ELSEVIER

Medical Image Analysis 6 (2002) 313–320

MEDICAL IMAGE ANALYSIS

www.elsevier.com/locate/media

Augmented reality guidance for needle biopsies: An initial randomized, controlled trial in phantoms☆,†

Michael Rosenthal[a,*], Andrei State[a], Joohi Lee[a], Gentaro Hirota[b], Jeremy Ackerman[a], Kurtis Keller[a], Etta D. Pisano[c], Michael Jiroutek[d], Keith Muller[d], Henry Fuchs[a]

[a]*University of North Carolina at Chapel Hill, Department of Computer Science, CB 3175, Chapel Hill, NC 27599-3175, USA*
[b]*nVidia Corporation, 2701 San Tomas Expressway, Santa Clara, CA 95050, USA*
[c]*University of North Carolina at Chapel Hill, Department of Radiology, CB 7510, Chapel Hill, NC 27599-7510, USA*
[d]*University of North Carolina at Chapel Hill, Department of Biostatistics, CB 7420, Chapel Hill, NC 27599-7420, USA*

Abstract

We report the results of a randomized, controlled trial to compare the accuracy of standard ultrasound-guided needle biopsy to biopsies performed using a 3D Augmented Reality (AR) guidance system. A board-certified radiologist conducted 50 core biopsies of breast phantoms, with biopsies randomly assigned to one of the methods in blocks of five biopsies each. The raw ultrasound data from each biopsy was recorded. Another board-certified radiologist, blinded to the actual biopsy guidance mechanism, evaluated the ultrasound recordings and determined the distance of the biopsy from the ideal position. A repeated measures analysis of variance indicated that the head-mounted display method led to a statistically significantly smaller mean deviation from the desired target than did the standard display method (2.48 mm for control versus 1.62 mm for augmented reality, $p<0.02$). This result suggests that AR systems can offer improved accuracy over traditional biopsy guidance methods.

Keywords: Augmented reality; Breast biopsy; Ultrasound guidance; Needle biopsy

1. Introduction

Our research group at the University of North Carolina has been working in the area of augmented reality (AR) visualization for ultrasound examinations and ultrasound-guided procedures for nearly a decade (Bajura et al., 1992; Fuchs et al., 1996; Garrett et al., 1996; Jacobs et al., 1997; State et al., 2001, 1994, 1996b). The vision for this project is to allow physicians to directly see into a patient, aided by real-time computer graphics and augmented reality technology. The notion of augmenting the view of one's surroundings with computer-generated images has its roots in Ivan Sutherland's seminal paper (Sutherland, 1968), which described a system with a head-mounted display (HMD) whose synthetic images the user could see optically overlaid on the view of the room around him. Many years of research, both in the general AR field (Azuma, 1997) as well as in specific medical AR applications (Edwards et al., 2000; Maurer et al., 2001; Stetten and Chib, 2001), have resulted in considerable improvement in each of the key technologies.

Using our biopsy guidance system in January 1996, a trained physician (Pisano) was able to guide a needle into a lesion within an artificial breast training phantom and report that the task was 'easy' (Fig. 1). A subsequent test with a human subject progressed to where the needle was partially inserted towards the target lesion, at which point the physician was forced to abandon the AR guidance and continue the intervention with conventional ultrasound guidance technology. During this and several subsequent experiments it slowly became clear that despite the tech-

☆Electronic Annexes available. See www.elsevier.com/locate/media.

*Corresponding author.
E-mail address: rosentha@cs.unc.edu (M. Rosenthal).

†A preliminary version of this paper was presented at the Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2001 conference in Utrecht. The Netherlands (Rosenthal et al., 2001)

1361-8415/02/$ – see front matter
PII: S1361-8415(02)00088-9

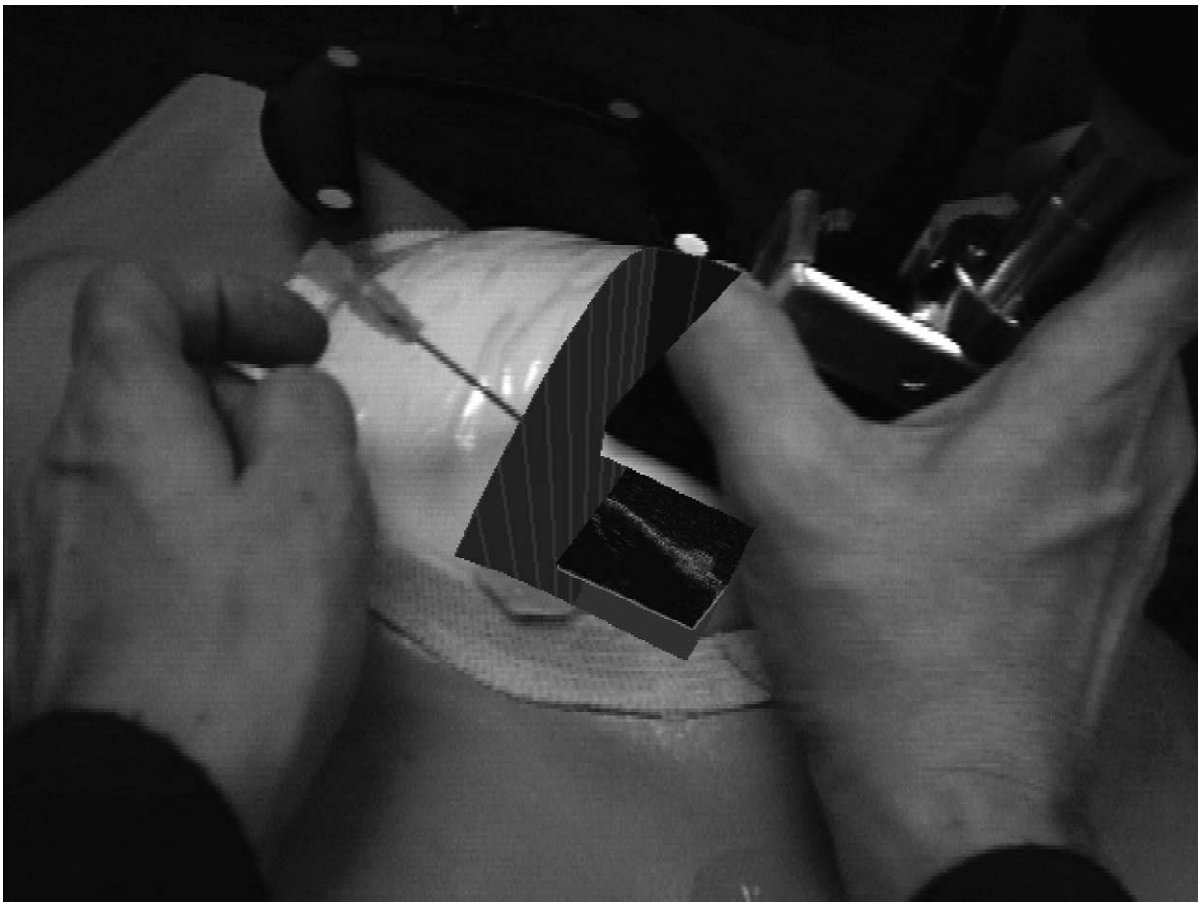

Fig. 1. View from the head-mounted display during a 1996 biopsy guidance experiment. The physician has inserted a cyst aspiration needle into a lesion within a breast phantom and holds the ultrasound transducer in her right hand. Correct ultrasound probe calibration and accurate tracking yield alignment between real needle and image of the needle in the ultrasound slice. The colored dots in the background are fiducials for head tracking correction (not used in our current system). A colour version of this figure is available at www.elsevier.com/locate/media (go to Electronic Annexes).

nological advancements effective patient studies were still not possible. This was mostly due to cumbersome equipment and inadequate tracking technology (Fuchs et al., 1996).

We have spent the intervening years developing an enhanced guidance system, which is now being used in live patient studies. In the following sections, we describe the new developments in our guidance system. We also describe the design and report the results of a randomized, controlled study to determine the relative effectiveness of our new AR system versus traditional ultrasound. We conclude with a description of our current and future work.

2. Materials and methods

Other papers have described our system design in detail (Fuchs et al., 1996; Garrett et al., 1996; Jacobs et al., 1997; State et al., 2001, 1996b). In the following sections, we provide a brief overview of our current system and the full design of our recent biopsy accuracy experiment.

2.1. Augmented reality guidance system

2.1.1. System overview

Our AR guidance system consists of four major components: an ultrasound imaging system, an instrument tracking system, a graphics and computation platform, and a head-mounted display (HMD). The ultrasound imaging system produces a live video stream that is digitized and fed into our computation platform. The computation platform also receives data on the position and orientation of the ultrasound scanner, the HMD, and the biopsy needle from our tracking system. This information is combined with our computational models of the instruments and patient to produce a ‘virtual’ procedure room. Selected features of this room are overlaid on the live video from our video-see-through HMD and fed back into the HMD, providing the user with an augmented view of the real world.

Additional details of our system, including our modeling and calibration methods, can be found in the references cited above.

2.1.2. Ultrasound imaging system

We are currently using a PIE Medical Ultrasound Scanner 350 to acquire ultrasound images during our experiments. The measurements on our ultrasound scanner have been tested to be within the manufacturer's specified error of less than 1 mm over a 20 mm span. The geometric transform between the ultrasound scanner images and our computational model is determined using a custom calibration phantom of interlaced silk threads.

2.1.3. Tracking system

We use an Image-Guided Technologies FlashPoint™ 5000 opto-electronic tracker in our system. The HMD, the ultrasound probe and the biopsy needle are all equipped with infrared LEDs. The FlashPoint delivers readings of the positions of these LEDs to the graphics computer with sub-millimeter accuracy. This HMD tracking technology is not quite as accurate as the closed-loop method used in our original 1996 system (State et al., 1996a), but it is superior to magnetic technologies and does not encumber the user's field of view (and the sterile operating field) with fiducials. The ultrasound probe is also tracked opto-electronically. It uses a specially developed 9-LED device that allows rotations up to 80° to any side without losing acquisition, thus freeing the physician to position and orient the probe in the most adequate way for a particular intervention.

2.1.4. Graphics and computation platform

The system runs on an SGI Onyx2 Reality Monster™ graphics computer equipped with multiple DIVO digital video input/output boards, allowing simultaneous capture of multiple video streams. The software runs at frame rates of 20–30 Hz in stereo on this platform. Fig. 2 shows imagery displayed by our system during an experiment with a breast training phantom in late 2000.

2.1.5. Head-mounted display

We have modified a stereoscopic Sony Glasstron LDI-D100 HMD[1] for use as our display system. This HMD provides full color, stereo, SVGA (800×600) resolution displays in a lightweight design. We have added an aluminum superstructure to hold two Toshiba IK-SM43H video cameras for image capture and three infrared LEDs for opto-electronic tracking. Fig. 3 shows the latest model

[1]Sony is no longer manufacturing the SVGA stereo version of their Glasstron HMD. Daeyang Corporation and Kaiser Electro-Optics manufacture HMDs that potentially could be modified for this application.

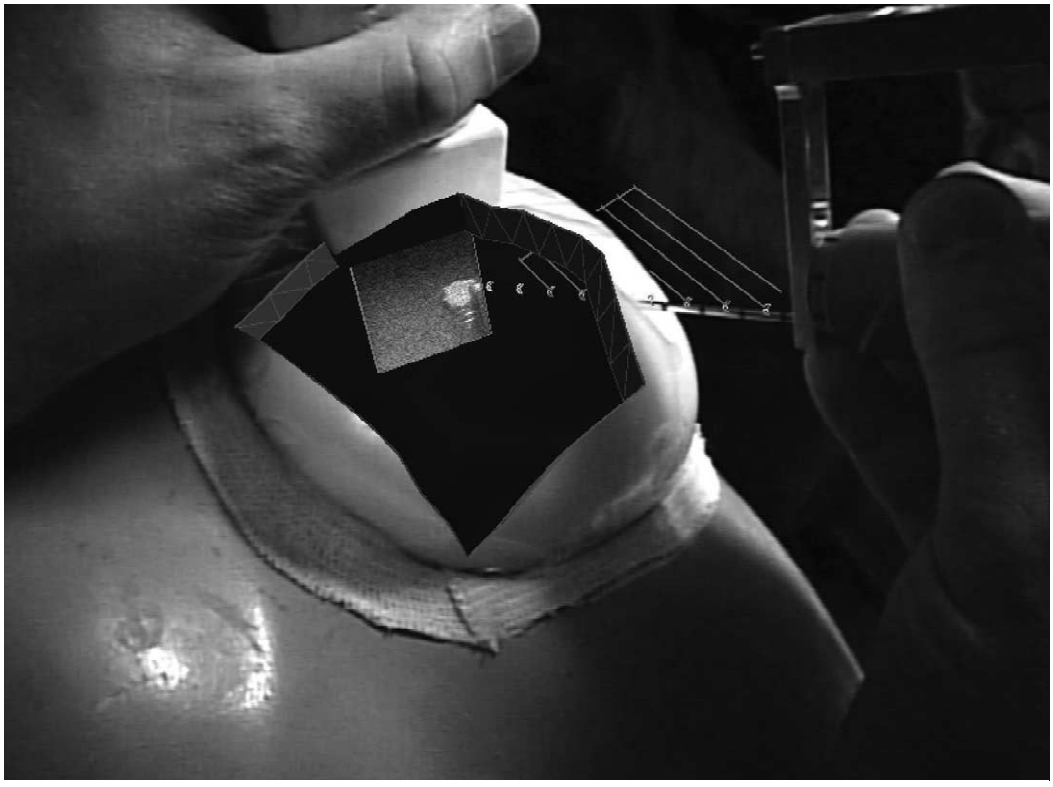

Fig. 2. Head-mounted display view during this phantom biopsy experiment. Both the ultrasound probe (left hand) and the biopsy needle (right hand) are tracked. The needle aims at the bright lesion visible in the ultrasound slice. The system displays the projection of the needle onto the plane of the ultrasound slice (blue lines) and also displays the projected trajectory of the needle if it were fired at this moment (yellow markers). A colour version of this figure is available at www.elsevier.com/locate/media (go to Electronic Annexes).

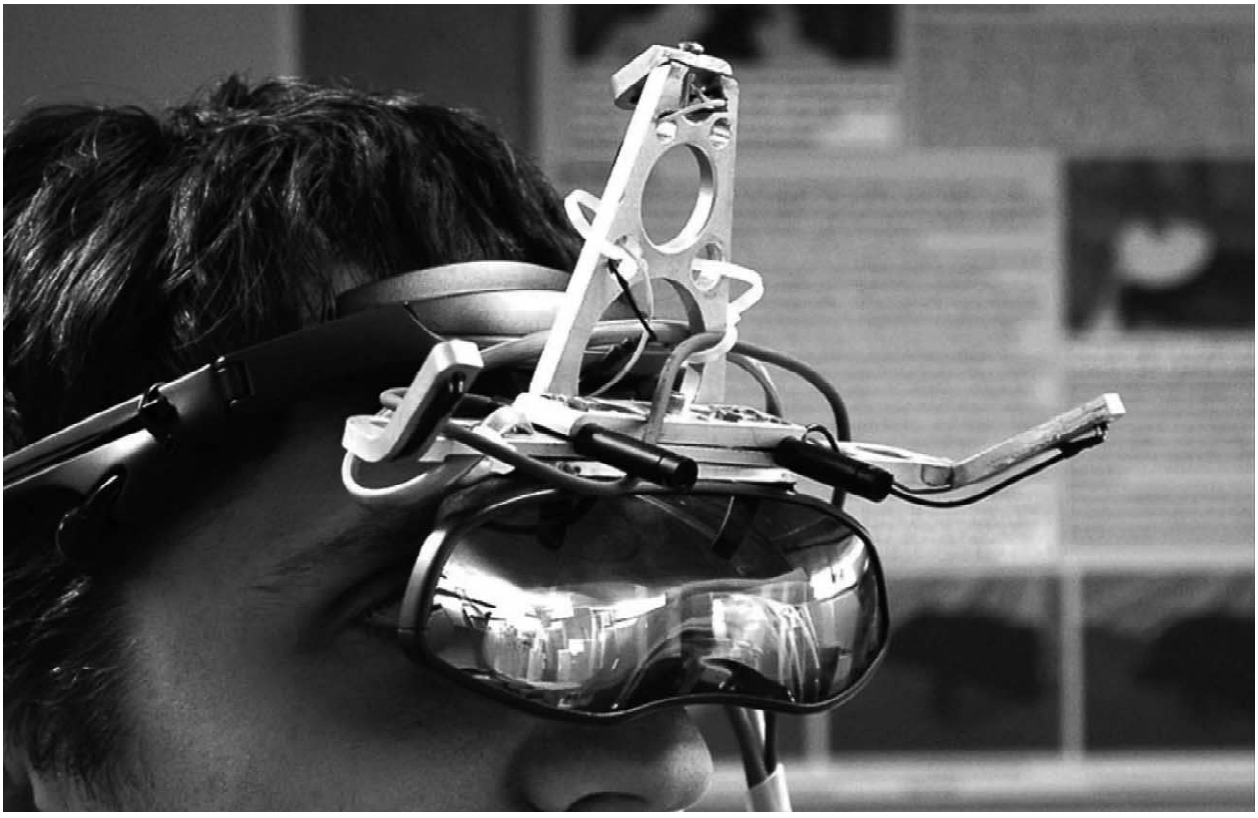

Fig. 3. Video-see-through augmented reality head-mounted display (HMD) built on the basis of a Sony Glasstron LDI-D100 device. The aluminum superstructure holds two miniature video cameras for image capture and three infrared LEDs for opto-electronic tracking of the HMD. A colour version of this figure is available at www.elsevier.com/locate/media (go to Electronic Annexes).

of our HMD. This 'video see-through' (Azuma, 1997) device and its operation are described in detail in (State et al., 2001).

2.2. Design of biopsy guidance study

We have performed an experiment to compare our AR guidance system to standard ultrasound guidance for the task of targeting needle biopsies in training phantoms. Our hypothesis was that the two guidance methods would be comparable in terms of needle placement accuracy for this task, indicating that it is safe to evaluate the AR system in humans.

The experimental component of this study was performed using the AR system described above. The control component was performed using only the PIE Medical Ultrasound Scanner 350 component of our system without any computer augmentation.

2.2.1. Biopsy task

Our task under evaluation was diagnostic biopsy of a simulated solid breast mass. Standard ultrasound training phantoms (Model 52 Biopsy Phantom, Computerized Imaging Reference Systems, Inc., Norfolk, VA) were used as our biopsy subjects. These phantoms each contained six tumor-like targets placed randomly throughout an ultrasound-compatible gel mold. The phantoms are approximately the size and shape of an average human breast. A new phantom was used whenever the radiologist felt that artifacts from previous biopsies were interfering with the current task.

Biopsies were performed using the standard procedure for breast biopsies in the US. In this approach, the lesion is located using the ultrasound scanner, and then the core biopsy needle is inserted in line with the plane of the ultrasound image. As the needle is advanced, the ultrasound probe is moved to keep the needle in full view at all times. To confirm the out-of-plane position, the needle is intermittently held steady while the ultrasound scanner is turned 90° and the needle's lateral placement is determined. This process is repeated until the biopsy needle is within range of the lesion. With the needle held in the plane of the ultrasound scanner, the biopsy mechanism is fired and the resulting biopsy position is confirmed in the standard plane and in the perpendicular ultrasound plane. This process is repeated to take biopsies from the center of the lesion and at the three, six, nine and twelve o'clock positions around the perimeter of the lesion (as viewed on

the plane orthogonal to the axis of the biopsy needle). Biopsy techniques are discussed in additional detail in (Parker et al., 1995; Pisano et al., 2001).

For each selected lesion in a phantom, five biopsies were performed using the method described above with a 14-gauge Monopty core biopsy needle (C.R. Bard, Inc., Covington, GA). The needle was withdrawn from the phantom after each biopsy attempt. The ultrasound video from each biopsy was reformatted and recorded directly from the ultrasound scanner to digital video tape for later evaluation.

The biopsies were all performed during a single experimental session.

2.2.2. Randomization and control scheme

This study was designed as a randomized, controlled trial in order to limit the effects of confounding factors. A single board-certified radiologist (Pisano) performed all of the biopsies in this experiment (Fig. 4). Ten targets within the phantoms were sequentially selected; five biopsies were performed on each lesion before selecting the next target. Randomization to the two guidance methods was performed by a coin flip before the selection of each biopsy target.

2.2.3. Evaluation of accuracy

Another board-certified radiologist (Cherie Kuzmiak, DO) evaluated the ultrasound video to determine the accuracy of each biopsy. The evaluator was blinded to the method of guidance for each biopsy. For each biopsy, she determined the geometric distance (in mm) between the ideal biopsy target point and the actual biopsy positions in the plane orthogonal to the needle. The evaluator also measured the dimensions of the lesions along the needle axis and along two perpendicular directions (approximately vertical and horizontal). The results were later entered into an Excel spreadsheet by another individual (MR) and associated with the corresponding guidance method.

The geometric distances mentioned above were measured on an NTSC display with respect to the ultrasound machine's displayed reference ruler. This is the standard clinical technique for anatomical measurements in breast biopsy. While it would be desirable to have confirmation from an independent modality, we know of no alternative 'gold standard' for accurately validating these measurements. We believe that the randomization and blinding scheme should minimize the impact of this uncertainty by minimizing the likelihood of a bias in measurement error in favor of either method.

2.2.4. Statistical analyses

Descriptive statistics (mean$\pm$S.D.) of the error distances were calculated. Separate and combined results were computed for the HMD and standard display methods for each location, mean error across locations and the mean of the maximum lesion dimension. The primary analysis was a repeated measures analysis of variance (REPM ANOVA) utilized to address the multiple locations targeted within each lesion (a within-'subject' repeated measures dimension). The SAS® procedure GLM was utilized.

To rule out lesion size bias as contributing to the effect attributed to display method in the primary analysis, we performed an exploratory full model in every cell (FMIC) REPM ANOVA analysis to show that the effect due to lesion size was not significant between the display methods. The FMIC was then reduced to a multivariate analysis of covariance (MANCOVA) model and reanalyzed. Maximum lesion dimension (in mm) was the measure we chose to represent lesion size.

3. Results

A total of 50 biopsies were performed: 25 in each of the AR guidance and standard guidance groups. The mean error distances for each of these groups are shown in Table 1 below, while Fig. 5 shows the distribution of errors for both groups. A repeated measures analysis of variance indicated that the HMD display method led to a statistically significantly smaller mean deviation from the desired target than did the standard display method (2.48 mm for control versus 1.62 mm for augmented reality, $p<0.02$). The biopsy location and the location-display combination did not yield statistically significant effects upon the accuracy.

The supportive FMIC ANOVA and MANCOVA analyses of the effects of lesion dimensions upon accuracy indicated that the maximum lesion dimension had no significant effect upon placement error ($p>0.05$ for the main effect and all combinations involving maximum lesion dimension). These results indicate that the guidance method was the only factor associated with a statistically significant difference in placement error.

The measured lesion dimensions were consistent with the manufacturer's specifications for the biopsy phantoms. The lesions were stated to be 6–12 mm in diameter by the manufacturer; our measurements ranged from 5 to 12.5 mm with a mean of 10.3 mm.

4. Conclusions

The results of the above study indicate that the AR guidance system yielded statistically improved accuracy as compared to the standard ultrasound guidance method for our expert user. In fact, we did not expect the AR system to be as good as the conventional guidance technique, especially for the expert user (Pisano). Our goal was to demonstrate the system's effectiveness on a procedure that is simple and not dangerous to the patient. The indication that the AR technique may be better in this comparison, where the advantage should go to the conventional ap-

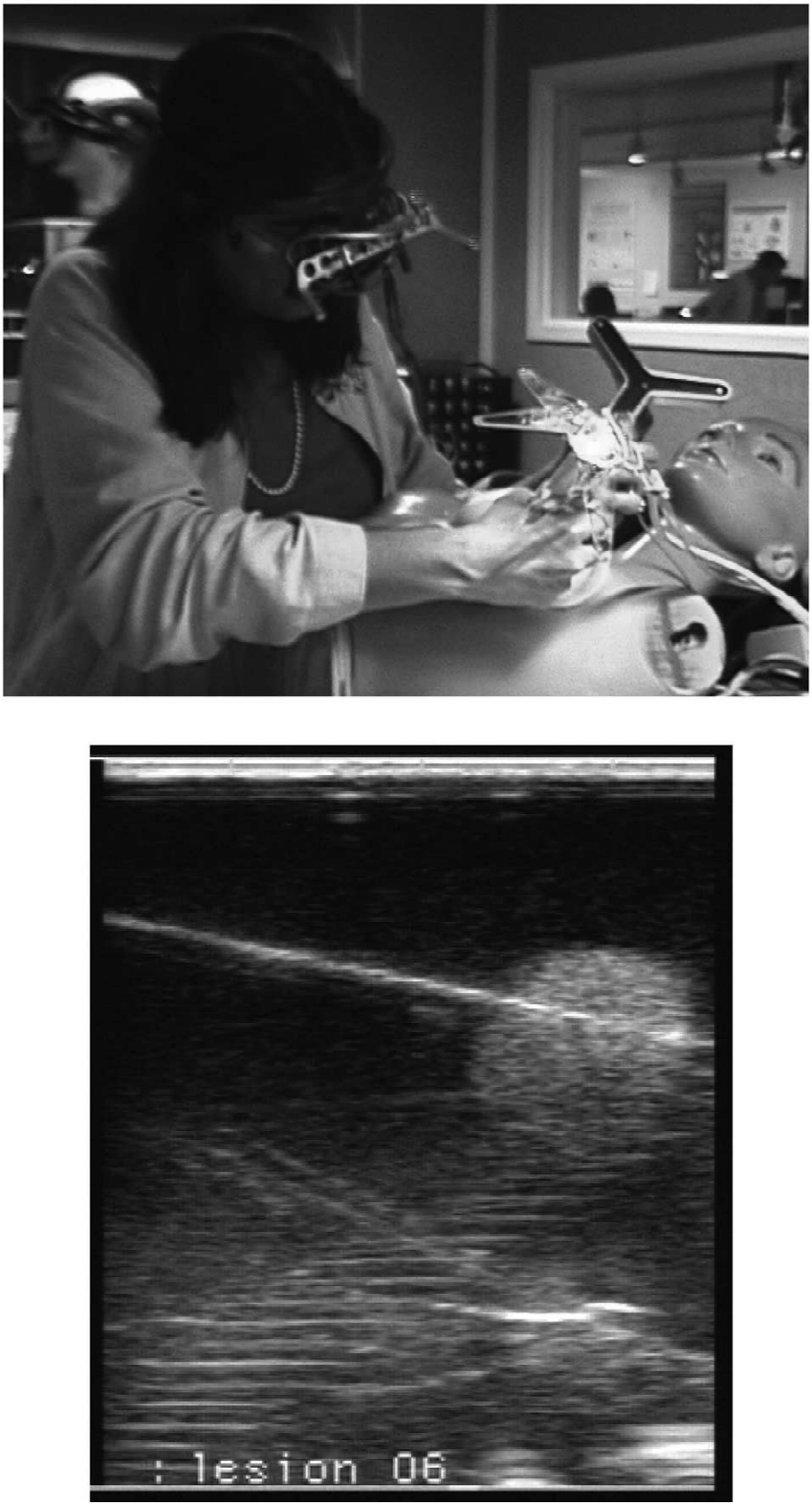

Fig. 4. Lab view (top) and ultrasound image (bottom) while the physician, wearing the Glasstron-based head-mounted display (HMD), performs a controlled study with the 2000 system. She holds the opto-electronically tracked ultrasound probe and biopsy needle in her left and right hands, respectively. Refer to Fig. 3 for an example of the view from the HMD. The horizontal streaks in the ultrasound image are common artifacts of acoustic reflection. A colour version of this figure is available at www.elsevier.com/locate/media (go to Electronic Annexes).

Table 1
Results from the phantom biopsy study

(All measures in mm, mean±S.D.)	Standard guidance	AR guidance
Error at center	4.20±1.92	1.50±1.41
Error at 3 o'clock	2.00±1.87	1.70±0.67
Error at 6 o'clock	1.20±0.84	0.90±1.02
Error at 9 o'clock	2.00±1.58	0.80±1.30
Error at 12 o'clock	3.00±2.00	3.20±2.05
Mean error across locations	2.48±0.44	1.62±0.48
Mean of maximum lesion dimension	10.50±3.26	12.00±2.09

proach, is both surprising and encouraging. We are hopeful that user studies with less experienced physicians may show an even greater improvement using an AR approach.

There are several limitations to this study that should be addressed in future investigations. We did not look for temporal or fatigue effects in our study due to the relatively short duration of the experiment (approximately 2 h) and limited number of data blocks (ten). A larger study will be needed to evaluate the possible trends in user performance over single sessions and over longer time periods. There may also be significant variability between users in accuracy and fatigue effects.

We did not measure the error in needle placement *along* the axis of the biopsy needle, the accuracy of which could be affected by differences in the display modalities. While the biopsy depth of the needle (~1 cm) is quite large relative to the size of the lesions and thus unlikely to miss in depth given the standard needle positioning technique, this should be measured in future studies. The structure of the distribution of errors in the study was also difficult to interpret. It is possible that there are multiple subpopulations of errors that should be explored through a larger study. The relatively noise-free images of training phantoms also may pose less of a challenge in an HMD than will the noisy images from human tissue. Finally, procedures on phantoms may be more advantageous for the new approach than procedures with live patients, since phantoms have simpler tissue characteristics and less noise in the ultrasound images.

Additional studies with human subjects are currently underway to confirm that these benefits translate to real improvements in medical care. Beyond that we are considering two possibly parallel paths of research: (1) exploring the AR approach for relatively simple medical tasks, such as cyst aspiration, for primary care physicians, and (2) investigating the AR approach for needle placement in more difficult areas of the body (e.g. liver), in which targets are in heavily vascular regions where avoidance of major vessels is a prime consideration. The incorporation of other imaging modalities, such as magnetic resonance angiograms or 3D ultrasound, could also expand the possible applications of this research.

While results reported here are preliminary and of limited scope, we believe that they suggest the potential of AR visualization to improve patient care. Further research will be needed to evaluate the usefulness of AR for each of the large host of possible applications. We hope that the next decade of research will continue to explore the

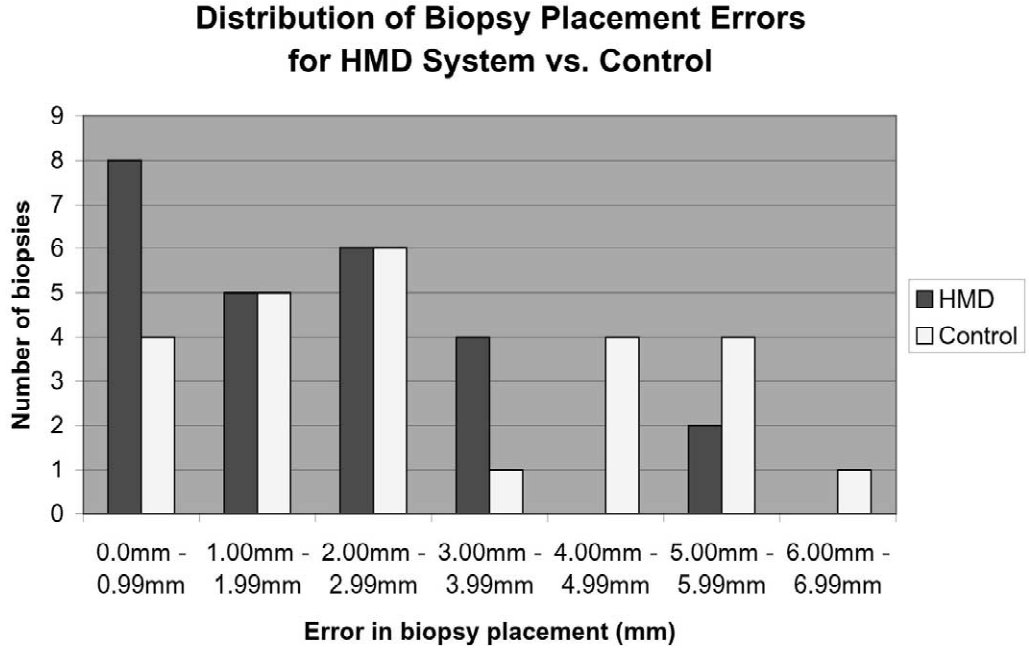

Fig. 5. Distribution of biopsy placement errors using head-mounted display (HMD; dark bars) and standard ultrasound guidance methods (Control; light bars). The HMD group showed a statistically significant reduction in mean biopsy placement error (1.62 mm for HMD versus 2.48 mm for control, $p<0.02$). A colour version of this figure is available at www.elsevier.com/locate/media (go to Electronic Annexes).

potential of augmented reality for both medical and non-medical applications.

Acknowledgements

Past and current researchers: Michael Bajura, Andrew Brandt, David T. Chen, D'nardo Colucci, Jessica R. Crawford, William F. Garrett, Jack Goldfeather, Arthur Gregory, Marco C. Jacobs, Mark A. Livingston, Michael J. North, Ryutarou Ohbuchi, Stephen M. Pizer, Paul Rademacher, Chris Tector, Mary C. Whitton.

Past and current medical collaborators: Nancy Chescheir MD, Mark Deutchman MD, Ricardo Hahn MD, Vern Katz MD, Cherie Kuzmiak DO, Matthew Mauro MD, Anthony A. Meyer MD, Melanie Mintzer MD.

Technical assistance: Samuel H. Drake, Caroline K. Green, David G. Harrison, John E. Thomas.

Support: NIH (CA 47982-10), NSF STC for Computer Graphics and Scientific Visualization (NSF Cooperative Agreement ASC-8920219), PIE Medical, and SGI.

References

Azuma, R.T., 1997. A survey of augmented reality. Presence: Teleoperators and Virtual Environments 6 (4), 355–385.

Bajura, M., Fuchs, H., Ohbuchi, R., 1992. Merging virtual objects with the real world: seeing ultrasound imagery within the patient. Computer Graphics, Proceedings of SIGGRAPH 1992, 26 (2), 203–210.

Edwards, P.J., King, A.P., Maurer, C.R.J., deCunha, D.A., Hawkes, D.J., Hill, D.L.G., Gaston, R.P., Fenlon, M.R., Jusczyzck, A., Strong, A.J., Chandler, C.L., Gleeson, M.J., 2000. Design and evaluation of a system for microscope-assisted guided interventions (MAGI). IEEE Trans. Med. Imag. 19 (11), 1082–1093.

Fuchs, H., State, A., Pisano, E., Garrett, W., Hirota, G., Livingston, M., Whitton, M., Pizer, S., 1996. Performing ultrasound-guided needle biopsies from within a head-mounted display. Proc. Visualizat. Biomed. Comput. 1996, 591–600.

Garrett, W., Fuchs, H., Whitton, M., State, A., 1996. Real-time incremental visualization of dynamic ultrasound volumes using parallel BSP trees. Proc. IEEE Visualizat. 1996, 235–240.

Jacobs, M., Livingston, M., State, A., 1997. Managing latency in complex augmented reality systems. In: ACM SIGGRAPH Proceedings of the Symposium of Interactive 3D Graphics 1997, Providence, RI, pp. 49–54.

Maurer, C.R.J., Sauer, F., Hu, B., Bascle, B., Geiger, B., Wenzel, F., Recchi, F., Rohlfing, T., Brown, C.M., Bakos, R.S., Maciunas, R.J., Bani-Hashemi, A., 2001. Augmented reality visualization of brain structures with stereo and kinetic depth cues: System description and initial evaluation with head phantom. Presented at Medical Imaging 2001: Visualization, Display, and Image-Guided Procedures. Proc. SPIE 4319, 445–456.

Parker, S.H., Stavros, A.T., Dennis, M.A., 1995. Needle biopsy techniques. Radiol. Clin. North Am. 33 (6), 1171–1186.

Pisano, E.D., Fajardo, L.L., Caudry, D.J., Sneige, N., Frable, W.J., Berg, W.A., Tocino, I., Schnitt, S.J., Connolly, J.L., Gatsonis, C.A., McNeil, B.J., 2001. Fine-needle aspiration biopsy of nonpalpable breast lesions in a multicenter clinical trial: results from the radiologic diagnostic oncology group V. Radiology 219 (3), 785–792.

Rosenthal, M.R., State, A., Lee, J., Hirota, G., Ackerman, J., Keller, K., Pisano, E., Jiroutek, M., Muller, K., Fuchs, H., 2001. Augmented reality guidance for needle biopsies: a randomized, controlled trial in phantoms. In: Niessen, W., Viergever, M. (Eds.), Proceedings of MICCAI 2001, Lecture Notes in Computer Science, 2208, pp. 240–248.

State, A., Chen, D., Tector, C., Brandt, A., Chen, H., Ohbuchi, R., Bajura, M., Fuchs, H., 1994. Case study: observing a volume-rendered fetus within a pregnant patient. In: Proceedings of IEEE Visualization 1994, Los Alamitos, CA, pp. 364–368.

State, A., Hirota, G., Chen, D., Garrett, W., Livingston, M., 1996a. Superior augmented-reality registration by integrating landmark tracking and magnetic tracking. In: ACM SIGGRAPH Proceedings of the Symposium of Interactive 3D Graphics 1997, Providence, RI, pp. 429–438.

State, A., Livingston, M., Hirota, G., Garrett, W., Whitton, M., Fuchs, H., Pisano, E., 1996b. Technologies for augmented-reality systems: realizing ultrasound-guided biopsies. Computer Graphics, Proceedings of SIGGRAPH 1996, pp. 439–446.

State, A., Ackerman, J., Hirota, G., Lee, J., Fuchs, H., 2001. Dynamic virtual convergence for video see-through head-mounted displays: maintaining maximum stereo overlap throughout a close-range work space. In: Proceedings of the International Symposium on Augmented Reality 2001, New York, NY, pp. 137–146.

Stetten, G., Chib, V., 2001. Overlaying ultrasound images on direct vision. J. Ultras. Med. 20 (1), 235–240.

Sutherland, I., 1968. A head-mounted three-dimensional display. AFIPS Conf. Proc. 33-I, 757–764.