided radio therapy (IGRT) has gained great attention because of its potential for more effective and safe radiotherapy treatments.[2] X-ray imaging is currently the most widely used IGRT approach. However, though it offers the possibility of visualizing internal structures, it does so by exposing patients to additional radiation doses. Ongoing investigations on the integration of radiotherapy devices with no dose, high soft tissue contrast imaging modalities are currently being developed.[3] Furthermore, many efforts are focused on optimizing the radiotherapy workflow by:

1. decreasing the time slot allocated to treating each patient;
2. potentially increasing the throughput of the radiotherapy device;
3. reducing the negative impact of long treatment times on the patient.

These two goals—the need for dose reduction in IGRT and workflow optimization—have stimulated the investigation of noninvasive surface-based positioning systems in order to:

- properly align and monitor patients in an expeditious manner;
- reduce the radiation dose by advantageous surface-based positioning or monitoring.

Motion management has also gained increasing attention in order to further maximize the tumor control rate and minimize the probability of normal tissue complications, especially in emerging hypofractionated regimens. Interfraction motion management is tightly linked to patient positioning and, to some extent, to the emerging adaptive radiotherapy (ART) techniques.[4] Intrafraction motion management consists of gating[5] and tracking. Tracking can be applied to the entire patient (patient tracking[6]—often known also as patient surveillance with the possibility of compensating for the patient movement), or to the tumor (tumor tracking[7]—using internal anatomycal images with a possible link to external information through correlation models[8]). Thus, a natural development of surface-based systems will be their evolution from patient positioning applications, to patient and tumor tracking. This implies, in turn, that there will be a particular emphasis on fast response systems.

Most of the available surface-based positioning systems are based on technologies like laser range scanners[9] or active stereo camera systems.[10] Representative commercially available systems are the laser range scanners Galaxy (LAP GmbH, Lueneburg, Germany), Sentinel (C-RAD AB, Uppsala, Sweden), and the stereo system AlignRT (Vision RT, London, UK). According to LAP, the Galaxy scanner has a patient positioning accuracy of less than 1 mm at a scan speed (for a single surface) of up to 5 s,[11] which can be critical in terms of real-time gating/tracking. The registration error of the Sentinel system evaluated on phantoms is reported as less than 0.5 mm for the translational and less than 0.5° for the rotational component.[12] The AlignRT system has been evaluated on both patients and phantoms. In real-time mode, AlignRT reaches a surface image capture rate of 0.1–0.3 frames per second (fps). Using a head-and-neck phantom, Peng *et al.* reported a mean registration error of 1.2 ± 0.7mm (mean $\pm$ standard deviation) in the translational and $0.7^\circ \pm 0.4^\circ$ in the rotational component.[10]. The evaluation on patients was performed on five subjects who received intracranial stereotactic radiotherapy (SRT). Thereby, the isocenter positioning accuracy was 0.2 ± 0.3 mm in the translational and $0.2^\circ \pm 0.6^\circ$ in the rotational component. As reference datasets for the registration they utilized previously recorded surface images of the phantom and the patients, respectively. In Bert *et al.*, the accuracy of the AlignRT system was evaluated with phantoms for relative translations of ± 1 cm and rotations of $\pm 5^\circ$.[13] As reference surface, the acquisition before the couch translation/rotation was used. They reported a mean Euclidean distance of 0.95 ± 0.58 mm and a maximum of 2.2 mm between the real and the computed translation. The mean rotational error was below 0.1°, but was not exactly specified.

Besides these acknowledged limitations regarding accuracy, available systems often exhibit the following additional drawbacks:

- Laser scanners acquire a scene consecutively: this is an essential drawback of laser scanners if dynamic scenes are acquired;
- Structured light systems involve a sensitive polyocular configuration. Small vibrations can slightly change their stereo base setup, which in turn affects the accuracy of the whole system. Consequently, polyocular systems are more sensitive to extraneous influences (e.g., forces, etc.), which may necessitate regular recalibration;[10,14]
- Both systems are currently expensive, which may prevent the large scale adoption of these solutions.

Thus, we propose an alternative to laser scanners and active stereo camera systems, which is based on time of flight (ToF) cameras. ToF sensors are a relatively novel technology for acquiring three-dimensional (3D) data in real-time. Thus, there is a limited body of work on the employment of ToF sensors in the medical field. The more closely related work is that by Schaller *et al.*[15] who first investigated the suitability of ToF technology for patient positioning and respiratory motion gating. They analyzed the positioning accuracy using both a body phantom and three human subjects. In their evaluation, the authors ignored rotations and considered ground truth translations of up to 10 cm at a working distance of 80 cm. They reported a mean registration error of 2.88 ± 1.84 mm for the phantom and 3.38 ± 2.00 mm for the human subjects. This first approach was further enhanced and evaluated on a rigid phantom for small translations of up to 10 mm.[16] For this "translation only" scenario, a registration accuracy of 0.74 ± 0.37 mm was achieved. Another interesting application involves the use of ToF sensors in the registration of organ surfaces with data extracted from a computed tomography volume. In that work, Müller *et al.* report a mean surface-to-surface distance of 1.70 ± 0.36 mm (after registration) on four porcine livers acquired with a ToF camera at a camera-object distance of 60 cm.[17]

In this paper, we present a surface-based registration framework using a time-of-flight (ToF) camera. ToF technology allows a single shot acquisition and is less susceptible to vibrations, since it is based on a single optic module. Currently, the price of ToF cameras starts at 1000 USD and will probably decrease in the future as the sensors become more widely used, driven by the increased demand in automotive and consumer electronic industries. ToF cameras emit intensity-modulated eye-safe near-infrared light. By calculating the phase shift between the emitted and reflected light, ToF sensors measure distances at each pixel. Current cameras provide more than 40 000 3D points at about 25 fps. In addition to the distance data, ToF cameras also provide a two-dimensional (2D) gray-scale image of the scene, which is called amplitude image. Details about the principle of operation of ToF cameras can be found in Ref. 18.

Some other medical applications for ToF cameras have been recently published, e.g., inverse C-arm positioning,[19] building and tracking of root shapes[20] and quantitative 3D endoscopy.[21] In the field of respiratory motion analysis, ToF cameras are used for gating,[22] the creation of a patient specific model based on 4D computer tomography (CT) Data[23] and 3D respiratory motion detection.[24,25] Furthermore, Fayad *et al.* investigated the correlation of respiratory motion between the external patient surface acquired by a ToF camera and internal anatomical landmarks.[26]

Despite their increasing popularity, ToF cameras still suffer from several sources of noise and systematic distance deviations. Systematic distance errors include the "wiggling" error, the intensity-related distance error, the temperature-related distance error, as well as the integration-time-dependent distance error.[27] "Flying pixels," motion artifacts, and multiple reflections are additional sources of noise.[9] According to the manufacturers, the per-pixel accuracy of recent ToF cameras without postprocessing is about 1 cm.

The aim of this work is to develop and validate a method for increasing the accuracy of a ToF camera based system to a degree that is acceptable for radiotherapy applications (i.e., less than 2 mm (Ref. 28) for emerging techniques like stereotactic body radiotherapy—SBRT—or in the range of approximately 1.5–11 mm (Refs. 29 and 30) for coarse skin-marker laser-based patient positioning in conventional radiotherapy). By using just a single ToF camera for fast patient motion applications, the implementation of the solution becomes easier and less expensive. The ToF camera may be the primary source of motion information for those cases where surface data are considered sufficient for driving radiotherapy applications (e.g., in the whole breast), while it may provide complementary information when surface information alone is inadequate. In particular, it may be used to replace laser-based coarse patient positioning, allowing for a fast, automatic positioning workflow.

Patient positioning and tracking in radiotherapy require the determination of transformation parameters between a reference dataset and the currently acquired data of the patient. Therefore, we evaluated the patient positioning accuracy/robustness and also started to investigate the dynamic properties of the proposed solution for future applications to motion management. In gating applications, the respiratory signal could be extracted over a selected region of interest in the acquired field of view. In patient tracking applications, the acquired field of view is used to perform a real-time correction based on rigid registration (e.g., by the table or by multileaf collimator leaf positions), with the additional constraint that, if the displacement exceeds a given threshold, a beam off signal is sent to the linear accelerator (linac). Tumor tracking is beyond the scope of the current investigation, but it is one of the future directions of this work, especially developing a correlation model between internal markers and the external surface:[20] the possibility of using the entire or a part of the acquired surface to create a robust and accurate correlation model holds a significant potential over the current marker based correlation.

II. MATERIALS AND METHODS

The proposed framework is capable of reducing the ToF measurement noise by means of distance calibration and temporal Kalman filtering of the distance data. After these preprocessing steps, a pair of mutually rotated and translated surface datasets are matched using a two-step registration algorithm: rotation and translation invariant features are used for preregistration, followed by a fast version of the iterative closest point (ICP) algorithm.

II.A Processing pipeline

An overview of the processing pipeline is shown in Fig. 1. The input consists of two datasets denoted as a dynamic *live dataset* (acquired, e.g., during daily treatment delivery) and a static *reference dataset* (acquired, e.g., in the first treatment session or a surface automatically extracted from planning/pretreatment tomographic images). In this work, we have used only ToF inputs. Before the ToF datasets are registered, the data quality is improved by applying a distance correction followed by Kalman filtering. Furthermore, patient and background regions are separated in a body segmentation step.

II.A.1 Distance correction

In order to reduce systematic errors, a per-pixel correction of the distance values is performed *a priori* to improve the spatial accuracy. The distance correction is independent from the final position of the camera in the treatment room. However, it should be regarded, that the calibrated distance range (here: 1.0–1.4 m) approximately corresponds to the range of distance values in the area of interest in the treatment room. For the proposed calibration, range images of eight large boards with different infrared reflectivity were acquired at three known distances ({1.0; 1.2; 1.4} m) with a fixed integration time. The eight calibration boards were fabricated with manufacturing accuracy sufficient for correcting

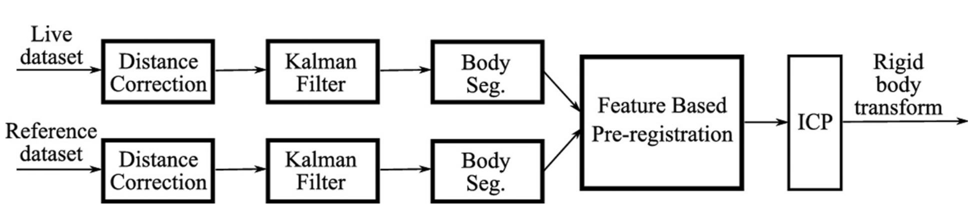

FIG. 1. Processing pipeline of the registration framework.

distance deviations in the submillimeter range. The systematic distance deviations were determined by comparing the ideal distance values (given by the calibration boards) to the ToF distance measurements in every pixel. The detected distance offsets were stored in lookup tables depending on the three distances, the eight intensities and all the pixel indices. During runtime, the errors are corrected by means of a Kochanek-Bartels spline interpolation[31] in the intensity domain and a linear interpolation in the distance domain. We also tried linear interpolation in the intensity domain. However, in our experiments the accuracy of the corrected distance values using spline interpolation was on average 0.5 mm better than the results obtained by linear interpolation.

In order to quantitatively measure the quality of the distance calibration, we acquire images of a flat board at three arbitrary distances lying between the distance samples used for the calibration. In each of the three acquisitions, the N measured 3D plane coordinates u_i that correspond to the flat board are approximated by a plane P_c using linear regression. Based on these measurements, a quality criterion G_c is computed as the mean of absolute distance (MAD) between the measured points and the approximated plane

$$G_c = \frac{1}{N}\sum_{i=1}^{N} |d(P_c, u_i)|. \qquad (1)$$

II.A.2 Kalman filtering

Quantum noise in the distance data is further reduced via Kalman filtering.[32] By employing this technique, we are capable of reducing the distance noise, while preserving possible motion, with very small delays. In our Kalman filtering step, we are processing the distance data pixel-wise over time according to an *a priori* acquired, bivariate measurement noise model, which depends on the amplitude and distance values. We use a 3D state vector, which incorporates the current distance, speed and acceleration (first and second derivatives of the distance). The derivatives with higher orders ($\geq$3) are considered process noise, which is adjustable and provides a trade-off between processing delay and noise reduction.

Unlike the most typically used noise removal techniques (e.g., median filtering or a filter with finite impulse response), Kalman filtering belongs to the class of prediction filters. As such, it is subject to an additional risk of overcompensating, especially when applied to very noisy data (like the distance values in a ToF pixel). However, the Kalman filter is regarded as an optimal linear filter, if the input data are corrupted with white, Gaussian distributed noise.[33] In our analyses of the PMD CamCube ToF camera (PMD Technologies, Siegen, Germany), we could measure approximately white, Gaussian noise in the raw distance values provided by the camera. This, coupled together with the time-domain filtering which is native to prediction filters, makes Kalman filtering particularly well-suited for this application. Figure 2 illustrates the effect of applying a Kalman filter to one pixel on the abdomen of a male, breathing patient: in general, Kalman filters do not involve a fixed delay in the output signal. Nevertheless, if the cross correlation between the unfiltered and the filtered signal is measured, we could approximate a mean delay of about two camera frames (ca. 65 ms). Low delays play a key role in real-time applications like gating/tracking, which demand a fast dynamic response of the system.

II.A.3 Body segmentation

After the temporal filtering step, all image points related to the patient are segmented in 3D for both the live and the reference datasets. In order to distinguish between patient and non-patient image areas, the treatment table is used as a separation plane. The table surface and coordinate system are detected in a one-time calibration step by rotating and shifting the treatment couch and observing with the ToF camera a checkerboard pattern lying on the table. Corresponding checkerboard corners are detected in the amplitude images in order to estimate the coordinate system of the table.

At runtime, the height of the separation plane is adapted according to the varying table height given by the couch

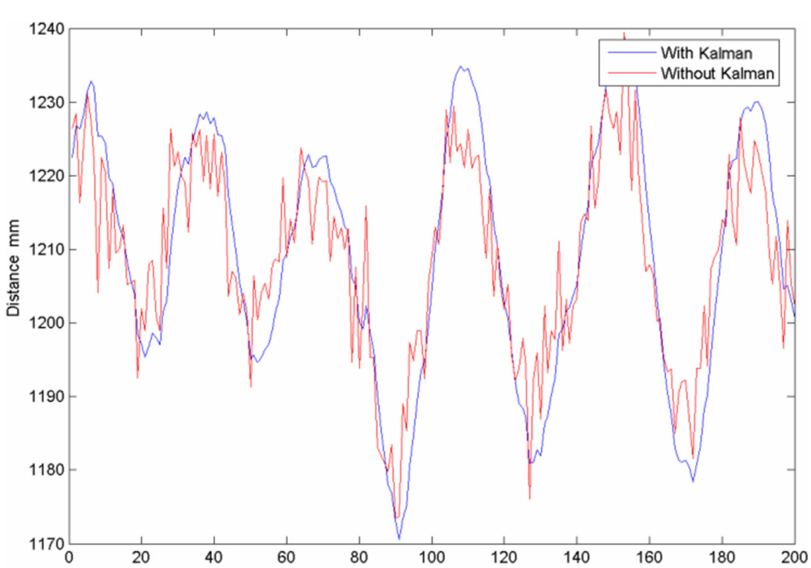

FIG. 2. Effects of Kalman filtering in the distance domain: Unfiltered signal and signal after Kalman filtering.

controller. Points lying below the table are assigned to the background. Points lying above the table plane are considered as body pixels, since no other objects (except the patient) should be positioned in the area between the ToF camera and the treatment couch. The 3D vertices corresponding to the body pixels are triangulated by subdividing blocks of four adjacent pixels into two triangles. Vertex normals are computed by averaging over the eight neighboring triangle normals of the current pixel.

II.A.4 Surface registration algorithms

Having segmented the body surfaces in the ToF datasets, a two-stage surface registration algorithm is performed. First, an initial registration estimate is obtained based on new translation- and rotation-invariant features. During the 2nd stage, the registration parameters are gradually refined through an accelerated ICP algorithm.

Step 1: A key component of our methodology is the introduction of novel surface features computed in the frequency domain. These features are particularly suitable for handling registration problems with large rotational/translational displacements and slightly overlapping surfaces. The feature computation for the live and the reference dataset differ in one aspect: in order to reduce the complexity, the features in the live dataset are only computed on a regularly subsampled grid in the index space (we use a reduction factor of 8 in both image dimensions). For the reference dataset, all points are considered. A single feature vector in the live/reference dataset is extracted as follows:

(i) Define a local neighborhood around the current point $\boldsymbol{p}_j$ by starting a region growing procedure on the surface mesh (with $\boldsymbol{p}_j$ as seed point). The region growing algorithm runs over all points within a fixed, maximum distance r_{max} to the seed point. In our implementation, a spatial threshold $r_{max} = 60$ mm is used.

(ii) Compute the mean normal $\boldsymbol{n}_z$ at the current point by means of an area weighted averaging over all triangle normals in the local neighborhood.

(iii) Define a local 3D coordinate system with the current point $\boldsymbol{p}_j$ as origin and the average normal $\boldsymbol{n}_z$ as z-axis. The x- and y-axis ($\boldsymbol{n}_x$, $\boldsymbol{n}_y$) are chosen arbitrarily, but perpendicular to each other and the z-axis. Transform all points in the local neighborhood into the new coordinate system. The new z-values can now be handled as the function values of a bivariate (x, y) function.

(iv) Resample the local function in polar coordinates [signal $g(r, \varphi)$] for several different radii r_n. In our system, we use a radius stepsize of 10 mm. The previous steps (i)–(iv) are defining a proper, local coordinate system (see Fig. 3 for a schematic illustration).

(v) Transform the circular, isoradial 1D signals [i.e., $g(r_k, \varphi)$] into Fourier space. Construct the feature vector for the current point by including a fixed number of low Fourier magnitude coefficients as well as a fixed

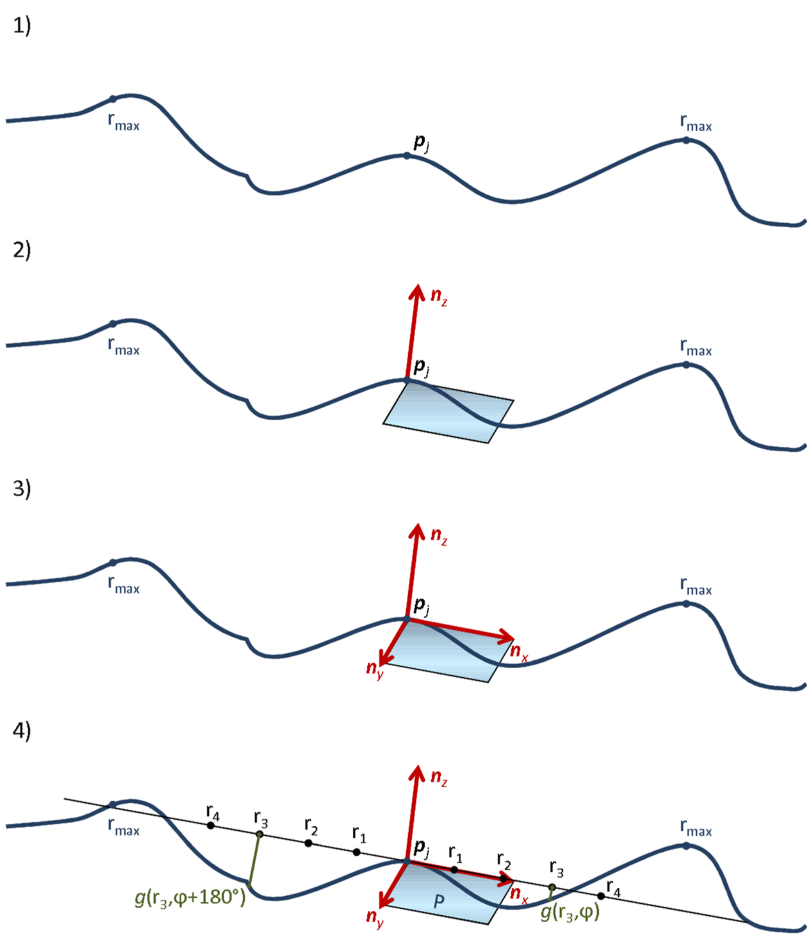

FIG. 3. Visualization of the steps (i)-(iv) of the feature computation algorithm: (1) Definition of local neighborhood, (2) normal computation, (3) coordinate system definition, (4) resampling in polar coordinates.

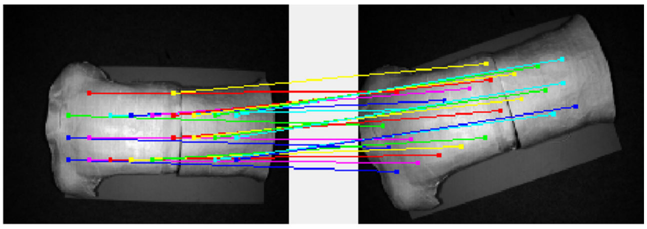

FIG. 4. Corresponding coordinates in live (right) and reference (left) dataset identified using the introduced features.

number of differential phase coefficients. For our evaluation, we chose 24 Fourier magnitude coefficients and their corresponding differential phase coefficients (that means 48 coefficients in total). The transformation into Fourier space is required to keep the features invariant to rotations around the normal axis of the surface patch. A translation of the radial 1D signal results in a linear phase in the Fourier domain.

With these feature vectors, corresponding points in the live and the reference dataset are computed by minimizing the n-dimensional Euclidean distance of a feature vector in the live dataset to all feature vectors of the reference dataset. An example output of the correspondence-finding step is illustrated in Fig. 4. False correspondences (e.g., due to high noise or occlusion effects) are removed by analyzing the internal, pairwise distances of corresponding points in the two datasets P (live dataset) and Q (reference dataset)

$$d_{\mathrm{RMS}}(j,P,Q)=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\|\boldsymbol{p}_i-\boldsymbol{p}_j\|-\|\boldsymbol{q}_i-\boldsymbol{q}_j\|)^2}, \quad (2)$$

where $(\boldsymbol{p}_j, \boldsymbol{q}_j)$ denotes the jth 3D point correspondence ($j \in \{1,\ldots,N\}$), N is the number of found correspondences and $\|.\|$ represents the Euclidean norm. Poor matches are eliminated by iteratively removing the correspondences with the worst d_{RMS} values, as long as all correspondences fulfill (2) with a threshold of 10 mm. Good correspondences (usually >80%) are then used to compute an initial rigid body transform according to

$$(\boldsymbol{R}_0,\boldsymbol{t}_0)=\operatorname*{argmin}_{\boldsymbol{R},\boldsymbol{t}}\sum_{j=1}^{N}\|\boldsymbol{R}\boldsymbol{p}_j+\boldsymbol{t},\boldsymbol{q}_j\|^2. \quad (3)$$

$\boldsymbol{R}_0$ and $\boldsymbol{R}$ represent 3×3 rotation matrices, $\boldsymbol{t}_0$ and $\boldsymbol{t}$ are 3D translation vectors.

Step 2: The initial transformation computed by means of the surface features is then refined with a modified version of the ICP algorithm. This second stage is based on the technique described in Ref. 34. Typically, space partitioning trees or projection strategies are used to determine closest points. In our framework, we precompute closest points on a dense voxel grid enveloping the whole reference dataset. Thereby, the complexity of determining the closest points is reduced to a single memory lookup. For the precomputation of closest points, the Euclidean distance is used as distance measure. In order to attenuate the influence of outliers, the computation of $\boldsymbol{R}$ and $\boldsymbol{t}$ includes a weighting $w_j^{(i)}$

$$w_j^{(i)}=\frac{1}{1+\left\|\boldsymbol{p}_j^{(i)}-\boldsymbol{q}_j\right\|^2}, \quad (4)$$

where $\boldsymbol{p}_j^{(i)}$ and $\boldsymbol{q}_j$ form a point correspondence in iteration i. Correspondences containing a point lying on a surface border are rejected to avoid problems with not completely overlapping datasets. Contrary to the step described in Eq. (2), a point-to-plane distance (d_{pl}) optimizer as in Eq. (4) is used in detecting the final rigid body transform between the point sets P and Q

$$(\boldsymbol{R}^{(i)},\boldsymbol{t}^{(i)})=\operatorname*{argmin}_{\boldsymbol{R},\boldsymbol{t}}\sum_{j=1}^{N}w_j^{(i)}d_{\mathrm{pl}}^2\left(\boldsymbol{R}\boldsymbol{p}_j^{(i)}+\boldsymbol{t},\boldsymbol{q}_j\right). \quad (5)$$

For small rotation angles (which are ensured by the preregistration), a linear estimator is available for detecting the rigid body transform.[35] The plane used in the distance measure d_{pl} is defined by the current reference point $\boldsymbol{q}_j$ and its surface normal. In order to find the closed form solution to Eq. (4), we implemented two variants that match the degrees of freedom of the most commonly-used treatment couches: one variant uses six degrees of freedom (three rotation angles around the table axes and a 3D translation vector) and one that employs four degrees of freedom (a rotation angle around the table normal and a 3D translation vector). The output parameters can be used to move the patient into the correct position.

II.B. Evaluation setup

The accuracy of the introduced surface registration system was evaluated on an ONCOR table (Siemens Healthcare, Erlangen, Germany) using a CamCube2 ToF camera with a field of view of $40\times40°$. In our test room the background illumination was kept constant. Nevertheless, the system shows low sensitivity regarding changing background light (if it is not direct sunlight), since the CamCube2 offers a feature called "suppression of background illumination" (SBI).[36]

For the evaluation a rigid plaster cast body phantom with infrared reflectivity similar to skin was put on the table.[37] The small gap between the thorax and the abdomen is due to the phantom's capability to simulate thoracic and abdominal respiration. The static phantom was moved along the three table axes and rotated around the normal axis of the table. The ground truth table rotation and translations provided by the table control were compared to the values computed by the registration framework. Since the utilized ONCOR treatment table allows only a rotation around the table's normal axis (isocentric rotations), we calculated the transformation between the two datasets using four degrees of freedom (rotation α around the **z**-axis, 3D translation $\boldsymbol{t}$). The accuracy of the ground truth values was 0.1° for the rotational component and 1 mm for each of the three translation parameters. For each table movement the target registration error (TRE) was computed. We define the TRE as the Euclidean distance between the true/computed translational component on the one hand [Eq. (5)] and the absolute angle difference [Eq. (6)] on the other hand

$$TRE_{\text{translation}} = \left\| \boldsymbol{t}_{\text{computed}} - \boldsymbol{t}_{\text{true}} \right\|, \tag{6}$$

$$TRE_{\text{rotation}} = \left\| \alpha_{\text{computed}} - \alpha_{\text{true}} \right\|. \tag{7}$$

Two modes were investigated: the positioning mode (simulating patient positioning phase) and the tracking mode (to evaluate the dynamic properties of the system to be used in future works on gating/tracking). After having found the correct transformation by the preregistration algorithm (positioning mode), the system switches to a "guided" tracking mode, where the rotation and translation parameters of the last frame are used as initializations for the ICP.

The ToF camera was mounted on the ceiling above the treatment table. It observed the table region between the linear accelerator gantry at an incident angle of 78° and at a mean working distance of 1.2 m. The mean working distance and incident angle in the evaluation was not varied because of the limited space in our linac test room. Rotations and translations along the table axis were analyzed independently of each other. For the feature based preregistration, 150–250 feature vectors (48-dimensional) were used (neighborhood radius: 6 cm, four sample radii at $\{1; 2; 3; 4\}$ cm).

In order to show the benefit of the introduced preregistration method, we compared the registration output of the processing pipeline (Fig. 1) to the output of another pipeline, where the feature based preregistration is replaced by centroid matching of the two surfaces. For this evaluation we applied transformations consisting of physical translation (i.e., moving the table) and virtual rotation (i.e., transforming the live dataset with a rotation matrix). Furthermore, we also analyzed the effect of cropping the body along a sagittal plane, so that only the right half of the live dataset goes into the registration with the complete reference dataset: this evaluates the robustness of the method to partial surface view.

FIG. 5. Color image of the plaster cast phantom used for the evaluation.

To investigate the dynamic properties of the proposed system, we applied a translational movement to the phantom in the lateral direction equal to 30 mm. We then evaluated the time needed for the system to detect the motion and to calculate the correct transformation compensating for the movement.

II.C Robustness to deformations

In order to test the robustness of our surface registration method to deformations, we virtually deformed the surface of our phantom (Fig. 5) by simulating both thoracic and abdominal respiration. We introduce a simple deformation model, which transforms the original surface vertices of the body according to the following equation:

$$z_{\text{trafo},1/2} = z_{\text{original}} + a_{\text{breath}} \cdot \cos\left(\min\left(\left\| \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{original}} - \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{thorax/abdomen}} \right\| \cdot \frac{1}{d_{\text{neighborhood}}} \cdot \frac{\pi}{2}, \frac{\pi}{2}\right)\right), \tag{8}$$

where $(x_{\text{original}}, y_{\text{original}}, z_{\text{original}})$ denote the original 3D coordinates of the surface and $z_{\text{trafo},1/2}$ represents the transformed coordinate after thoracic (number 1) or abdominal respiration (number 2), respectively. The x- and y-coordinates are left unchanged. The centroid of the thorax and abdomen are denoted by $(x_{\text{thorax}}, y_{\text{throax}}, z_{\text{thorax}})$ and $(x_{\text{abdomen}}, y_{\text{abdomen}}, z_{\text{abdomen}})$. All points within a maximum vertex-centroid distance $d_{\text{neighborhood}}$ are transformed, whereby points farther away from the thorax/abdomen centroid are less influenced by the deformation. The current amplitude of the respiration is represented by a_{breath}. In our evaluation we configured the virtual deformation by setting $d_{\text{neighborhood}} = 20$ cm and a_{breath} within the range [−20 mm, 20 mm]. We also evaluated combined thoracic and abdominal respiration, whereby the deformation is constructed by adding the z-deformation for the thorax and the abdomen

$$z_{\text{trafo,combined}} = z_{\text{trafo},1} + z_{\text{trafo},2}. \tag{9}$$

III. RESULTS

III.A Distance correction benefit

Figure 6 compares the quality criterion G_c defined in Sec. II A 1 with and without the application of distance correction. For the sake of completeness, the distance correction results are also visualized for the case of linear interpolation in the intensity domain (blue curve). Without distance correction G_c shows a mean (and standard deviation) of 1.90 ± 0.08 mm, with linear interpolation the mean is 0.94 ± 0.05 mm and with spline interpolation we achieve a value of 0.44 ± 0.02 mm.

As can be seen in Fig. 6, after distance correction the measured board points are on average around 1.5 mm closer

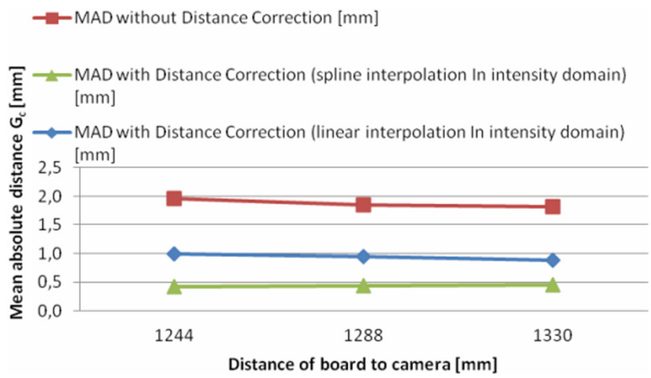

FIG. 6. Quality criterion G_c with and without distance correction.

to a perfect plane than before distance correction. That corresponds to an improvement of approx. 200%. Although the distance space is sampled quite coarsely for calibration (at 20 cm intervals), the distance correction improves the depth accuracy, so that the MAD between the approximated plane and the measured points is below 0.5 mm for each of the three board measurements.

III.B. Patient positioning accuracy and robustness

The detailed registration results are listed in Table I–V. Our registration method performs best for rotations around the z-axis as well as x- and y-translations. The z-axis accuracy was somewhat lower, since systematic distance deviations have the most influence on that coordinate axis. Our proposed distance correction (see Sec. II A) improves the z-accuracy of the positioning system (see Tables IV and V). On average, the Euclidean error decreased by 20.5%. Considering all measurements, we achieved a mean *TRE* of 1.62 ± 1.08 mm in the translational and $0.07° \pm 0.05°$ in the rotational component. An example visualization of the registration result is shown in Fig. 7.

The results of comparative evaluation with and without feature based preregistration as well as of partial surface view are summarized in Table VI and Fig. 8.

As shown in Table VI, centroid matching is not a sufficient preregistration technique for very large rotations (larger than 80°) and cropped datasets. The introduced feature based preregistration, however, allows for a consistently fast convergence of an accurate registration by means of ICP (25 iterations take about 1 s with our current configuration).

TABLE I. *TRE* for rotations around the isocentric table axis.

Ground truth	Euclidean error (mm)	Angular error (deg)
−9.0°	1.98	0.15
−6.0°	1.50	0.13
−3.0°	1.99	0.13
3.0°	0.75	0.01
6.0°	1.46	0.05
9.0°	1.96	0.11
Mean	1.61	0.10
SD	0.49	0.05

TABLE II. *TRE* for translations along the lateral table axis (x).

Ground truth	Euclidean error (mm)	Angular error (deg)
−90 mm	1.57	0.18
−60 mm	1.12	0.11
−30 mm	0.83	0.05
30 mm	0.53	0.07
60 mm	0.46	0.08
90 mm	1.42	0.10
Mean	0.99	0.10
SD	0.46	0.05

Furthermore, the output parameters of the successful registrations were exactly equal to the artificially applied parameters, representing the ground truth. For a rotation of 0° and centroid matching, the failed registration in the half body case appears like an outlier. However, this is due to the very low overlap after the centroid matching for 0°. When the dataset is slightly rotated, the overlap in the upper right body part increases and the registration succeeds for small rotations.

III.C. Robustness to deformations

Before applying the deformation described in Eq. (8), we applied a virtual rotation of 90° to our test live dataset, which is also translated by 60 mm in the y-direction with respect to the reference dataset. In order to further test the robustness of our methodology, we additionally limited the live dataset to the abdomen and thorax area, respectively. Even though the live dataset is modified via cropping, translation, rotation, and deformation, more than 50% of the correspondences we obtain from our feature based preregistration are good, i.e., they satisfy Eq. (2) with a threshold of 30 mm. All further processing steps are performed using only these well-corresponded points. Figure 9 shows an example registration result for a respiration amplitude a_{breath} of −20 mm in the live dataset and 0 mm in the reference dataset for both thoracic and abdominal respirations.

In a quantitative evaluation we compared the registration results using four different configurations of the live dataset: no respiration (static full torso), abdomnial respiration (abdomen only), thoracic respiration (thorax only), and combined abdominal and thoracic respirations (dynamic full torso). As evaluation criteria, we used the mean surface-to-surface distance (after applying the ground truth transformation and after

TABLE III. *TRE* for translations along the longitudinal table axis (y).

Ground truth	Euclidean error (mm)	Angular error (deg)
−90 mm	1.04	0.03
−60 mm	1.11	0.04
−30 mm	0.29	0.02
30 mm	0.95	0.02
60 mm	1.16	0.04
90 mm	1.16	0.04
Mean	0.95	0.03
SD	0.33	0.01

TABLE IV. *TRE* along table normal (z) without distance calibration.

Ground truth	Euclidean error (mm)	Angular error (deg)
−90 mm	5.43	0.07
−60 mm	3.76	0.15
−30 mm	1.70	0.01
30 mm	1.52	0.04
60 mm	3.62	0.12
90 mm	5.95	0.01
Mean	3.66	0.07
SD	1.83	0.06

the transformation computed by our registration system) and the numeric distance to the ground truth transformation parameters. The results are listed in Table VII.

In all cases, the surface-to-surface distance with the computed transformation outperforms its corresponding gain value using the ground truth parameters (0 mm, 60 mm, 0 mm, 90°). The reason for this effect is the partial compensation of the deformation through the rigid body transformation parameters.

We further evaluated the effects of missing the correct respiratory phase with abdominal respiration by using the deformation simulation itself. Therefore, we measured both the translational and the angular *TRE*s between the reference (respiration amplitude $a_{\mathrm{breath}} = 0$ mm) and the live datasets for different respiration amplitudes a_{breath}. The results are visualized in Fig. 10 (see also videos 1–4).

It can be observed, that large deformations in the abdomen (e.g., due to respiration) lead to a combination of translational and rotational deviations. For deformations within an amplitude range a_{breath} of [−20 mm, 20 mm], the final *TRE* is always below 25 mm in the translational (euclidean) and 0.45° in the rotational component.

III.D. Runtime aspects: Positioning and tracking modes

In positioning mode, the whole framework processed two surface datasets (both ≈10 000 points) in 561 ms (not parallelized) on a 2.26 GHz Intel Core2Duo CPU. More specifically, the processing time is broken down to 24 ms for preprocessing, 473 ms for feature based preregistration and 64 ms for ICP (involving ten iterations).

In "guided" tracking mode, the processing time for one frame is reduced to around 100 ms. Accordingly, in such a mode our system is running within 40% of the maximum possible framerate, since the utilized ToF camera provides a framerate of 25 fps. The system needed about 100 ms to detect motion and about 2 s to calculate the correct transformation compensating for the movement.

TABLE V. *TRE* along table normal (z) with distance calibration.

Ground truth	Euclidean error (mm)	Angular error (deg)
−90 mm	4.64	0.03
−60 mm	3.14	0.08
−30 mm	1.46	0.05
30 mm	1.14	0.03
60 mm	2.85	0.06
90 mm	4.24	0.09
Mean	2.91	0.06
SD	1.42	0.03

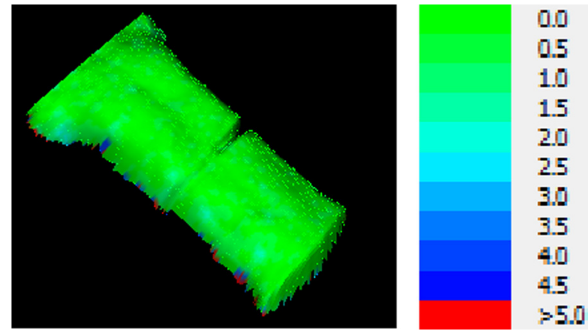

FIG. 7. Color coded distance between two registered body surfaces (mm).

IV. DISCUSSION

IV.A. Comparative evaluation

The proposed approach has several advantages over the previously published work on patient positioning using ToF technology. The approaches shown in Refs. 15 and 16 use ToF-ToF surface registration for patient positioning but do not involve distance calibration and preregistration of the surfaces. Our method improves the registration accuracy for the phantoms shown in Ref. 15 from 3.38 ± 2.00 to 1.62 ± 1.08 mm, even though the working distance is increased from 80 to 120 cm. Our rotations and translations were much larger than 1 cm. Thus, the registration accuracy of 0.74 ± 0.37 mm cited in Ref. 16 is not directly comparable to our results—the accuracy of the registration parameters decreases with increasing translations/rotations due to distance deviations of the ToF camera.

The most referenced surface features in literature are based on curvature[38,39] and spherical integral operations[40,41] When using such techniques, only a fractional amount of the local surface information is considered. More sophisticated approaches like *spin images*[42] or *shape contexts*[43] utilize very high dimensional vectors to represent the local geometry. Compared to these techniques, the features introduced in this work have the following key advantages. Contrary to *spin images*, no topological information in the signal is lost due to the computation of histograms. Compared to *shape contexts*, the rotational invariance of features is ensured by robust transformations instead of the more noise-sensitive estimation of a local reference axis.

Although the sensor specific, systematic distance deviations negatively influence the registration, we are still able to improve the positioning results by applying a distance calibration to the ToF camera. By defining a proper, automatic calibration protocol (e.g., by using the treatment couch to place the calibration boards in a number of different positions), the distance correction samples become denser, and thus, the registration accuracy is further improved.

TABLE VI. Comparison between feature based preregistration and centroid matching registrations when using virtual rotation in combination with physical translation of the phantom considering the full surface as well as half of the body surface.

Virtual rotation (deg)	Physical Y-translation (mm)	Only right half of body	Success with centroid matching (number of ICP iterations)	Success with feature based preregistration
0	60	no	yes (≈10 iterations)	yes (≈10 iterations)
0	60	yes	no	yes (≈10 iterations)
10	60	no	yes (≈15 iterations)	yes (≈10 iterations)
10	60	yes	yes (≈60 iterations)	yes (≈10 iterations)
20	60	no	yes (≈25 iterations)	yes (≈10 iterations)
20	60	yes	yes (≈45 iterations)	yes (≈10 iterations)
30	60	no	yes (≈35 iterations)	yes (≈10 iterations)
30	60	yes	yes (≈70 iterations)	yes (≈10 iterations)
40	60	no	yes (≈65 iterations)	yes (≈10 iterations)
40	60	yes	yes (≈ 00 iterations)	yes (≈10 iterations)
50	60	no	yes (≈ 85 iterations)	yes (≈10 iterations)
50	60	yes	yes (≈100 iterations)	yes (≈10 iterations)
60	60	no	yes (≈110 iterations)	yes (≈10 iterations)
60	60	yes	yes (≈100 iterations)	yes (≈10 iterations)
70	60	no	yes (≈145 iterations)	yes (≈10 iterations)
70	60	yes	yes (≈135 iterations)	yes (≈10 iterations)
80	60	no	yes (≈175 iterations)	yes (≈10 iterations)
80	60	yes	yes (≈200 iterations)	yes (≈10 iterations)
85	60	no	yes (≈215 iterations)	yes (≈10 iterations)
85	60	yes	no	yes (≈10 iterations)
86	60	no	no	yes (≈10 iterations)
86	60	yes	no	yes (≈10 iterations)
87	60	no	no	yes (≈10 iterations)
87	60	yes	no	yes (≈10 iterations)
88	60	no	no	yes (≈10 iterations)
88	60	yes	no	yes (≈10 iterations)
90	60	no	no	yes (≈10 iterations)
90	60	yes	no	yes (≈10 iterations)

As we use the patient surface for registration, respiratory motion is completely reflected by the transformation parameters returned by the registration pipeline. In the current work, we focus on the evaluation of rigid motion using ToF cameras as a new modality and its potential for patient positioning. Our evaluation results are comparable to existing systems like AlignRT, Galaxy, and Sentinel, which are also evaluated assuming rigid motion, and to recently published results on optical systems in radiotherapy/radiosurgery.[44] Interfraction (e.g., loss of weight) and intrafraction (e.g., respiration related) nonrigid components of the patient surface are likely to be present and to affect the results of positioning and tracking in real applications. Nevertheless, nonrigid information cannot be currently exploited, since even modern

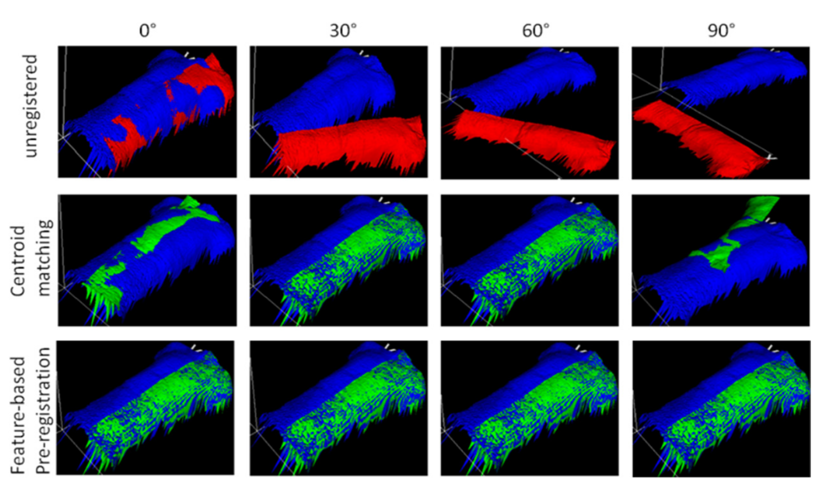

FIG. 8. Registration results for the half body case using rotation angles 0°, 30°, 60°, and 90° (unregistered live datasets in top row, centroid matching based registration in middle row, feature-based registration in bottom row). The unregistered and registered live datasets cover half of the torso, whereas the reference dataset covers the full torso.

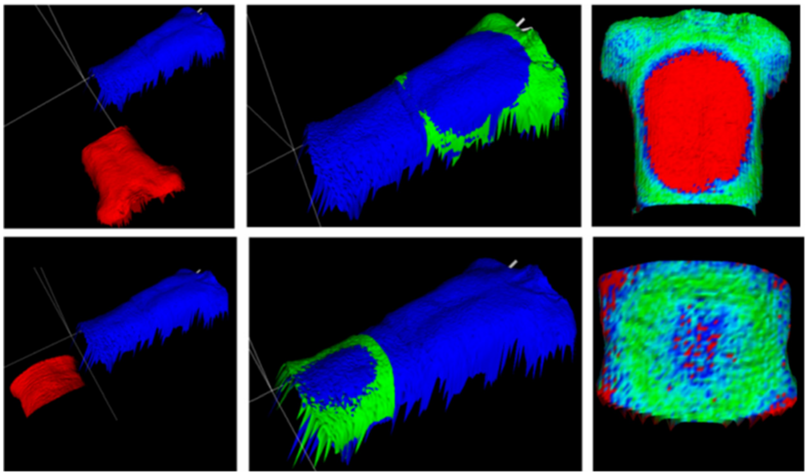

FIG. 9. Registration in the presence of thoracic (first row) and abdominal deformations (second row). Left: Untransformed live dataset of thorax/abdomen and reference surface; Middle: Transformed live dataset of thorax/abdomen and reference surface; Right: Distance map between live and reference surface after registration (same color code as used in Fig. 7).

radiotherapy devices are only able to deal with six degrees of freedom. Furthermore, advanced techniques for respiratory tracking like the Synchrony system (Accuray Inc., Sunnyvale, US) currently compensate only through rigid body transformation. Irregularities of the breathing cycle are not taken into account in our investigation. Future work will focus on the possibility of acquiring two time series of surface: one at the simulation phase and another prior to the treatment, calling for a series-to-series surface registration to minimize the effect of breathing irregularities. Also, in exploiting nonrigid components, it would be necessary to propagate the patient surface deformation to the internal structure surrounding the target, which is the actual focus of the radiotherapy treatment. This could be achieved via, for example, finite element method techniques. This approach is clearly interesting, but it is beyond the scope of the current investigation. For a more qualitative evaluation of the nonrigid error, we proposed a color coded visualization (see Figs. 6 and 7), where errors larger than 5 mm are highlighted in red: this provides a tool for possible large mismatch visualization.

In practice, potential deformations of the patient surface need to be considered to evaluate possible target misalignment. By simulating possible thoracic and abdominal deformations (e.g., generated by respiration) and their combination, we could show that the developed feature based preregistration still detects a sufficient number of good point correspondences that can act as an adequate initialization for the refining ICP step. This is also valid, even if the input dataset covers only a part of the reference dataset (in our evaluation only the abdomen) and is additionally rotated and translated. Nevertheless, similar results are also expected when using other optical systems.

To limit the effect of surface deformation, it could be useful to restrict the surface acquisition to a limited area. On the other hand, especially in smooth surface areas with little details like the abdominal region, registration accuracy is linked to the extension of the surfaces. Thus, the optimal surface area to be considered is a trade-off between the stability of the registration parameters and the management of local surface deformations. In our experiments, minimal surface extension should not be less than 500 cm^2.

Given that the ground truth of this investigation was provided by the table coordinates of our test linac table consisting of only four degrees of freedom (no pitch and roll), it was not possible to investigate the effect of the other two degrees of freedom on the surface registration accuracy. Nevertheless, these additional degrees of freedom are significant in current radiotherapy treatments,[45] and should be considered in further investigations.

IV.B. Clinical applications

The proposed method will likely play an important role in radiotherapy, similarly to other surface-based techniques.[1] The experimental evaluation using the cropped half body surface with a table rotation of 90° was designed for testing the possibility of monitoring the position of the patient when noncoplanar treatments are applied (e.g., couch 90°, gantry 35°). These treatments are likely to become more common due to the current trend toward hypofractionated treatments, where noncoplanarity will play an important role. The goal

TABLE VII. Quantitative evaluation of the registration output in the presence of deformations produced by virtual respiration and comparison to corresponding gain values of the ground truth transformation.

Respiration type a_{breath} Region	None 0 mm Static full torso	Abdominal −20 mm Abdomen	Thoracic −20 mm Thoracic	Combined −20 mm Dynamic full torso
Mean surface distance after applying ground truth transformation (mm)	1.15	7.34	7.50	7.91
Mean surface distance after applying computed transformation (mm)	0.99	2.46	5.57	4.71
Distance to ground truth translation after registration (x, y, z) (mm)	(−0.60,0.58,0.63)	(2.18,0.55, 14.69)	(−1.99, 4.08, 3.71)	(−0.19, −8.18, 8.66)
Absolute distance to ground truth rotation after registration (°)	0.06°	0.38°	0.16°	0.01°

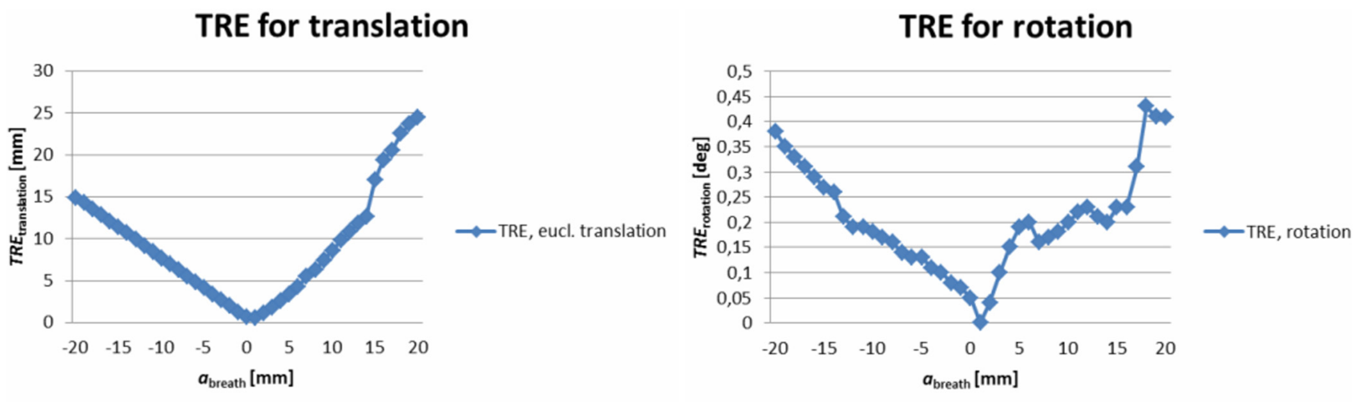

FIG. 10. Effects of different respiration amplitudes on the translational and rotational TRE (enhanced online). Video 1 [URL: http://dx.doi.org/10.1118/1.3664006.1]; Video 2 [URL: http://dx.doi.org/10.1118/1.3664006.2]; Video 3 [URL: http://dx.doi.org/10.1118/1.3664006.3]; Video 4 [URL: http://dx.doi.org/10.1118/1.3664006.4].

would then be to correlate the position of the patient at the couch at 90° with the corresponding one at the couch at 0°. Unfortunately, CT images of patients are available only with the couch at 0° (with the couch at 90° no CT image can be acquired due to a gantry/couch collision). Thus, the possibility of correlating the surface of the patient when the couch is at 90° to the CT images where the couch is at 0° will be an important tool in assuring that the target is still in the correct position. This correlation can be based on the two surfaces acquired when the couch is set to 0° and 90°. Patient surface could be acquired directly with the ToF camera or automatically extracted from the CT images. Note that, when the couch is at 90°, the ToF camera will be able to capture a partial surface view of the patient only from one side, so it is essential to test the accuracy and the robustness of the method in such an extreme condition. The success of the feature based registration over the ICP-only approach indicates the importance of using such a method for large displacements like in noncoplanar treatments. The constistently fast convergence (always less than 1 s) when using the feature based preregistration allows the interbeam time to be minimized, decreasing in turn the probability of patient movement. Furthermore, the exhibited accuracy in the experiments of the artificially applied rotations and translations demonstrates the precise correlation of the two patient positions when the couch is at 0° and 90°.

One could use a patient positioning system based on CT surface extraction and a 90° rotation. However, CT acquisition is usually performed before the treatment starts and assumes no movement during treatment delivery. Thus, the use of a ToF surface acquisition at 0° and 90° angles allows for a more accurate control of patient position prior to the couch rotation. Also, the most widely used cone beam CT (CBCT) acquisitions used in IGRT suffer from respiration artefacts which affect the accuracy of the registration. Conversely, rapid ToF acquistions are able to smooth out these artefacts, allowing also for a series-to-series 4D registration.

Physiological deformations, which may happen due to weight loss over the treatment course or respiration, are an important source of error and must be carefully considered. Our results show that even when simulating large deformations and partial surface view, the system is still competitive, in terms of accuracy, over laser-based coarse positioning in a large range of cases. The color coded visualization supports the user in highlighting the mismatch and triggering action toward complementary image acquisition such as CT. The series-to-series registration we will investigate in our future work will help minimize the error related to different respiration phases.

The dynamic properties of the proposed system suggest promising applications to gating/tracking. In fact, the accuracy of the system (TRE of 1.62 ± 1.08 mm in the translational component) as well as the very small time delay (about two frames) support the acquisition of surfaces with sufficient spatial and temporal resolutions. The short time delay needed to detect motion allows for a fast beam off switch during radiation therapy dose delivery, minimizing possible errors related to patient motion: this minimization becomes essential due to currently implemented fast treatments. After motion detection, the 2 s time needed to find the correct transformation appears a reasonably short time frame to perform patient tracking. Another possible implementation could consist of decreasing the dose rate at the time of motion detection waiting for the evaluation of the motion correction needed: if larger than a given threshold, then, the beam is switched off; otherwise the beam stays on at low dose rate and reaches again the maximum dose rate value when the motion compensation has been applied. Nevertheless, in clinical applications patients are immobilized through masks or vacuum cushions to minimize large movements. In addition to patient tracking, the dynamic information provided by the proposed system can be used to either extract the respiratory signal for gating or tumor tracking creating a correlation between internal markers and external surfaces (or points selected on a surface). This will ultimately allow the development of a motion management system like Synchrony or Real-time Position Management (Varian Medical Systems, Palo alto, US), with the additional advantange of either selecting points across the surface or using the entire surface itself.

In general, for applications like brain[46] or breast[47] radiotherapy as well as thoraco-abdominal radiotherapy, surface-based techniques are likely to play an important role[10] for patient positioning and motion management. In more general radiotherapy applications, the proposed system would be mainly used as an additional modality that improves the accuracy of other imaging data. It also has the potential of possibly reducing the number of images that are needed and therefore the patient dose. Although our framework was primarily developed for radiotherapy, it may also be applicable to other medical fields requiring mono-modal or multimodal surface registration.

V. CONCLUSIONS

We have developed and evaluated a ToF based system for automatic surface registration in medical applications, with a particular focus on radiotherapy. The ToF surface data are improved by the incorporation of a distance calibration as well as Kalman filtering. Furthermore, we introduced a new, feature based method for preregistering two surface datasets. This preregistration is used as a very accurate initialization for the subsequent ICP optimization. This approach is likely to have a significant impact on cases involving large displacements in patient positioning, like in noncoplanar treatments. We evaluated the presented framework on a typical patient positioning application. The proposed system achieved an accuracy of 1.62 ± 1.08 mm in the translational component and $0.07^\circ \pm 0.05^\circ$ in the rotational component. Whereby, both of these uncertainties are caused by a combination of the imaging system (ToF camera with preprocessing) and the automated registration process. The system is characterised by promising dynamic properties which make it a good candidate for use in gating/tracking applications. Deformations may play an important role in surface-based target misalignment and particular care must be taken to assure an appropriate use of optical systems in clinical routine. Last, but not least, the introduced Fourier features themselves can also be computed on surfaces extracted from different modalities (e.g., CT or MRI), which will be one of the goals of our future work.

ACKNOWLEDGMENTS

The work was supported by the International Max Planck Research School for Optics and Imaging (Erlangen, Germany). The evaluation on the linear accelerator was gratefully approved by Siemens Healthcare (Erlangen, Germany).

[a]Electronic mail: simon.placht@metrilus.de

[1]D. P. Gierga, M. Riboldi, J. C. Turcotte, G. C. Sharp, S. B. Jiang, A. G. Taghian, and G. T. Y. Chen, "Comparison of target registration errors for multiple image-guided techniques in accelerated partial breast irradiation," Int. J. Radiat. Oncol., Biol., Phys. **70**(4), 1239–1246 (2008).
[2]D. R. Olsen and D. I. Thwaites, "Now you see it… Imaging in radiotherapy treatment planning and delivery," Radiat. Oncol. **85**(2), 173–175 (2007).
[3]J. J. Lagendijk, B. W. Raaymakers, A. J. Raaijmakers, J. Overweg, K. J. Brown, E. M. Kerkhof, R. W. van der Put, B. Hårdemark, M. van Vulpen, and U. A. van der Heide, "MRI/linac integration," Radiother. Oncol. **86**(1), 25–29 (2008).
[4]M. Ghilezan, D. Yan, and A. Martinez, "Adaptive radiation therapy for prostate cancer," Semin. Radiat. Oncol. **20**(2), 130–137 (2010).
[5]P. Keall, S. Vedam, R. George, C. Bartee, J. Siebers, F. Lerma, E. Weiss, and T. Chung, "The clinical implementation of respiratory-gated intensity-modulated radiotherapy," Med. Dosim. **31**(2), 152–162 (2006).
[6]T. H. Wagner, S. L. Meeks, F. J. Bova, W. A. Friedman, T. R. Willoughby, P. A. Kupelian, and W. Tome, "Optical tracking technology in stereotactic radiation therapy," Med. Dosim. **32**(2), 111–120 (2007).
[7]Y. Suh, E. Weiss, H. Zhong, M. Fatyga, J. V. Siebers, and P. J. Keall, "A deliverable four-dimensional intensity-modulated radiation therapy-planning method for dynamic multileaf collimator tumor tracking delivery," Int. J. Radiat. Oncol., Biol., Phys. **71**(5), 1526–1536 (2008).
[8]C. Ozhasoglu, C. B. Saw, H. Chen, S. Burton, K. Komanduri, N. J. Yue, S. M. Huq, and D. E. Heron, "Synchrony—Cyberknife respiratory compensation technology," Med. Dosim. **33**(2), 117–123 (2008).
[9]D. M. Cash, T. K. Sinha, R. L. Galloway, M. I. Miga, W. C. Chapman, H. Terawaki, and B. M. Dawant, "Incorporation of a laser range scanner into image-guided liver surgery: Surface acquisition, registration, and tracking," Med. Phys., **30**(7), 1671–1682 (2003).
[10]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. **37**(10), 5421–5433 (2010).
[11]http://www.lap-america.com/pdf_files/Galaxy0707.pdf.
[12]http://www.c-rad.info/document/pallotta.pdf.
[13]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., **32**(9), 2753–2762 (2005).
[14]Y. F. Li and S. Y. Chen, "Automatic recalibration of an active structured light vision system," IEEE Trans. Rob. Autom., **19**(2), 259–268 (2003).
[15]C. Schaller, A. Adelt, J. Penne, and J. Hornegger, "Time-of-flight sensor for patient positioning," *Proceedings of SPIE on Medical Imaging*, edited by M. I. Miga, and K. H. Wong (SPIE, Lake Buena Vista, Florida, 2009), Vol. 7261, No. 8.
[16]S. Placht, C. Schaller, M. Balda, A. Adelt, C. Ulrich, and J. Hornegger, "Improvement and evaluation of a time-of-flight based patient positioning system," *Bildverarbeitung für die Medizin 2010*, edited by H.-P. Meinzer, T. M. Deserno, H. Handels, and T. Tolxdorff (Springer Berlin/Heidelberg, Aachen, Germany, 2010), pp. 177–181.
[17]K. Müller, S. Bauer, J. Wasza, and J. Hornegger, "Automatic multi-modal ToF/CT organ surface registration," *Bildverarbeitung für die Medizin 2011*, edited by H. Handels, J. Ehrhardt, T. M. Deserno, H. P. Meinzer, and T. Tolxdorff, (Springer Berlin/Heidelberg, Lübeck, Germany, 2011), pp. 154–158.
[18]Z. Xu, R. Schwarte, H. Heinol, B. Buxbaum, and T. Ringbeck, "Smart pixel—Photometric mixer device (PMD)/new system concept of a 3D-imaging-on-a-chip," *5th International Conference on Mechatronics and Machine Vision in Practice* (Nanjing, China, 1998), pp. 259–264.
[19]C. Schaller, C. Rohkohl, J. Penne, M. Stürmer, and J. Hornegger, "Inverse C-arm positioning for interventional procedures using real-time body part detection," *MICCAI 2009, Part I, LNCS 5761*, edited by G.-Z. Yang, D. Hawkes, D. Rueckert, A. Noble, and C. Taylor, (Springer Berlin/Heidelberg, London, UK, 2009), pp. 549–556.
[20]J. Jacq, C. Schwartz, V. Burdin, C. Lefevre, O. Rémy-Néris, C. Roux, and R. Gerard, "Building and tracking root shapes," IEEE Trans. Biomed. Eng., **57**(3), 696–707 (2010).
[21]J. Penne, K. Höller, M. Stürmer, T. Schrauder, A. Schneider, R. Engelbrecht, H. Feußner, B. Schmauss, and J. Hornegger, "Time-of-flight 3-D endoscopy," *MICCAI 2009, Part I, LNCS 5761*, edited by G.-Z. Yang, D. Hawkes, D. Rueckert, A. Noble, and C. Taylor, (Springer Berlin/Heidelberg, London, UK, 2009), pp. 467–474.
[22]C. Schaller, J. Penne, and J. Hornegger, "Time-of-flight sensor for respiratory motion gating," Med. Phys. **35**(7), 3090–3093 (2008).
[23]H. Fayad, T. Pan, C. Roux, C. Cheze-Le Rest, O. Pradier, J. F. Clement, and D. Visvikis, "A patient specific model based on 4D CT data and a time of flight camera (TOF)," *IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS-MIC)* (IEEE, Orlando, Florida, 2009), pp. 2594–2598.
[24]J. F. Clement, H. Fayad, M. Lamard, O. Pradier, C. Cheze-Le Rest, and D. Visvikis, "Time of flight camera (TOF) for contact-less and marker-less 3D respiratory motion detection," *Society of Nuclear Medicine Annual Meeting* (Society of Nuclear Medicine, Toronto, Canada, 2009).
[25]T. Wentz, H. Fayad, J. F. Clement, J. Savéan, M. Hatt, and D. Visvikis, "Extraction and evaluation of anatomical patient surface and associated respiratory motion with a time-of-flight (ToF) camera," *Society of Nuclear*

Medicine Annual Meeting (Society of Nuclear Medicine, San Antonio, Texas, 2011).

[26]H. Fayad, T. Pan, J. F. Clement, and D. Visvikis, "Correlation of respiratory motion between external patient surface and internal anatomical landmarks," Med. Phys., **38**(6), 3157–3164 (2011).

[27]A. Kolb, E. Barth, R. Koch, and R. Larsen, "Time-of-flight sensors in computer graphics," *Eurographics (State-of-the-Art Report)*, edited by M. Pauly, and G. Greiner, (Eurographics, Munich, Germany, 2009), pp. 119–134.

[28]S. H. Benedict, K. M. Yenice, D. Followill, J. M. Galvin, W. Hinson, B. Kavanagh, P. Keall, M. Lovelock, S. Meeks, L. Papiez, T. Purdie, R. Sadagopan, M. C. Schell, B. Salter, D. J. Schlesinger, A. S. Shiu, T. Solberg, D. Y. Song, V. Stieber, R. Timmerman, W. A. Tomé, D. Verellen, L. Wang, and F. F. Yin, "Stereotactic body radiation therapy: The report of AAPM Task Group 101," Med. Phys. **37**(8), 4078–4101 (2010).

[29]M. Wang, R. Rohling, C. Duzenli, B. Clark, and N. Archip, "Evaluation of targeting errors in ultrasound-assisted radiotherapy," Ultrasound Med. Biol. **34**(12), 1944–1956 (2008).

[30]C. F. Serago, S. J. Buskirk, T. C. Igel, A. A. Gale, N. E. Serago, and J. D. Earle, " Comparison of daily megavoltage electronic portal imaging or kilovoltage imaging with marker seeds to ultrasound imaging or skin marks for prostate localization and treatment positioning in patients with prostate cancer," Int. J. Radiat. Oncol., Biol., Phys. **65**(5), 1585–1592 (2006).

[31]D. H. U. Kochanek and R. H. Bartels, "Interpolating splines with local tension, continuity, and bias control," SIGGRAPH '84: Proceedings of the 11th annual conference on computer graphics and interactive techniques **18**(3), 33–41 (1984).

[32]G. Welch and G. Bishop, "An introduction to the Kalman filter," Technical Report 95-041 (Department of Computer Science, University of North Carolina at Chapel Hill, Chapel Hill, NC, 2006) 95–041.

[33]R. E. Kalman, "A new approach to linear filtering and prediction problems," Trans. ASME J. Basic Eng., **82**(1), 35–45 (1960).

[34]A. M. Bronstein, M. M. Bronstein, and R. Kimmel, "Numerical geometry of non-rigid shapes," 1st ed., *Monographs in Computer Science* (Springer, New York, 2008).

[35]S. Rusinkiewicz, M. Levoy, "Efficient variants of the ICP algorithm," *Proceedings of the Third International Conference on 3D Digital Imaging and Modeling* (IEEE Computer Society, Québec City, Canada, 2001), pp. 145–152.

[36]T. Möller, H. Kraft, J. Frey, M. Albrecht, and R.Lange, "Robust 3D Measurement with PMD Sensors," *Proceedings of the 1st Range Imaging Research Day at ETH*, Zürich (2005).

[37]C. Ulrich, C. Schaller, J. Penne, and J. Hornegger, "Evaluation of a time-of-flight based respiratory motion management system," *Bildverarbeitung für die Medizin*, edited by H.-P. Meinzer, T. M. Deserno, H. Handels, and T. Tolxdorff, (Springer Berlin/Heidelberg, Aachen, Germany, 2010), pp. 152–156.

[38]G. Agam and X. Tang, "A sampling framework for accurate curvature estimation in discrete surfaces," IEEE Trans. Vis. Comput. Graph., **11**(5), 573–583 (2005).

[39]M. Meyer, M. Desbrun, P. Schröder, and A. Barr, "Discrete differential geometry operators for triangulated 2-manifolds," *International Workshop on Visualization and Mathematics*, edited by H.-C. Hege, and K. Polthier, (Springer Berlin/Heidelberg, Berlin, Germany, 2002).

[40]N. Gelfand, N. J. Mitra, L. J. Guibas, and H. Pottmann, "Robust global registration," in *SGP '05, Proceedings of the third Eurographics symposium on Geometry processing*, edited by M. Desbrun, and H. Pottmann, (Eurographics Association, Vienna, Austria, 2005), pp. 197–206.

[41]S. Manay and B.-W. Hong, "Integral invariants for shape matching," IEEE Trans. Pattern Anal. Mach. Intell. **28**(10), 1602–1618 (2006).

[42]A. E. Johnson and M. Hebert, "Using spin images for efficient object recognition in cluttered 3D scenes," IEEE Trans. Pattern Anal. Mach. Intell., **21**(5), 433–449 (1999).

[43]S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," IEEE Trans. Pattern Anal. Mach. Intell., **24**(4), 509–522 (2002).

[44]P. G. Maxim, B. W. Loo, Jr, J. D. Murphy, K. P. Chu, A. Hsu, and O. T. Le, "On-board imaging validation of optically guided stereotactic radiosurgery positioning system for conventionally fractionated radiotherapy for paranasal sinus and skull base cancer," Int. J. Radiat. Oncol., Biol., Phys. **81**, 1153 (2011).

[45]M. F. Spadea, G. Baroni, M. Riboldi, R. Luraschi, B. Tagaste, C. Garibaldi, G. Catalano, R. Orecchia, and A. Pedotti, "Benefits of six degrees of freedom for optically driven patient set-up correction in SBRT," Technol. Cancer Res. Treat. **7**(3), 187–195 (2008).

[46]G. Li, Å. Ballangrud, L. Kuo, H. Kang, A. Kirov, M. Lovelock, Yo. Yamada, J. Mechalakos, and H. Amols, "Motion monitoring for cranial frameless stereotactic radiosurgery using video-based three-dimensional optical surface imaging," Med. Phys. **38**(7), 3981–3994 (2011).

[47]M. Riboldi, G. Baroni, R. Orecchia, and A. Pedotti, " Enhanced surface registration techniques for patient positioning control in breast cancer radiotherapy," Technol. Cancer Res. Treat. **3**(1), 51–58 (2004).