A Novel Mixed Reality Navigation System for Laparoscopy Surgery

Jagadeesan Jayender[1,2]([✉]), Brian Xavier[3], Franklin King[1],
Ahmed Hosny[3], David Black[4], Steve Pieper[5], and Ali Tavakkoli[1,2]

[1] Brigham and Women's Hospital, Boston, MA 02115, USA
jayender@bwh.harvard.edu
[2] Harvard Medical School, Boston, MA 02115, USA
[3] Boston Medical School, Boston, MA 02115, USA
[4] Fraunhofer MEVIS, Bremen, Germany
[5] Isomics, Inc., Boston, MA 02115, USA

Abstract. OBJECTIVE: To design and validate a novel mixed reality head-mounted display for intraoperative surgical navigation. DESIGN: A mixed reality navigation for laparoscopic surgery (MRNLS) system using a head mounted display (HMD) was developed to integrate the displays from a laparoscope, navigation system, and diagnostic imaging to provide context-specific information to the surgeon. Further, an immersive auditory feedback was also provided to the user. Sixteen surgeons were recruited to quantify the differential improvement in performance based on the mode of guidance provided to the user (laparoscopic navigation with CT guidance (LN-CT) versus mixed reality navigation for laparoscopic surgery (MRNLS)). The users performed three tasks: (1) standard peg transfer, (2) radiolabeled peg identification and transfer, and (3) radiolabeled peg identification and transfer through sensitive wire structures. RESULTS: For the more complex task of peg identification and transfer, significant improvements were observed in time to completion, kinematics such as mean velocity, and task load index subscales of mental demand and effort when using the MRNLS ($p < 0.05$) compared to the current standard of LN-CT. For the final task of peg identification and transfer through sensitive structures, time taken to complete the task and frustration were significantly lower for MRNLS compared to the LN-CT approach. CONCLUSIONS: A novel mixed reality navigation for laparoscopic surgery (MRNLS) has been designed and validated. The ergonomics of laparoscopic procedures could be improved while minimizing the necessity of additional monitors in the operating room.

Keywords: Mixed-reality · Surgical navigation · Laparoscopy surgery
Audio navigation · Visual navigation · Ergonomics

1 Introduction

For several years now, surgeons have been aware of the greater physical stress and mental strain during minimally invasive surgery (MIS) compared to their experience with open surgery [1, 2]. Limitations of MIS include lack of adequate access to the anatomy, perceptual challenges and poor ergonomics [3]. The laparoscopic view only provides surface visualization of the anatomy. The internal structures are not revealed on white light laparoscopic imaging, preventing visualization of underlying sensitive structures. This limitation could lead to increased minor or major complications. To overcome this problem, the surrounding structures can be extracted from volumetric diagnostic or intraprocedural CT/MRI/C-arm CT imaging and augmented with the laparoscopic view [4–6]. However, interpreting and fusing the models extracted from volumetric imaging with the laparoscopic images by the surgeon intraoperatively is time-consuming and could add stress to an already challenging procedure. Presenting the information to the surgeon in an intuitive way is key to avoiding information overload for better outcomes [7]. Ergonomics also plays an important role in laparoscopic surgery. It not only improves the performance of the surgeon but also minimizes the physical stress and mental demand [8]. A recent survey of 317 laparoscopic surgeons reported that an astonishing 86.9% of MIS surgeons suffered from physical symptoms of pain or discomfort [9]. Typically, during laparoscopic surgery, the display monitor is placed outside the sterile field at a particular height and distance, which forces the surgeon to work in a direction not in line with the viewing direction. This causes eye-strain and physical discomfort of the neck, shoulders, and upper extremities. Continuous viewing of the images on a monitor can lead to prolonged contraction of the extraocular and ciliary muscles, which can lead to eye-strain [9]. This paper aims to address the problem of improving the image visualization and ergonomics of MIS procedures by taking advantage of advances in the area of virtual, mixed and augmented reality.

2 Mixed Reality Navigation for Laparoscopy Surgery

A novel MRNLS application was developed using the combination of an Oculus Rift Development Kit 2 virtual reality headset, modified to include two front-facing pass-through cameras, navigation system, auditory feedback and a virtual environment created and rendered using the Unity environment.

2.1 Mixed-Reality Head Mounted Display (HMD)

The Oculus Rift Development Kit 2 (DK2) is a stereoscopic head-mounted virtual reality display that uses a 1920 × 1080 pixel display (960 × 1080 pixels per eye) in combination with lenses to produce a stereoscopic image for the user with an approximately 90° horizontal field of view. The headset also features 6 degrees of freedom rotational and positional head tracking achieved via gyroscope, accelerometer, magnetometer, and infrared LEDs with an external infrared camera. A custom fitted mount for the DK2 was designed and created to hold two wide-angle fisheye lens cameras, as

shown in Fig. 1. The cameras add the ability to provide a stereoscopic real-world view to the user. The field of view for each camera was set to 90° for this mixed reality application. The double-camera mount prototype was 3D printed allowing for adjusting the interpupillary distance as well as the angle of inclination for convergence between the 2 cameras. These adjustments were designed to be independent of one another. Camera resolution was at 640×480 pixels each. It was found that the interpupillary distance had the greatest contribution to double vision - and was hence adjusted differently from one user to another. The prototype was designed to be as lightweight and stable as possible to avoid excessive added weight to the headset and undesired play during head motion respectively. An existing leap motion attachment was used to attach the camera mount to the headset.

Fig. 1. (left) CAD model showing the camera attachment. (right) 3D printed attachment on the Oculus Rift.

2.2 Mixed Reality Navigation Software

A virtual environment was created using Unity 3D and rendered to the Oculus Rift headset worn by the user (Fig. 2). As seen in Fig. 3, a real-world view provided by the mounted cameras is virtually projected in front of the user. Unlike the real-world view, virtual objects are not tethered to the user's head movements. The combination of a real-world view and virtual objects creates a mixed reality environment for the user. Multiple virtual monitors are arranged in front of the user displaying a laparoscope camera view, a navigation view, and diagnostic/intraprocedural images.
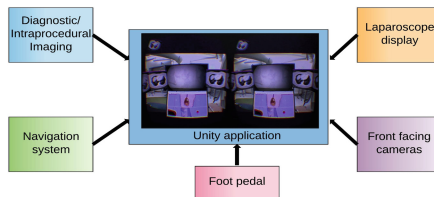

Fig. 2. Software layout of the mixed reality navigation for laparoscopic surgery.

Diagnostic/Intraprocedural Images. A custom web server module was created for 3D Slicer allowing for external applications to query and render DICOM image data to the headset. Similar to the VR diagnostic application [ref-withheld], we have developed a web server module in 3D Slicer to forward volume slice image data to the MR

application, created using the Unity game engine. The Unity application created a scene viewable within the HMD and query the 3D Slicer Web Server module for a snapshot of image slice windows, which is then displayed and arrayed within the Unity scene. The Unity application renders the scene stereoscopically with distortion and chromatic aberration compensating for the DK2's lenses. At startup, image datasets were arrayed hemispherically at a distance allowing for a quick preview of the image content, but not at the detail required for in-depth examination. Using a foot pedal while placing the visual reticule on the images brings the image window closer to allow for in-depth examination.

Surgical Navigation Module (iNavAMIGO). The iNavAMIGO module was built using the Wizard workflow using Qt and C++. The advantage of this workflow is that it allows the user to step through the different steps of setting up the navigation system in a systematic method. The Wizard workflow consists of the following steps – (a) Pre-operative planning, (b) Setting up the OpenIGTLink Server and the Instruments, (c) Calibration of the tool, (d) Patient to Image Registration, (e) Setting up Displays, (f) Postoperative Assessment, and (g) Logging Data.

Setting up the OpenIGTLink Server and the Instruments. In this step, an OpenIGTLink server is initiated to allow for the communication with the EndoTrack module. The EndoTrack module is a command line module that interfaces to the electromagnetic tracking system (Ascension Technologies, Vermont, USA) to track the surgical instruments in real-time. Further an additional server is setup to communicate with a client responsible for the audio feedback. Visualization Toolkit (VTK) models of the grasper and laparoscope are created and set to observe the sensor transforms. Motion of the sensor directly controls the display of the instrument models in 3D Slicer.

Calibration and Registration. Since the EM sensors are placed at an offset from the instrument tip, calibration algorithms are developed to account for this offset. The calibration of the instruments is performed using a second sensor that computes the offset of the instrument tip from the sensor location. Although the iNavAMIGO module supports a number of algorithms to register the EM to imaging space, in this work we have used fiducial-based landmark registration algorithm to register the motion of the instruments with respect to the imaging space.

Displays. The display consists of three panes – the top view shows the three-dimensional view of the instruments and the peg board. This view also displays the distance of the grasper from the target and the orthogonal distance of the grasper from the target. The bottom left view shows the virtual laparoscopic view while the bottom right view shows the three-dimensional view from the tip of the grasper instrument. The instrument-display models and the two bottom views are updated in real-time and displayed to the user. The display of the navigation software is captured using a video capture card (Epiphan DVI2PCI2, Canada) and imported into the Unity game development platform. Using the VideoCapture API in Unity, the video from the navigation software is textured and layered into the Unity Scene. The navigation display pane is placed in front of the user at an elevation angle of −30° within the HMD (Fig. 3 (right)).

Laparoscopic and Camera View. Video input from both front-facing cameras mounted on the HMD was received by the Unity application via USB. The video input was then projected onto a curved plane corresponding to the field of view of the webcams in order to undistort the image. A separate camera view was visible to each eye creating a real-time stereoscopic pass-through view of the real environment from within the virtual environment. Laparoscopic video input was also received by the Unity application via a capture card (Epiphan DVI2PCI2, Canada). The laparoscopic video appears as a texture on an object acting as a virtual monitor. Since the laparoscopy video is the primary imaging modality, this video is displayed on the virtual monitor placed 15° below the eye level at 100 cm from the user. The virtual monitor for the laparoscopy video is also be placed directly in line with the hands of the surgeon to minimize the stress on the back, neck and shoulder muscles, see Fig. 3 (right).

2.3   Audio Navigation System

The auditory feedback changes corresponding to the grasper motion in 3DOFs. In basic terms, up-and-down (elevation) changes are mapped to the pitch of a tone that alternates with a steady tone so that the two pitches can be compared. Changes in left-and-right motion (azimuth) are mapped to the stereo position of the sound output, such that feedback is in both ears when the grasper is centered. Finally, the distance of the tracked grasper to the target is mapped to the inter-onset interval of the tones, such that approaching the target results in a decrease in inter-onset interval; the tones are played faster. The synthesized tone consists of three triangle oscillators, for which the amplitude and frequency ratios are 1, 0.5, 0.3 and 1, 2, and 4, respectively. The frequency of the moving base tone is mapped to changes in elevation. The pitches range from note numbers 48 to 72 on the Musical Instrument Digital Interface (MIDI). These correspond to a frequency range of 130.81 Hz to 523.25 Hz, respectively. Pitches are quantized to a C-major scale. For the y axis (elevation), the frequency f of the moving base tone changes as per the elevation angle. The pitch of the reference tone is MIDI note 60 (261.62 Hz). Thus, the moving tone and reference tone are played in a repeating alternating fashion, so that the user can compare the pitches and manipulate the pitch of the moving tone such that the two pitches are the same and elevation $y = 0$. Movement along the azimuth (x-axis) is mapped to the stereo position of the output synthesizer signal. Using this mapping method, the tip of the grasper is imagined as the 'listener,' and the target position is the sound source, so that the grasper should be navigated towards the sound source.

3   Experimental Methods

A pilot study was conducted to validate the use of the head mounted device based mixed reality surgical navigation environment in the operating room simulated by a FLS skills training box. IRB approval was waived for this study.

Fig. 3. (left) User with the MRNLS performing the trial (right) view provided to the user through the HMD. Virtual monitors show the laparoscopy view (panel a - red hue) and the navigation system display (panel b, c, d). The surrounding environment (label e) can also be seen through the HMD.

Participants were asked to complete a series of peg transfer tasks on a previously validated FLS skills trainer, the Ethicon TASKit - Train Anywhere Skill Kit (Ethicon Endo-Surgery Cincinnati, OH, USA). Modifications were made to the Ethicon TASKit to incrementally advance the difficulty of the tasks as well as to streamline data acquisition (see Fig. 4 (left)). Two pegboards were placed in the box instead of one to increase the yield of each trial. The pegboards were placed inside a plastic container that was filled with water, red dye, and cornstarch to simulate decreased visibility for the operator and increased reliance on the navigation system. Depending on the task, visualization and navigation would be performed using laparoscopic navigation with CT imaging (LN-CT, standard of care) or mixed reality navigation (MRNLS).

Tasks 1 and 2 - Peg Transfer. Using standardized instructions, participants were briefed on the task goals of transferring all pegs from the bottom six posts to the top six posts and then back to their starting position. This task was done on two pegboards using the LN-CT (task 1) and then repeated using the head mounted device (task 2). No additional information or navigation system was given to the participants while wearing the head mounted device other than the laparoscopic camera feed. To determine time and accuracy of each trial, grasper kinematics were recorded from the grasper sensor readings, including path length, velocity, acceleration, and jerk.
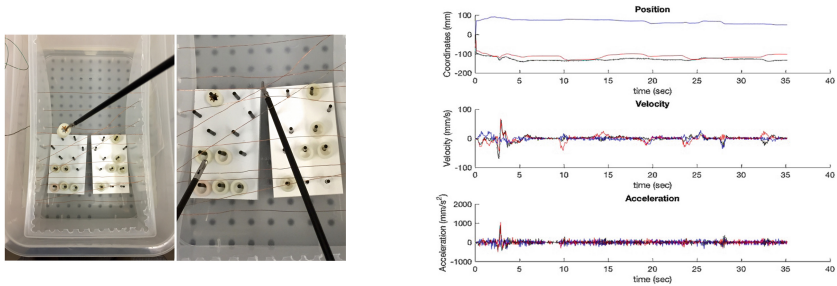

Fig. 4. (left) Example trajectory of the grasper as recorded by the EM sensor.

Tasks 3 and 4 - "Tumor" Peg Identification and Transfer. Tasks 3 and 4 were designed as a modified peg transfer with a focus on using the navigation system and all information to identify and select a target "tumor" peg from surrounding normal pegs,

which were visually similar to the "tumor" peg but distinct on CT images. Participants were instructed to use the given navigation modalities to identify and lift the "tumor" peg on each pegboard and transfer it to the last row at the top of the pegboard. Task 3 had participants use the standard approach of laparoscopy and CT guidance (LN-CT), whereas task 4 was done with the laparoscopic feed, audio navigation, and 3D renderings integrated on the mixed reality HMD environment, i.e., the MRNLS. Metrics recorded included time to completion, peg drops, incorrect peg selections, and probe kinematics such as path length, velocity, acceleration, and jerk.

Tasks 5 and 6 - "Tumor" Peg Identification and Transfer Through Sensitive Structures. For the final two tasks, modifications were made to the laparoscopic skills trainer box to stress the navigation system and recreate possible intraoperative obstacles such as vasculature, nerves, and ducts. Using a plastic frame and conductive wire, an intricate structure was made that could easily be attached for tasks 5 and 6. The structure held the conductive wire above the pegboards in three random, linear tiers (Fig. 4 (left)). A data acquisition card (Sensoray S826, OR, USA) was used to asynchronously detect contact with the wires by polling the digital input ports at a sampling rate of 22 Hz. Contact between the grasper and the wires could then be registered and tracked over time. Operators were asked to identify the radiolabeled "tumor" peg and transfer this peg to the last row on the pegboards. However, in this task they were also instructed to do so while minimizing contact with the sensitive structures. In task 5, participants used the current standard approach of LN-CT, while in task 6, they used the proposed MRNLS system with fully integrated audio feedback, 3D render-based, and image guided navigation environment viewed on the HMD.

Participants. A total of 16 surgeons with different experience levels in laparoscopic surgery volunteered to participate in the study and were assigned to novice or experienced subject groups. Novice surgical subjects included participants who performed more than 10 laparoscopic surgeries as the secondary operator but less than 100 laparoscopic surgeries as the primary operator. Experienced subjects were those who performed more than 100 laparoscopic surgeries as primary operator.

Questionnaire and Training Period. Following each task, participants were asked to complete a NASA Task Load Index questionnaire to assess the workload of that approach on six scales: mental demand, physical demand, temporal demand, performance, effort, and frustration.

Statistical Analysis. The Wilcoxon signed-rank test for non-parametric analysis of paired sample data was used to compare the distributions of metrics for all participants by task. The Mann-Whitney U test was used to compare distributions in all metrics between novice and expert cohorts. $P < 0.05$ was considered statistically significant.

4   Results and Discussion

Figure 4 (right) shows an example trajectory of one of the trials, from which the kinematic parameters have been derived.

Tasks 1 and 2

On the initial baseline peg transfer task with no additional navigational modalities, participants took longer to complete the task when viewing the laparoscopic video feed on the mixed reality HMD, as part of the MRNLS (standard: 166.9 s; mixed reality: 210.1 s; P = 0.001). On cohort analysis, expert participants showed higher significance in time to completion than novices (P = 0.004, P = 0.011). Additionally, there was no difference in number of peg drops or kinematic parameters such as the mean velocity, mean acceleration, and mean jerks per subject amongst all participants or by expertise. During these baseline tasks, mental demand, physical demand, and frustration were significantly increased (P < 0.05) when using the mixed reality HMD environment with mildly significant decrease in perceived performance (P = 0.01). However, effort and temporal demand showed no significant differences amongst all subjects nor novices and experts.

Tasks 3 and 4

Compared to the standard LN-CT in task 3, all participants showed significant decrease in time to completion with the aid of the MRNLS (decrease in time = −20.03 s, P = 0.017). When comparing the addition of the MRNLS in task 4 to the standard approach in novice and expert participants, novice participants showed significant improvements in mean velocity, mean acceleration, and mean jerks between tasks 3 and 4, compared to only mean velocity in experts. Mental demand was significantly decreased when combining the results of both novice and expert participants (P = 0.022) and there was near significance for performance (P = 0.063) and effort (P = 0.089) for the MRNLS.

Tasks 5 and 6

Tasks 5 and 6 were designed to compare the standard LN-CT and proposed MRNLS on a complex, modified task. These final tasks again demonstrated significantly faster time to completion when using the MRNLS in task 8 (100.74 s) versus the LN-CT in task 7 (131.92 s; P = 0.044.) All other kinematic metrics such as average velocity, acceleration, jerks, as well as time in contact with sensitive wire structures, peg drops, or incorrect selections showed no significant difference between navigation modalities for all participants, novices, or experts. Amongst novice participants, there was a decrease in the means of time to completion (−45.5 s), time in contact (−14.5 s), and path length (−432.5 mm) while amongst experts there was a smaller decrease in these metrics (−20.1 s, 2.12 s, −163.1 mm) for the MRNLS. Novices were twice as likely to make an incorrect selection using LN-CT versus MRNLS, however, and experts were 3 times as likely. According to the NASA Task Load Index values, the effort that participants reported to complete the task was significantly lower using the MRNLS compared to the LN-CT (Difference of 1.375, P = 0.011). Upon analysis by expert group, this significance is present among the novice participants but not among expert participants (Novices: −2.57, P = 0.031; Experts: −0.44; P = 0.34). There was a similar result for frustration that was near significance (All participants: −1.38, P = 0.051; Novices: −2.43, P = 0.063; Experts: −0.22, P = 1).

5   Conclusion

We have validated the use of a novel mixed reality head mounted display navigation environment for the intraoperative surgical navigation use. Although further studies are warranted, we find the use of this novel surgical navigation environment proves ready for in-vivo trials with the objective of additionally showing added benefits with respect to surgical success, complication rates, and patient-reported outcomes.

References

1. Patkin, M., Isabel, L.: Ergonomics, engineering and surgery of endosurgical dissection. J. R. Coll. Surg. Edinb. **40**(2), 120–132 (1995)
2. Kant, I.J., et al.: A survey of static and dynamic work postures of operating room staff. Int. Arch. Occup. Environ. Health **63**(6), 423–428 (1992)
3. Keehner, M.M., et al.: Spatial ability, experience, and skill in laparoscopic surgery. Am. J. Surg. **188**(1), 71–75 (2004)
4. Fuchs, H., et al.: Augmented reality visualization for laparoscopic surgery. In: Wells, W.M., Colchester, A., Delp, S. (eds.) MICCAI 1998. LNCS, vol. 1496, pp. 934–943. Springer, Heidelberg (1998). https://doi.org/10.1007/BFb0056282
5. Mountney, P., Fallert, J., Nicolau, S., Soler, L., Mewes, Philip W.: An augmented reality framework for soft tissue surgery. In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (eds.) MICCAI 2014. LNCS, vol. 8673, pp. 423–431. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10404-1_53
6. Shuhaiber, J.H.: Augmented reality in surgery. Arch. Surg. **139**(2), 170–174 (2004)
7. Dixon, B.J., et al.: Surgeons blinded by enhanced navigation: the effect of augmented reality on attention. Surg. Endosc. **27**(2), 454–461 (2013)
8. Erfanian, K., et al.: In-line image projection accelerates task performance in laparoscopic appendectomy. J. Pediatr. Surg. **38**(7), 1059–1062 (2003)
9. Park, A., Lee, G., et al.: Patients benefit while surgeons suffer: an impending epidemic. J. Am. Coll. Surg. **210**(3), 306–313 (2010)