Towards A Clinical Stereoscopic Augmented Reality System for Laparoscopic Surgery

Xin Kang, Jihun Oh, Emmanuel Wilson, Ziv Yaniv, Timothy D Kane, Craig A Peters, Raj Shekhar

Sheikh Zayed Institute for Pediatric Surgical Innovation, Children's National Medical Center
{xkang,jhoh,ewilson,zyaniv,tkane,crpeters,rshekhar}@cnmc.org

Abstract. We describe our development of a complete real-time stereoscopic augmented reality system that overlays laparoscopic ultrasound (LUS) images on stereoscopic laparoscopic video for conventional laparoscopic surgery. The system was designed and developed to achieve near-term clinical evaluation as a primary goal. Special consideration was paid to system interactivity, accuracy and easy integration within the existing clinical workflow. Custom-designed fixtures for the two imaging devices were created to avoid their recalibration in the operating room and thus to minimize setup time. The system was assembled on a rolling cart to make it portable for the use in the operating room. Utilizing our optimized design and hardware-accelerated implementation, the system achieved a low system latency of approximately 150 ms. The LUS image-to-video registration accuracy, measured in terms of target registration accuracy and recorded separately for the left and right eye channels of the stereoscopic camera, was 2.38±1.15 mm and 3.66±1.42 mm.

Keywords : Stereoscopic augmented reality (AR), real-time AR, clinical prototype, laparoscopic surgery

1 Introduction

In minimally invasive laparoscopic surgery, the laparoscopic camera is currently the primary means to provide real-time visual information on the surgical field. However, conventional laparoscopes have two significant limitations. First, they provide only a flat 2D representation of the 3D surgical field, introducing ambiguity in depth perception. Second, they are incapable of providing information on internal structures and cannot visualize surgical targets located beneath the exposed organ surfaces.

Several research groups have developed systems and methods to provide combined surface and internal anatomy information by overlaying pre- and intra-operative tomographic images on the laparoscopic video. These augmented reality (AR) efforts have been reported for both robotic surgery [1-2] and conventional surgery [3-7].

The systems and methods using pre- and intra-operative CT and MR images have many limitations that make them less robust and reliable and thus less desirable for operating room (OR) use. These limitations arise from 1) the inability of pre-operative imaging to properly and accurately describe the ever-deforming anatomy during sur-

gery, 2) the fact that the currently employed registration procedures are mostly rigid when the soft-tissue organs deform non-rigidly throughout the surgery; and 3) the subjective and non-reproducible accuracy of manual or semi-automatic registration.

For continuous and automatic updates of the surgical field, Shekhar et al. [7] acquired low-dose non-contrast CT continuously throughout a surgical procedure. The investigators also acquired a standard contrast CT scan immediately before starting surgery. Applying high-speed deformable registration, the vasculature data was transferred from the initial contrast CT to intra-operative non-contrast CT. Although the method does track continuous deformation of the surgical anatomy and provides accurate CT-to-video registration, the risk of high radiation exposure to the patient and the surgical team renders this approach clinically impractical in the foreseeable future.

Laparoscopic ultrasound (LUS) can provide real-time intra-operative images without any ionizing radiation. An advantage of using LUS in combination with live laparoscopic video is that soft-tissue deformation does not need to be modeled, and accurate image-to-video registration can be achieved with standard tracking techniques. Leven et al. [1] took this approach to create a module for the da Vinci robotic surgical system (Intuitive Surgical, Sunnyvale, CA) that superimposes LUS images on stereoscopic laparoscopic video. The rigid LUS probe is tracked by means of a vision-based method that localizes a distinctive pattern situated close to the LUS transducer.

Cheung et al. [3] have developed a platform that uses a stereoscopic laparoscope and a flexible-tip LUS. The two devices were tracked using an electromagnetic (EM) tracking system. The wired EM sensors were affixed to the LUS transducer and the laparoscope. A phantom study was performed to mimic minimally invasive partial nephrectomy [4] in the laboratory. The platform may not be suitable for clinical use since it was designed mainly for feasibility study and its usability was not reported.

In this paper, we describe our development of a complete stereoscopic AR system for conventional laparoscopic surgery. The system overlays LUS images on stereo laparoscopic video in real time. Unlike prior developments that were limited to research prototypes and laboratory testing, our system has been designed and developed by a team of biomedical engineers and minimally invasive surgeons with near-term clinical demonstration as a primary goal. Tested thus far in the laboratory and in animal studies, the system offers acceptable accuracy, low latency, minimal setup time, and minimal changes to the existing surgical workflow for clinical use.

2 Method

2.1 System Design

Our stereoscopic AR system uses two FDA-approved imaging devices: a vision system (VSII, Visionsense Corp., New York, NY) with a stereo laparoscope and an LUS scanner (flex Focus 700, BK Medical, Herlev, Denmark). The stereo laparoscope has a 70-degree field of view, a fixed focal length of 2.95 mm, and an interpupillary distance (IPD) of 1.04 mm. It further features an integrated light source and automatic white balance. The LUS transducer has an operating frequency range from 5 MHz to 10 MHz with a maximum allowable scan depth of 13 cm. The LUS system is capable

of gray-scale B-mode and color Doppler mode scanning. An optical tracker (Polaris, Northern Digital Inc., Waterloo, Canada) is used to track the pose (location and orientation) of the stereo laparoscope and the LUS transducer in real time.

Utilizing the optical tracking data, the LUS images are registered and then overlaid on the stereoscopic video through hardware-accelerated image processing and image fusion. Consequently, two ultrasound-augmented video streams, one for the left eye and the other for the right eye, are generated. Finally, the composite AR streams are rendered for interlaced 3D display. These functions were accomplished in the fusion module. Fig. 1 depicts the architecture of this module. The module is implemented on a 64-bit Windows 7 PC with an 8-core 3.2 GHz Intel CPU, 12 GB memory, and an NVidia Quadro 4000 graphics card.

To acquire images from the two imaging devices in a fast manner, the stereoscopic video and LUS images are streamed to the fusion module over gigabyte Ethernet from the two imaging devices. A custom software library, based on OEM Ethernet communication protocol, was developed to communicate with the LUS scanner. Using this library, our system can fetch LUS images and query imaging parameters (image size, pixel size, imaging mode, and imaging depth). Stereoscopic video images are similarly streamed from the vision system using its Ethernet OEM interface.

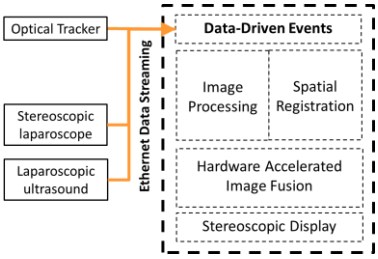

Fig. 1. Architecture of the fusion module.

2.2 System Calibration

For a successful clinical AR system, the two types of images must be spatially registered with sufficient accuracy. This necessitates accurate calibration of the two imaging devices because any calibration errors will lead to misalignment in the stereoscopic AR visualization.

For the stereo laparoscope calibration, a custom-designed calibration phantom was used. It consists of a checkerboard of alternating 5-mm black and white squares in the central region surrounded by a border that is 1 mm higher than the central region (Fig. 2). Four divots (in red and green in Fig. 2) with a depth of 1 mm were created near the four corners of the border. The size of the squares was chosen to ensure that the entire checkerboard stayed within the stereo laparoscope's field of view at the working distance of 5 cm–10 cm. The phantom was printed using a 3D printer (Objet500 Connex, Stratasys, Eden Prairie, MN) with sub-millimeter accuracy. The method proposed by Zhang [8] was used for intrinsic parameters estimation. The transformation from the laparoscope to the reference frame attached to it was determined from the coordinates

of the corners of the checkerboard. By locating the four divots using a tracked stylus, the locations of the checkerboard corners with respect to the reference frame were determined. The left- and right-eye channels of the stereo laparoscope are treated as two standalone cameras and calibrated separately. Although the cameras have a fixed spatial configuration, refinement using homography was not adopted as we empirically found that doing so led to a less accurate calibration, due to the small IPD.

For LUS calibration, we extended the ultrasound imaging research library PLUS [9] by incorporating our data streaming library. Thus, streamed LUS images rather than screenshots of the scanner display formed the input for calibration. A calibration phantom with known geometry for PLUS was 3D printed with sub-millimeter accuracy. Two wires forming two "N" shapes (called N-wires) with known geometry related to the phantom reference frame were made. The intersecting points of the two N-wires with the LUS imaging plane were used for calibration. We used 0.2 mm suturing wires as it produced clear intersecting points with little artifacts in LUS images.

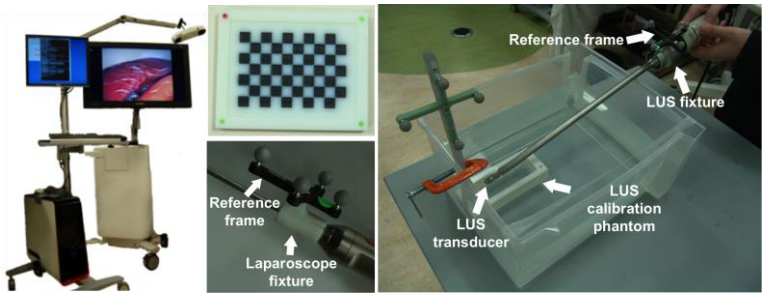

Fig. 2. The system on a rolling cart, the camera calibration phantom, the ultrasound calibration phantom and the custom-designed fixtures for optical tracking.

2.3 Easy Setup for Clinical Use

To track the stereo laparoscope and the LUS transducer optically, reference frames with reflective spheres need to be affixed on them. Because sterilization must precede OR use, an easy mechanism to detach the reference frames from the imaging devices and then re-attach in the OR is desirable. System calibration could be repeated in the OR after re-attaching the reference frames, however, doing so would consume expensive OR time and require an extra technician in the surgical team.

We avoid performing calibration in the OR by re-using laboratory calibration results. This is achieved using our custom-designed mechanical fixtures (see Fig. 2) that are attached to the stereo laparoscope and the LUS transducer uniquely and serve as mounts for the reference frames. In this manner, the reference frames are connected to the devices in exactly the same position as they were before dismounting. This strategy maintains a fixed geometric relationship between the reference frames (and reflecting markers) and the imaging devices before and after sterilization. The fixture for the stereo laparoscope was printed on a 3D printer and made of a synthetic resin. The fixture for the LUS transducer was machined from an aluminum alloy. Both fixtures can withstand the standard autoclave sterilization process.

To make the system portable for OR use, all the components were assembled on a custom-designed rolling cart with an articulated arm for mounting the optical tracker.

3 Experiments and Results

3.1 System Latency

The processing time of the fusion module was measured by imaging a high-resolution digital clock. The difference between the actual time and the time seen in the output image of our fusion module is the processing latency. Note that this measurement included the time of streaming video data from the vision system. The fusion module ran in full-operation mode when measuring the processing latency. The processing latency was 144 ± 19 ms. Independently, we also estimated the latency of streaming LUS images using PLUS and found it to be 230 ± 12 ms.

3.2 System Accuracy

The LUS image-to-video registration accuracy and calibration accuracies of the two imaging devices were measured. The standard target registration error (TRE) metric was used to quantify these accuracies. To measure the LUS image-to-video registration TRE, a target point was imaged using the LUS and its pixel location was identified in the overlaid LUS images. Aiming the laparoscope to the target point from different viewpoints, the 3D location of the target point was calculated using triangulation and compared with its actual location obtained from a tracked pointer. In our experiments, the LUS image-to-video registration TREs were 2.38 ± 1.15 mm and 3.66 ± 1.42 mm for the left-eye and right-eye channels, respectively.

For the stereo laparoscope calibration TRE, images of the calibration pattern were acquired from different viewpoints. Then, the 3D location of the pattern corners were computed using triangulation and compared with their actual locations. Treating the stereo laparoscope as two standalone cameras, the calibration TREs were 0.93 ± 0.18 mm and 0.93 ± 0.19 mm for the left- and right-eye channels, respectively.

For the LUS calibration TRE, a target point was imaged using the LUS. Its 3D location was estimated using the calibration result and compared with the actual location obtained from a tracked pointer. The TRE of the LUS calibration was 1.51 ± 0.39 mm. The error was larger than what is previously reported in the literature using PLUS. This is mainly due to the long shaft of the LUS (Fig. 2). The distance from the reference frame to the LUS transducer is approximately 287 mm. This is considerably larger when compared to non-laparoscopic probes.

3.3 Phantom and Animal Studies

The system was tested in a phantom study and two animal studies involving swine. In the phantom study, an intraoperative abdominal ultrasound phantom (IOUSFAN, Kyoto Kagaku Co. Ltd., Kyoto, Japan), created specifically for laparoscopic applica-

tions, was used. It includes realistic models of the liver, spleen, kidneys, pancreas, biliary tract, and detailed vascular structures, and simulated lesions such as biliary stones cysts, and solid tumors. A stereoscopic AR video snapshot (left-eye channel) recorded during the phantom study is shown in the top row of Fig. 3.

The animal studies were performed on two 40-kg Yorkshire swine by minimally invasive laparoscopic surgeons. Using the stereoscopic AR system, right kidney, liver, and biliary structures were examined by the surgeons with the real-time LUS images superimposed on the stereo laparoscopic video to provide internal anatomical details of the organs. Stereoscopic AR video snapshots (left-eye channel) recorded during the animal studies are shown in the bottom row of Fig. 3.

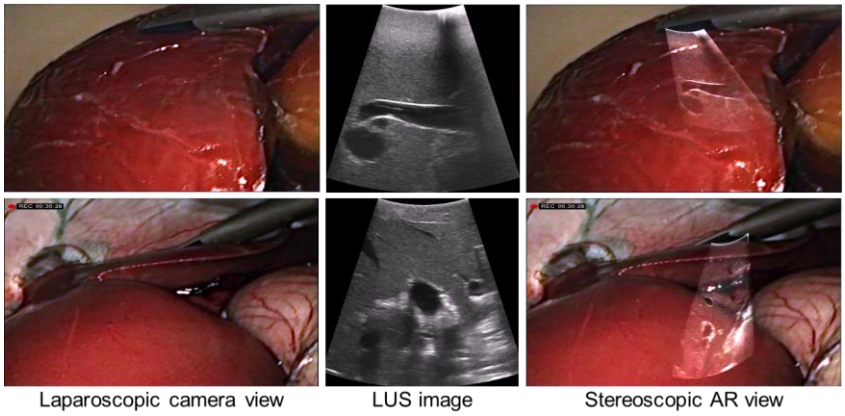

Fig. 3. Two stereoscopic AR video snapshots (left-eye channel) recorded during the phantom (top) and animal (bottom) studies. Each row shows the original stereo laparoscopic camera image (left column), the original LUS image (middle column), and the stereoscopic AR image generated by our system (right column).

4 Discussion

Developing an AR system ready for clinical evaluation requires that the system be interactive, accurate, and integrate easily within the existing clinical workflow. In the present work, in addition to carrying out technical development, we paid considerable attentions to these design criteria and obtained acceptable performance for each.

The processing time including stereoscopic video streaming and stereoscopic AR processing was approximately 144 ms. This low latency comes from optimized system design and hardware-accelerated implementation. In our system, data streaming via Ethernet OEM interfaces significantly reduced communication overhead. Additionally, texture mapping, alpha blending and interlaced display were performed using OpenGL on a quad-buffered graphics processing unit (GPU), which significantly reduced the AR processing time. As for the latency of LUS image streaming, our experiments attribute it to data acquisition and transfer on the LUS scanner side. To improve this, a more powerful backend processor on the scanner side may be needed.

Due to different delays in streaming data from the two imaging devices, attention needs to be placed to data synchronization for AR visualization. One could synchronize the two data streams before performing the stereoscopic AR processing. If using this scheme, the system would have to delay video frames and the net latency would be 374 ms (sum of the three processing times in Fig. 4). The overall result would be a noticeable delay and visual disconnect between AR visualization and the real action related to the surgical procedure. Our system processes the two streaming data in a parallel manner (Fig. 4), i.e., the stereoscopic video is acquired and processed while acquiring the LUS data. In the final step of AR processing, the latest LUS image is overlaid on the prepared stereoscopic video and the video is rendered for 3D display. In this scheme, an 86-ms time lag results between the video and LUS imaging data but it is not noticeable. Our stereoscopic AR system thus uses the most up-to-date video and LUS images to produce real-time stereoscopic AR visualization.

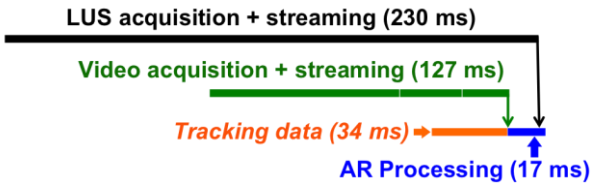

Fig. 4. The scheme for performing temporal synchronization.

The LUS image-to-video registration accuracy is critical from the standpoint of surgical safety. In our system, the primary determinant of accuracy is: how accurately the two imaging devices are calibrated. The stereo laparoscope calibration gave desirable result with a TRE of approximately 1 mm. It is intuitive to treat the stereo laparoscope as a stereo-camera system. However, this led to larger TRE than when treating it as two standalone cameras. The large TRE in the stereo-camera mode can be attributed to small separation between the left-eye and right-eye channel (1.04-mm IPD). Hence, a small error in camera calibration or localization of an image feature leads to a large TRE for the target point far from the reference frame. In our clinical applications, the desired LUS image-to-video registration accuracy is 2.5 mm. Future developments will attempt to further improve the accuracy.

The LUS calibration accuracy, TRE of approximately 1.5 mm, was acceptable using the current design. The long shaft of the LUS (Fig. 2) reduces calibration accuracy using optical tracking because, compared with traditional non-laparoscopic ultrasound probes, the LUS transducer is much farther from the reference frame affixed on its handle. Embedding an EM sensor close to the LUS transducer and replacing optical tracking with EM tracking may improve the accuracy of LUS calibration. No line-of-sight requirement of EM tracking is advantageous in the surgical setting as well.

In clinical practice, easy setup of the system in the OR is desirable to lower adoption barrier and to improve eventual success. The custom-designed fixtures can be easily snapped on to the two imaging devices by the current surgical team. These eliminated system recalibration as described. The rolling cart makes the system porta-

ble without the need for separately setting up the optical tracking system in the OR. These simplify the system setup and hence save expensive OR time.

5 Conclusion

Aiming to bring the stereoscopic AR technology into the clinic, we have developed a complete real-time stereoscopic AR visualization system for conventional laparoscopic surgery. Our ongoing work has addressed key issues of system latency, accuracy and portability. Our future work will include replacing optical tracking with EM tracking and further reducing the size and footprint of the system.

It is expected that the full development and clinical adoption of real-time stereoscopic AR visualization will make minimally invasive laparoscopic surgeries more precise and safer.

References

1. Leven, J., Burschka, D., Kumar R., Zhang, G., Blumenkranz, S., Dai, X.D., Awad M., Hager, G.D., Marohn, M., Choti, M., Hasser, C., Taylor, R.H.: DaVinci canvas: A telerobotic surgical system with integrated, robot-assisted, laparoscopic ultrasound capability. MICCAI, LNCS 3749, 8:811–818 (2005).
2. Su, L-M, Vagvolgyi, B.P., Agarwal, R., Reiley, C.E., Taylor, R.H., Hager, G.D.: Augmented reality during robot-assisted laparoscopic partial nephrectomy: Toward real-time 3D-CT to stereoscopic video registration. Urology 73(4), 896-900 (2009).
3. Cheung, C.L., Wedlake, C., Moore, J., Pautler, S.E., Ahmad, A., Peters, T.M.: Fusion of stereoscopic video and laparoscopic ultrasound for minimally invasive partial nephrectomy. Proceedings of SPIE 7261, 726109–10 (2009).
4. Cheung, C.L., Wedlake, C., Moore, J., Pautler, S.E., Peters, T.M.: Fused video and ultrasound images for minimally invasive partial nephrectomy: a phantom study. MICCAI, LNCS 6363, 408-415 (2010).
5. Teber, D., Guven, S., Simpfendörfer, T., Baumhauer, M., Güven, E.O., Yencilek, F., Gözen, A.S., Rassweiler, J.: Augmented reality: A new tool to improve surgical accuracy during laparoscopic partial nephrectomy? Preliminary in vitro and in vivo results. European Urology 56(2), 332-8 (2009).
6. Simpfendörfer, T., Baumhauer, M., Müller, M., Gutt, C.N., Meinzer, H.P., Rassweiler, J.J., Guven, S., Teber, D.: Augmented reality visualization during laparoscopic radical prostatectomy. J. Endourology 25(12), 1841-5 (2011).
7. Shekhar, R., Dandekar, O., Bhat, V., Philip, M., Lei, P., Godinez, C., Sutton, E., George, I., Kavic, S., Mezrich, R., Park, A.: Live augmented reality: A new visualization method for laparoscopic surgery using continuous volumetric computed tomography. Surg. Endosc. 24(8), 1976-1985 (2010).
8. Zhang, Z.: A flexible new technique for camera calibration. IEEE Trans. Pattern Anal. Machine Intell. 22(11), 1330-1334 (2000).
9. 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. MICCAI - Systems and Architectures for Computer Assisted Interventions, 1-12 (2012).