Stereoscopic Augmented Reality using Ultrasound Volume Rendering for Laparoscopic Surgery in Children

Jihun Oh, Xin Kang, Emmanuel Wilson, Craig A Peters, Timothy D Kane, Raj Shekhar*
Sheikh Zayed Institute for Pediatric Surgical Innovation, Children's National Medical Center, Washington, DC, USA

ABSTRACT

In laparoscopic surgery, live video provides visualization of the exposed organ surfaces in the surgical field, but is unable to show internal structures beneath those surfaces. The laparoscopic ultrasound is often used to visualize the internal structures, but its use is limited to intermittent confirmation because of the need for an extra hand to maneuver the ultrasound probe. Other limitations of using ultrasound are the difficulty of interpretation and the need for an extra port. The size of the ultrasound transducer may also be too large for its usage in small children. In this paper, we report on an augmented reality (AR) visualization system that features continuous hands-free volumetric ultrasound scanning of the surgical anatomy and video imaging from a stereoscopic laparoscope. The acquisition of volumetric ultrasound image is realized by precisely controlling a back-and-forth movement of an ultrasound transducer mounted on a linear slider. Furthermore, the ultrasound volume is refreshed several times per minute. This scanner will sit outside of the body in the envisioned use scenario and could be even integrated into the operating table. An overlay of the maximum intensity projection (MIP) of ultrasound volume on the laparoscopic stereo video through geometric transformations features an AR visualization system particularly suitable for children, because ultrasound is radiation-free and provides higher-quality images in small patients. The proposed AR representation promises to be better than the AR representation using ultrasound slice data.

Keywords: Laparoscopic surgery, augmented reality, volumetric ultrasound, stereoscopic video, surgical visualization

1. INTRODUCTION

Minimally invasive surgery (MIS), including the laparoscopic approach, is known for improving surgical outcomes, leaving fewer scars, and leading to significantly faster patient recovery. Visual information is critical for the safety and efficacy of laparoscopic surgeries in which the visualization is constrained because of the remote center of motion and the lack of tactical feedback. Live laparoscopic video, the primary imaging technique, provides a visualization of exposed organ surfaces, but is unable to show anatomical structures beneath those organ surfaces. Static pre- and intra-operative tomographic images have been used to add information on the internal structures; however, these approaches have limited applicability when deformable soft-tissue organs are the surgical targets. To deal with soft-tissue organs, continuous intra-operative computerized tomography has been suggested as an imaging option[1], with the compensation for increased radiation exposures. In current clinical practice, laparoscopic ultrasound is used to visualize internal structures in real time and the method poses no radiation risk. Several groups, including ours, have proposed methods to overlay laparoscopic ultrasound images on the laparoscopic video.[2] However, overlaying conventional two-dimensional (2D) laparoscopic ultrasound image has its own limitations. The laparoscopic ultrasound image is only a 2D cross-section of the three-dimensional (3D) surgical anatomy whereas the laparoscopic video image gives a perspective view of the same anatomy. An overlay of these images with fundamentally different image formats is non-intuitive and thus makes the interpretation of the resulting augmented reality (AR) image difficult for laparoscopic surgeons. In addition, the use of laparoscopic ultrasound may require an additional port and an extra pair of hands to operate the transducer.

In this paper, we report a new AR visualization system prototype that features continuous hands-free volumetric ultrasound scanning coupled with video imaging by a stereoscopic laparoscope. The limitation of depth perception associated with conventional laparoscopes is addressed through the use of stereoscopy. The volumetric ultrasound provides a comprehensive real-time visualization of the surgical anatomy. Overall, the proposed prototype could enable

* Email: rshekhar@childrensnational.org

Medical Imaging 2014: Image-Guided Procedures, Robotic Interventions, and Modeling,
edited by Ziv R. Yaniv, David R. Holmes III, Proc. of SPIE Vol. 9036, 90360Y
 · doi: 10.1117/12.2043929

surgeons to perceive true depth and gain improved understanding of 3D spatial relationship among visible anatomical structures.

The volumetric ultrasound is acquired using a mechanically steered robotic "drop-in" transducer (a transducer mounted on the tip of a flexible cable) designed for intraoperative imaging. This permits continuous volumetric scanning of the patient anatomy. The availability of volumetric data allows presenting volume rendered views similar to the laparoscope view in format to the surgeons. Using the spatial tracking data, a 3D rendering of the ultrasound volume from the perspective of the stereoscopic laparoscope is overlaid on the live stereoscopic video of the surgical field. Compared with the prior approaches that overlay cross-sectional images that are different in format from the laparoscopic view, the representation produced by the described system can potentially ease the interpretation of overlaid ultrasound images. The proposed extracorporeal ultrasound imaging method eliminates the need for an additional port or an extra pair of hands to operate laparoscopic ultrasound. The proposed system is particularly suitable for children because it can provide higher-quality ultrasound images in small patients without any radiation hazard.

2. METHODS

2.1 System Design

Our surgical AR prototype consists of a vision system (VSII, Visionsense Corp., New York, NY) with a 5-mm stereo laparoscope, an ultrasound machine (flexFocus 700, BK Medical, Herley, Denmark) with a robotic "drop-in" transducer, a custom-designed mechanical ultrasound scanner, an optical tracker (Polaris Vicra, Northern Digital, Waterloo, Ontario, Canada) and a workstation, as shown in Figure 1. The vision system features a novel stereoscopic laparoscope with a viewing angle of 70 degrees, a working distance of up to 5 cm, and a fixed focal length of 2.95 cm. The drop-in transducer has an operating frequency range of 5 MHz–12 MHz with a maximum scan depth of 13 cm. The custom-designed hands-free volumetric ultrasound scanner is described in the next section.

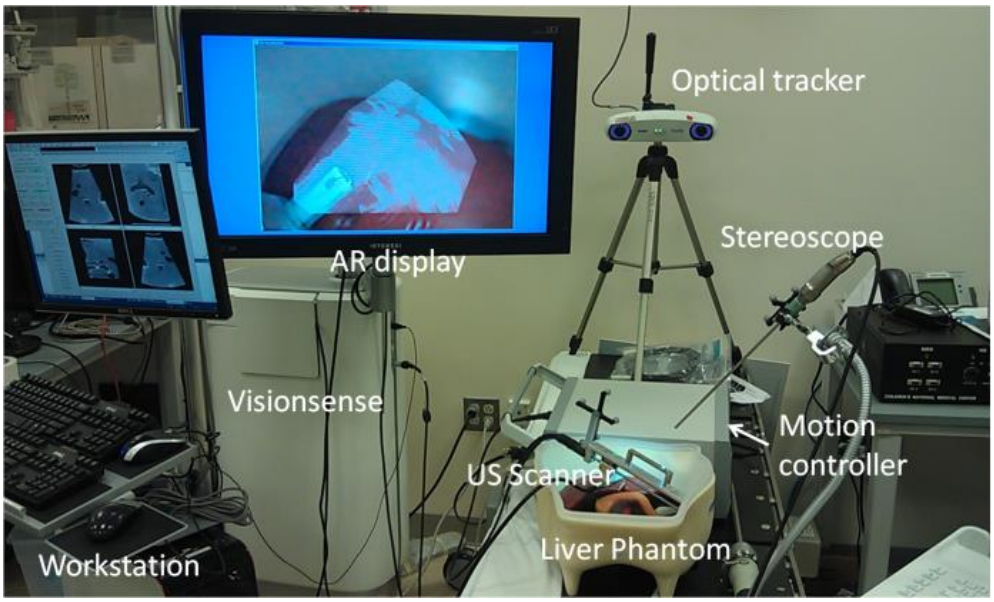
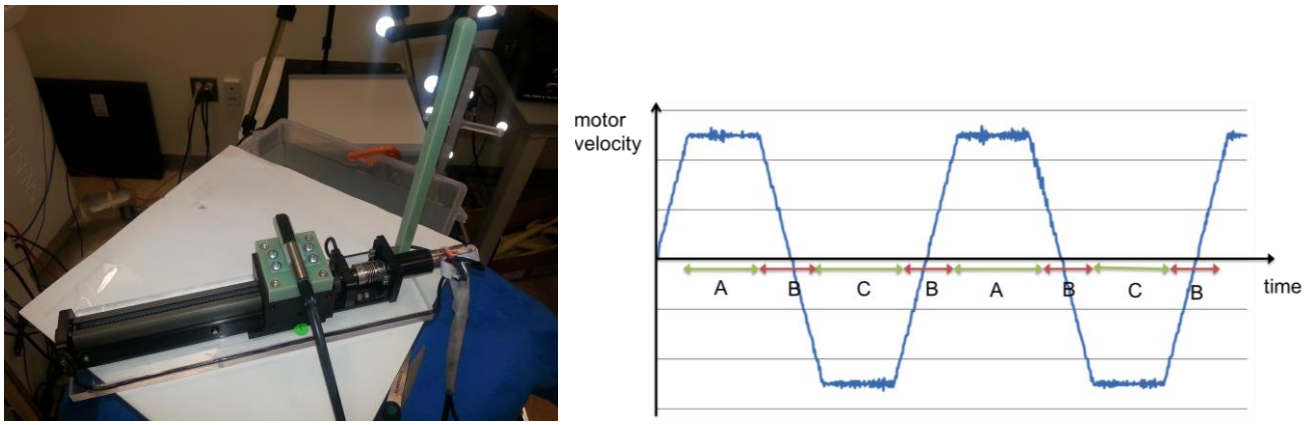


Figure 1. Hardware components of the proposed AR system.

A workstation (64-bit Window 7 with 8 core 3.2 GHz, 12 GB memory, NVidia Quadro 4000) is used for data acquisition from the vision system, the ultrasound machine, the optical tracker, and the motion controller. The ultrasound images are streamed from the ultrasound machine over Gigabyte Ethernet using a proprietary communication library. The tracked position of the transducer on the linear stage is acquired using the library provided by the manufacturer of the motion controller. The optical tracker and passive reflective markers on the dynamic reference frame (DRF) fixed to tracking targets are used to track the 3D pose (location and orientation) of the stereoscopic laparoscope and the mechanical ultrasound scanner in real time. The workstation also performs volume reconstruction, image processing, volumetric rendering, image fusion, and stereoscopic visualization.

2.2 Hands-free Volumetric Ultrasound Acquisition

Figure 2 (a) shows the prototype of our custom-designed ultrasound scanner using a drop-in transducer. In its current implementation, the drop-in transducer (ProART, BK Medical, Herley, Demark) is mounted on a linear stage (154 mm in length) made of aluminum and actuated using a brushed DC motor (Maxon RE16 3.2 W, Sachsein, Switzerland) and a motion controller (DMC-41x3, Galil Motion Control, Rocklin, CA). As an early prototype, the scanner (including the drop-in transducer and the linear stage) is not fully enclosed. For scanning purposes, the imaging specimen was submerged in a water tank and the 3D ultrasound scanner assembly was held inverted by the operator with the ultrasound transducer grazing the water surface. The next-generation prototype will have the transducer located within a fluid-filled compartment with an acoustically transparent imaging surface. The imaging surface will make contact with the patient skin and be flexible enough to adapt to the body contour.



(a) (b)

Figure 2. (a) Current implementation of the 3D ultrasound scanner that is mounted on a linear stage a drop-in transducer whose movement is controlled by a motor. (b) The velocity profile of the motor.

A similar concept for enclosing an ultrasound transducer in a fluid-filled chamber has been reported in which a joystick was used to manually control the position of the ultrasound transducer.[3] This design does not completely free up the surgeon from acquiring ultrasound image during the surgery. In addition, this design cannot produce volumetric images. On the contrary, our scanner is designed to keep the transducer automatically moving back and forth on the linear stage. With linear scanning, parallel and equidistant 2D images are acquired and stacked to create a volume at each ends of the linear excursion of the transducer.

Figure 2(b) shows the velocity profile of the motor. During intervals A and C, image series are stacked to build the ultrasound volume except that in interval C the stack is reversed as the motor moves in the opposite direction. During interval B, the time when the motor is ramping up and down, the stack is written to VTK image data, representing a geometrical structure including a volume, for visualization. The speed of the motor is determined by the physical spacing desired between consecutive slices and the streaming rate (motor speed [mm/s] = spacing [mm] x streaming rate [Hz]). The workstation communicates with the motion controller through a USB port to control the speed, acceleration, and linear excursion.

2.3 System Calibration

In our prototype AR system, as in most other systems, device calibration is critical for accurately overlaying different types of images. For stereoscopic laparoscope calibration, a 9-by-6 checkerboard of alternating 5-mm black and white squares was used. The size of the square was selected to ensure that the entire checkerboard always lies within the field of view of the laparoscope at the recommended working distance. The transformation from the laparoscope to the DRF attached to it was determined from the 3D coordinates defined by the four corners of the checkerboard pointed to by a tracked pointer. The intrinsic parameters of the laparoscopic camera were estimated using Zhang's method.[4]

For ultrasound calibration, the calibration phantom of PLUS[5] was used. As shown in Figure 3, suture wires were used to form three "N" shapes with known geometry related to the phantom. Because position tracking data and image data are not synchronized, the time difference between them was recorded and taken into account in calibration. During

calibration, the transducer is fixed at the starting position with the motor turned off. The nine intersecting points of the N-wires in the ultrasound frames continuously imaged at varying positions were used for calibration.

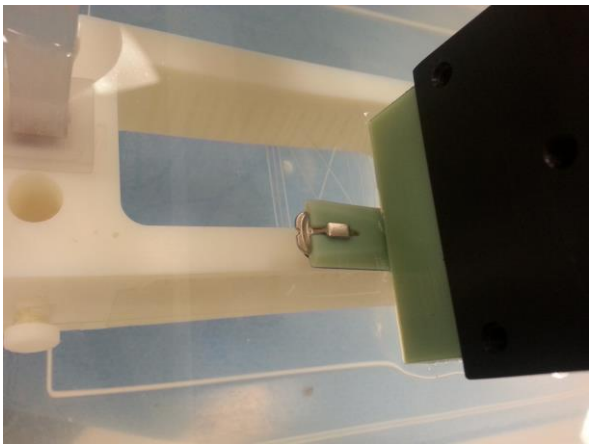


Figure 3. Ultrasound calibration using the PLUS phantom with three N wires.

2.4 Fusion of Volume Rendering and Stereoscopic Video

As a rendering technique, maximum intensity projection (MIP) has been found helpful in the visualization of the vasculature, tumors, and lesions.[6] The MIP is a volume rendering method that projects the voxel with the highest intensity value in the volume along each ray-case direction. To overlay MIP of the ultrasound volume onto laparoscopic video, the tracked position of the laparoscope camera with respect to the optical tracker is transformed to the coordinate of the ultrasound volume as

$$P'(\vec{x}) = (T^{LUS}_{LUS_{DRF}})^{-1} * (T^{LUS_{DRF}}_{tracker})^{-1} * T^{LAP_{DRF}}_{tracker} * T^{camera}_{LAP_{DRF}} * P(\vec{0}),$$

where $P'(\vec{x})$ is the position of the laparoscopic camera in the coordinates of the ultrasound volume, $T^{camera}_{LAP_{DRF}}$ is the transformation from the laparoscopic camera to the laparoscope DRF, $T^{LAP_{DRF}}_{tracker}$ is the transformation from laparoscope DRF to the optical tracker, $T^{LUS_{DRF}}_{tracker}$ is the transformation from the ultrasound DRF to the tracker, and $T^{LUS}_{LUS_{DRF}}$ is the transformation from transducer to the ultrasound DRF. Note that the scale in the rotation matrix of ultrasound should be excluded from consideration. Among camera parameters, the focal point can be computed by finding the intersection point of the line extending from the laparoscope DRF to the camera with ultrasound imaging plane passing the center of the volume. The up vector for visualization is obtained by computing the vector of the laparoscope DRF with respect to the local coordinates. Therefore, a perspective transformation is obtained using intrinsic parameters encompassing focal length and principal point. The above procedure is repeated for left and right channels of the stereoscopic laparoscope. Next, two MIPs are generated for the two channels from the ultrasound volume using VTK (Kitware Inc., Clifton Park, NY, USA). Finally, these two MIPs are composited with the stereoscopic video images using alpha blending and displayed on a stereo monitor.

Usually the proximal and distal areas of the ultrasound images are bright because of the ring-down effect and strong reflections from the bottom surface of the water tank. It is therefore desirable to crop the top and bottom parts of the ultrasound volume before creating MIPs. We call this technique thin-slab MIP in which the user can select the ultrasound slab thickness and adjust the slab location in the original volume before volume rendering.

3. RESULTS

3.1 System Evaluation

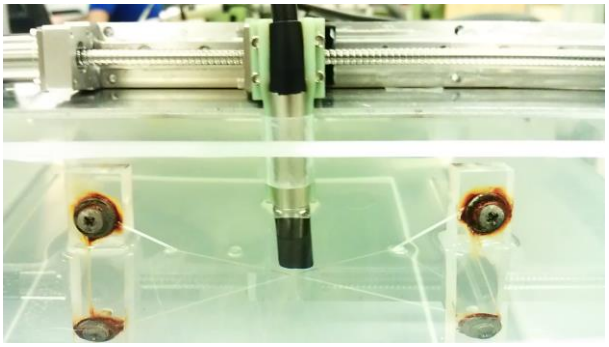
High ultrasound image-to-laparoscopic video registration accuracy and low system latency are desirable for a successful clinical AR visualization system. Any registration error will cause misalignment of anatomical structures between the two imaging modalities involved in AR visualization and make surgeries based on such visualization unsafe. A long time delay between the motion of the ultrasound transducer and the corresponding images presented to the surgeons will affect the real-time interaction and visual feedback necessary for surgical efficiency and effectiveness.

The accuracy of the overall AR system relies on the calibration accuracy of the stereoscopic laparoscope as well as the drop-in ultrasound transducer, and the accuracy of the stereoscopic AR visualization. The target registration error (TRE) metric was used to measure the accuracies of these components. We conducted experiments using a cross-wire phantom (Figure 4(a)), with the intersection of the two wires as the target point. The TRE is defined as the Euclidean distance between the actual 3D location known by touching the intersection point with a tracked pointer and the computed 3D location determined through triangulation of the identical points in a pair of images. The calibration accuracies of the stereoscopic laparoscope, the drop-in ultrasound transducer, and the overall AR visualization accuracy were measured to be 0.93 $\pm$0.21 mm, 3.52 $\pm$1.16 mm and 4.04 $\pm$1.00 mm, respectively.

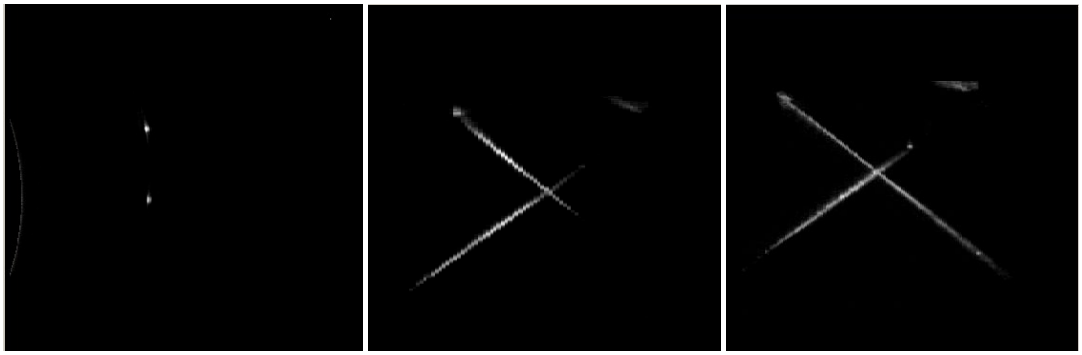
The system latency was measured by imaging a high-resolution (millisecond) digital clock. The difference between the actual clock time and the time seen in the AR output image of our system determines the system latency. To account for all delays, the stereoscopic AR system was operated in the full-function mode with simultaneous laparoscopic ultrasound imaging. The overall system latency was measured to be 195 $\pm$10 ms. This includes the latencies of the laparoscopic video and ultrasound imaging systems as well as those of data streaming through Ethernet, volume reconstruction and stereoscopic AR visualization computation.

3.2 Volume Reconstruction

Reconstruction of the ultrasound volume consists of reconstruction from the stack of parallel, sequentially acquired ultrasound slices. Because linear scanning in our 3D ultrasound system is predefined and precisely controlled by the motion controller, the relative position of the acquired 2D ultrasound slices is known. As shown in Figure 4(b), we conducted experiments on a conventional cross-wire phantom (70 mm wide, 50 mm high) submerged in a water tank. At a streaming rate of 30 Hz, ultrasound images were acquired with a resolution of 780 x 800 with a pixel size of 0.07 mm x 0.07 mm.



(a)



(b)

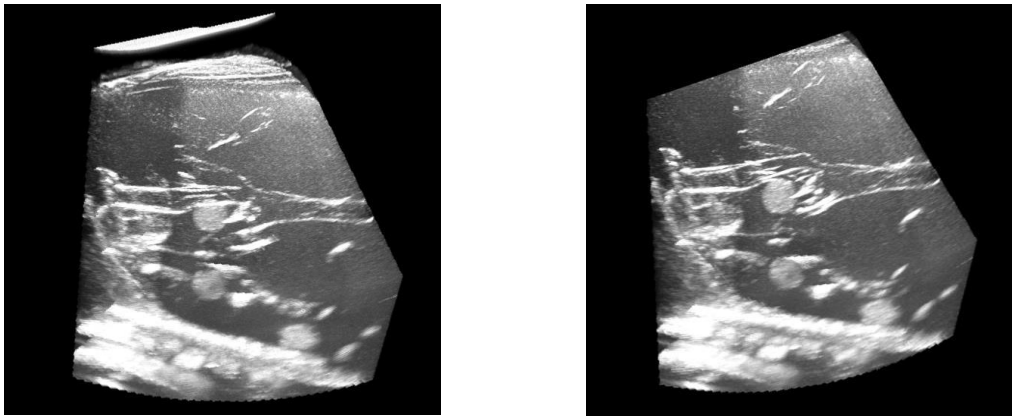
Figure 4. (a) The picture of a cross-wire phantom used for 3D ultrasound imaging to study volume reconstruction quality. (b) A native ultrasound slice (transverse view) showing cross-sections of the two wires (left), reconstructed longitudinal view at slice spacing of 0.2 mm (middle), and reconstructed longitudinal view at a slice spacing of 0.1 mm (right).

Our tests included two different slice spacing settings of 0.2 mm and 0.1 mm with the goal of assessing the quality of volume reconstruction by observing the presence or absence of any staircase effect in the reconstructed volume. For the two settings, the corresponding motor speeds were 6 mm/s and 3 mm/s and it took 12 s and 23 s to complete 3D

scanning of the phantom. The resulting longitudinal view of the reconstructed 3D ultrasound volume exhibited good reconstruction quality (i.e., wires appeared as straight lines). The smaller slice spacing (0.1 mm), as expected, yielded better quality with no staircase artifact.

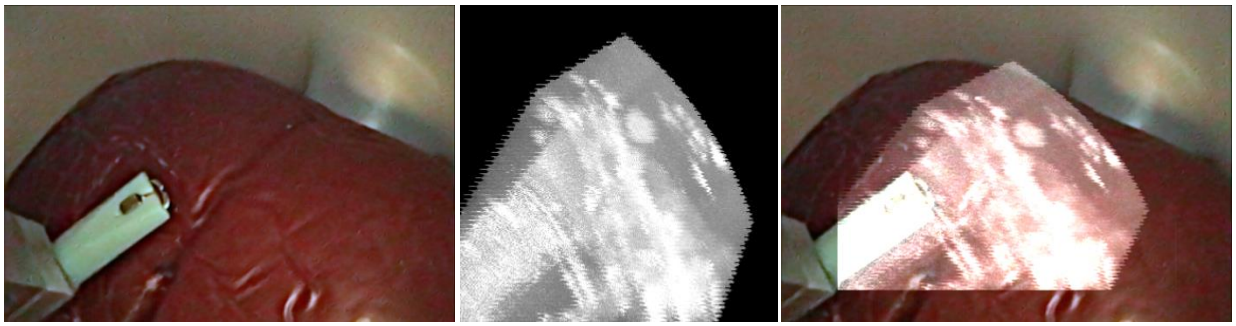
3.3 Phantom study

A phantom study was performed using a realistic intraoperative abdominal ultrasound phantom (IOUSFAN, Kyoto Kagaku Co. Ltd., Kyoto, Japan). The ultrasound volume was reconstructed with a slice spacing of 0.2 mm. Accordingly, a scanned volume of 104 mm x 101 mm x 71 mm was acquired, including anatomical structures such as vasculatures and tumors. With the streaming rate of 30 Hz and the motor speed of 6 mm/s, approximately 12 s were needed to complete a volume sweep. The thin-slab MIP image was created so as to remove the top of the volume as that portion obscured the view of the critical anatomy (i.e., the anatomy of interest) as illustrated in Figure 5. A snapshot of stereoscopic AR video (left-eye channel) recorded during the phantom study is shown in Figure 6. The resulting overlay of the MIP of the ultrasound volume with the stereoscopic video exhibited good alignment.



(a) (b)

Figure 5. (a) Original MIP and (b) thin-slab MIP of the reconstructed ultrasound volume of the abdominal phantom.



(a) (b) (c)

Figure 6. Two stereoscopic AR video snapshots (left-eye channel) recorded during the abdominal phantom studies: (a) the original laparoscopic image, (b) the MIP, and (c) the overlay image generated by our prototype.

4. CONCLUSION

We have developed an early prototype of a surgical AR visualization system that overlays the MIP of a reconstructed ultrasound volume on the video obtained using a stereoscopic laparoscope. The preliminary results show that MIP is an effective volume rendering technique for the present application and that the described mode of AR visualization may be better at interpreting ultrasound data and understanding surgical anatomy than when cross-section ultrasound images are overlaid on laparoscopic video. In our future work, we will convert the current hands-free 3D ultrasound scanner to a fully enclosed version that can then function as a tabletop scanner. Another future direction is to investigate additional volume rendering methods other than MIP. It is expected that the proposed AR system will provide detailed information

on the internal anatomical structures in an easy-to-use and easy-to-interpret manner, and potentially add to the visual information currently available during laparoscopic surgeries.

ORIGINALITY

The work presented in this paper is original and has not been submitted to any other conference or journal for consideration.

REFERENCES

[1] Shekhar, R., Dandekar, O., Bhat, V., Philip, M., Lei, P., Godinez, C., Sutton, E., George, I., Kavic, S., Mezrich, R., Par,k A., "Live Augmented Reality – A new visualization method for laparoscopic surgery using continuous volumetric CT," Surgical Endoscopy 24(8), 1976-1985 (2010).

[2] Kang, X., Azizian, M., Wilson, E., Wu, K., Martin, A.D., Kane, T.D., Peters, C.A., Cleary, K., Shekhar, R., "Stereoscopic Augmented Reality for Laparoscopic Surgery," Surgical Endoscopy (2014) (in press).

[3] Gumprecht, J.D.J., Bauer, T., Stolzenburg, J-U., Lueth, T.C., "A robotics-based flat-panel ultrasound device for continuous intraoperative transcutaneous imaging," Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2152–2155 (2011).

[4] Zhang, Z., "A flexible new technique for camera calibration," IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11), 1330–1334 (2000).

[5] Lasso. A., Heffter, T., Pinter, C., Ungi, T., Fichtinger, G., "Implementation of the PLUS open-source toolkit for translational research of ultrasound-guided intervention systems," Medical Image Computing and Computer-Assisted Intervention (MICCAI) Workshop—Systems and Architectures for Computer Assisted Interventions, 1–12 (2012).

[6] Wilson, S.R., Jang, H-J., Kim, T.K., Iijima, H., Kamiyama, N., Burns, P.N., "Real-Time Temporal Maximum-Intensity-Projection Imaging of Hepatic Lesions with Contrast-Enhanced Sonography," American Journal of Roentgenology 190(3), 691-695 (2008).