Inverse C-arm positioning for interventional procedures using real-time body part detection

Christian Schaller[1], Christopher Rohkohl[1], Jochen Penne[1], Michael Stúrmer[1], and Joachim Hornegger[1]

Chair of Pattern Recognition and Erlangen Graduate School in Advanced Optical Technologies (SAOT), Friedrich-Alexander University Erlangen-Nuremberg, Martensstr. 3, 91058 Erlangen, Germany
`christian.schaller@informatik.uni-erlangen.de` *

Abstract. The automation and speedup of interventional therapy and diagnostic workflows is a crucial issue. One way to improve these workflows is to accelerate the image acquisition procedures by fully automating the patient setup. This paper describes a system that performs this task without the use of markers or other prior assumptions. It returns metric coordinates of the 3-D body shape in real-time for inverse positioning. This is achieved by the application of an emerging technology, called Time-of-Flight (ToF) sensor. A ToF sensor is a cost-efficient, off-the-shelf camera which provides more than 40,000 3-D points in real-time. The first contribution of this paper is the incorporation of this novel imaging technology (ToF) in interventional imaging. The second contribution is the ability of a C-arm system to position itself with respect to the patient prior to the acquisition.

We are using the 3-D surface information of the patient to partition the body into anatomical sections. This is achieved by a fast two-stage classification process. The system computes the ISO-center for each detected region. To verify our system we performed several tests on the ISO-center of the head. Firstly, the reproducibility of the head ISO-center computation was evaluated. We achieved an accuracy of (x: 1.73±1.11 mm/y: 1.87±1.31 mm/z: 2.91±2.62 mm). Secondly, a C-arm head scan of a body phantom was setup. Our system automatically aligned the ISO-center of the head with the C-arm ISO-center. Here we achieved an accuracy of ± 1 cm, which is within the accuracy of the patient table control.

1 Introduction and Related Work

Workflow optimization is an important task in clinical procedures. The combination of increased life expectancy and advancements in the field of medicine

have resulted in a constantly increasing number of patients. This in-turn creates a heavier workload for hospitals and healthcare systems in general. Thus, it is becoming progressively more important to develop efficient healthcare procedures without any decline on the quality of patient care. In the future, either healthcare costs will increase dramatically or smart cost-effective solutions for optimizing current workflows have to be found.

With that last goal in mind we examined the image acquisition workflows, in general, and analyzed them in terms of bottlenecks. In this paper we introduce a system to shorten the probably most important and for sure most time consuming part of this procedure: the initial patient positioning step. Image acquisition procedures have constantly been optimized within the past years. Today the scanning time for CT and C-arm systems is basically negligible. What is still very time consuming within the workflow of image acquisition procedures is the patient-dependent setup procedure. Each patient is physiologically relatively unique. There is also a diverse number of clinical procedures. As a result, the patient setup process cannot be easily generalized. Nevertheless, the correct positioning of the patient is a crucial parameter for the quality of image acquisition. Also for retrospective hybrid-image acquisitions it is important to acquire images at the correct position. Therefore, it is important that patient positioning is performed accurately, which can make it a very time consuming task. Especially in interventional procedures correct positioning of the image acquisition device is mandatory. Todays C-arm systems are already very intuitive to use. Strobel et al. [1] describe the setup procedure and the associated time consuming steps. Several fluoroscopic images have to be taken before the required body part is within the ISO-center of the C-arm. These steps are both very time consuming and applying additional radiation doses to the patient.

To automate and speed-up the setup procedure for interventional image acquisition procedures we suggest a system based on a ToF sensor. The proposed system partitions the whole body into several 3-D bounding boxes. It operates marker-lessly and does not rely on any other assumptions. The C-arm system receives exact metric coordinates from the proposed system which can then be used in automatically adjusting the C-arm's position with respect to the patient and for data acquisition. We call this *inverse positioning*, as the system is positioned with respect to the target and not the target with respect to the system. Furthermore, the C-arm knows exactly the bounds of the object to scan and an instantaneous collision detection can be provided. Last but not least, the proposed solution is very cost-effective.

The importance of a solution to these problems is also reflected by prior work. Grimson et al. [2] introduced a system wich supports optical tracking of patient and instrument locations using surface data. Navab et al. [3] proposed a system for intraoperative positioning and repositioning of mobile C-arms using a camera-augmented mobile C-arm. This system speeds up the whole procedure and also reduces the radiation of the patient as it decreases the number of images which have to be acquired for positioning the patient. Using an optical camera and X-ray/optical markers the system supports guidance for C-arm reposition-

Fig. 1. Examples for ToF sensor models. On the left side: CamCube from PMDtec GmbH, Germany. On the right side: SR4000 from MESA Imaging, AG, Switzerland (images taken from manufactor websites)

ing. C-arm systems also suffer from their narrow field-of-view. Wang et al. [4] suggest a stitching algorithm for scaning long bones using a C-arm system. They, too use an additional optical camera to augment different views. Ladikos et al. [5] provide a method for collision detection for a C-arm environment. An array of optical cameras is used to generate a 3-D representation of the operation room. All of these applications are either based on markers and/or use multiple cameras/systems. The proposed system is independent of markers and only needs a single sensor.

2 Materials and Methods

2.1 Time-of-Flight Sensor

ToF sensors provide a direct way to acquire 3-D surface information [6]. ToF sensors have several advantages over other 3-D surface acquisition techniques. The most promising advantage is that ToF is on its way to become a component of mass markets like consumer electronics and the automotive industry. Currently a high-end ToF sensor is available for about USD 7,000 (see Fig. 1). A target price for ToF sensors of a few hundred dollars can be expected in the near future. ToF sensors also render calibration steps, which are mandatory for stereo based systems, unnecessary. This is based on the monocular all-solid-state architecture of ToF sensors. Such an architecture also enables a high portability of the system and a variety of integration prospects in existing systems. Recent ToF sensors provide data rates up to 25 frames per second with an lateral resolution of up to 204×204 pixels. Each of these 3-D points provides precise metric information in the sensor coordinate system. Recently several systems for medical applications like respiratory motion or patient positioning using a ToF sensor were proposed [7, 8]. Further information about the principle of ToF sensors can be found in Xu et al. [6]. A detailed discussion about the advantages of ToF sensors over other 3-D surface acquisition techniques can be found in Schaller et. al [7].

2.2 Body part detection

In this section we will give an overview of how a ToF sensor enables inverse real-time positioning. We will introduce a generalized, very fast and fairly simple

algorithm to solve this task. The proposed algorithm for body part detection consists of two main parts which require a preprocessing and segmentation of the ToF sensor data. The first part describes an effective and robust two-stage classification procedure. At the end of this first subtask, the 3-D body surface is subdivided into multiple anatomically meaningful regions of interest. The second part computes bounding boxes for each of these regions and their corresponding ISO-center. Furthermore, a fairly simple calibration method for aligning the computed ISO-center with the C-arm ISO-center is introduced.

The ToF sensor is rigidly mounted above the patient table and the whole patient is within the field of view of the ToF sensor. We denote $\mathbf{P}$ the $M \times N$ 3-D points of interest acquired by a ToF sensor.

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

Typically ToF data is affected by noise. To reduce this noise we apply both, a bilateral filter [9] and a temporal averaging filter. The averaging filter returns the average 3-D point cloud using data from the last n accumulated 3-D point clouds. To reduce the amount of data and to identify 3-D points belonging to the body we detect the patient table and compute a virtual plane. For this task we use a Hough-Transform like method based on surface normals described in Schaller et al. [7]. Knowing the virtual table plane we can discard all 3-D points which do not belong to the patient. As a result for further computations only relevant 3-D points $\hat{\mathbf{P}} \subseteq \mathbf{P}$ ($\subseteq$ denotes a subset of points) belonging to the body have to be considered. To speed-up the classification process, we apply a normalization on the remaining 3-D points $\hat{\mathbf{P}}$. Therefore a Karhunen-Loeve-Transformation (PCA) is performed. As a result the origin of the coordinate system is placed in the centroid μ of the point cloud $\hat{\mathbf{P}}$ and the axes are aligned with the axial, sagittal and coronal plane of the patient. The main advantage we achieve is parallelism of the bounding boxes to the axes. This results in a more computationally efficient run-time for the algorithm. The three principal axes computed by the transformation are denoted as ρ_0, ρ_1, ρ_2 (see Fig. 2). We have prior knowledge about the shape of the human body. Therefore, we can assume that ρ_0 corresponds to the height of the patient, ρ_1 to the width and ρ_2 to the depth.

One can then compute five intersections (ι_k, where $k \in \{0, 1, .., 5\}$) along the first principle axis (see Fig. 2). The outer most intersection points ι_0 and ι_5 are defined by the outermost 3-D points on ρ_0. Again, we utilize prior knowledge about the shape of the human body. For a coarse first stage initialization of the classification we compute three golden intersections γ_0, γ_1, γ_2 along ρ_0 (see Fig. 2). The golden section γ_2 roughly indicates the position of the neck. γ_1 is located near the diaphragm and γ_0 at the knees of the person. These points constrain the search space for the exact position of the remaining intersection points.

To refine these coarse initial values we introduce a second stage in our classification. Two histograms H_0 and H_1 along ρ_0 are computed. The vertical axis of H_0 bins the silhouette of the body along the positive p_1 axis, while the vertical axis of H_1 bins it along the positive p_2 direction. The horizontal axis of H_0
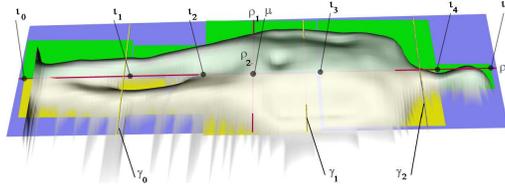
Fig. 2. Overview: The red lines indicate the three principal axes ρ_0,ρ_1,ρ_2. The three golden sections γ_0, γ_1, γ_2 are shown as yellow lines. Furthermore, all intersection points ι_k and the centroid μ are also depicted.
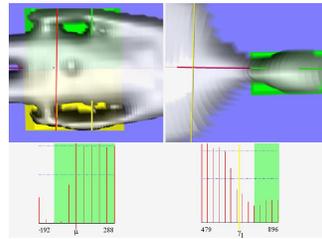

Fig. 3. Refinement of the bounding boxes. Left: Detection of the pelvic region with the corresponding histogram below. μ is the origin of the coordinate system. The search interval d_1 is 30 cm in each direction with a bining size of 4.8 cm. Right: Detection of the neck with the corresponding histogram below. γ_1 is the upper golden section intersection. The search interval d_0 is 10 cm with a bining size of 3.2 cm.

ranges from $\gamma_2 - d_0$ to $\gamma_2 + d_0$. The horizontal axis of H_1 ranges from $\gamma - d_1$ to $\gamma + d_1$. d_0 and d_1 are emirically determined. Based on these histograms, ι_4 and ι_5 are the minima of H_0 and H_1 accordingly. ι_3 is then defined as the mid-point between ι_4 and ι_5, while ι_1 is set to γ_0.

We can then use these intersections as an input for the second part of the algorithm. This part computes bounding boxes using $\iota_0, .., \iota_5$. With the exception of the head, left and right bounding boxes are computed for each body segment (e.g. left and right abdomen), where ρ_0 acts as a delimiter. In addition to the boundary values, the ISO-center and the volume is computed for each box. Figure 4 shows the full body part segmentation.

Before we are able to position a target, a fairly simple calibration step has to be performed. The following has to be done only once for the whole system. A coin is placed on a box on the patient table (see Fig. 5). The ISO-center I_{Carm} of the C-arm is manually aligned with the center of the coin. We use two fluoroscopic images, one from $0°$ and one from $90°$ to do this. After determining I_{Carm}, the corresponding 3-D coordinate (ISO-center) of the center of the coin I_{ToF} in the ToF coordinate system has to be identified. These two points, I_{Carm} and I_{ToF} can be considered as the origins of each of the corresponding coordinate

BODY PART INFORMATION:
head:
  V: 4313.82 (cm³)
  I: (-114, 776, 2774) (mm)
  B: ((+614 | +809) (+705 | +836) (+2737 | +2906)) (mm)
left thorax:
  V: 28131.56 (cm³)
  I: (45, 436, 2779) (mm)
  B: ((+282 | +772) (+437 | +717) (+2696 | +2901)) (mm)
right thorax:
  V: 31033.83 (cm³)
  I: (-236, 425, 2779) (mm)
  B: ((+000 | +490) (+142 | +425) (+2697 | +2920)) (mm)
left abdomen:
  V: 30042.78 (cm³)
  I: (76, -52, 2812) (mm)
  B: ((-179 | +314) (-052 | +251) (+2723 | +2923)) (mm)
right abdomen:
  V: 35592.83 (cm³)
  I: (-230, -64, 2818) (mm)
  B: ((-487 | +006) (-374 | -064) (+2730 | +2962)) (mm)
upper left leg:
  V: 7079.52 (cm³)
  I: (39, -445, 2865) (mm)
  B: ((-507 | -220) (-417 | -272) (+2815 | +2985)) (mm)
upper right leg:
  V: 8866.00 (cm³)
  I: (-157, -451, 2862) (mm)
  B: ((-703 | -414) (-628 | -466) (+2806 | +2995)) (mm)
lower left leg:
  V: 18596.99 (cm³)
  I: (43, -811, 2858) (mm)
  B: ((-941 | -507) (-811 | -634) (+2736 | +2977)) (mm)
lower right leg:
  V: 21398.85 (cm³)
  I: (-143, -813, 2856) (mm)
  B: ((-1116 | -692) (-1007 | -813) (+2729 | +2989)) (mm)

patient info:
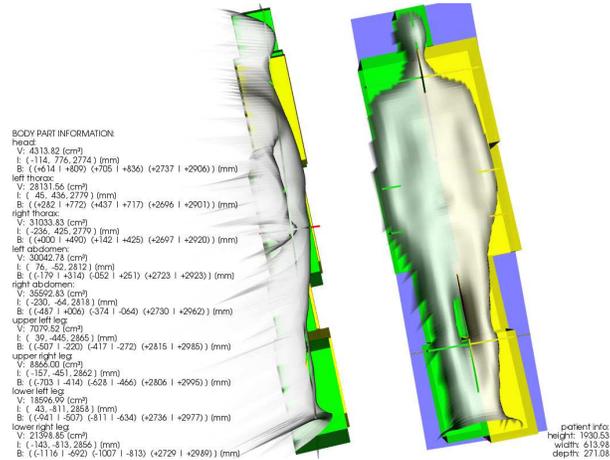height: 1930.53
width: 613.98
depth: 271.08

Fig. 4. Full body acquisition using a ToF sensor including bounding boxes (side view and top view). On the left side information about the different bounding boxes is shown. V shows the volume of the bounding box in cm^3, I the ISO-center of the bounding box in mm and B the bounds of the actual bounding box in mm. On the lower right side, some basic information about the dimensions of the patient is displayed.

systems. To position a bounding box ISO-center, the 3-D coordinate of the ISO-center is shifted to I_{ToF}.

3 Experiments and Results

For the evaluation we used a SR-3000 ToF sensor from MESA Imaging, AG, Switzerland and a C-arm system of the Artis zee family from Siemens AG, Healthcare Sector, Germany. We rigidly mounted the ToF sensor on the ceiling above the patient table so that the patient can be fully visible. The ToF sensor has a resolution of 176x144 pixels with a field of view of 47.5° and 39.6° respectively. The depth accuracy after preprocessing the data is below 1 mm per pixel. We first examined the reproducibility of the ISO-center computation. Without loss of generality, this was done for the head ISO-center. For the other ISO-centers this could be done in a similar manner. The head ISO-center of three persons was computed 70 times per person. We took the mean x- y- z- coordinate as our reference ISO-center. We again computed 140 head ISO-centers for each of these persons while they were in the same position. The mean squared errors between those measurments and the reference in x- y- z- direction (corresponding to ρ_0,ρ_1,ρ_2) are: x: 1.73±1.11 mm, y: 1.87±1.31 mm, z: 2.91±2.62 mm. Furthermore, we computed the inter-subject standard deviation on these datasets: x: 3.16 mm y: 2.85 mm z: 4.42 mm. To evaluate our approach within a clinical setup a body phantom was placed in a C-arm environment. We were again interested in the ISO-center of the phantom head. The phantom was shifted in all three
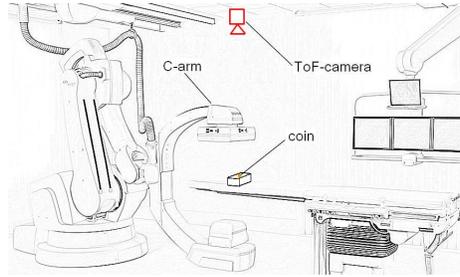
Fig. 5. Coin calibration: A coin is placed on a box on the patient table. The coin is aligned in the ISO-center of the C-arm. Afterwards, the corresponding 3-D coordinate of the position of the coin in the ToF point cloud can be determined.
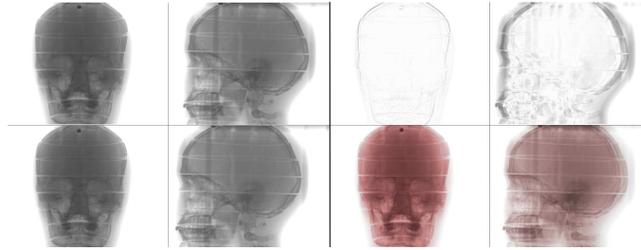

Fig. 6. Phantom evaluation: Two upper left images: $0°$ and $90°$ projections of manually aligned head (gold standard). Two lower left images: $0°$ and $90°$ projections of automatically aligned head using the proposed method. Two upper right images: $0°$ and $90°$ projection difference images of manually and automatically aligned head. Two lower right images: $0°$ and $90°$ projections of manually and automatically aligned head (automatically aligned projections are superimposed in red).

room dimensions multiple times. After computing the ISO-center I_h of the head bounding box with respect to the previously computed I_{ToF} we could directly compute the translation to align I_h with I_{Carm}. A gold standard was defined by manually aligning the ISO-center of the head using the same technique as for the previously described coin calibration (see Fig. 6). This is also the way it is done in hospitals today. We compared the table position of the gold standard alignment and the automatic alignment. In all cases we were able to automatically position the head in the x and the y direction within the accuracy of the patient table. According to the manufactor the accuracy is 1 cm. In the z-direction, we had an error of $\pm$ 1 cm (see Fig. 6) because the z-coordinate of the ISO-center heavily depends on the quality of the table segmentation. On a standard CPU (2.0 GHz dual-core) our algorithm has an execution time of 65 ms. Including data acquisition, preprocessing and segmentation, the overall execution time is about 143 ms.

4   Conclusion

Automation of time consuming steps is a key competency of future clinical procedures and workflow aware hospitals. We have shown that our system for fully automatic patient setup is able to identify anatomical regions for image acquisition in real-time. This information can be used to align the ISO-center of anatomical targets with the ISO-center of a C-arm system fully automatically. For this task, a 3-D surface acquired by a ToF sensor is analyzed. A two-stage classification process is introduced to identify body parts and to compute the corresponding ISO-centers. The extracted information can be directly used to control any image acquisition devices. This results in an inverse positioning, where the image acquisition device can be positioned according to the patient. Furthermore, several patient dependent metric information are provided by the system. This information can be used for e.g., initial positioning, collision detection, dose estimation or respiratory motion correction [7].

References

1. Strobel, N., Meissner, O., Boese, J., Brunner, T., Heigl, B., Hoheisel, M., Lauritsch, G., M., N., Pfister, M., Rúhrnschopf, E.P., Scholz, B., Schreiber, B., Spahn, M., Zellerhoff, M., Klingenbeck-Regn, K. In: 3D Imaging with Flat-Detector C-Arm Systems. 3 edn. Springer Berlin Heidelberg (2009) 33–51
2. Grimson, W.E.L., Leventon, M.E., Ettinger, G.J., Chabrerie, A., Ozlen, F., Nakajima, S., Atsumi, H., Kikinis, R., Black, P.M.: Clinical experience with a hich precision image-guided neurosurgery system. In: MICCAI. (1998) 63–73
3. Navab, N., Wiesner, S., Benhimane, S., Euler, E., Heining, S.M.: Visual servoing for intraoperative positioning and repositioning of mobile c-arms. In: MICCAI (1). (2006) 551–560
4. Wang, L., Traub, J., Heining, S.M., Benhimane, S., Euler, E., Graumann, R., Navab, N.: Long bone x-ray image stitching using camera augmented mobile c-arm. In Metaxas, D.N., Axel, L., Fichtinger, G., Szkely, G., eds.: MICCAI (2). Volume 5242 of Lecture Notes in Computer Science., Springer (2008) 578–586
5. Ladikos, A., Benhimane, S., Navab, N.: Real-time 3d reconstruction for collision avoidance in interventional environments. In: MICCAI (2). (2008) 526–534
6. Xu, Z., Schwarte, R., Heinol, H., Buxbaum, B., 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. (1998) 259–264
7. Schaller, C., Penne, J., Hornegger, J.: Time-of-Flight Sensor for Respiratory Motion Gating. Medical Physics **35**(7) (2008) 3090–3093
8. Schaller, C., Adelt, A., Penne, J., Hornegger, J.: Time-of-flight sensor for patient positioning. In: Proceedings of SPIE Medical Imaging 2009: Visualization, Image-guided Procedures and Modeling. (2009) online
9. Paris, S., Durand, F.: A fast approximation of the bilateral filter using a signal processing approach. Int. J. Comput. Vision **81**(1) (2009) 24–52