Time-of-Flight sensor for patient positioning

Christian Schaller, Andre Adelt, Jochen Penne, Joachim Hornegger

Chair of Pattern Recognition, Department of Computer Science, Friedrich-Alexander University Erlangen-Nuremberg, Martensstr. 3, 91058 Erlangen, Germany

ABSTRACT

In this paper we present a system that uses Time-of-Flight (ToF) technology to correct the position of a patient in respect to a previously acquired reference surface. A ToF sensor enables the acquisition of a 3-D surface model containing more than 25,000 points using a single sensor in real time. One advantage of this technology is that the high lateral resolution makes it possible to accurately compute translation and rotation of the patient in respect to a reference surface. We are using an Iterative Closest Point (ICP) algorithm to determine the 6 degrees of freedom (DOF) vector. Current results show that for rigid phantoms it is possible to obtain an accuracy of 2.88 mm and 0.28° respectively. Tests with human persons validate the robustness and stability of the proposed system. We achieve a mean registration error of 3.38 mm for human test persons. Potential applications for this system can be found within radiotherapy or multimodal image acquisition with different devices.

Keywords: Therapy Planning, Patient Positioning, Surface Matching, Surface Registration, Time-of-Flight camera, ToF, Radiotherapy

1. INTRODUCTION

Patient positioning is a crucial issue in the field of radiotherapy and multimodal imaging. The number of annually diagnosed cancer cases has been continuously growing since many years. Due to this increasing number of patients there is a demand for effective and economic treatment processes. In a common workflow, a planning CT scan of the patient is acquired first in order to plan the treatment process. This planning CT is usually acquired a few days or even weeks prior to the treatment and provides the basis for the treatment plan. This treatment plan created at the beginning of the therapy has to be valid for each of the following therapy sessions. Physicians define so called target volumes in order to plan the radiation procedure. In cases where a fractional treatment has to be applied the patient has to be positioned accurately before every treatment session. Otherwise, the treatment plan would not be valid anymore.

Recently, a very interesting technology originally driven by the automotive industry has risen. This technology is called Time-of-Flight (ToF). As the technical possibilities of this emerging technology are quite promising we suggest a system to provide an innovative solution to position patients using a ToF camera sensor. ToF sensors enable the direct acquisition of 3-D surface information of more than 25,000 3-D points in real-time ($>$ 15 Hz).[1] Lately, applications like gesture recognition[2] or tracking applications[3] are using ToF sensors in the field of e.g.,gaming-, security-, automotive industry. We are convinced that ToF sensors can also contribute to improve medical applications by using the sensor as an additional imaging device. Though until now, 3-D endoscopy[4] and respiratory motion gating[5] are the only medical application where ToF sensors are supposed to be deployed. For showing further benefits of ToF sensors within medical applications we suggest a system using a ToF sensor to correct the position of a patient in respect to a previously acquired reference surface.

ToF sensors emit an incoherent light signal to actively illuminate the patient, where the light signal is modulated by a cosine-shaped signal of frequency f. The emitted light is usually part of the non-visible area near the infrared spectral range at about 780 nm. This light signal is travelling with the constant speed of light through the surrounding medium and is reflected by the patient. For computing the distance d, the estimated phase-shift ϕ (in rad) between the emitted and the reflected light signal is used:

Further author information: (Send correspondence to Christian Schaller)
Christian Schaller: E-mail: christian.schaller@informatik.uni-erlangen.de, Telephone: +49 9131 85 27275

Figure 1. Left: ToF camera MESA Imaging AG. Right: ToF camera PMDTec GmbH.

$$d = \frac{c}{2f} \cdot \frac{\phi}{2\pi} \quad (1)$$

Due to the periodicity of the cosine-shaped modulation signal this equation is only valid for distances smaller than $c/(2f)$. Common modulation frequencies for todays available ToF sensors (see Fig. 1) are about 20 MHz, which causes an upper limit for the observable distance of this sensors of approx. 7.5 m. This distance is sufficient for the application we want to address. Additionally, ToF sensors also provide intensity values, representing the amount of light sent back from a specific point. Detailed information about the working principle of ToF cameras can be found in Xu et al.[1]

In the following we will briefly emphasize on the key benefits of ToF cameras. The most important advantage of ToF cameras over other common used technologies like e.g., stereo vision is that ToF sensors are all-solid-state off-the-shelf technology. This is mostly due to the automotive and consumer electronics industry are heavily interested in this emerging technology. As both of these industries are mass markets there soon will be a market for mass production of ToF sensors. Currently a typical ToF camera is available for about 5,000 Euros (approximately 7,500 USD). As some applications for ToF cameras are already established in mass markets there will be a significant decrease in manufacturing costs for such cameras. Furthermore, there are also no calibration steps at all necessary to operate a ToF based system. Such systems can be used as a Plug and Play application. ToF cameras are highly portable and very compact and small and can easily be used and integrated in different existing systems without any complex setup procedures.

Currently, there are a lot of different solutions available to help solving the task of patient positioning, but each of these systems has some major drawbacks. Some of these systems are presented in the following.

2. STATE OF THE ART

Today most commonly used and wide spread positioning systems in clinical workflows are based on wall mounted laser systems. These kind of systems rely on skin tattoos which are manually aligned with laser pointers. The laser pointers are projected on the patient skin and the treatment table has to be aligned manually. As these systems are installed in almost every hospital, people are quite familiar in dealing with such a system as it is well established. Also the costs for such a system are relatively low. The major drawback of a system based on laser pointer is the lack of accuracy. The system is very inaccurate and a mean positioning error of 4-8 mm[6] is quite common. Another disadvantage is the setup time to manually align each patient.

Other quite common systems are based on external fiducial or index markers. Such systems are using about 5-10 reflective markers as surface landmarks. One or more cameras are used to detect and track these markers. These markers can also be used to track respiratory motion. An example for such a system is the NOVALIS system from BrainLAB, Germany*. One drawback of this method is the setup procedure which is vulnerable to setup errors.

Systems for image guided radiotherapy are using images to verify the position of the patient. Usually a very good accuracy can be achieved by this systems, but additional irradiation is necessary to acquire the

*http://www.brainlab.com

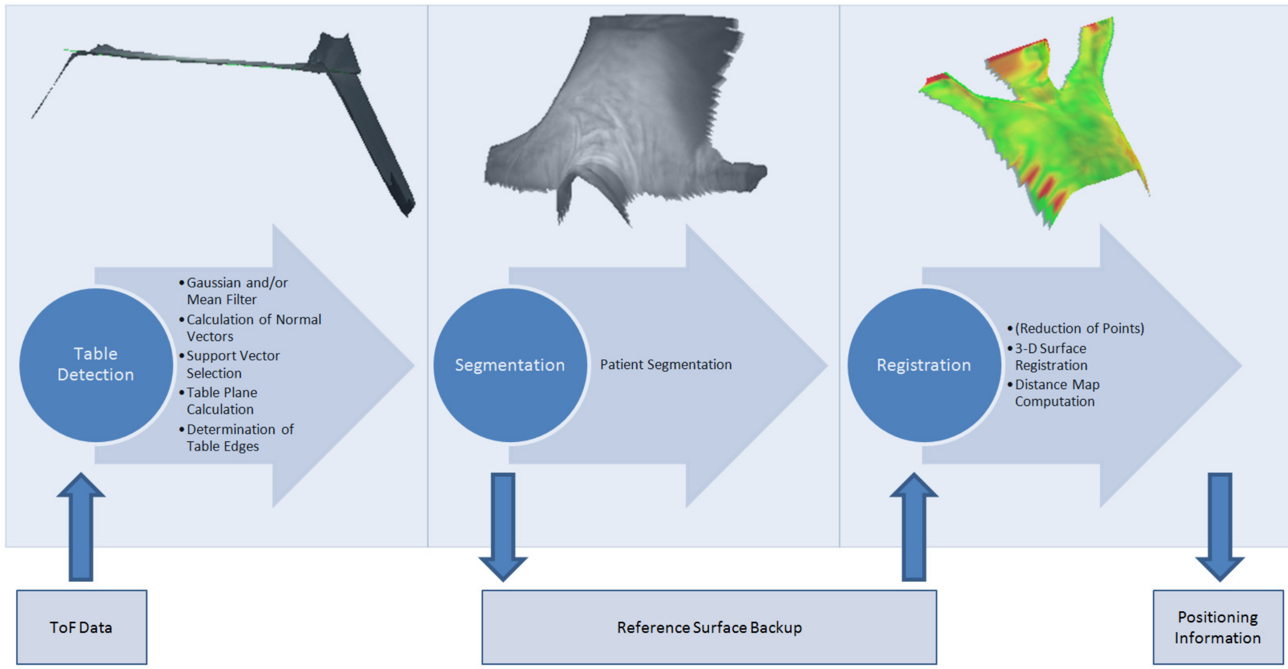

Figure 2. Schematic overview of the proposed method. From left to right: determination of the treatment table plane (side view), segmentation of the patient, registration result of the remaining body surface (the color-coding shows the error, ranging from 0 mm green to 10 mm red)

images. While those conventional methods are either inaccurate or require additional exposure to the patient, latest developments are based on optical measurements providing good results.[7] An example for such an optical surface based technology is the VisionRT system from VisionRT, United Kingdom[†].

Surface based solutions are usually quite expensive, because special hardware, like lasers or very accurate stereo vision cameras are needed. Furthermore, these systems have to be calibrated very accurately. To address this issue we present a method to determine patient translation and rotation in 3-D by the acquisition of the patients body surface during the planning CT and again right before or during each treatment session with the help of a ToF camera. By using a ToF camera, off-the-shelf hardware is used to solve the task of patient positioning, therefore a solution for a possible system can be realized quite cost effective.

3. METHODS

In the following we will give a brief overview how a system for patient positioning can be successfully accomplished by using a ToF sensor. Figure 2 shows the overall process in detail. Please note, that in the following indices can be considered to be integer values. We also assume that the ToF camera is rigidly mounted above the patient table and the table normal is perpendicular to the viewing direction of the camera.

We denote $\mathbf{P}$ as the $K \times L$ 3-D points of interest acquired by a ToF camera.

$$\mathbf{P} = [\mathbf{p}_{i,j}], i \in \{0, 1, .., K-1\}, j \in \{0, 1, ..L-1\} \tag{2}$$

As the data acquired directly from the ToF camera is affected by noise and the quality would not be sufficient for the proposed task some preprocessing steps are applied. We apply a Gaussian low pass filter with a standard deviation $\sigma = 2$ as default preprocessing step. Instead of the gaussian filter it is also possible to apply a mean filter with a default kernel size of 7×7 to the ToF raw data.

[†]http://www.visionrt.com

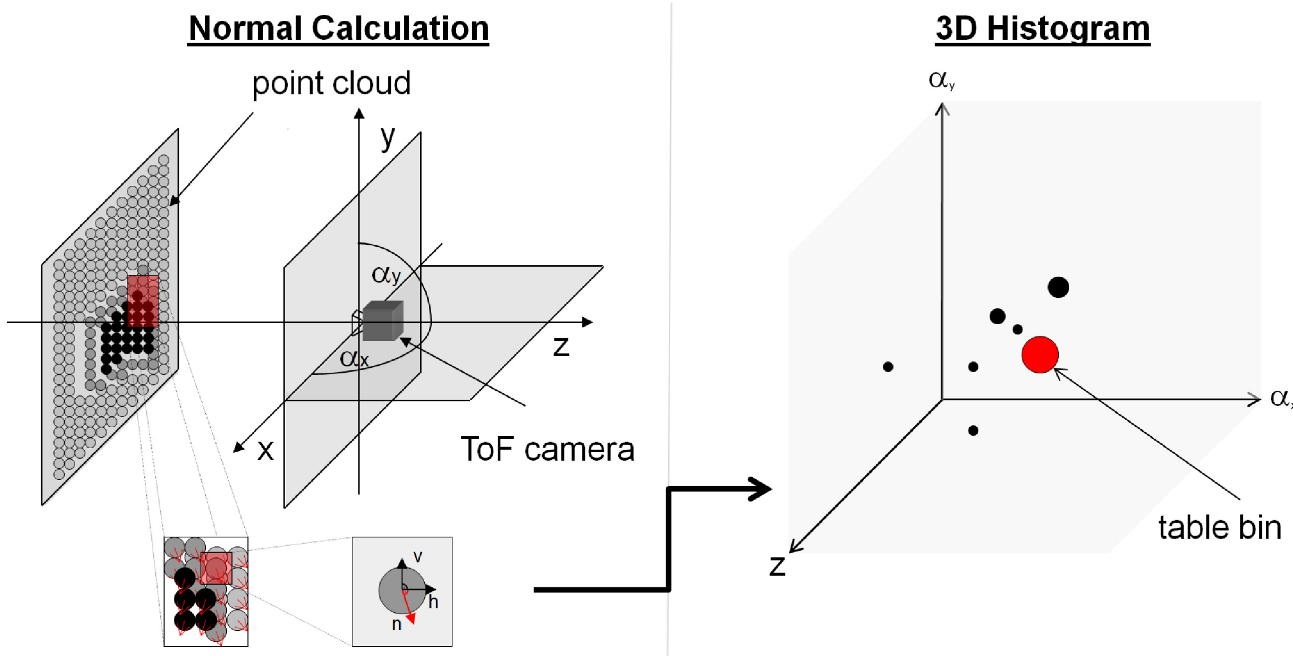

Figure 3. Schematic overview of the normal computation and the 3-D histogram.

In order to segment the body from the background, we determine a so called best fitting plane through all 3-D points representing the table. Therefore, we compute a set $\mathbf{N}$ (see equation 5) of unit normals (see Figure 3) using equation 3 and equation 4, where $i \in \{1, 2, .., K-2\}, j \in \{1, 2, .., L-2\}$:

$$\mathbf{h}_{i,j} = \mathbf{p}_{i+1,j} - \mathbf{p}_{i-1,j} \tag{3}$$

$$\mathbf{v}_{i,j} = \mathbf{p}_{i,j+1} - \mathbf{p}_{i,j-1} \tag{4}$$

$$\mathbf{N} = \{\mathbf{n}_{i,j} | \mathbf{n}_{i,j} = \frac{\mathbf{h}_{i,j} \times \mathbf{v}_{i,j}}{||\mathbf{h}_{i,j} \times \mathbf{v}_{i,j}||}\}. \tag{5}$$

For grouping all normals and identifying the most present one, the normals are inserted into a three-dimensional histogram (see Figure 3). The axes of the histogram represent two quantized normal angles, namely those between x coordinate and z axis (α_x) and y coordinate and z axis (α_y), and the pixel's quantized z coordinate as third dimension. Each histogram bin contains a list of all associated normal locations, defining their positions in the lateral camera grid. Using setup dependent histogram bin sizes, it can be assured that only points of the treatment table are considered for further computations. By selecting the maximum histogram entry (table bin) the most dominant normals are selected to belong to the table plane. The associated points are defined by the histogram bins normal list. They are all supposed to exclusively belong to the table surface. Given all those points, we compute a best fitting plane through them to minimize the squared distance of all points to the resultant plane.

In the second step, we use this best fitting plane (furthermore called table plane) to segment the patient, which is lying on the table, from the background by neglecting all 3-D points which are located beyond the table plane. Both steps can still be considered as preprocessing because they have to be performed before both the reference surface acquisition and the surface acquisition of the surface we want to register.

The final step is to register the previously recorded reference surface with a currently acquired one to determine the patient's misalignment. The registration itself utilizes the standard ITK Iterative Closest Point (ICP)[8]

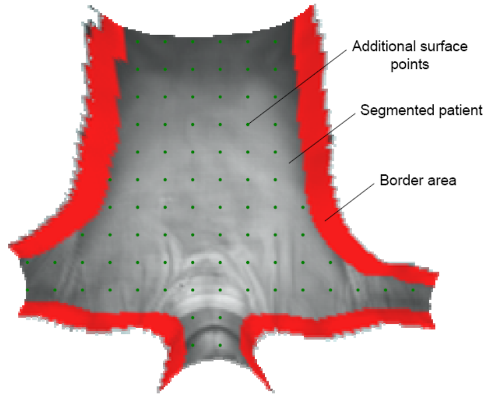

Figure 4. A segmented patient with colored border area. Only points of the border area (red) and additional points on the surface (green, optional) are used for registration.

implementation. In general a reference (fixed) point set $\mathbf{Q} = \{q_i\}$ and a moving point set $\mathbf{W} = \{w_i\}$ with identical cardinality $N_Q = N_W$ are given, where $i \in \{0, 1, .., N_Q - 1\}$. A rotation matrix $\mathbf{R}$ and a translation vector $\mathbf{t}$ have to be computed. Therefore a 6-dimensional objective function (see equation 6) has to be minimized.

$$f(\mathbf{R}, \mathbf{t}) = \frac{1}{N_W} \sum_{i=1}^{N_W} \|q_i - \mathbf{R} \cdot w_i - \mathbf{t}\|^2 \tag{6}$$

This implies, that the algorithm is processed iteratively and can be described briefly as follows:

1. Initialize $W_0 = W, k = 0$, where k denotes the iteration step.
2. (a) Compute for each point $W^{(k)}$ the set $Y^{(k)}$ of the nearest neighbors in Q
   (b) Compute the rotation matrix $\mathbf{R}^{(\mathbf{k})}$ and the translation vector $\mathbf{t}^{(\mathbf{k})}$ using a least square expression to map $Y^{(k)}$ to W_0
   (c) Apply the registration to each point of $W : w_i' = \mathbf{R}^{(\mathbf{k})} \cdot w_i + \mathbf{t}^{(\mathbf{k})}$. The set of transformed points is then $W^{(k+1)} = \{w_i'\}$
3. Set $k = k + 1$.
   Terminate iteration, if the change in mean suqare error falls below a defined threshold or if the number of iterations k exceeds a maximum number of iterations.

The computation time for computing the pose parameters of the current surface relative to the reference surface grows quadratically with the number of points in the point sets used for the registration. In order to speed up the registration process we reduce the number of points by selecting just the border of the segmented point cloud (see Fig. 4). Additionally, we can also add subsampled points inside the segmented point cloud to the points used for the registration process (see Fig. 4). A border width of 5 pixels turned out as a proper value.

4. RESULTS

In order to evaluate the quality of our method we attached a body phantom to a robot arm. The robot arm allows precise shifts in all three dimensions with a precision of 0.1 mm and also rotations in $\frac{1}{10}^\circ$ steps. A reference surface was acquired and then the phantom was shifted arbitrarily into all three room dimensions. We performed this test with different positions and in a distance of 80 cm and 170 cm (see Table 1). The Euclidean distance between desired shift and the shift obtained by the registration process was used to measure the error of our method. For judging the rotation, the mean absolute rotational error of all rotations around the axes was computed. The mean registration error was 2.88 mm for translations and 0.28° for rotations.

camera distance $[cm]$	center subsampling	n	euclidean distance $\mu_{\Delta ED}$ $[mm]$	euclidean distance $\sigma_{\Delta ED}$ $[mm]$	RMS μ_{RMS} $[mm]$	rotation error $\mu_{\bar{\varphi}}$ $[°]$	rotation error $\sigma_{\bar{\varphi}}$ $[°]$	time μ_t $[s]$
translation evaluation								
170	none	22	9.47	6.61	8.26	0.54	0.37	6.41
	4x	8	6.71	4.69	7.66	0.45	0.39	8.21
80	none	8	2.88	1.84	3.36	0.28	0.13	22.12

Table 1. Phantom evaluation results for border width $bw = 5$. Shows mean and standard deviation of all available values (count: n) in each group. Longer mean execution times μ_t are due to higher numbers of points in the point sets.

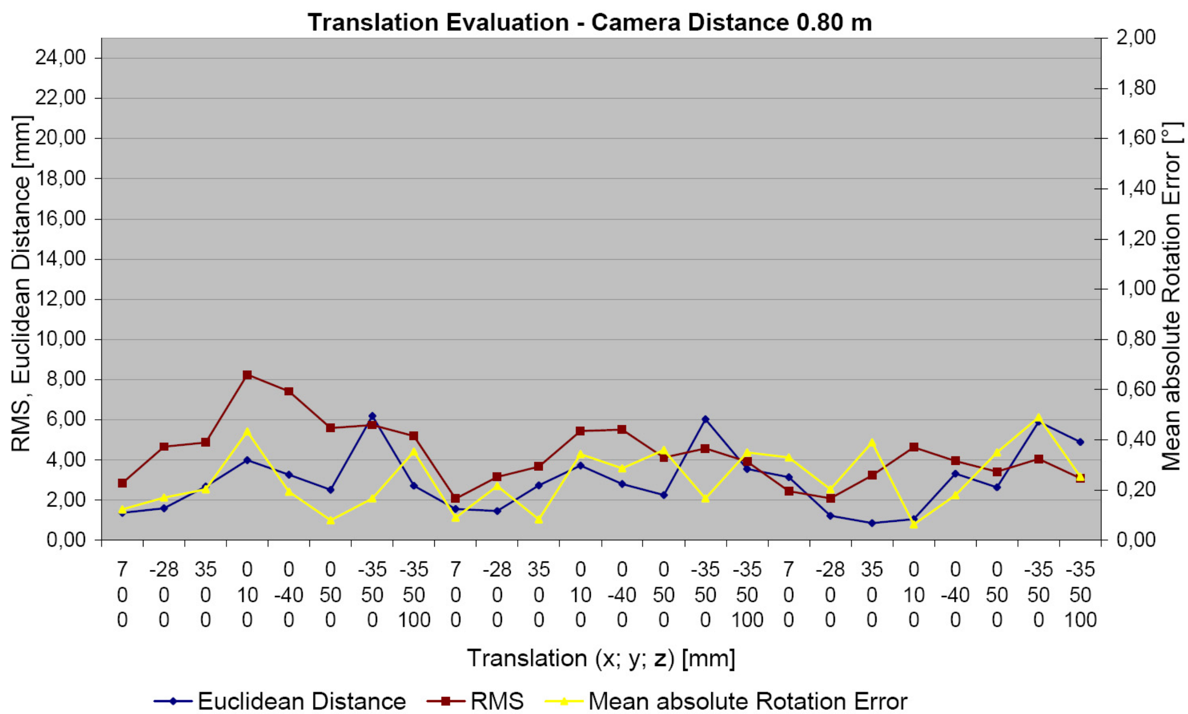

Figure 5. Translation evaluation for a camera distance of 80 cm. Shows ΔED, RMS and $\bar{\varphi}$ for different given translations. This diagram shows the results for the border widths $bw_1 = 1$, $bw_2 = 2$ and $bw_3 = 5$. (left to right).

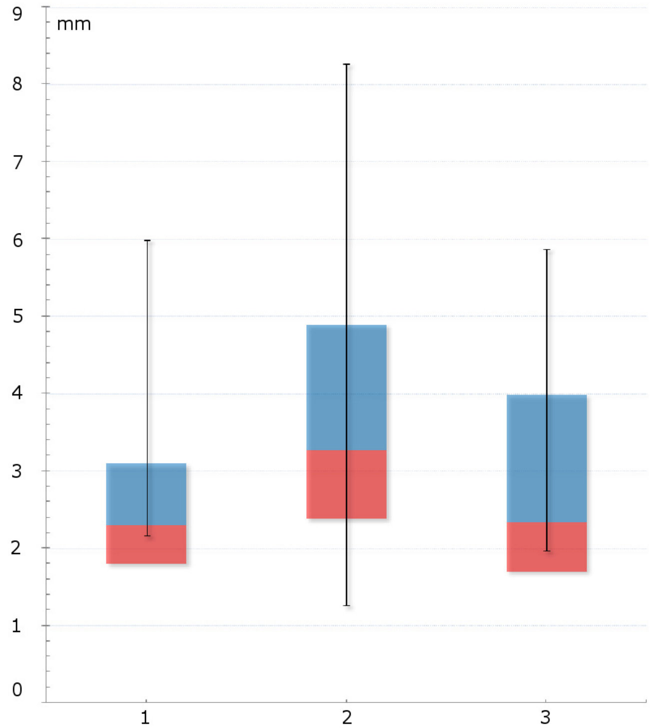

Figure 6. Left: The box plot of test person evaluation shows the 2nd quartile (red) and the 3rd quartile (blue) including the error bars for three persons. Right: Test setup showing the 4 markers, three persons were positioned three times each. On the registered surfaces the Euclidean distance between the markers was computed and plotted in the left diagram.

Figure 7. Left: The box plot of test person evaluation shows the 2nd quartile (red) and the 3rd quartile (blue) including the error bars for three persons. Right: Test setup showing the 4 markers, three persons were positioned three times each. On the registered surfaces the Euclidean distance between the markers was computed and plotted in the left diagramm.

Furthermore, the algorithm was tested with three test candidates of different height and size. In order to evaluate the accuracy of our method, four reflective marker were put on the patients' body (see Figure 7). These markers can easily be seen in the ToF acquired data. Once a reference surface was acquired of each person, they were required to change their position thrice. After each change, the displacement vector was computed. Using this vector both, the reference and the actual surface were overlaid and the euclidean distances between each of the four corresponding markers was computed. In figure 6 a box plot of the results showing the 2nd and 3rd quartile, including error bars is presented. The figure also illustrates the setup of the markers on the skin of the test candidate. Respiratory motion was not considered within this evaluation. The test candidates were only required to fully exhale while the surfaces were acquired. The overall mean registration error was 3.38 mm with a standard deviation of 2.00 mm. For the human test candidates, the camera distance was about 80 cm to the object.

5. CONCLUSION

We propose a system that uses a ToF sensor to correct the position of a patient regarding a reference surface. The algorithm includes a patient table detection using a clustering approach for similar surface normals. Based on the segmented part of the body an ICP algorithm is used to register both surfaces. We compute a 6 DOF displacement vector, in order to realign the patient according to the translation and rotation described by this vector. The method was intensely tested with a rigid torso phantom and with three human test candidates. We could show that our method has a mean registration error of 2.88 mm and 0.28° for the torso phantom and 3.38 mm for humans.

ACKNOWLEDGMENTS

This work was supported by the International Max Planck Research School for Optics and Imaging (IMPRS-OI), Erlangen, Germany, the Erlangen Graduate School in Advanced Optical Technologies (SAOT) and the softgate GmbH, Erlangen.

REFERENCES

[1] Xu, Z., Schwarte, R., Heinol, H., Buxbaum, B., and Ringbeck, T., "Smart Pixel - Photometric Mixer Device (PMD) / New System Concept of a 3D-Imaging-on-a-Chip," in [*5th International Conference on Mechatronics and Machine Vision in Practice*], 259–264 (1998).

[2] Soutschek, S., Penne, J., Hornegger, J., and Kornhuber, J., "3-D Gesture-Based Scene Navigation in Medical Imaging Applications Using Time-Of-Flight Cameras," in [*IEEE Computer Society Conference on Computer Vision, Omnipress*], DVD Proceedings (2008).

[3] Hansen, D. W., Hansen, S. M., Kirschmeyer, M., and Larsen, R., "Cluster Tracking with Time-of-Flight Cameras," in [*IEEE Computer Society Conference on Computer Vision, Omnipress*], DVD Proceedings (2008).

[4] Penne, J., Höller, K., Krüger, S., and Feussner, H., "NOTES 3D: Endoscopes learn to see 3-D; Basic Algorithms for a Novel Endoscope," in [*VISAPP 2007, Second international Conference on Computer Vision Theory and Applications*], 134–139 (2007).

[5] Schaller, C., Penne, J., and Hornegger, J., "Time-of-Flight Sensor for Respiratory Motion Gating," *Medical Physics* **35**(7), 3090–3093 (2008).

[6] Gierga, D. P., Riboldi, M., Turcotte, J. C., Sharp, G. C., Jiang, S. B., Doppke, K. P., Katz, A., Taghian, A. G., and Chen, G. T. Y., "Target Registration Error with Three-Dimensional Surface Imaging in Setup of Image-Guided Partial Breast Irradiation," *International Journal of Radiation Oncology Biology Physics* **70**(4), 1239–1246 (2008).

[7] Schöffel, P. J., Harms, W., Sroka-Perez, G., Schlegel, W., and Karger, C. P., "Accuracy of a commercial optical 3D surface imaging system for realignment of patients for radiotherapy of the thorax," *Physics in Medicine and Biology* **52**(13), 3949–3963 (2007).

[8] Besl, P. J. and McKay, N. D., "A Method for Registration of 3-D Shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14**(2), 239–256 (1992).