IEEE/ACM International Symposium on Augmented Reality, October 2001, New York City, pp. 11-19

Real Time Tomographic Reflection: Phantoms for Calibration and Biopsy

George Stetten, M.D., Ph.D., Vikram Chib, Daniel Hildebrand, Jeannette Bursee
Department of Bioengineering, University of Pittsburgh
The Robotics Institute, Carnegie Mellon University
www.stetten.com

Abstract

We aim to validate Real Time Tomographic Reflection (RTTR) as an image guidance technique for needle biopsy. RTTR is a new method of in situ visualization, which merges the visual outer surface of a patient with a simultaneous ultrasound scan of the patient's interior using a half-silvered mirror. The ultrasound image is visually merged with the patient, along with the operator's hands and the invasive tool in the operator's natural field of view. Geometric relationships are preserved in a single environment, without the tool being restricted to lie in the plane of the ultrasound slice. The present experiment illustrates the effectiveness of needle biopsy using RTTR on a phantom consisting of an olive embedded in a turkey breast and discusses several prototypes of calibration phantoms.

1. Introduction

The innovation described in this paper derives from an extensive body of prior work whose goal has been to look directly into the human body in a natural way. From the discovery of X-rays over a century ago, clinicians have been presented with an amazing assortment of imaging modalities capable of yielding maps of localized structure and function within the human body. Continual advances are being made in magnetic resonance (MR), computerized tomography (CT), positron emission tomography (PET), single photon emission computerized tomography (SPECT), ultrasound, confocal microscopy, optical coherence tomography (OCT), and ultrasound backscatter microscopy (UBM). Each of these is a *tomographic* imaging modality, meaning that the data are localized into voxels, rather than projected along lines of sight, as they are in conventional X-ray images. Tomographic images, with their unambiguous voxels, are essential for our present work.

New techniques to display tomographic images in such a way as to extend human vision into the body have lagged behind the development of the imaging modalities themselves. In the practice of medicine, the standard method of viewing an image is still to examine a film or screen rather than look directly into the patient. Previous experimental approaches to fuse images with direct vision have not met with widespread acceptance, in part, because of their complexity. Our approach is simpler, and thus, we hope, more likely to find its way into clinical practice. If so, the technique could have a broad impact on the use of imaging in the interventional diagnosis and treatment of disease. We begin this paper with a review of our development of RTTR using ultrasound.

2. Real Time Tomographic Reflection

Although our research may eventually be adapted to a wide variety of imaging modalities, our roots are in ultrasound. Ultrasound is appealing because it is non-ionizing, real-time, and relatively inexpensive. The transducer is small and easily manipulated permitting rapid control of the location and orientation within the patient of the illuminated slice. Ultrasound is thus well suited for guiding invasive procedures, but difficulties arise in determining the spatial relationships between anatomical structures, the invasive tool, and the ultrasound slice itself.

Percutaneous ultrasound-guided intervention encompasses a wide range of clinical procedures [1-3]. In such procedures, the needle may be manipulated freehand with the ultrasound transducer held in the other hand or by an assistant. Alternatively, a needle may be constrained by a guide attached to the transducer so that the entire length of the needle remains visible within the plane of the ultrasound scan. In either case, the operator must look away from the patient at the ultrasound display and employ a displaced sense of hand-eye coordination. The difficulty in mastering these skills has motivated research into developing a more natural way to visually merge ultrasound with the perceptual real world.

Two distinct approaches have generally been explored in what is called *augmented reality* using a Head Mounted Display (HMD). In the first approach, small half-silvered mirrors mounted in the HMD are used to optically combine a direct view with a graphical image [4]. Alternatively, direct vision can be replaced with miniature video cameras in the HMD, displaying video and graphical images merged on miniature monitors in the HMD. Fuchs et al. have developed an ultrasound augmented reality system following the second approach. Replacing direct vision with video permits greater control of the display, although it introduces significant reduction in visual resolution [5-7]. The HMD and the ultrasound transducer must be tracked so that an appropriate perspective can be computed for the ultrasound images. Head-mounted displays, in general, restrict the operator's peripheral vision and freedom of motion.

In related work at the Medical Robotics and Computer Aided Surgery (MRCAS) laboratory at the CMU Robotics Institute, DiGioia, et al., have merged real-world images with CT data while achieving a reduction in the total apparatus that the operator must wear [8, 9]. In their system, called *image overlay*, a large half-silvered mirror is mounted above the patient with a flat panel monitor fixed above the mirror. Images of the CT data on the monitor are reflected by the mirror and superimposed on the operator's view of the patient through the mirror. The operator needs only to wear a small head-tracking optical transmitter, so that the proper three-dimensional perspective of the CT data can be rendered from his/her particular viewpoint. A second tracking device is attached to the patient to achieve proper registration between the rendered CT data and the patient. Special glasses are needed only if stereoscopic visualization is desired.

We have modified DiGioia's approach, applying it to ultrasound with significant simplification. By restricting ourselves to a single tomographic slice in real time (i.e. ultrasound), and strategically positioning the transducer, the mirror, and the display, we have eliminated the need for tracking either the observer or the patient. This is possible because we are actually merging the *virtual image* in 3D with the interior of the patient. The word "virtual" is used here in its classical sense: the reflected image is optically indistinguishable from a light-emitting slice suspended in space within the patient.

Ultrasound produces a *tomographic* slice representing a set of 3D locations that lie in a plane. The image of that tomographic slice, displayed at its correct size on a flat panel display, may be reflected to occupy the same physical space as the actual slice within the patient. If a half-silvered mirror is used, the patient may be viewed through the mirror with the reflected image of the slice superimposed, independent of viewer location. The reflected image truly occupies its correct location within the patient and does not require a particular perspective to be rendered for each viewpoint. We have adopted the term *tomographic reflection* to convey this concept [10-13].

Masamune et al have previously demonstrated the concept of tomographic reflection, which he calls *slice display*, on CT data, although not in real time [14]. A slice from a stored CT data set is displayed on a flat panel monitor and reflected by a half-silvered mirror to occupy its proper location within the patient. Since the CT scanner is not physically a part of this apparatus, independent registration of the patient's location is still required. Furthermore, the static data does not change during a procedure. Using a real time imaging modality such as ultrasound eliminates these restrictions, resulting in what we call *real time tomographic reflection* (RTTR) [TM].

To accomplish RTTR, certain geometric relationships must exist between the slice being scanned, the monitor displaying the slice, and the mirror. As shown in Fig. 1, the mirror must bisect the angle between the slice and the monitor. On the monitor, which must be a flat monitor, the image must be correctly translated and rotated so that each point in the image is paired with a corresponding point in the slice to define a line segment perpendicular to, and bisected by, the mirror. By fundamental laws of optics, the ultrasound image will thus appear at its physical location, independent of viewer position. The actual apparatus we have constructed is sketched in Fig. 2.

Superimposing ultrasound images on human vision using RTTR may improve an operator's ability to find targets while avoiding damage to neighboring structures, while generally facilitating interpretation of ultrasound images by relating them spatially to external anatomy. As such, it holds promise for increasing accuracy, ease, and safety during percutaneous biopsy of suspected tumors, amniocentesis, fetal surgery, brain surgery, insertion of catheters, and many other interventional procedures.

In Fig 3, a human hand is seen with the transducer pressed against the soft tissue between the thumb and index finger. While not a common target for clinical ultrasound, the hand was chosen because it clearly demonstrates successful alignment. The external surfaces of the hand are located consistent with structures within the ultrasound image. The photograph cannot convey the strong sense, derived from stereoscopic vision, that the reflected image is located within the hand. This sense is intensified with head motion because the image remains properly aligned from different viewpoints. To one experiencing the technique in person, ultrasound targets within the hand would clearly be accessible to direct percutaneous injection, biopsy or excision. We have named the device the *Sonic Flashlight* ™.

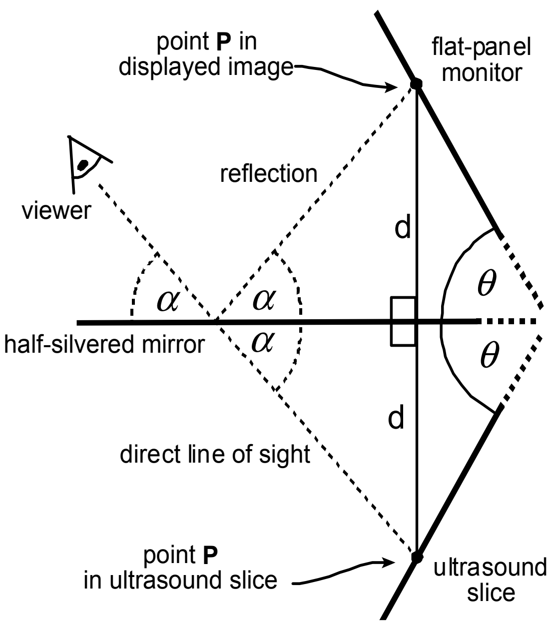

Fig. 1 The half-silvered mirror bisects the angle between the ultrasound slice (within the target) and the flat-panel monitor. Point P in the ultrasound slice and its corresponding location on the monitor are equidistant from the mirror along a line perpendicular to the mirror (distance = d). Because the angle of incidence equals the angle of reflectance (angle = α) the viewer (shown as an eye) sees each point in the reflection precisely at its corresponding physical 3D location.

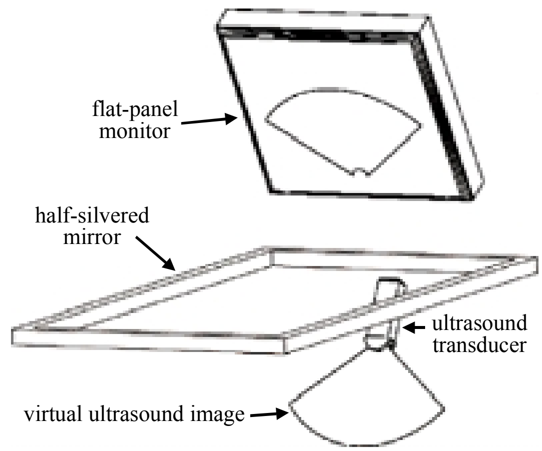

Fig. 2 Schematic representation of the apparatus. A flat-panel monitor and an ultrasound transducer are placed on opposite sides of a half-silvered mirror such that the mirror bisects the angle between them.

Fig. 3 Photograph, from the viewpoint of the operator, showing a scan of a hand using the apparatus in Fig. 2. The reflected ultrasound image is merged with the direct visual image.

IEEE/ACM International Symposium on Augmented Reality, October 2001, New York City, pp. 11-19

Even as the initial prototype in Fig. 2 was being constructed, it was clear that the ultrasound transducer must be free to move in the clinician's hand. One way to accomplish this is to mount the mirror and flat-panel display directly on the ultrasound transducer. It is envisioned that this handheld version of the sonic flashlight is likely to be one of its most useful forms in clinical practice, facilitating the insertion of intravenous catheters and other common procedures. A hand-held version of the Sonic Flashlight is shown in Fig. 4.

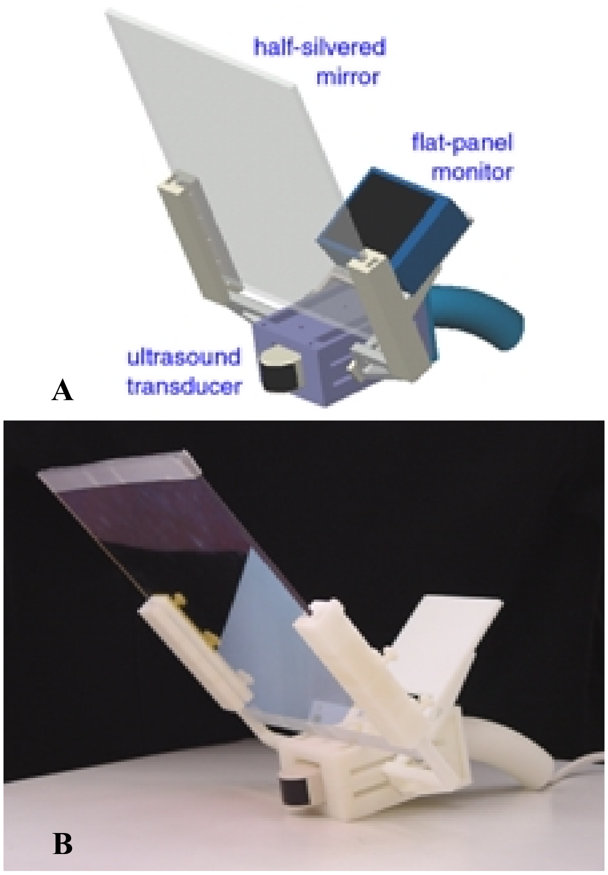

Fig. 4 **A**. Computer aided design (CAD) drawing of hand-held Sonic Flashlight. **B.** Rapid prototyping version using Fused Deposition Modeling (flat panel monitor not shown).

3. Calibration of the RTTR Display

Clearly, a method for accurately calibrating the RTTR display is required. Without calibration, the RTTR system would be unable to guide interventional procedures, because the virtual image and the actual slice would not be registered. Calibration of an RTTR system requires careful consideration of the degrees of freedom in the registration process. The challenge is to make each pixel in the virtual image occupy, and therefore appear to emanate from, its actual 3D location in the slice, as was shown in Fig. 1.

3.1 Linear Transform Approximation

We first consider only the *geometric transform* of a rigid body, i.e., we assume that the ultrasound slice is displayed without distortion at its correct scale on a perfectly flat monitor. The geometric transform required to superimpose the virtual image onto the slice can be represented as two sets of translations and rotations, each of which has 3 degrees of freedom.

Fig. 5A depicts 3 degrees of freedom (2 rotations and 1 translation) that allow the flat panel display to be moved into its correct plane, making the virtual image coplanar with the actual ultrasound slice. We can achieve this by physically moving the display, the ultrasound transducer, or the mirror.

Fig 5B shows the remaining 3 degrees of freedom (two translations and one rotation) that can be achieved simply by adjusting the image of the ultrasound slice on the flat panel monitor.

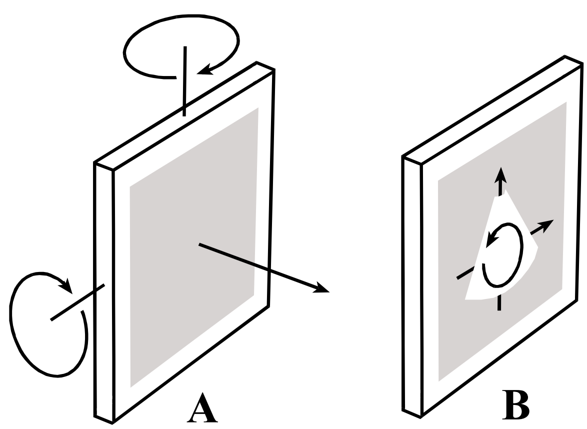

Fig. 5 Geometric transforms for **(A)** physically moving the display, and **(B)** moving the image on the screen.

The assumption that both the ultrasound slice and the flat panel monitor are truly *flat* is generally accurate, and makes a rigid-body (geometric) transform sufficient to move the monitor into the correct plane

(Fig. 5A). However, to correctly adjust the image on the screen actually requires more degrees of freedom than those shown in Fig. 5B, if only because the image must be scaled to its proper size. Extending the geometric transform to an *affine transform* allows us to scale the ultrasound slice to its correct size, as well as adjust its aspect ratio and correct for skewing. Certain computers (e.g., Silicon Graphics O2) are capable of mapping a video image through an affine transform in real time using *2D texture memory*. This permits the displayed ultrasound slice to be translated, rotated, and anisotropically scaled. The calibration process becomes a matter of finding the optimal parameters for that affine transform.

Mapping location (x,y) to (x',y') with an affine transform is accomplished by multiplying the homogeneous form of (x,y) by a 3 x 3 matrix **A**. As demonstrated by the equation below, an affine transform is capable of mapping any triangle to any other triangle. If the apices of both triangles are known (and not colinear), we can find an explicit solution for the 6 unknown elements of the matrix **A**.

$$\begin{bmatrix} x_1' & x_2' & x_3' \\ y_1' & y_2' & y_3' \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ 1 & 1 & 1 \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ 0 & 0 & 1 \end{bmatrix}$$

3.2 Water Tank Calibration Phantom

Based on this theory, we have constructed a "bead" phantom with 3 *training* beads suspended in roughly an equilateral triangle on threads in the plane of the ultrasound slice (see Fig. 6). The threads are held in the plane of the slice by a frame attached directly to the transducer. An uncalibrated ultrasound image of this phantom in a water tank can be captured and displayed on the flat panel monitor, and the water subsequently drained to avoid diffraction at the air-water interface during visual inspection. The translational errors in the location of each bead can then be used to solve for the affine transform **A**. Assuming the virtual image is coplanar with the slice, the calibration process is independent of viewer location.

Further calibration beyond the affine transform may be required to correct for non-linear geometry in both the imaging system and the display by warping the image with a non-linear transform.

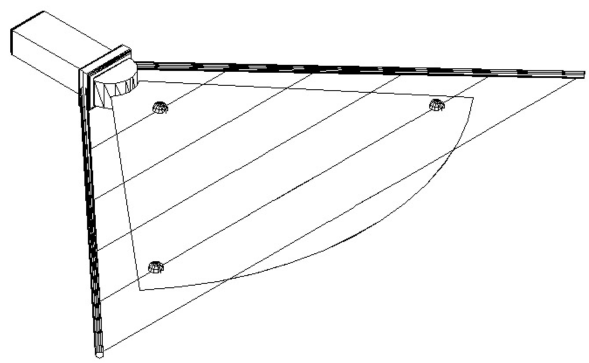

Fig. 6 Water tank phantom: 3 beads suspended on strings in the plane of the ultrasound slice by a frame mounted directly on the transducer.

To the extent that the slice geometry does not change with tissue type, and therefore remains constant as the transducer is moved relative to the target object, calibration of the system may only need to be performed initially using a phantom such as that just described. This initial calibration will suffice for that part of image geometry which is fully definable relative to the scanner. Further changes in image geometry due to tissue properties may be due to changes in the speed of sound in different tissue types and to refraction of the ultrasound beam at discontinuities in acoustic impedance. It may be possible to correct for some of these using image analysis techniques that rely on tissue characterization.

3.3 Opaque Calibration Phantom

The bead phantom in a water tank described above has the obvious disadvantage that it requires the water to be drained from the tank to avoid refraction at the air-water interface. More fundamentally, it does not simulate real clinical targets very well. Patients are virtually opaque to human vision. Unlike water, living tissue has a consistency that generally limits lateral motion of an invasive tool such as a biopsy needle once it is partially inserted. For training and calibration purposes, a variety of gels have traditionally been used to fill ultrasound phantoms. We have developed a phantom with the particular needs of RTTR in mind that can be used to calibrate RTTR displays despite an inability to visually inspect the phantom's interior.

Our approach to building an opaque calibration phantom for RTTR is shown in Fig 7, with the hand-held version of the Sonic Flashlight scanning through a latex diaphram (not shown) in the wall of the phantom. The gel-filled phantom contains 3 small beads fixed on thin rods. Three corresponding hollow tubes ("virtual line-of-sight tubes") are mounted on the exterior of the phantom such that each tube is pointed at its corresponding bead. The tubes are in place before the insertion of the gel permitting visual inspection of the beads through the tubes. A small LED is then inserted into the base of each tube so that, even without seeing the bead, the viewer can determine that he is looking directly at the bead by seeing the LED through the tube. The ultrasound images of the beads can then be aligned, each bead with its corresponding LED, and the affine transform found as before.

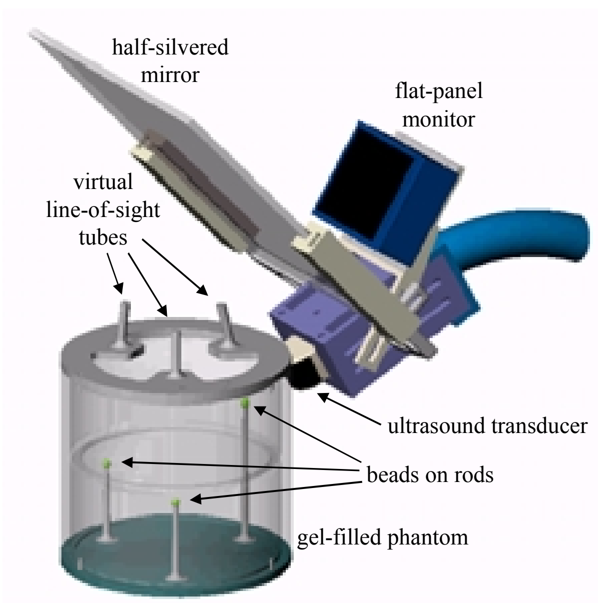

Fig. 7 Gel phantom: small spherical targets are suspended in the gel with *virtual line of site* tubes emitting light colinear with each. target.

Photographs of the opaque calibration phantom in operation are shown in Figs 8 and 9.

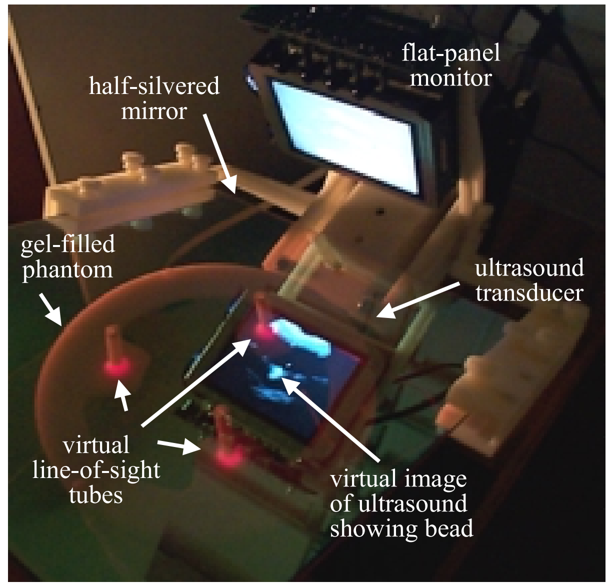

Fig. 8 Looking through the half-silvered mirror at gel-filled phantom with ultrasound image reflected to merge, using apparatus in Fig. 7. Viewpoint is not colinear with virtual line-of-sight tubes.

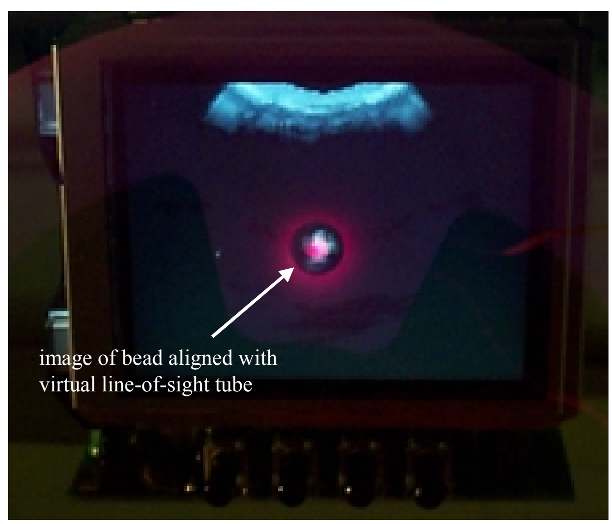

Fig. 9 View through half-silvered mirror, as in Fig. 8, but colinear with a virtual line-of-sight tube, showing alignment with ultrasound image of corresponding bead.

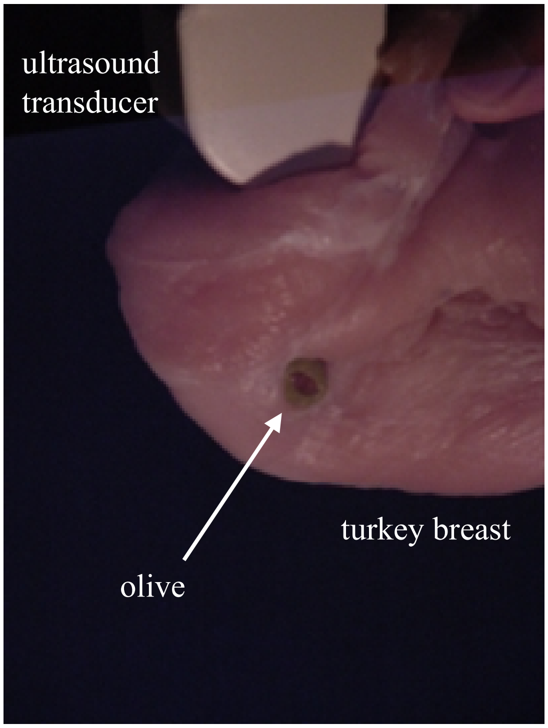

Fig. 10 View through half-silvered mirror of turkey breast phantom with olive inserted (ultrasound display is off).

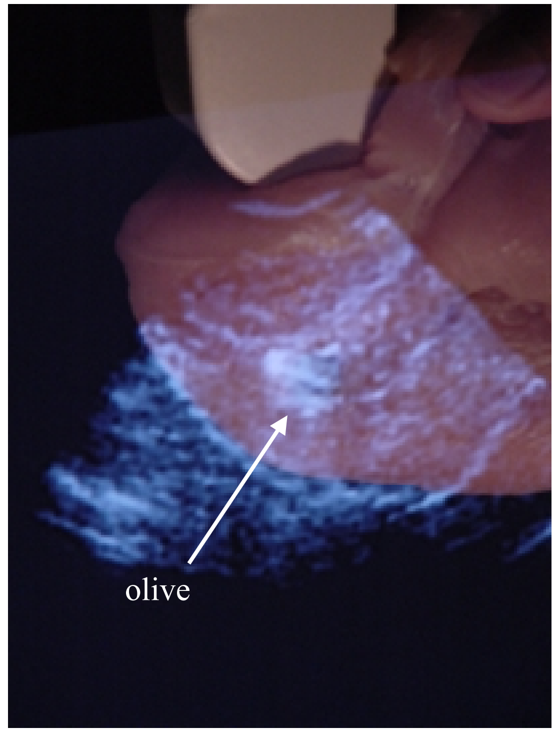

Fig. 11 Same view as in Fig. 10, but with ultrasound display turned on to show olive and border of turkey.

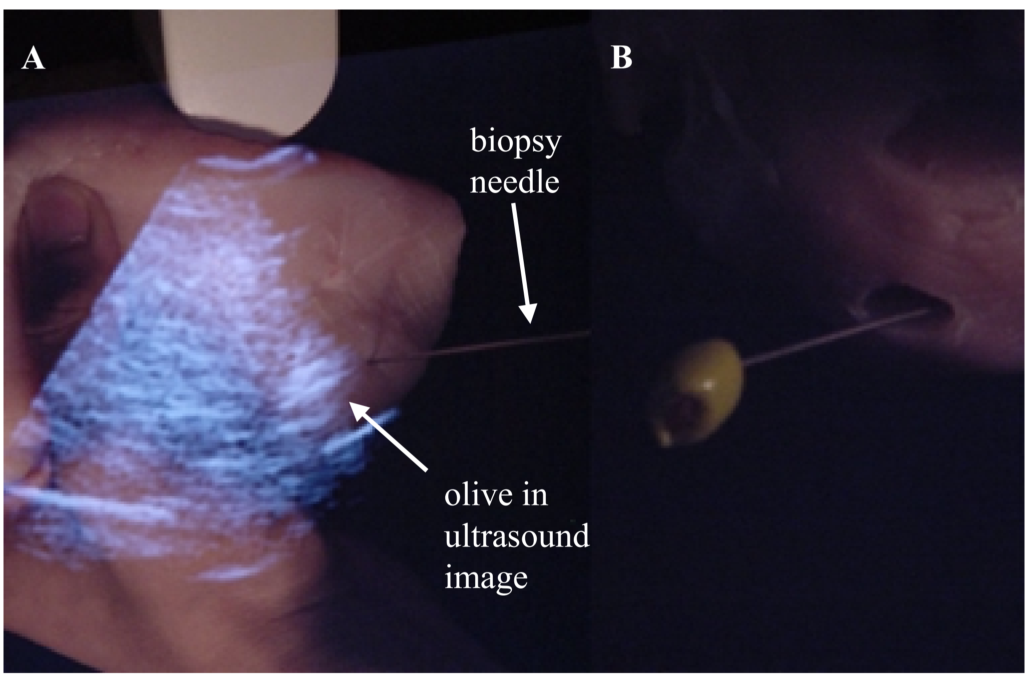

Fig. 12 A. View through half-silvered mirror as in Fig 11, but with turkey breast turned around so that olive is not visible except by ultrasound. Biopsy needle was inserted using hand-eye coordination to aim for the ultrasound image of the olive. **B.** Olive was successfully hit by the biopsy needle and pushed out the other side of the turkey breast.

4. Turkey and Olive Biopsy Phantom

An olive in a turkey breast is a commonly used phantom for breast biopsy. We conducted an initial test under RTTR guidance as follows: A small incision was made from one side of a turkey breast that did not penetrate to the other side. An olive was inserted into the incision so as to be still visible from that side the turkey breast (Fig. 10). This made it easy to identify the olive in the superimposed ultrasound image (Fig. 11). When the turkey breast was subsequently turned around so that the olive could not be seen visually, an accurate biopsy was still performed using the superimposed ultrasound image (Fig. 12). The needle was visible up to where it entered the phantom, permitting natural visual extrapolation to the target in the ultrasound image. While far from being a statistical study of the procedure, it was clear, from this one case, that RTTR enabled effective guidance of the needle biopsy using direct hand-eye coordination.

5. Conclusions

We have demonstrated working phantoms for calibration and biopsy using what we believe to be an important new method of image guidance. Much work remains to truly validate RTTR and to determine the limits of its accuracy, before it is ready for clinical use. However, the simplicity of RTTR and its ability to permit a unified intuitive environment for the clinician are important enough features to justify further research in this area.

Acknowledgments

This research was supported through a seed-fund grant from Carnegie Mellon University and a grant from the Whitaker Foundation to the Department of Bioengineering at the University of Pittsburgh. Valuable assistance and advice was received from Robert Tamburo, Costa Nikou, Richard LaBarca, Robert Kucler, Reed McManigle, Jim DiPasquale, Mark Peterson, Juels Sumkin, Mindy Michalerya, Stephanie Foster, Ethicon Endosurgery, Inc., the Kresge Rapid Manufacturing Laboratory and the Medical Robotics and Computer Assisted Surgery (MRCAS) Laboratory.

References

1. Dodd, G., C. Esola, D. Memel, A. Ghiatas, K. Chintaplli, E. Paulson, R. Nelson, J. Ferris, and R. Baron, *Sonography: the undiscovered jewel of interventional radiology.* Radiographics, 1996. **16**: p. 1271-88.

2. Sheafor, D., E. Paulson, C. Simmons, D. DeLong, and R. Nelson, *Abdominal percutaneous interventional procedures: comparison of CT and US guidance.* Radiology, 1998. **207**(3): p. 705-10.

3. Holm, H. and B. Skjoldbye, *Interventional ultrasound.* Ultrasound Med Biol, 1996. **22**(7): p. 773-89.

4. Azuma, R., *A Survey of Augmented Reality.* A Survey of Augmented Reality, in Presence: Teleoperators and Virtual Environments, 1997. **6**(4): p. 355-385.

5. State, A., M. Livingston, W. Garret, G. Hirota, M. Whitton, E. Pisano, and H. Fuchs. *Technologies for Augmented Reality Systems: Realizing Ultrasound-Guided Needle Biopsies*. in *ACM SIGGRAPH*. 1996, New Orleans, LA: p. 439-446.

6. Fuchs, H., A. State, E. Pisano, W. Garret, G. Hirota, M. Livingston, M. Whitton, and S. Pizer. *Towards Performing Ultrasound-Guided Needle Biopsies from within a Head-Mounted Display*. in *Visualization in Biomedical Computing*. 1996, Hamburg, Germany: p. 591-600.

7. Fuchs, H., M. Livingston, R. Raskar, D. Colucci, K. Keller, and A. State. *Augmented Reality Visualization for Laproscopic Surgery*. in *MICCAI*. 1998, Massachusetts Institute of Technology, Cambridge, MA, USA.

8. DiGioia, A., B. Colgan, and N. Koerbel, *Computer-Aided Surgery*, in *Cybersurgery*, R. Satava, Editor. 1998, Wiley: New York. p. 121-139.

9. Blackwell, M., F. Morgan, and A. DiGioia, *Augmented Reality and Its Future in Orthopaedics.* Clinical Orthopaedic and Related Research, 1998. **345**: p. 111-122.

10. Stetten, G., V. Chib, and R. Tamburo. *System for Location-Merging Ultrasound Images with Human Vision*. in *Applied Imagery Pattern*

Recognition (AIPR) Workshop. 2000, Washington, DC: p. 200-205 IEEE Computer Society.

11. Stetten, G. and V. Chib, *Real Time Tomographic Reflection with ultrasound: Stationary and hand-held implementations*, . 2000: CMU Robotics Insititute Technical Report # CMU-RI-TR-00-28.

12. Stetten, G. and V. Chib, *Overlaying ultrasound images on direct vision.* Journal of Ultrasound in Medicine, 2001. **20**(1): p. 235-240.

13. Stetten, G. and V. Chib. *Magnified real-time tomographic reflection (submitted)*. in *Medical image Computing and Computer-Assisted Intervention (MICCAI)*. 2001, Lecture Notes in Computer Science.

14. Masamune, K., Y. Masutani, S. Nakajima, I. Sakuma, T. Dohi, H. Iseki, and K. Takakura. *Three-dimensional slice image overlay system with accurate depth perception for surgery*. in *Medical image Computing and Computer-Assisted Intervention (MICCAI)*. 2000, Pittsburgh, **1935**: p. 395-402 Springer.