An Augmented Reality Approach for Initializing 2D/3D Registration

Ren Hui Gong, Özgür Güler, John Lovejoy, and Ziv Yaniv

The Sheikh Zayed Institute for Pediatric Surgical Innovation, Children's National Medical Center, Washington DC 20010, USA

Abstract. We describe a technique for intraoperative initialization of 2D/3D registration. The technique uses a tracked tool that is already available in the operating room, as part of an image-guided navigation system, to establish the transformation between the preoperative volume and the intraoperative patient. Initialization is performed in two phases: volume-tool pose planning in the virtual world, and patient-tool pose mimicking in the physical world. Depending on the requirements for accuracy and interaction time, the second phase can be done using either instant, coarse, initialization or Augmented Reality (AR) based interactive initialization. The former method is fast and requires no intraoperative modalities, while the latter uses intraoperative x-rays as a guide to continuously refine the initialization. The proposed technique is appropriate for intraoperative 2D/3D initialization as it is contactless, fast, and uses devices already available as part of the navigation system. Evaluation was done using three publicly available reference data sets. The instant, coarse, initialization was able to provide a mean Target Registration Error (mTRE) of 28-40mm, with the majority of the error associated with errors in translation. The AR-based initialization was able to achieve a mTRE on the order of 5-10mm with an average interaction time of 40-60sec.

1 Introduction

The ability to register preoperative 3D images, CT or MR, to the intraoperative setting is a prerequisite of the majority of image-guided navigation systems. Currently, this is primarily performed using fiducials or anatomical landmarks and surfaces which are digitized intraoperatively. An alternative approach is to perform 2D/3D anatomy-based rigid registration, aligning the volumetric data using x-ray images. This subject has been studied extensively resulting in a large number of published algorithms, as surveyed in [1]. These algorithms vary by modality, anatomical structure and algorithmic approach. While the differences are many, all of them have one characteristic in common, they are iterative and require initialization.

In practice, the majority of 2D/3D registration algorithms have not been able to transition from bench to bedside, except in the domain of radiation therapy. The distinguishing feature of this domain as compared to the operating room

C.A. Linte et al. (Eds.): MIAR/AE-CAI 2013, LNCS 8090, pp. 117–126, 2013.

(OR) is that a good initial estimate of the registration parameters is available via accurate patient positioning using other means. In the laboratory setting various approaches to initialization of 2D/3D registration have been used [1,2], including: (1) knowledge of the spatial relationships associated with the clinical setup; (2) coarse paired point registration using skin adhesive fiducials or anatomical landmarks; and (3) manual, keyboard and mouse based, initialization via interactive positioning of the volumetric data using visual comparison between the medical images and the virtual images.

These approaches are often less applicable in a general clinical setting. Knowledge of the clinical setup to estimate an initial transformation is often not sufficiently accurate. Coarse paired-point registration is not always applicable as it either requires placement of fiducials prior to imaging, making the clinical workflow more cumbersome, or requires digitizing anatomical landmarks which may not be accessible. Finally, use of a keyboard and mouse to perform initialization does not fit well in the OR environment due to the requirement for sterility and the fact that the clinical setting is already physically cramped.

We present an Augmented Reality (AR) approach to initializing 2D/3D anatomy-based registration. Our method uses a tracked tool (e.g. a pointer tool) to augment the physical x-ray image with a virtual, volume rendered, image of the anatomy. The user interactively positions and orients the tool so that the virtual and physical images overlap. This approach uses existing hardware found in any navigation system and can be used for all anatomical volumetric imaging modalities. The method was assessed using publicly available reference data sets for evaluation of 2D/3D registration.

2 Method

Our initialization approach is based on the use of a tracked tool to interactively overlay a volume rendering of the anatomy onto the x-ray images. Using multiple AR views, the user manipulates the tool in physical space until the volume rendering of the anatomy overlaps with the corresponding anatomical structures in all x-ray images.

The approach consists of two steps, planning in the *virtual* world, and interaction in the *physical* world. In the planning step the user places a virtual representation of a physical tool next to the volumetric representation of the anatomical structure obtained from CT or MR. In the interaction step, the user mimics the plan in the physical world. That is, they attempt to position the tool in the same pose relative to the anatomy as was done in the virtual world. Figure 1(a) illustrates the concept of this approach.

In Figure 1(b) we present all of the coordinate systems used by our approach. We assume that the transformation from the tool model to its Dynamic Reference Frame (DRF), $T_{toolmodel}^{toolDRF}$, and from the x-ray images to the tracker are known via tool calibration and camera calibration. The transformation from the volume coordinate system to the tool model coordinate system, $T_{volume}^{toolmodel}$, is specified by the user in the planning step, and the patient coordinate system corresponds

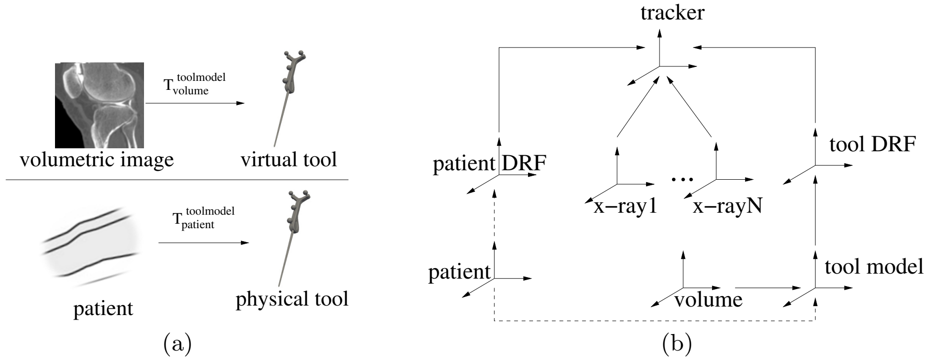


Fig. 1. (a) The user attempts to replicate the pose of the tool defined in the virtual world in the physical world. That is they attempt to position the tool in the physical world such that $T^{toolmodel}_{volume} = T^{toolmodel}_{patient}$. When this happens the volume and patient coordinate systems coincide and we can compute the desired transformation, $T^{patientDRF}_{patient}$. (b) Coordinate systems involved in the registration, solid lines denote known transformations, dashed denote unknown transformations.

to the correct volume pose in the OR, which coincides with the physical location of the patient.

The transformation we seek is given by:

$$\begin{aligned} T^{patientDRF}_{patient} &= (T^{tracker}_{patientDRF})^{-1} T^{tracker}_{toolDRF} T^{toolDRF}_{toolmodel} T^{toolmodel}_{patient} \\ &\Downarrow \\ T^{patientDRF}_{patient} &= T^{patientDRF}_{toolmodel} T^{toolmodel}_{patient} \end{aligned}$$

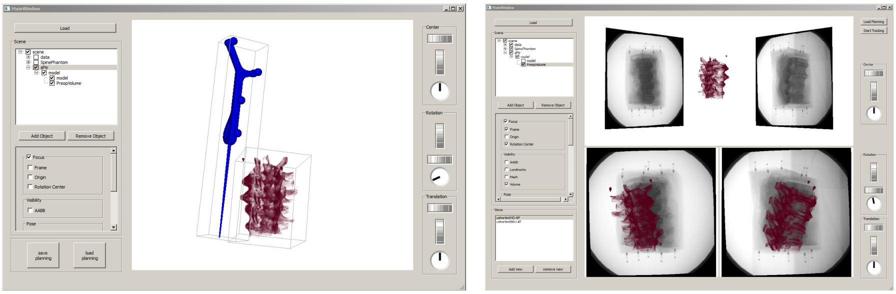
If the planned transformation is mimicked accurately in the OR, we have $T^{toolmodel}_{volume} = T^{toolmodel}_{patient}$, which gives us the desired transformation by substitution into the previous equation to yield.

$$T^{patientDRF}_{patient} = T^{patientDRF}_{toolmodel} T^{toolmodel}_{volume} \qquad (1)$$

Depending on the accuracy requirements imposed by the subsequent registration algorithm one can use this approach in two ways: instant, coarse, initialization, and interactive AR based initialization. The former is applicable for a variety of intraoperative registration methods as it does not utilize any intraoperative images. The later does require availability of intraoperative images and is thus only applicable to procedures where intraoperative imaging is used.

2.1 Planning

The goal of planning is to define the relative pose between the *tool model* and the *preoperative volume*. This is performed using a graphical user interface (Figure 2(a)), within which the poses of the volume and the tool model can be manipulated individually or concurrently. The tool's pose should match its



(a) (b)

Fig. 2. Graphical user interfaces for pose planning in virtual world (a) and pose mimicking in physical world (b)

intended pose in the physical world and it is the user's responsibility to position it in a valid location. This means that the tool cannot overlap with anatomical structures and if using an optical tracking system, its planned position is expected to be visible in the OR.

We use the same approach for registration of CT and MR. Given that in x-ray images the visible structures are primarily bones, we require the user to manipulate the volume rendering transfer function so that these structures are visible to them. This does not imply that the transfer function is optimal, only that for the specific user it yields a visually clear set of anatomical structures.

It should be noted that our approach imposes several requirements on the design of the tracked tool. It must provide six degrees of freedom so that we can manipulate the volume pose in the physical world, and it should not be symmetrical so that the user can visually distinguish between different tool poses. That is, a cylindrical tool such as a needle is best avoided as it defines an infinite number of poses which only differ in rotation about the needle axis. One can design a specific tool based on these requirements, but this is most often not necessary. In our case we utilize a pointer probe which is available as part of the navigation system.

2.2 Instant, Coarse, Initialization

The coarse initialization approach consists of a single step. The planned tool pose is mimicked in the OR by placing the tool besides the patient as planned and initiating the initialization with a foot switch. No further user interaction is required, and $T_{patient}^{patientDRF}$ is estimated instantly. Obviously the accuracy of the result depends on the difference between $T_{volume}^{toolmodel}$, the transformation we use, and $T_{patient}^{toolmodel}$ the correct transformation.

While this method is simple and fast, the initialization accuracy depends on how accurately the planned transformation can be replicated in the OR. By using anatomical landmarks one can plan easy-to-reproduce poses which can provide relatively accurate initializations. As no intraoperative modality is involved, the method can potentially be used to initialize other forms of registration (e.g. point cloud/surface).

2.3 Interactive Augmented Reality Based Initialization

The AR initialization approach is iterative and based on visually guiding the user to the correct pose. We achieve this by real-time direct volume rendering which is overlaid onto the x-ray images. In our case, we perform hardware accelerated volume rendering in parallel for 2-3 images (Figure 2(b)). It should be noted that the camera parameters used to perform the rendering are specific to each x-ray image and are obtained from accurate calibration of the clinical imaging system.

To use the AR based approach, the user starts by performing a coarse initialization as described above. This is required so that there is a reasonable overlap between the rendered image and the x-ray. Then, the user translates and rotates the tracked tool based on the AR views with the goal of maximizing the visual similarity between the overlaid volume rendering and underlying x-ray images. The process continues until a good overall overlay between the x-rays and the corresponding renderings is achieved. The maximal overlap is obtained when $T_{volume}^{toolmodel} = T_{patient}^{toolmodel}$, and the desired transformation is computed as described above. Again, the accuracy of the result depends on the difference between these two transformations.

3 Experiments

3.1 Data

We evaluate our initialization approach using three publicly available reference data sets for 2D/3D registration. The first data set [3] is from the Image Science Institute (ISI), Netherlands, and consists of images from a spine phantom containing three vertebra. The second data set [4] is from the University of Ljubljana, from a phantom consisting of five lumbar vertebra. The third data set [5] is from the Medical University of Vienna, and consists of a cadaver animal head. Unlike the previous two data sets, this data set contains a significant amount of soft tissue which is visible in the x-ray images.

For each of the data sets, we selected two x-ray images, one CT, one MR, and the reference transformations for the CT and MR. The reference transformations position the volumes in the “tracker”, common, coordinate frame to match the corresponding x-ray images. Figure 3 shows the x-ray images from all reference data sets.

3.2 Evaluation Scheme

In the reference data sets, only reference transformations with respect to their own “trackers”, common coordinate frame, were provided. We need to link these

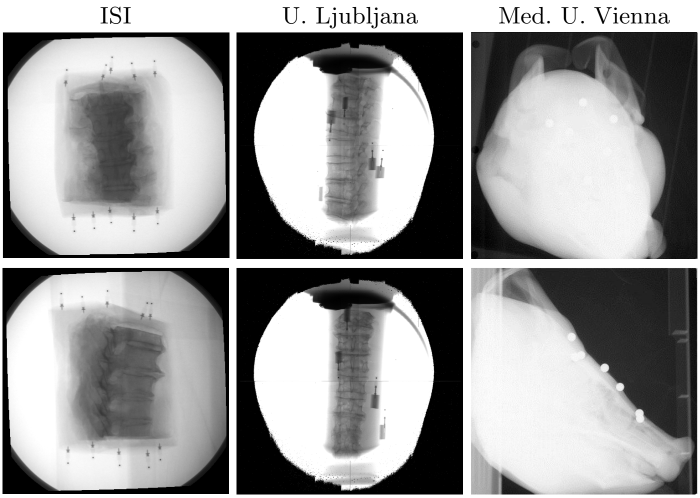
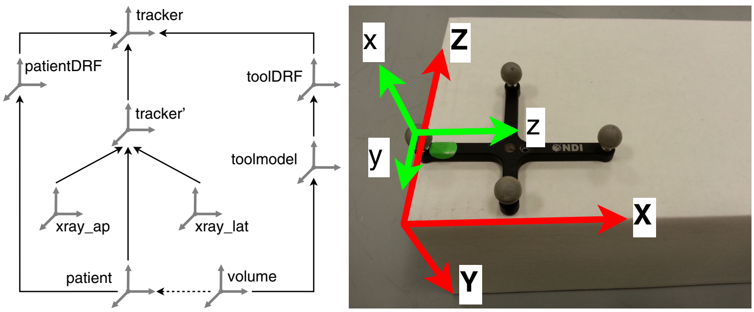


Fig. 3. X-ray images of the reference data sets. First row: AP view. Second row: lateral view.

"tracker" coordinate frames to our physical setup and tracking system. Figure 4(a) illustrates how the reference transformation is established, and how the error transformation is computed. $T_{patient}^{tracker'}$, $T_{xray\_ap}^{tracker'}$ and $T_{xray\_lat}^{tracker'}$ were provided as part of the reference data. The transformation we are interested in, $T_{patient}^{patientDRF}$, is unknown as we do not have the physical phantoms from which the reference data sets were created. We thus need to make an arbitrary choice, relating a physical, tracked, reference frame to the phantom. Once this



(a) (b)

Fig. 4. (a) Transformations involved in the validation of the initialization approach. The *tracker'* coordinate system is the common/world coordinate system used by the reference data set. (b) Definition of reference transformation. The patient and DRF coordinate frames are shown in red and green colors, respectively, and the transformation between the two was used as the reference of our experiments.

transformation is established we can compute the transformation $T_{tracker'}^{tracker}$ accordingly to obtain the reference transformations with respect to our tracker. Then the error transformation between the estimated and ground-truth pose is computed as

$$T_{volume}^{patient} = (T_{patientDRF}^{tracker} T_{patient}^{patientDRF})^{-1} T_{toolDRF}^{tracker} T_{toolmodel}^{toolDRF} T_{volume}^{toolmodel} \quad (2)$$

Note that when the user is able to exactly mimic the planned tool position in the physical world we have $T_{volume}^{patient} = I$.

In our experiments, $T_{patient}^{patientDRF}$ was chosen based on the bounding box of the volume. First, a cardboard box, Figure 4(b), was used to represent the physical patient. The box roughly matches the volume's bounding box in size, and its coordinate frame is aligned with the volume's coordinate frame. Then the patient DRF was placed at the lower-left corner of the xz-surface of the box. Finally, we obtained the coordinates of three known points on the box in the DRF's coordinate system. Thus we have the coordinates of the same points in the patient coordinate system and in the DRF coordinate system. From this setup, $T_{patient}^{patientDRF}$ is readily available via paired point rigid registration [6].

We used the Polaris Vicra optical tracking system from Northern Digital Inc. (Waterloo, ON, Canada) to evaluate our approach. Initialization accuracy is evaluated using the mean Target Registration Error (mTRE):

$$mTRE(e;S) = \frac{1}{N} \sum_{i=1}^{N} \| T_{volume}^{patient} p_i \|, \quad (3)$$

where $T_{volume}^{patient}$ is the error transformation with parameters e, computed as Eq. 2, p_i is a point on our target bone surface S, and N is the total number of surface points.

For each data set, combination of x-ray/MR and x-ray/CT, we created ten plans. We thus had 60 planned tool poses in the virtual world. For each of the sixty plans we had two users perform initialization using the coarse and AR based approaches. For the AR-based initialization approach we also recorded the interaction time.

3.3 Results

Table 1 summarizes the results for the coarse initialization. This method resulted in a relatively high mTRE (28-40mm). However it should be noted that the rotational errors are relatively low, while the translational errors are high. This fact can benefit registration algorithms as correcting rotation errors is more challenging than correcting translation errors. We observed that for each data set the rotational errors are more dominant along one axis. Not surprisingly this axis corresponds to the long axis of our tool. As we noted in section 2.1, the uncertainty in orientation when using a cylindrical like tool is higher around its main axis. This theoretical observation is reflected in practice by our results. We also note that the translational errors have large standard deviations. This is

Table 1. Experimental results for coarse initialization (summary of all 120 initialization trials)

	ISI	U. Ljubljana	Med. U. Vienna
mTRE (mm)	28.1 ± 15.4	29.8 ± 14.0	39.9 ± 19.9
θ_x (°)	0.8 ± 3.6	-0.9 ± 1.7	4.4 ± 8.2
θ_y (°)	-6.3 ± 6.6	-0.3 ± 1.3	2.0 ± 1.9
θ_z (°)	2.0 ± 2.5	-5.1 ± 5.0	-0.5 ± 2.8
t_x (mm)	-10.3 ± 13.3	-25.4 ± 17.8	-8.1 ± 12.5
t_y (mm)	7.3 ± 10.1	7.5 ± 15.2	2.6 ± 17.7
t_z (mm)	-10.4 ± 16.3	8.4 ± 11.8	11.7 ± 6.7

primarily due to the variations in pose planning. In our current implementation we did not provide quantitative feedback (i.e. distances between the tool and anatomy), thus the user visually judges the distance in the virtual world and attempts to mimic it in the physical world. This variability can potentially be minimized by allowing the user to measure distances in the virtual world, which they can then mimic more accurately in the physical world.

Figure 5 summarizes the results for the AR-based interactive initialization. With an average interaction time of 40-60 seconds, an average mTRE of 5-10mm can be achieved. These numbers satisfy the requirements of most 2D/3D registration applications.

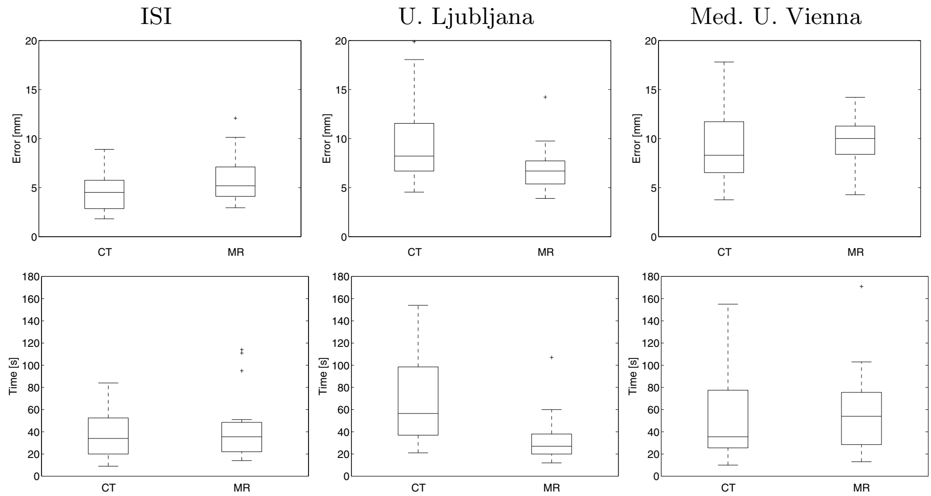


Fig. 5. Experimental results for AR-based interactive initialization. Results from the two users were combined: (top row) mTRE, and (bottom row) interaction time.

3.4 Discussion

The subject of 2D/3D rigid registration is often considered a solved problem. One would assume this is the case given the large number of solutions presented in the literature. Unfortunately, the majority of these methods assume reasonably accurate initialization, mTRE < 10mm, is available, yet they do not specify how it is obtained. While there are specific clinical settings where such an initialization is available, this is not the case in general. As a consequence 2D/3D registration has not been able to transition from bench to bedside.

We propose an AR based solution to initialization which is applicable for image-guided navigation when x-ray images are available. Our method was inspired by an observation made in [7] in the context of segmentation which is equally relevant for registration: Humans are highly adept at determining the presence and rough location of an object of interest in an image. In the context of registration, we provide an intuitive interaction approach which allows us to take advantage of the operators recognition abilities. As a result, the initialization is quick and robust to occlusions in the x-ray image.

To evaluate our approach we used three publicly available reference data sets. On the one hand, this enables a fair comparison between our approach and other methods, all evaluated on the same data sets. On the other hand, these reference data sets do not fully reflect the complexity of the clinical setting. This is primarily visible in the spine data sets which have much less soft tissue than their clinical counterparts. We do not expect this to significantly effect the accuracy of our approach as the operator will implicitly compensate for these differences while interactively setting the volume transfer function during the planning phase. An additional difference between our setting and the OR is that the anatomy of interest may not be visible due to sterile drapes. This can potentially have a significant effect on the results of the coarse registration phase as it solely relies on visually positioning the tool relative to the anatomy without acquiring any intraoperative images. We do not expect this to have a significant effect on our overall results. Our only requirement from the coarse initialization is that the resulting transformation enables us to augment the x-ray images. That is, the volume rendering should have some overlap with the x-ray images which is the case even when the transformation we use is far from the correct one. The user then manipulates the tracked tool based on the AR view which is not effected by the draping. As a consequence we expect similar accuracy in the OR to that obtained in our phantom studies.

4 Conclusion

We described an AR-based approach for initializing 2D/3D registration as part of an image-guided navigation system. The approach does not require any additional equipment and uses a tracked tool which is already part of the navigation system. As our approach is interactive it is equally applicable to registration of CT or MR. In addition, our initialization is based on visual alignment of anatomical structures which simplifies the clinical workflow as there is no requirement for placement of fiducials prior to imaging.

Acknowledgement. We would like to thank E. Wilson for his help creating the mesh model of the tracked tool.

References

1. Markelj, P., Tomaževič, D., Likar, B., Pernuš, F.: A review of 3D/2D registration methods for image-guided interventions. Medical Image Analysis 16(3), 642–661 (2012)
2. Yaniv, Z.: Rigid registration. In: Peters, T., Cleary, K. (eds.) Image-Guided Interventions: Technology and Applications. Springer (May 2008)
3. van de Kraats, E.B., Penney, G.P., Tomaževič, D., van Walsum, T., Niessen, W.J.: Standardized evaluation methodology for 2-D–3-D registration. IEEE Trans. Med. Imag. 24(9), 1177–1189 (2005)
4. Tomaževič, D., Likar, B., Pernuš, F.: Gold standard data for evaluation and comparison of 3D/2D registration methods. Computer Aided Surgery 9(4), 137–144 (2004)
5. Pawiro, S., Markelj, P., Pernus, F., Gendrin, C., Figl, M., Weber, C., Kainberger, F., Nöbauer-Huhmann, I., Bergmeister, H., Stock, M., Georg, D., Bergmann, H., Birkfellner, W.: Validation for 2D/3D registration I: A new gold standard data set. Med. Phys. 38(3), 1481–1490 (2011)
6. Horn, B.K.P.: Closed-form solution of absolute orientation using unit quaternions. Journal of the Optical Society of America A 4(4), 629–642 (1987)
7. Udupa, J.K., et al.: A framework for evaluating image segmentation algorithms. Comput. Med. Imaging Graph. 30(2), 75–87 (2006)