An Augmented Reality & Virtuality Interface for a Puncture Guidance System: Design and Validation on an Abdominal Phantom

S. Nicolau[1], J. Schmid[1], X. Pennec[2], L. Soler[1], and N. Ayache[2]

[1] IRCAD-Hopital Civil, Virtual-surg, 1 Place de l'Hopital, 67091 Strasbourg Cedex
{stephane.nicolau, jerome.schmid, luc.soler}@ircad.u-strasbg.fr
http://www.virtual-surg.com, Phone number: +33 388119095
[2] INRIA Sophia, Epidaure, 2004 Rte des Lucioles, F-06902 Sophia-Antipolis Cedex
{Stephane.Nicolau, Xavier.Pennec, Nicholas.Ayache}@sophia.inria.fr
http://www-sop.inria.fr/epidaure/Epidaure-eng.html

Abstract. In order to design an augmented reality system applied to liver punctures, we devised previously the algorithms that permit to obtain quickly an accurate patient to model registration. In this article we tackle the interface design of the system. The main constraints are the speed and accuracy with which an expert can position correctly a needle into a predefined target. Moreover, to ensure the system safety, the interface has to inform the expert when a registration failure occurs. We present here our interface that allows to fulfill the intervention requirements, by combining the two classical concepts: Augmented Reality and Augmented Virtuality. A validation, on an abdominal phantom, showed evidence that an expert can reach very accurately and quickly the predefined targets inside the phantom.

1 Introduction

Fusion of intra- or pre-operative data with the reality becomes a common tool in the fields of neurosurgery and orthopaedic surgery. This fusion enables the medical expert to see through the patient and to guide his gesture with respect to the additional information provided. Generally, the fusion is made thanks to a registration between the two reference frames in which are localized the patient and the operative data.

To design such a system, two main issues have to be tackled. Firstly, it is mandatory, for security reasons, to assess experimentally the *registration accuracy* between the two reference frames. Indeed, if the accuracy provided does not fulfill the constraints needed by the intervention, the medical expert is guided by a biased information, that can lead to dangerous gesture for the patient. Secondly, we have to evaluate the efficiency and safety of the guidance interface used by the medical expert. The interface has to allow the expert to reach the registered target with an accuracy (called here *guidance accuracy*) and a duration time compatible with the intervention constraints. Moreover, the system has to

G.-Z. Yang and T. Jiang (Eds.): MIAR 2004, LNCS 3150, pp. 302–310, 2004.

enable the expert to detect quickly any failure before and during the intervention (bad registration, incorrect tool tracking...).

Our purpose is to build an augmented reality system to guide liver punctures during interventional radiology (preliminary works are described in [9,8]). According to surgeons, the *overall accuracy* (resp. the guidance step duration) of this system has to be better than 5 mm (resp. shorter than 10 minutes) to provide significant help. In our setup, we stick radio-opaque markers on the patient skin and acquire a CT-scan of his abdomen just before the intervention. Then, an automatic 3D-reconstructions of his skin, his liver and the target is performed [10]. Two cameras (jointly calibrated) view the patient skin and a square marker attached to the needle. This marker enables to locate the needle position in the cameras reference frame. The patient is intubated during the intervention, so the volume of gas in his lungs can be controlled and monitored. Then, it is possible to fix the volume at the same value during a few seconds repetitively and to perform the needle manipulation almost in the same volume's condition than the one obtained during the preliminary CT-scan. Balter [1] and Wong [11] indicates that the mean tumor repositioning at exhalation phase in a respiratory-gated radiotherapy context is under 1 mm. Thus, it is reasonable to assume that a rigid registration of the markers, visible in both CT and video images, is sufficient to register accurately the 3D-model extracted from the CT in the cameras reference frame. A quantitative validation study on a phantom, carried out in [9], showed that a mean registration accuracy σ_r of 2 mm (RMS) was reached within the liver.

In this paper, we focus on the interface design of our system, devoted to percutaneous liver punctures. In our context, knowing that $\sigma_r = 2$ mm, we can afford at most a guidance accuracy of $\sqrt{5^2 - \sigma_r^2} \simeq 4.5$ mm in order to reach the 5 mm of overall accuracy. In addition, we need a quick targeting guidance (shorter than the 10 minutes routinely needed for this kind of intervention). Eventually, the software has to enable the expert to check quickly the correctness of the model registration. Classically, there are two types of interface used in existing medical computer-aided systems. One type, so called Augmented Reality, superimposes intra- or pre-operative data on an image of the reality [4,3]. The other type, called Augmented Virtuality, displays the tool position in the reference frame of the operative data [6]. We argue in Sec. 3 that each of them presents individually advantages and drawbacks, and that an interface integrating both approaches will provide the best efficiency.

In the sequel, we first recall in Sec. 2 how we register automatically the reconstructed model and how we find in real time the needle location in the camera frame. Then, we present our interface, and we show in Sec. 3 how the double approach allows us to obtain an excellent accuracy and to secure the system during the intervention.

2 Principles of Our Guidance System

The overall purpose of our system is to guide the needle manipulated by the expert toward a predefined target. This section deals with the first steps: the computation of the transformation T relating the operative data to the camera frame, and the localization of the needle. To find T, we use the fiducials that are automatically extracted from the CT and the video images. After a matching process, a 3D/2D point-based registration is performed to relate the model and the patient in the same reference frame.

2.1 Automated Localization and Matching of Markers

The principle of the marker localization in the video images is based on a HSV color analysis, followed by a component size and shape thresholding, and the assumption that the skin takes up the main surface. The markers in the CT-image are extracted by a top-hat characterization that emphasizes small singularities on the skin surface.

The matching between the video markers is realized thanks to epipolar geometry, and, the correspondences between video and CT markers is carried out by a prediction/verification algorithm. A validation carried out in [8] showed that these algorithms are robust and that the overall computation time of the extraction, matching and registration process is below 120 sec.

2.2 Registration of the Virtual Model in the Cameras Frame

We choose a 3D/2D points registration approach to provide the rigid transformation that relates scanner frame and cameras frame. The classical choice is to optimize the SPPC criterion (see [9]):

$$SPPC(T) = \sum_{k=1}^{S} \sum_{i=1}^{N} \xi_i^k \cdot \frac{\| \tilde{m}_i^{(k)} - P^{(k)}(T \star \tilde{M}_i) \|^2}{2 \cdot {\sigma_{2D}}^2}$$

where S (resp. N) is the number of cameras (resp. markers), $\tilde{m}_i^{(k)}$ is the observed 2D coordinates of the ith markers in the kth video image, $\tilde{M}_i$ is the observed 3D coordinates of the ith markers in the CT-image, $P^{(k)}$ the projective function, ξ_i^k is a binary variable equal to 1 if the ith marker is visible in the kth video image and 0 if not, and T the seeked transformation. However, this criterion considers that noise only corrupts the 2D data and that 3D data are exact. In our context, this assumption is erroneous as the markers extraction from the CT-image is corrupted by noise as well.

A more realistic statistical hypothesis is that we are measuring noisy versions $\tilde{M}_i$ of the unknown exact 3D points M_i (more details are given in [9]). Moreover, we can now safely assume that all 2D and 3D measurements are independent. A

ML estimation of the transformation T and the *auxiliary variables* M_i leads to minimize the *Extended Projective Points Criterion* (EPPC):

$$EPPC(T, M_1, \dots M_N) = \sum_{i=1}^{N} \frac{\| \tilde{M}_i - M_i \|^2}{2 \cdot {\sigma_{3D}}^2} + \sum_{k=1}^{S} \sum_{i=1}^{N} \xi_i^k \cdot \frac{\| \tilde{m}_i^{(k)} - m_i^{(k)} \|^2}{2 \cdot {\sigma_{2D}}^2}$$

The minimization procedure is consequently modified into an alternated minimization w.r.t. the seeked transformation T, and w.r.t. the M_i.

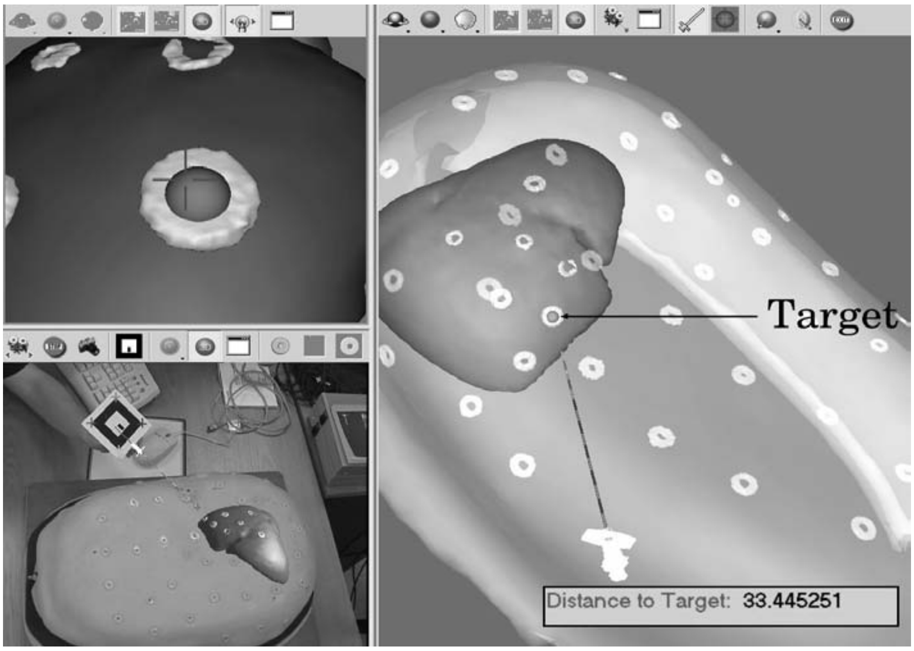
2.3 Needle Tracking

We have to track the needle location and orientation in the camera reference frame. To realize it, we attach an oriented square marker whose corners are automatically localized on video images in real-time using an adapted version of the ARTkit library [5]. Then, knowing the size of the square, we are able to localize it in the camera reference frame by minimizing the classical 3D/2D SPPC criterion. Calibrating the relative needle position w.r.t. the square marker with the pivot method [7], we are finally able to superimpose the virtual model on the real one on video images.

3 A Secured and Ergonomic Guidance Interface

Our interface has to be designed and adapted for our particular application: liver punctures. In the field of craniotomy, Grimson *et al* [4] superimpose the reconstructed model on external video image of the patient skull. This approach allows the surgeon to check instantly the validity of the registration: if the registration is false, the superimposition will be visually incorrect. However, this kind of interface provides a view that does not correspond to the surgeon natural field of view. Realized and visualized movements can be inverted. Therefore, it needs an important interpretation effort. Moreover, since the focal lengths of the cameras are fixed, no zoom of the area of interest is available.

In the context of laparoscopy guidance, Lango [6] registers the 3D reconstructed model with the patient by pointing with a tracked tool (Polaris$^{\text{TM}}$) several radio-opaque markers stick on the patient skin. He proposed an interface that showed the tool position with respect to the model. Moreover, he displays the 3 CT-slices where the tip of the laparoscope lies. This approach is very useful to understand the relative position of the tool with respect to the model, since the user can choose his angle of view and an appropriated zoom. Nevertheless, since there is no camera, it is not possible to display the 3D model on an external video view of the patient. Then, the quality of the registration cannot be assessed quickly during the intervention. Indeed, this can only be done interactively, at a given time point, by pointing some reference points on the patient skin. Therefore, if the patient moves after the registration has been done, it will undergo a bias. This analysis lead us to realize an interface that provides the information of both approaches.



Distance to Target: 33.445251

Fig. 1. Three screens guidance interface. The bottom left image corresponds to the augmented reality view, in which are displayed the 3D reconstruction of the liver and the virtual needle. The top left image displays the virtual needle view (oriented toward a marker stick on the liver surface). The right image shows the main virtual view, in which one can see the relative position of the needle w.r.t. the phantom. We indicate in its corner the virtual distance in mm that separates the tip needle to the target (in this case, a marker center).

3.1 A Three Screens Interface

Our interface (showed on Fig. 1) is divided into three screens described below. Their features and properties have been optimized with surgeons, in order to provide them a clear and intuitive tool. Each of the action associated to each screen can be done by another operator with a mouse action only (no keyboard action). These considerations should reduce time consuming manipulation.

The Augmented Reality View (Bottom Left Image in Fig. 1) In this screen, one of the two video images returned by our cameras is displayed. The user can switch between both views, enable or disable the real time superimposition of the 3D model on the video images, choose the transparency level of its different elements and display the real-time extraction of the markers. Furthermore, the user can superimpose the virtual needle on the tracked real needle and monitor the real-time tracking of the square marker attached on it. Finally, the user can check visually the registration quality by superimposing virtual elements. If he considers that this is not acceptable (which can occur if the patient has slightly moved during the intervention), a new extraction of the markers is done in order to update the registration.

The Virtual Needle View (Top Left Image in Fig. 1) In order to direct a tool toward a target, Carrat *et al* [2] proposed three crosses displayed on a screen, that have to be superimposed. The optimal trajectory is represented by a static central cross-hair. The tool tip and axis are projected dynamically on a view orthogonal to this trajectory, and are represented by two different cross-hairs. Although this interface enables the user to reach a correct orientation, it is not very intuitive as the user looses any representation of the reality.

In the virtual needle screen, we propose to display a view that corresponds to what would see a camera positioned on the tip needle and oriented along its axis. This view was created to facilitate the orientation of the needle toward the target point. In our interface, it is represented by a green sphere of 2 mm of diameter. This view is easily understood by surgeons since it is very similar to an endoscopic view they are used to. To keep a good visibility when the needle goes through organs, the classical actions of 3D model visibility and transparency are available.

The Virtual Exterior View (Right Image in Fig. 1) In this screen, the 3D virtual scene, composed by the 3D reconstruction and the tool representation, is rendered from a viewpoint controlled by the user. Like in a classical viewer, he can rotate, translate and zoom the elements and define their properties (visibility and transparency). Moreover, it is possible to display as well the CT-scan from which the reconstruction is made, and navigate through its slices. The contrast can be enhanced like in the usual radiological viewer. When the 3D reconstructions of the liver and tumors are available the medical expert guides the needle to the tumor center, using the 3D visualization of the tumor and tip needle relative position. If the reconstructions are not available, for time or technical reasons, the expert can visualize the 3D CT-slices instead of the 3D reconstruction. Then, he can define the target position on a specific CT slice by a mouse click (cf. Fig. 2). Since it is difficult to assess visually the distance between the tip needle and the target, we print it inside the virtual exterior view.

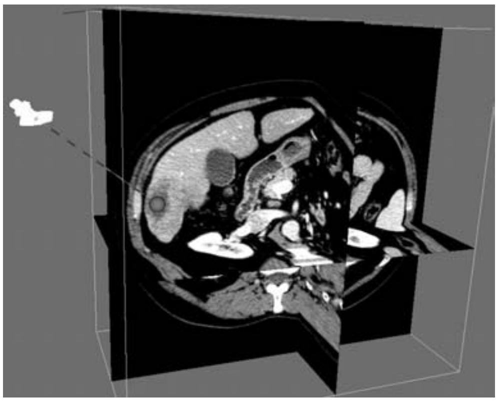


Fig. 2. Patient CT image displayed in the virtual exterior view. One can see a green sphere target that was put by the user.

3.2 Evaluation of the Overall System

The purpose of the experiment was to assess the accuracy of the needle targeting obtained by several surgeons and engineers, using our AR guidance system. Four targets were modeled with radio-opaque markers stuck on the fake liver inside the phantom. Seven participants each performed 10 consecutive needle targetings of the model tumors (cf. Fig. 3 a). During the positioning, the operator placed the needle and stopped his movement when he thought that he had reached the tumor center. After each trial, the time required to position the needle was recorded, and the accuracy of the hits was verified by an independent observer using an endoscopic camera introduced into the phantom and focusing on the targets (cf. Fig. 3 b). Accuracy and time results are shown in Table 1.

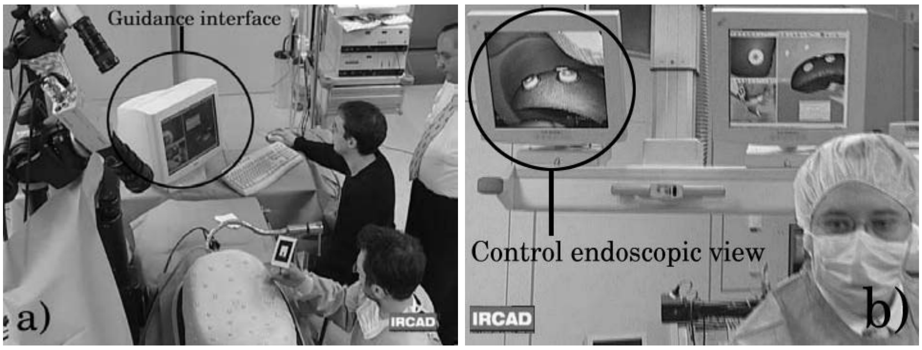


Fig. 3. a) Setup of the experiment: the user is positioning the needle, tracked by a stereoscopic system, thanks to the guidance interface. b) An endoscopic view is displayed behind the user. It enabled to assess visually the correctness of each needle targeting.

3.3 An Intuitive and Powerful Interface

The results indicate that the worst average accuracy obtained is below 3 mm, which clearly fulfills our accuracy constraint (5 mm). In addition, the system allows to reach the target very quickly (average time under 30 sec.) with respect to the usual time needed for a standard percutaneous intervention (10 minutes).

A previous experiment (see [8]), in which the user was guided only by an augmented reality screen, provided less accurate results, and more importantly longer manipulation times. It confirms the fact that the information complementarity given by the three different screens is a powerful aspect of our interface.

It has to be noticed that each one of the three screens was used intuitively at the same stage during the needle positioning by each person involved in our experiment. The *augmented reality view* has been used at the beginning of the

	Average distance (mm) ± std.	Minimum distance	Maximum distance	Average time (sec.) ± std.
Engineer 1	0.95 ± 0.67	0	2	25 ± 5.6
Engineer 2	1.7 ± 0.97	0	3	14 ± 2.0
Engineer 3	1.8 ± 0.84	0	3	18 ± 5.5
Surgeon 1	2.2 ± 0.57	1	3	32 ± 12.2
Surgeon 2	2.9 ± 1.25	0	5	22 ± 3.1
Surgeon 3	1.3 ± 1.16	0	3	32 ± 3.7
Professor	0.84 ± 0.48	0	1	32 ± 6.4
All	1.8 ± 0.7	-	-	23.8 ±7.3

Table 1. Accuracy and time results obtained by each user. The average distance, which is always below 3 mm, fulfills our accuracy constraints (5 mm). Moreover, the time needed is, by far, shorter than 1 minute, whereas an expert needs routinely 10 minutes for such intervention.

needle insertion. Firstly, it was used to check the automatic skin fiducials detection, the visual quality of the skin registration, and the tool superimposition. Secondly, it allowed to define a rough estimation of a correct skin entry point and needle orientation. During the insertion, the *virtual needle view* was always used. Indeed, it seems really adapted to needle orientation problem, since the user has only to keep the target under the cross displayed on the view: this act seemed very intuitive to everybody. Finally, the user swapped his attention to the *virtual exterior view* when the tip needle was very close to the target (below 3 mm). At this moment, a little variation of the needle position produces a big virtual view displacement. As it could make disappear the target from the *virtual needle view*, each user carried out the fine positioning with the *virtual exterior view*. At this step, he was helped by another operator that zoomed on the interest zone.

4 Conclusion

In order to design an augmented reality system devoted to liver punctures, we developed in [9,8] the procedures that allow to register accurately and quickly a patient CT model to video images. The present article deals with the interface design of our system. To overcome the constrains of this intervention (overall targeting accuracy below 5 mm, and guidance duration shorter than 10 minutes), the interface has to enable the expert to reach quickly and accurately the predefined target. Moreover, to ensure the system safety, it has to provide the expert the possibility to check visually the model registration quality during the intervention.

To fulfill these requirements, we propose a three screens interface. Its main advantage over classical augmented reality system, is that it provides two complementary kind of view: a view of the reality on which are superimposed the patient 3D model and the virtual needle, and a virtual view of the 3D model, in which is displayed the current needle position.

This double approach enables the expert to check continuously the model registration quality, and to choose the best angle of view during the needle insertion. A validation experiment on an abdomen phantom, realized with both engineers and surgeons, proved that our interface is very intuitive and permits the user to reach the planned targets with an excellent accuracy with respect to the intervention requirements. Moreover, the average time needed for a correct needle positioning is by far smaller than the routinely intervention duration (less than 40 sec. against 10 minutes).

In the immediate future, we plan to carry out our first validation on a patient. In addition, we will adapt the current system to laparoscopic interventions. Our interface will optimize the laparoscopic tool positioning before the intervention, and it will help the surgeon by merging the 3D patient model into the endoscopic video image.

References

1. J.M. Balter, K.L. Lam, C.J. McGinn, T.S. Lawrence, and R.K. Ten Haken. Improvement of CT-based treatment-planning models of abdominals targets using static exhale imaging. *Int. J. Radiation Oncology Biol. Phys.*, 41(4):939–943, 1998.
2. L. Carrat, J. Tonetti, P. Merloz, and J. Troccaz. Percutaneous computer-assisted iliosacral screwing: Clinical validation. In Springer Verlag, editor, *MICCAI'00*, volume LNCS 1935, pages 1229–1237, 2000.
3. J. Feldmar, N. Ayache, and F. Betting. 3d-2d projective registration of free-form curves and surfaces. *Journal of Comp. Vis. and Im. Under.*, 65(3):403–424, 1997.
4. W. Grimson, G. Ettinger, S. White, W. Wells T. Lozano-Perez, and R. Kikinis. An automatic registration method for frameless stereotaxy, image-guided surgery and enhanced reality visualization. *IEEE TMI*, 15(2):129–140, April 1996.
5. Hiro. Human interface technology laboratory, http://www.hitl.washington.edu/.
6. T. Lango, B. Ystgaard, G. Tangen, T. Hernes, and R. Marvik. Feasibility of 3d navigation in laparoscopic surgery., Oral presentation at the SMIT (Society for Medical Innovation and Technology) Conference. September 2002. Oslo. Norway.
7. S. Lavalle, P. Cinquin, and J. Troccaz. *Computer Integrated Surgery and Therapy: State of the Art*, chapter 10, pages 239–310. IS Press, Amsterdam, NL, in C. Roux and J.L. Coatrieux edition, 1997.
8. S. Nicolau et al. An Augmented reality system to guide radio-frequency tumor ablation. In *Journal of Computer Animation and Virtual World*, 2004. In Press.
9. S. Nicolau, X. Pennec, L. Soler, and N. Ayache. An accuracy certified augmented reality system for therapy guidance. In *European Conference on Computer Vision (ECCV'04)*, LNCS 3023, pages 79–91. Springer-Verlag, 2004.
10. L. Soler et al. Fully automatic anatomical, pathological, and functional segmentation from CT scans for hepatic surgery. *Comp. Aided Surg.*, 6(3), Aug. 2001.
11. J. Wong et al. The use of active breathing control (abc) to reduce margin for breathing motion. *Int. J. Radiation Oncology Biol. Phys.*, 44(4):911–919, 1999.