Augmented Reality Visualization for Laparoscopic Surgery

Henry Fuchs[1], Mark A. Livingston[1], Ramesh Raskar[1], D'nardo Colucci[1], Kurtis Keller[1], Andrei State[1], Jessica R. Crawford[1], Paul Rademacher[1], Samuel H. Drake[3], and Anthony A. Meyer, MD[2]

[1] Department of Computer Science, University of North Carolina at Chapel Hill
[2] Department of Surgery, University of North Carolina at Chapel Hill
[3] Department of Computer Science, University of Utah

Abstract. We present the design and a prototype implementation of a three-dimensional visualization system to assist with laparoscopic surgical procedures. The system uses 3D visualization, depth extraction from laparoscopic images, and six degree-of-freedom head and laparoscope tracking to display a merged real and synthetic image in the surgeon's video-see-through head-mounted display. We also introduce a custom design for this display. A digital light projector, a camera, and a conventional laparoscope create a prototype 3D laparoscope that can extract depth and video imagery.

Such a system can restore the physician's natural point of view and head motion parallax that are used to understand the 3D structure during open surgery. These cues are not available in conventional laparoscopic surgery due to the displacement of the laparoscopic camera from the physician's viewpoint. The system can also display multiple laparoscopic range imaging data sets to widen the effective field of view of the device. These data sets can be displayed in true 3D and registered to the exterior anatomy of the patient. Much work remains to realize a clinically useful system, notably in the acquisition speed, reconstruction, and registration of the 3D imagery.

1 Introduction

1.1 Challenges in Laparoscopic Surgery

The success of laparoscopy as a surgical technique stems from its ability to give the surgeon a view into the patient's internal spaces with only small incisions in the skin and body wall. Surgery done through such minimally invasive techniques leads to reduced trauma, shorter hospitalization, and more rapid return to normal activity. Although laparoscopy is a powerful visualization and intervention tool, it suffers from some visual limitations which we believe our proposed system will ameliorate.

- **The imagery is 2D and not 3D**. The surgeon can only estimate the depth of structures by moving the camera (to achieve motion parallax) or

by physically probing the structures. While stereo laparoscopes and stereo displays ameliorate this problem, they still separate the camera from the physician's point of view and fail to provide head-motion parallax.

- **The laparoscope has a small field of view**. The surgeon must frequently adjust the camera position and orientation, which requires skilled coordination with the assistant. A surgeon may opt to operate the camera himself to reduce the discoordination of the actual view with the desired view, but that limits him to only one hand with which to operate. Fixed camera holders can be used, but then the viewpoint and view direction are limited; this introduces risk due to the possible presence of important, vulnerable *structures* outside the viewing field.

- **The procedure requires significant hand-eye coordination**. The laparoscopic camera does not generally face the direction in which the surgeon is facing. This means that the instruments' on-screen movements will not match the surgeon's hand movements. It requires experience and hand-eye coordination for a surgeon to adjust to this disparity.

1.2 Benefits of Augmented Reality

Augmented reality (AR) refers to systems that attempt to merge computer graphics and real imagery into a single, coherent perception of an enhanced world around the user. Emerging AR technologies have the potential to reduce the problems caused by the visual limitations of laparoscopy. The AR system can display the resulting 3D imagery in the proper place with respect to the exterior anatomy of the patient. By acquiring depth information and rendering true 3D images of the structures visible in the laparoscopic camera, the AR system gives the physician most of the depth cues of natural vision. (Exceptions include focus and visual acuity.) The display of the laparoscopic data is not limited to the current viewpoint of the camera, but can include data acquired from a previous camera location (perhaps subject to a limit on the length of time the data is considered "current"). Thus objects not currently within view of the camera can still be displayed by the AR system.

We want to emphasize that this technology is fundamentally different than coupling a stereo laparoscope with a stereo display system. AR systems allow the surgeon to view the medical imagery from the natural viewpoint, use head-induced motion parallax (instead of hand-eye coordination and camera-induced motion parallax), allow the medical imagery to be visually aligned to the exterior anatomy of the patient, and incorporate proprioceptive (body-relative) cues.

The lack of depth perception in laparoscopic surgery might limit delicate dissection or suturing [Durrani95]. An AR display presents objects in correct perspective depth, assuming that the geometry has been accurately acquired. With an AR guidance system, a laparoscopic surgeon might be able to view the peritoneal cavity from any angle merely by moving his head, without moving the endoscopic camera. AR may be able to free the surgeon from the technical limitations of the imaging and visualization methods, recapturing much of the physical simplicity and direct visualization characteristic of open surgery.

2 Previous Work

2.1 Medical Augmented Reality Systems

The first medical application of AR was to neurosurgery [Kelly86]. Similar systems [Lorensen93,Grimson95] have been developed independently. AR has also been applied to otolaryngology [Edwards95]. These applications demand less of the AR system than laparoscopy for four reasons. The surgical field is small, the patient doesn't move, the view into the patient is from a single viewpoint and view direction, and this viewpoint is external to the patient (e.g. already suitable for a hand-eye coordination). This simplifies the difficult task of building an enhanced visualization system.

Our research on medical applications of AR has until recently concentrated on ultrasound-guided procedures such as fetal examination [Bajura92,State94] and breast biopsy [Fuchs96,State96]. In the latter system, the ultrasound data is captured as a video stream and registered to the patient in real time. The physician's head must be tracked in order to view the dynamic data from any direction. We calibrate the location of the ultrasound data with respect to the probe geometry and track the probe location. These two tasks enable registration of multiple discrete slices to each other and registration of the ultrasound data set to the patient. A virtual pit [Bajura92] within the patient's body provides proper occlusion cues for the registered ultrasound data. We base our proposed system to aid laparoscopic surgery on this system.

2.2 Depth Extraction

The major new technology needed for laparoscopic visualization is acquisition of the depth map associated with the image from the laparoscopic camera. Determination of 3D scene structure from a sequence of 2D images is one of the classic problems in computer vision [Faugeras93]. There are numerous techniques for computing 3D structure, including cues from motion, stereo, shading, focus, defocus, contours, and structured light. We chose structured light for several reasons. It is an efficient and direct computation. It is as robust to shading variations and repeating patterns as other methods (although no method is immune to some features, such as specular highlights) and can be dynamically tuned to increase robustness. It offers a large depth range and allows us to trade speed for spatial resolution in the acquisition.

Structured light has long been used in computer vision to acquire depth information [Besl89,Daley95]. A variety of patterns have been tried: points, lines, multiple points, multiple lines, grids, circles, cross-hairs, thick stripes, binary-coded patterns, color-coded stripes, and random textures. Pseudo-random binary arrays [Lavoie96] are grids with recognizable points based on a pattern of "large" and "small" intersection points. We initially chose binary-coded patterns, but switched to lines since our prototype system cannot acquire images of the pattern fast enough to support depth extraction from dynamic scenes. (See Section 7 for our future plans regarding this issue.)

3   Hardware Configuration

There are four primary hardware components to our system. Three are the standard tools of AR systems: an image generation platform, a set of tracking systems, and a see-through head-mounted display (HMD) that allows the user to see the real environment. The fourth component required for this application is a 3D laparoscope that can acquire both color and range data. As noted above, we have previously applied AR to in-place visualization of ultrasound imagery. The current system is similar to that system [State96,Fuchs96].

3.1   See-Through Head-Mounted Display

We believe that the depth cue of occlusion is vital to the physician in determining the 3D structure of the medical imagery. Video-see-through (VST) displays offer the possibility of complete occlusion of the real world by the computer-generated imagery, which in this case is the medical image data. (The other option, optical-see-through displays, cannot achieve complete occlusion of the real world.) Being unaware of any commercially available VST HMDs, we initially built a simple prototype VST HMD from commercial components [State96]. This device had numerous limitations [Fuchs96]. In response to our experience with that device, we designed and implemented a new VST HMD, which is described in Section 4.

3.2   Image Generation

Having chosen to build a VST system, we needed an image generation platform capable of acquiring multiple, real-time video streams. We use an Onyx Infinite Reality system from Silicon Graphics, Inc. equipped with a Sirius Video Capture™unit. This loads video imagery from the cameras on the VST HMD directly into the frame buffer. We augment this background image with a registered model of the patient's skin acquired during system calibration [State96]. We then render the synthetic imagery in the usual manner for 3D computer graphics. At pixels for which there is depth associated with the video imagery (e.g. the patient's skin), the depth of the synthetic imagery is compared. The synthetic imagery is painted only if it is closer. This properly resolves occlusion between the synthetic imagery and the patient's skin.

The video output capabilities of the Infinite Reality architecture allow us to output two VGA signals to the displays in the HMD and a high resolution video signal, which contains a user interface and is displayed on a conventional monitor. The system architecture is depicted graphically in Figure 1.

3.3   Tracking Systems

We use UNC's optoelectronic ceiling tracker [Welch96] for tracking the physician's head. It offers a high update rate, a high degree of accuracy, and a large range of head positions and orientations. The large range allows the physician

to move freely around the patient and to examine the patient from many viewpoints. We track the laparoscope with the FlashPoint™ 5000 [IGT97] optical tracker from Image-Guided Technologies, Inc. It also offers high accuracy, but over a small range. Since the laparoscope does not move much, this is suitable for our system. Its accuracy enables the registration between multiple laparoscope data images and between the laparoscopic data set and the patient.

3.4   3D Laparoscope

To properly display the 3D structure, the laparoscope must acquire depth information. To properly display the visual texture (e.g. color, shading), the laparoscope must acquire the usual 2D color video image. We can then texture the resulting 3D mesh with the color data. We designed a custom device, described in Section 5. The device requires input of structured light images and outputs images suitable for depth and color processing. We off-load the processing of these image streams to a Silicon Graphics O2, which outputs the structured light and acquires the camera video. After simple image processing, the O2 sends to the Onyx a list of lit pixels, from which the depth is computed.

4   Video-See-Through Head-Mounted Display

We use a miniature HMD custom-designed at the computer science laboratories at the University of North Carolina and University of Utah. This VST HMD has a miniature video camera mounted in each display optic in front of each eye (Figure 4). A pair of mirrors place the apparent centroid of the camera in the same location as the center of the eye when the HMD is properly fitted to the user. A 640 × 480 LCD mounted in the eyepiece is viewed through a prism assembly which folds the optical path within a small cube. This design reduces the problem of unequal depths for the user's visual and tactile senses. The HMD has two eyepieces mounted on a horizontal bar which provides one degree of translational freedom and one degree of rotational freedom. This allows the user to adjust the inter-camera distance and the convergence angle. The entire front bar can be moved out of the way (Figure 3). The complete HMD weighs only twelve ounces, compared to six pounds for our initial prototype.

5   3D Laparoscope

To extract 3D shape, we added a structured light to a conventional laparoscope. Our "structure" is a vertical line in the image plane of the projector. We calibrate the device by building a table of depth values, then use a simple extraction algorithm which interpolates through the table. This technique has performed well on simple geometry such as scenes with little or none of the surface occluded from the projector's view. It has not yet performed well on topologically complex models that have great discontinuities in depth, highly specular reflections, low reflectance, or large patches of surfaces hidden from the projector's viewpoint.

5.1 3D Laparoscope Design

The structured light 3D laparoscope design (Figure 7) uses a conventional laparo-scope in a rather unique way. Instead of being the both the illumination source and imaging device, it is only a projector—but of structured light patterns. A digital micromirror device [Hornbeck95] projector displays its image through a custom optic and through a standard laparoscope, projecting its image inside the patient. The image is the dynamic, calibrated structured light image. Along-side the projecting laparoscope is a miniature video camera mounted in a metal tube similar to a second laparoscope. This camera observes the structured light pattern on the scene and sends the image to the host. The two laparoscopes are mounted a fixed distance from each other for accurate and repeatable depth extraction.

5.2 Depth Calibration and Extraction

We measure the reflected light pattern for a set of known depths and store the results in a table. By imaging each potential light stripe from the projector onto a flat grid at a known depth, we can determine the 3D location of the point at each pixel in the camera image. With several depths, we can build a table indexed by the column number from the projector and the u and v-coordinates on the camera image plane. At each cell in the table is a 3D point. Simple thresholding determines which pixels in the camera image are illuminated by the light stripe. We find the centroid of the biggest and brightest 1D blob on each camera scanline. The 3D location of this point is interpolated from the table.

6 Experiments and Results

We have implemented two versions of this system. In the first prototype, we acquired depth via manual digitization. This implies a pre-operative acquisition of the 3D structure of the internal anatomy. Guided by real-time color images textured onto the 3D mesh, the surgeon (Meyer) successfully pierced a small foam target inside the abdominal cavity of a life-sized human model (Figures 2 and 7). This experiment showed the potential of our proposed paradigm for laparoscopic surgery. It also emphasized the importance of extracting the internal 3D structure in real time. For example, a manipulator inserted into the abdomen was severely distorted onto the surface mesh instead of appearing to be above the surface because only imagery was acquired in real time, not the 3D structure.

The second experiment was recently conducted with a system that imple-ments interactive depth extraction. The results of this system have been promis-ing (Figure 7). The augmented images shown to a moving HMD user clearly present the 3D structure. The computer-generated imagery of the internal struc-ture is visually aligned with the exterior patient anatomy.

7 Discussion

We are currently focusing on three issues. First, depth extraction is slow due to inconsistent delay between commanding the projector to emit a pattern and receiving the image of the pattern from the camera. (This is more complex than synchronizing the vertical refresh.) Second, gathering multiple views is difficult due to the rigid connection between the (bulky) projector and the laparoscope. Third, multiple depth images are misregistered due to poor depth calibration.

Our current solution to the slow speed is to wait for the delay to expire. We are developing a tighter, coupled control of the camera and projector. When this is in place, we will be able to extract new data at every frame. We can also return to using binary-coded patterns as the structured light. These algorithmic and hardware improvements, along with a higher-speed projector and camera, will enable us to incrementally update an entire range image with each new video image of the pattern, thus extracting depth from a larger area of the surgical field at each time step. We will investigate methods of adaptive depth acquisition to increase accuracy and resolution in regions of particular concern to the surgeon. For gathering multiple views, we are working with fiberoptic cables and miniature cameras and displays to make the 3D laparoscope smaller and easier to maneuver into multiple positions. By improving depth calibration and merging multiple range images [Turk94], we hope to provide a more complete view of the interior scene than visible from a single laparoscope location, approaching the wide-area surgical field in open surgery. In the future, by registering pre-operative images (e.g. MRI or CT), surgical planning data, and intra-operative (e.g. ultrasound), we hope to provide a more comprehensive visualization of the surgical field than even open surgery.

We postulate that viewing laparoscopic images with our augmented reality paradigm, from outside the body, as if there were an opening into the patient, will be more intuitive than observing laparoscopic imagery on a video monitor or even viewing images from stereo laparoscopes on a stereo video monitor. We expect that the physician will still choose to move the laparoscope frequently (closer to view structures of interest or farther away to view of the entire intervention site), but with our system such movements will not cause confusing changes in the viewpoint requiring mental adaptation. Rather they will change the level of detail and update the visualization of structures that become visible to the laparoscope. We expect the physician's use of the laparoscope to be somewhat akin to exploring a dark room with a flashlight, with the added benefit of visual persistence of the regions of the scene that were previously illuminated.

We hope that our proposed system will eventually offer the following specific benefits. It could reduce the average time for the procedures (benefiting both physician and patient), reduce training time for physicians to learn these procedures, increase accuracy in the procedures due to better understanding of the structures in question and better hand-eye coordination, reduce trauma to the patient through shorter and more accurate procedures, and increase availability of the procedures due to ease of performing them.

References

[Bajura92] Bajura, M., Fuchs, H., and Ohbuchi, R. (1992). Merging virtual objects with the real world: Seeing ultrasound imagery within the patient. In *Computer Graphics (SIGGRAPH '92 Proceedings)*, volume 26, pages 203–210.

[Besl89] Besl, P. J. (1989). Active optical range imaging sensors. In *Advances in Machine Vision*, pages 1–63. Springer-Verlag.

[Daley95] Daley, R. C., Hassebrook, L. G., Stanley C. Tungate, J., Jones, J. M., Reisig, H. T., Reed, T. A., Williams, B. K., Daugherty, J. S., and Bond, M. (1995). Topographical analysis with time modulated structured light. *SPIE Proceedings*, 2488(5):396–407.

[Durrani95] Durrani, A. F. and Preminger, G. M. (1995). Three-dimensional video imaging for endoscopic surgery. *Computers in Biological Medicine*, 25(2):237–247.

[Edwards95] Edwards, P., Hawkes, D., Hill, D., Jewell, D., Spink, R., Strong, A., and Gleeson, M. (1996). Augmentation of reality in the stereo operating microscope for otolaryngology and neurosurgical guidance. *Journal of Image-Guided Surgery*, 1(3).

[Faugeras93] Faugeras, O. (1993). *Three-Dimensional Computer Vision: A Geometric Viewpoint*. MIT Press.

[Fuchs96] Fuchs, H., State, A., Pisano MD, E. D., Garrett, W. F., Hirota, G., Livingston, M. A., Whitton, M. C., and Pizer, S. M. (1996). Towards performing ultrasound-guided needle biopsies from within a head-mounted display. In *Visualization in Biomedical Computing 1996*, pages 591–600.

[Grimson95] Grimson, W., Ettinger, G., White, S., Gleason, P., Lozano-Pérez, T., Wells III, W., and Kikinis, R. (1995). Evaluating and validating an automated registration system for enhanced reality visualization in surgery. In *Proceedings of Computer Vision, Virtual Reality, and Robotics in Medicine '95 (CVRMed '95)*,.

[Hornbeck95] Hornbeck, L. J. (1995). Digital light processing and MEMS: Timely convergence for a bright future. In *Micromachining and Microfabrication '95*.

[IGT97] Image-Guided Technologies, Inc. (1997). *FlashPoint™ Model 5000 3D Localizer User's & Programmer's Manual*. Boulder, CO.

[Kelly86] Kelly MD, P. J., Kall, B., and Goerss, S. (1986). Computer-assisted stereotaxic resection of intra-axial brain neoplasms. *Journal of Neurosurgery*, 64:427–439.

[Lavoie96] Lavoie, P., Ionescu, D., and Petriu, E. M. (1996). 3-D object model recovery from 2-D images using structured light. In *IEEE Instrument Measurement Technology Conference*, pages 377–382.

[Lorensen93] Lorensen, W., Cline, H., Nafis, C., Kikinis, R., Altobelli, D., and Gleason, L. (1993). Enhancing reality in the operating room. In *Proceedings of IEEE Visualization '93*.

[State94] State, A., Chen, D. T., Tector, C., Brandt, A., Chen, H., Ohbuchi, R., Bajura, M., and Fuchs, H. (1994). Case study: Observing a volume-rendered fetus within a pregnant patient. In *Proceedings of IEEE Visualization '94*, pages 364–368.

[State96] State, A., Livingston, M. A., Hirota, G., Garrett, W. F., Whitton, M. C., and Fuchs, H. (1996). Technologies for augmented-reality systems: Realizing ultrasound-guided needle biopsies. In *SIGGRAPH 96 Conference Proceedings*, Annual Conference Series, pages 439–446. ACM SIGGRAPH, Addison Wesley.

[Turk94] Turk, G. and Levoy, M. (1994). Zippered polygon meshes from range images. In *Proceedings of SIGGRAPH '94*, Computer Graphics Proceedings, Annual Conference Series, pages 311–318.

[Welch96] Welch, G. F. (1996). *Single-Constraint-At-A-Time Tracking*. Ph.D. Dissertation, University of North Carolina at Chapel Hill.
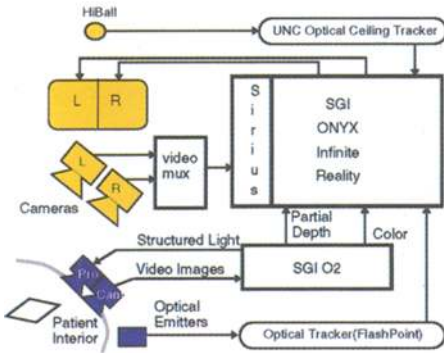
Fig. 1. Diagram of the hardware configuration of the prototype system. The VST HMD consists of two cameras, two displays, and a HiBall tracking sensor.

Fig. 2. The physician (Meyer) uses the system in the preliminary experiment (Dec 96). The mechanical arm he holds is unnecessary in the current implementation. The colored circular landmarks on the "body" surface assist the head tracking subsystem.

Fig. 3. (*Above*) Custom-designed video-see-through head-mounted display for augmented reality applications. The lightweight unit can be flipped up and down.
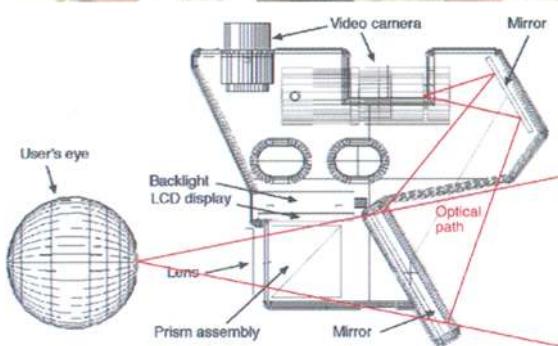
Fig. 4. (*Left*) The design of one eyepiece of the VST HMD. The optical paths from the camera to the world and from the user's eye to the LCD are folded in order to match the lengths.
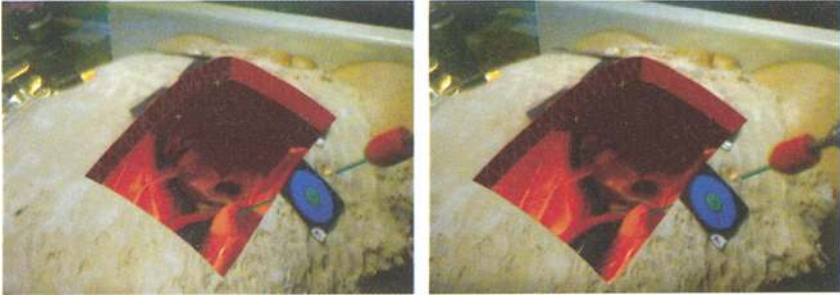
Fig. 5. Wall-eyed stereo pair of images the physician sees in the HMD. We manually digitized the interior structure prior to this experiment (Dec 96).
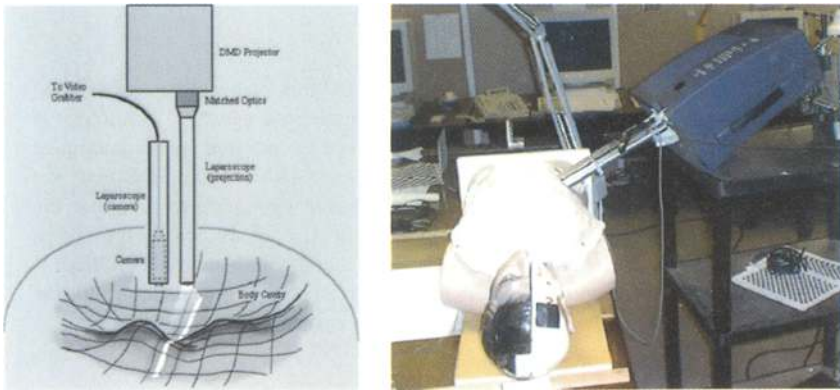

Fig. 6. Our prototype 3D laparoscope combines a conventional laparoscope, a projector emitting structured light in the form of vertical stripes, and a camera to create a laparoscope that acquires depth and color data.
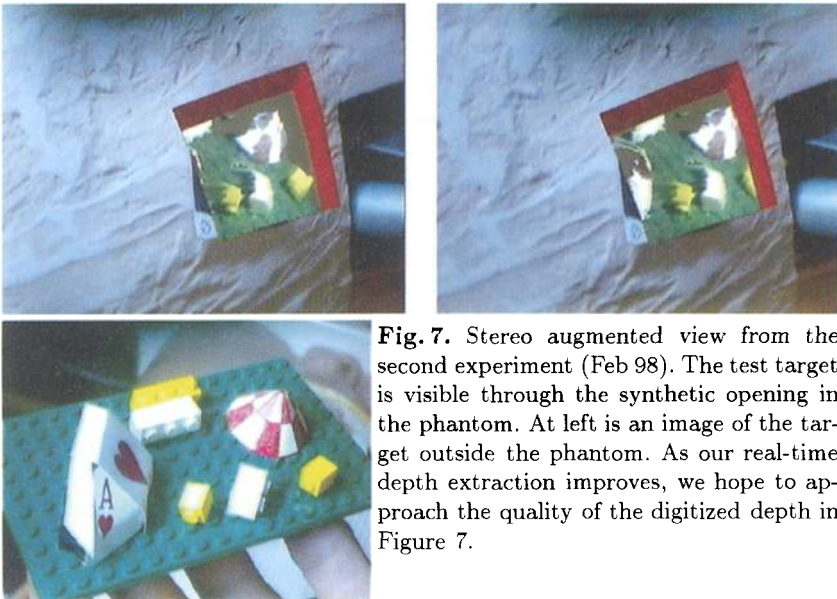

Fig. 7. Stereo augmented view from the second experiment (Feb 98). The test target is visible through the synthetic opening in the phantom. At left is an image of the target outside the phantom. As our real-time depth extraction improves, we hope to approach the quality of the digitized depth in Figure 7.