

An Augmented Reality Navigation System with a Single-Camera Tracker: System Design and Needle Biopsy Phantom Trial

F. Sauer, A. Khamene, and S. Vogt

Imaging & Visualization Dept, Siemens Corporate Research,
755 College Road East, Princeton, NJ 08540, USA
{sauer, khamene, vogt}@scr.siemens.com

Abstract. We extended a system for augmented reality visualization to include the capability for instrument tracking. The original system is based on a video-see-through head-mounted display and features single-camera tracking. The tracking camera is head-mounted, rigidly fixed to a stereo pair of cameras that provide a live video view of a workspace. The tracker camera includes an infrared illuminator and works in conjunction with a set of retroreflective markers that are placed around the workspace. This marker frame configuration delivers excellent pose information for a stable overlay of graphics onto the video images. Using the single camera also for instrument tracking with relatively small marker clusters, however, encounters problems of marker identification and of noise in the pose data. We present a multilevel planar marker design, which we used to build a needle placement phantom. In this phantom, we achieved a stable augmentation; the user can see the location of the hidden target and the needle without perceptible jitter of the overlaid graphics. Piercing the needle through a foam window and hitting the target is then intuitive and comfortable. Over a hundred users have tested the system, and are consistently able to correctly place the needle on the 6mm target without prior training.

1 Introduction

Image guidance systems help the physician to establish a mapping between a patient's medical images and the physical body. In conventional systems, a pointer or an instrument is tracked and the location visualized in the medical images. In contrast, augmented reality (AR) image guidance maps the medical data onto the patient's body. Anatomical structures are being perceived in the location where they actually are – the patient becomes transparent to the physician.

Augmented reality for medical applications has first been suggested in [1], and various groups have since been working on realizing augmented reality systems, based on overlaying graphics onto video streams [2-5], on “injecting” graphics overlays into operating microscopes [6-8], or simply by using semitransparent graphics display configurations through which the user observes the real world. Reference [9] compares some of the efforts.

We built an AR system [10] that makes use of a video-see-through head-mounted display (HMD) similar to the one described in [1]. Two miniature color video cam-

eras are mounted on the HMD as the user's artificial eyes. The two live video streams are augmented with computer graphics and displayed on the HMD's two screens in realtime. With the HMD, the user can move around and explore the augmented scene from a variety of viewpoints. The user's spatial perception is based on stereo depth cues, and also on the kinetic depth cues that he receives with the viewpoint variations. Our system has been put into a neurosurgical context [11], adapted to an interventional MRI operating room [12,13], and has also been integrated with an ultrasound scanner [14,15].

Tracking is an essential enabling technology both for conventional and AR navigation systems. Commercial systems either employ optical or magnetic trackers. Optical trackers achieve a higher accuracy, with the requirement of an unobstructed line-of-sight between the tracker camera and the tracked markers. The commercial optical tracking systems are all multi-camera systems. They find the 2D marker locations in the cameras' images and determine their 3D location by triangulation. The most popular optical tracking system in the medical arena is a stereo camera system.

Our AR system's special feature is the use of single camera tracking with a head-mounted tracking camera, which is rigidly attached to the two cameras that capture the stereo view of the scene. Originally we used this tracking camera only in combination with a set of markers framing a workspace. In the current paper, we describe how we extended our single-camera tracking to include instrument tracking with marker clusters. We achieved stable tracking with a cluster that extends only over a small area in the tracker camera's image.

We built a needle placement phantom, where we simultaneously track the phantom with a frame of markers and the needle with the cluster of markers. The tracking works in a very stable manner; targets and needle can be visualized graphically in the augmented view without perceivable jitter. More than one hundred users tried the needle experiment and consistently were able to correctly hit a chosen 6 mm target with the needle. No training was required to succeed with the needle placement. The AR guidance was experienced as very intuitive and comfortable.

In section 2, we present technical details of our AR system. Section 3 describes the needle placement phantom. The paper then concludes with a summary in section 4.

2 AR System Details

2.1 System Overview

The centerpiece of the system is a head-mounted display that provides the user with the augmented vision. Figures 1 and 2 show how three miniature cameras are rigidly mounted on top of the HMD. A stereo pair of color cameras captures live images of the scene. They are focused to about arm's length distance and are tilted downward so that the user can keep his head in a comfortable straight pose. The third camera is used for tracking retroreflective markers in the scene. This black-and-white camera is sensitive only for near infrared wavelengths. It is equipped with a wide angle lens and a ring-shaped infrared LED flash. The flash is synchronized with the tracking camera and allows us to select a fast speed of its electronic shutter. The short exposure time of only 0.36 ms efficiently suppresses background light in the tracker camera's images, even when the scene is lit with strong incandescent or halogen lamps.

Mounting the tracking camera on the user's head helps with the line-of-sight restriction of optical tracking; the user cannot step into the tracker camera's way (though he still can, of course, occlude markers with his hands). Placing a forward looking tracking camera on the head is optimal for the perceived accuracy of the augmentation, as the tracker camera's sensitivity to registration errors is matched to the user's sensitivity to perceive these errors. Furthermore, this configuration makes good use of the tracker camera's field of view. Tracking is only required when the user actually looks at the workspace. And then the tracker camera is automatically looking at the workspace markers. For this reason, the markers can extend over a sizeable part of the tracker camera's image, yielding good tracking accuracy.



Fig. 1. Video-see-through HMD with mounted tracking camera



Fig. 2. Camera triplet with a stereo pair of camera to capture the scene and a dedicated tracking camera with infrared LED flash

Display and cameras are connected to two PCs. One SGI 540 processes the tracker camera images and renders the augmented view for the left eye, an SGI 320 renders the augmented view for the right eye. Both PCs communicate over an Ethernet connection to exchange information concerning camera pose, synchronization, and choice of graphics objects to be used for augmentation. Table 1 lists the particular hardware components that we are using.

Table 1. Hardware Components

HMD	Kaiser Proview XL35, XGA resolution, 35° diagonal FOV
Scene cameras	Panasonic GP-KS1000 with 15mm lens, 30° diagonal FOV
Tracker camera	Sony XC-77RR with 4.8mm lens, 90° horizontal FOV
Computers	SGI 540 and 320 with Windows 2000

2.2 Single Camera Tracking

We want to render a computer generated 3D object onto a real-world video sequence in a way that the 3D graphics object is accurately aligned with respect to some real object seen in the video sequence. For this, we need to know the relative location and orientation of video camera and objects of interest. Or in other words, we need to

know the relationship between two coordinate systems, one attached to the camera, the other attached to the object. Registration initially establishes this relationship in terms of translation and rotation. Tracking denotes the process of keeping track of it.

Single camera tracking is possible when the geometry of the tracked object is known and the internal camera parameters have been pre-determined in a calibration procedure. We fabricated objects for camera calibration and for tracking with retro-reflective disc shaped markers. We then base our system calibration on 3D-2D point correspondences. The 3D coordinates of the markers we measured with commercial stereo system made by the German company A.R.T. GmbH, the 2D positions we determine from the images we take with the camera. We follow Tsai's calibration algorithm [8,9], benefiting from an implementation that is available as freeware at <http://www.cs.cmu.edu/~cil/v-source.html>. The camera calibration object contains over one hundred markers [10], which allows us to estimate the internal camera parameters with sufficient accuracy. The marker sets for tracking then need to provide us with at least seven point correspondences so that we can calculate the external pose, i.e. translation and rotation, for the given camera's internal camera parameters.

For the calibration of our camera triplet (Fig. 2), we determine the internal camera parameters for all three cameras, and the relative external pose between the tracker camera and the two scene cameras. In the realtime tracking mode, we then deduce the pose parameters of the two scene cameras from the measured pose of the tracking camera, which allows us to augment the scene camera images with correctly registered graphics.

2.3 Marker Configuration Design

In our original tabletop system [10], we placed seventeen markers around a workspace. The markers were all lying in the same plane, framing the workspace in three straight lines on the sides and on the top. This marker configuration provided very stable pose estimation in conjunction with the head-mounted tracking camera. The augmented views did not show any perceivable jitter. One main reason for the good results was that the marker frame extended over large part of the tracking camera's image, providing very good leverage for precise pose estimation.

A subsequent system was designed for a neurosurgical setting. A curved frame of markers was fitted onto a head clamp [12,13]. In the first version, the marker locations on this frame were still all coplanar. The resulting pose estimation was in general also still very good. For some viewpoints, however, some jitter could now be perceived. We assume that the slight performance deterioration was at least partially to blame on the reduced number of markers. We added two markers on little posts, sticking out of the plane. This increased the number of markers; at the same time, it turned the co-planar marker configuration into a 3D marker configuration. Now, the tracking was again "perfect", i.e. we could again not perceive any jitter in the augmented views.

We used the same marker frame design for the needle placement phantom that we describe in section 3. Fig. 3 shows a photo of this marker frame.

We were designing our marker configurations mainly based on heuristic reasoning, not with strict mathematical simulation. One relationship seemed obvious: the larger the extent of the marker body in the camera image, the more precise the result of the pose determination in regard to the rotation. Large marker configurations are fine as

workspace frames. For instrument tracking, however, large marker configurations are not practical. We want to use small marker clusters, which do not get into the way when handling the instrument, and which we can keep apart from the markers that frame the workspace. We found that we do not obtain stable pose estimation from small clusters when the markers are distributed in a coplanar fashion. For a reliable estimation of the rotation, we need to distribute the markers in 3D.

Fig. 4 shows a biopsy needle, and attached to it a marker cluster design that we found to be efficient: it provides good pose results, and it is simple to fabricate at the same time. Flat disc-shaped markers are arranged in a multilevel planar fashion. For a given lateral extent of the marker body, there is a trade-off between its depth extent, and the range of viewing angles for which the markers are seen as separate entities in the tracking camera's image. Therefore, one wants to spread the markers out evenly. In our design, one marker is placed in the center, the other markers are arranged on a circle around it. The marker body shown in Fig. 4 measures about 8 cm in diameter, and is built from 6 mm thick material. The markers are arranged on several depth levels: The central marker sits two levels (1.2 cm) below the main level, three of the peripheral markers are placed two respectively three levels (1.2 cm and 1.8 cm) above the main level. "High" and "low" markers are mostly alternated in neighboring positions. The tracking camera can reliably locate the individual markers while tilting the marker body within an angle range of about 45° from the normal direction (i.e. the direction where the marker body directly faces the camera). As can be seen on Fig. 4, we attach the marker body to the needle in a tilted way, so that the markers look towards the head-mounted tracking camera when the user holds the needle in a comfortable standard position.

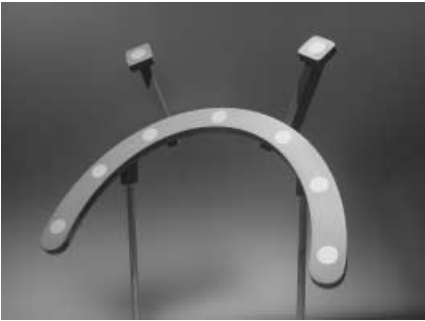


Fig. 3. Marker frame

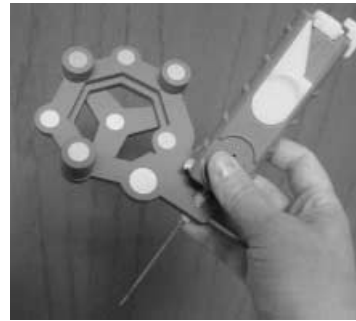


Fig. 4. Multilevel planar marker cluster attached to biopsy needle

2.4 System Performance

Our AR video system is running at the full standard video rate of 30 frames per second. We synchronize video and graphics, eliminating any time lag between the real and the virtual objects. The virtual objects do not lag behind, neither does one see them swim or jitter with respect to the real scene. As the augmented view shows the graphics firmly anchored in the real scene, the user can assess the information in a comfortable way. Overall, there is a time delay of about 0.1 seconds between an actual event and its display to the user.

We measured the overlay accuracy of our original system. Evaluating a set of augmented video images, we found the mismatch between calibration marks and their overlaid graphical counterparts to be typically smaller than 1 mm in object space, going up to 2 mm at the edges of the images. We do not have measurements for the needle placement configuration described in the present paper, but expect the accuracy to be in the same range. This is supported by simple visual inspection of the real needle as it appears in the video image and the virtual needle that is overlaid onto it. There is no apparent jitter in the overlay, so that such accuracy estimation can be performed with ease.

3 Needle Placement Phantom

3.1 Design

For a needle placement experiment, we designed a box with a set of mechanical push-buttons. The pushbuttons are like small pistons (Fig. 5) with a head diameter of 6 mm. Pushing down a piston in turn depresses a key of an underlying USB keypad. The keypad is connected to the computer and allows us to provide feedback to the user when he or she correctly places the needle onto one of the piston targets.

The targets are accessible through a round window on the slanted top face of the box (Fig. 6). A foam pad covers the window to hide the targets from the user's direct view. We chose a 5 cm thickness for the foam pad so that it provides mechanical resistance to the needle insertion. The targets lie about 7 cm below the top surface of the foam pad.

Fig. 7 shows the box for the needle placement experiment with the foam pad in place. It is sitting on a platform with a marker frame that contains seven coplanar markers on a half circle plus two additional ones that stick out on little posts. We also put retroreflective markers onto the heads of the piston targets. This allowed us to acquire the location of all the targets with respect to the marker frame coordinate system in a single measurement, using our stereo camera system ARTtrack.

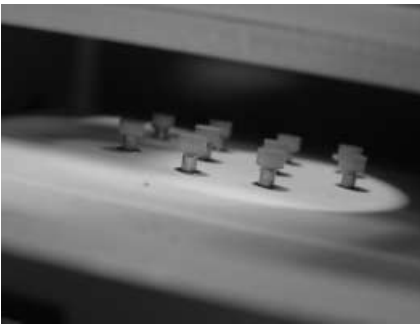


Fig. 5. Piston targets for needle placement



Fig. 6. View through window onto targets

3.2 Visualization

We visualize the top surface of the targets as flat discs. We surround each virtual target disc with a ring, rendered as a shaded torus. This torus helps with the 3D perception of the target location. We show the disc-torus target structure in a red color, which switches to a green color when the needle is pointing towards the target. The needle itself is visualized as a blue wireframe cylinder. A yellow wireframe cylinder marks the extrapolation of the needle path. Observing the intersection of the path cylinder intersects with the disc target, the user can easily see whether the needle is correctly pointing towards the target. Fig. 8 shows an example of an augmented view that guides the user. The needle is already partially inserted through the foam window, positioned about 1 cm above and correctly pointing to one of five targets shown.

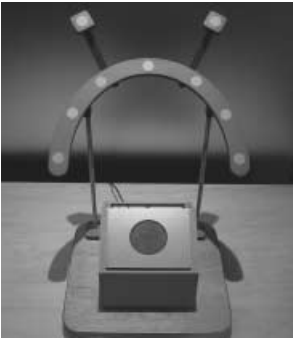


Fig. 7. Phantom box with foam window and frame of markers



Fig. 8. Augmented view for needle guidance

3.3 Needle Placement Experiment

Over one hundred users have tested the needle experiment. We usually slide the cover aside to show the real targets to the user, before we hide them again with the foam window. We then explain the needle visualization so that the user understands how to judge correct orientation and insertion depth of the needle.

The user is now asked to perform three needle placements. Initially, three targets are being shown. When the user correctly hits one of the targets (i.e. depresses the real piston with the needle), an audio signal sounds and the corresponding virtual target fades away.

Consistently, the users were able to hit the targets. In fact, the visualization is so intuitive and the visual feedback so conclusive, that one is basically not able to miss once one understands the basic concept. Most people grasped the concept immediately, some after a bit of experimentation. Even for the latter group the learning curve was below a minute. A couple of test users with a competitive attitude could successfully and repeatedly perform the three 7cm-deep needle placements at a rate of one per second, after only a few training trials.

The most common initial problem the test users had was to hold the needle in a way that the markers face towards the head-mounted tracking camera. In our opinion, the fact that we can track the marker body only over about $\pm 45^\circ$ away from the nor-

mal does not really represent a practical limitation for the needle placement. The user just needs to be aware not to turn the needle around its axis away from the tracking camera.

4 Summary and Conclusions

We developed an augmented reality system based on a stereoscopic video-see-through head-mounted display. Looking at the patient, the user can perceive medical images *in-situ*, e.g. see a virtual representation of a tumor in the location of the actual tumor. We extended this original system to include instrument tracking with our head-mounted tracking camera. For this, we designed a marker body in a multilevel planar configuration that provided very stable pose estimation results for single-camera tracking.

Making use of the new capability of instrument tracking, we designed a phantom box for a needle placement experiment. The user has to insert a needle through a foam pad and hit an underlying mechanical target. He or she is guided by the stereoscopic video view that is augmented with a graphics overlay showing the hidden target and the needle, including a forward extrapolation of the needle as an aiming aid. The user sees where the needle path intersects with the target, and can easily bring the needle into correct alignment. The user interface was experienced as very intuitive, and among a group of over one hundred test users, all were able to consistently succeed with the needle placement.

Augmented Reality guidance may be especially helpful when the user encounters complex anatomy, where vital structures like nerves or blood vessels have to be avoided while the needle is advanced towards a target like a tumor. Our system not only gives intuitive access to understanding the 3D geometry of the anatomy, it also provides a comfortable and believable augmented reality experience, where the graphical structures appear firmly anchored in the video scene. They do not jitter or swim, nor do they exhibit any time lag to the real objects in the video images. Currently, we are working towards testing the system in a clinical context.

References

1. M. Bajura, H. Fuchs, and R. Ohbuchi. "Merging Virtual Objects with the Real World: Seeing Ultrasound Imagery within the Patient." Proceedings of SIGGRAPH '92 (Chicago, IL, July 26-31, 1992). In *Computer Graphics* 26, #2 (July 1992): 203-210
2. .Andrei State, Mark A. Livingston, Gentaro Hirota, William F. Garrett, Mary C. Whitton, Henry Fuchs, and Etta D. Pisano, "Technologies for Augmented Reality Systems: realizing Ultrasound-Guided Needle Biopsies, " Proceed. of SIGGRAPH (New Orleans, LA, August 4-9, 1996), in *Computer Graphics Proceedings, Annual Conference Sereis1996*, ACM SIGGRAPH, 439-446.
3. Michael Rosenthal, Andrei State, Joohee Lee, Gentaro Hirota, Jeremy Ackerman, Kurtis Keller, Etta D. Pisano, Michael Jiroutek, Keith Muller, and Henry Fuchs, "Augmented Reality Guidance for Needle Biopsies: A Randomized, Controlled Trial in Phantoms," Proceedings of Medical Image Computing and Computer-Assisted Intervention – MICCAI 2001 (Utrecht, The Netherlands, October 14-17, 2001), *Lecture Notes in Computer Science* 2208, W. Niessen and M. Viergever (Eds.), Springer Berlin, Heidelberg, New York, pages 240-248.

4. Henry Fuchs, Mark A. Livingston, Ramesh Raskar, D'ardo Colucci, Kurtis Keller, Andrei State, Jessica R. Crawford, Paul Rademacher, Samuel H. Drake, and Anthony A. Meyer, MD. "Augmented Reality Visualization for Laparoscopic Surgery," Proceedings of Medical Image Computing and Computer-Assisted Intervention – MICCAI '98 (Cambridge, MA, USA, October 11-13, 1998), 934-943.
5. W. Eric L. Grimson, Ron Kikinis, Ferenc A. Jolesz, and Peter McL. Black, "Image-Guided Surgery," *Scientific American*, June, 1999, 62-69.
6. P.J. Edwards, D.J. Hawkes, DLG Hill, D. Jewell, R. Spink, A. Strong, and M. Gleeson, "Augmentation of Reality in the Stereo Operating Microscope for Otolaryngology and Neurosurgical Guidance," *Computer Aided Surgery* 1:172-178, 1995.
7. King AP, Edwards PJ, Maurer CR, de Cunha DA, Gaston RP, Clarkson M, Hill DLG, Hawkes DJ, Fenlon MR, Strong AJ, Cox TCS, Gleeson, MJ, "Stereo augmented reality in the surgical microscope," *Presence: Teleoperators and virtual environments* 9:360-368 2000.
8. W. Birkfellner, K. Huber, F. Watzinger, M. Figl, F. Wanschitz, R. Hanel, D. Rafolt, R. Ewers, and H. Bergmann, "Development of the Variscope AR, a See-through HMD for Computer-Aided Surgery," *IEEE and ACM Int. Symp. On Augmented Reality – ISAR 2000* (Munich, Germany, October 5-6, 2000), pages 54-59.
9. J.P. Rolland and H. Fuchs, "Optical versus Video See-Through Head-Mounted Displays in Medical Visualization," *Presence* (Massachusetts Institute of Technology), Vol. 9, No. 3, June 2000, pages 287-309.
10. F. Sauer, F. Wenzel, S. Vogt, Y. Tao, Y. Genc, and A. Bani-Hashemi, "Augmented Workspace: Designing an AR Testbed," *IEEE and ACM Int. Symp. On Augmented Reality – ISAR 2000* (Munich, Germany, October 5-6, 2000), pages 47-53.
11. Calvin Maurer, Frank Sauer, Chris Brown, Bo Hu, Benedicte Bascle, Bernhard Geiger, Fabian Wenzel, Robert Maciunas, Robert Bakos, and Ali Bani-Hashemi, "Augmented Reality Visualization of Brain Structures with Stereo and Kinetic Depth Cues: System Description and Initial Evaluation with Head Phantom," talk presented at SPIE's Int. Symp. on Medical Imaging 2001 (San Diego, CA, February 2001).
12. Frank Sauer, Ali Khamene, Benedicte Bascle, and G.J. Rubino, "A Head-Mounted Display System for Augmented Reality Image Guidance: Towards Clinical Evaluation for iMRI-guided Neurosurgery," *Proceedings of Medical Image Computing and Computer-Assisted Intervention – MICCAI 2001* (Utrecht, The Netherlands, October 14-17, 2001), *Lecture Notes in Computer Science* 2208, W. Niessen and M. Viergever (Eds.), Springer Berlin, Heidelberg, New York, pages 707-716.
13. Frank Sauer, Ali Khamene, Benedicte Bascle, Sebastian Vogt, and Gregory J. Rubino, "Augmented Reality Visualization in iMRI Operating Room: System Description and Pre-Clinical Testing," to appear in *SPIE Proceed. of Medical Imaging*, San Diego, February 2002.
14. Frank Sauer, Ali Khamene, Benedicte Bascle, Lars Schimmang, Fabian Wenzel, and Sebastian Vogt, "Augmented Reality Visualization of Ultrasound Images: System Description, Calibration, and Features," *IEEE and ACM Int. Symp. On Augmented Reality – ISAR 2001* (New York, NY, October 29-30, 2001), pages 30-39.
15. Frank Sauer, Ali Khamene, Benedicte Bascle, and Sebastian Vogt, "An Augmented Reality System for Ultrasound Guided Needle Biopsies," *Medicine Meets Virtual Reality MMVR 02/10* (Newport Beach, CA, January 2002), J.D. Westwood et al. (Eds.), IOS Press, 2002, pages 455-460.
16. Roger Y. Tsai, "A versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses", *IEEE Journal of Robotics and Automation*, Vol. RA-3, No. 4, August 1987, pages 323-344.
17. <http://www.cs.cmu.edu/~cil/v-source.html>: freeware implementation of the Tsai algorithm.