THE INTERNATIONAL JOURNAL OF MEDICAL ROBOTICS AND COMPUTER ASSISTED SURGERY
Int J Med Robotics Comput Assist Surg 2006; **2**: 279–285.
Published online in Wiley InterScience (www.interscience.wiley.com). **DOI:** 10.1002/rcs.102

ORIGINAL ARTICLE

3D ultrasound in robotic surgery: performance evaluation with stereo displays†

Paul M. Novotny[1]
Stephen K. Jacobsen[1]
Nikolay V. Vasilyev[2]
Daniel T. Kettler[1]
Ivan S. Salgo[3]
Pierre E. Dupont[4]
Pedro J. Del Nido[2]
Robert D. Howe[1]*

[1]*Division of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA*

[2]*Department of Cardiovascular Surgery, Children's Hospital Boston, Boston, MA, USA*

[3]*Ultrasound Division, Philips Medical, Andover, MA, USA*

[4]*Department of Aerospace and Mechanical Engineering, Boston University, Boston, MA, USA*

*Correspondence to:
Robert D. Howe, Division of Engineering and Applied Sciences, Harvard University, 29 Oxford Street, Cambridge, MA 02138, USA.
E-mail: howe@deas.harvard.edu

†No conflict of interest was declared.

Accepted: 18 July 2006

Abstract

Background The recent advent of real-time 3D ultrasound (3DUS) imaging enables a variety of new surgical procedures. These procedures are hampered by the difficulty of manipulating tissue guided by the distorted, low-resolution 3DUS images. To lessen the effects of these limitations, we investigated stereo displays and surgical robots for 3DUS-guided procedures.

Methods By integrating real-time stereo rendering of 3DUS with the binocular display of a surgical robot, we compared stereo-displayed 3DUS with normally displayed 3DUS. To test the efficacy of stereo-displayed 3DUS, eight surgeons and eight non-surgeons performed *in vitro* tasks with the surgical robot.

Results Error rates dropped by 50% with a stereo display. In addition, subjects completed tasks faster with the stereo-displayed 3DUS as compared to normal-displayed 3DUS. A 28% decrease in task time was seen across all subjects.

Conclusions The results highlight the importance of using a stereo display. By reducing errors and increasing speed, it is an important enhancement to 3DUS-guided robotics procedures. Copyright © 2006 John Wiley & Sons, Ltd.

Keywords 3D ultrasound; robotic surgery; performance evaluation; stereo display

Introduction

Real-time three-dimensional ultrasound (3DUS) has been demonstrated as a viable tool for guiding surgical procedures (1). This visualization method enables a range of new minimally invasive techniques in cardiac and fetal surgery. For example, beating heart intracardiac procedures are now possible with 3DUS and minimally invasive instruments (2,3): 3DUS permits visualization through the opaque blood pool in the heart, and the advent of real-time 3DUS overcomes difficulties with 3D spatial perception inherent to conventional 2D ultrasound (1). Furthermore, these beating heart procedures eliminate the need for a cardiopulminary bypass and its well documented adverse effects (4–6).

Initial animal trials highlighted several obstacles to clinical implementation of ultrasound-guided intracardiac surgery (2,3). Two such limitations are decreased dexterity and limited spatial perception, making navigation difficult. The first limitation, decreased dexterity, is caused by the necessity of inserting surgical instruments through ports in the heart wall, thereby preventing blood loss and air entry into the heart. These ports limit the range of motion for instruments that are accessing the target site within the heart. A master/slave surgical robot, such as the da Vinci (Intuitive Surgical Inc., Sunnyvale, CA), is ideally suited for this situation. These robots improve dexterity with an actuated wrist that increases the range of motion when compared to minimally invasive instruments. However, these robots are used with endoscopic guidance, with no reported work with 3DUS. The second limitation, limited spatial perception, is caused by the distorted appearance of instruments and tissue under 3DUS, due to high noise levels, shadowing and a variety of artifacts. In addition, depth information is lost when the intrinsically three-dimensional (3D) data is projected onto a two-dimensional (2D) monitor, making it difficult to visualize instruments and tissue. This loss of depth perception exacerbates the already difficult task of performing surgical procedures with 3DUS guidance. However, research into 3D endoscopes has shown that stereo displays improve surgical performance in laparoscopic (7–9) and robotic surgery (10).

In this paper we examine the importance of depth information in 3DUS-guided surgical procedures. We begin by outlining a new fast volume renderer developed for rendering 3DUS volume data on a stereo display. The renderer uses real-time data from a 3DUS machine to render the volume in stereo for a surgeon controlling a surgical robot. The stereo display is tested in tank tests by evaluating subjects' performances while completing surgical tasks with the robot. Eight surgeons and eight non-surgeons performed tasks using stereo-endoscopic visualization, stereo-displayed 3DUS and 2D-displayed 3DUS. Performance is evaluated in terms of task completion times and number of errors.

Materials and methods

To address the inability to visualize depth in 3DUS-guided surgical procedures, we developed a system that displays 3DUS volumes on the stereo display of the da Vinci surgical robot (Figure 1). Stereoscopic viewing provides the surgeon with realistic depth perception and spatial interpretation of the 3DUS data. To produce stereoscopic images of 3DUS data, we developed a custom real-time renderer (11). For real-time visualization, the system must handle and render 30 MB of data every second. This is accomplished by harnessing the computation power of consumer-level graphics processing units (GPUs). Driven by entertainment applications, the computational GPU capacity is moving beyond current CPU capacity, allowing visualization of large volumetric datasets at interactive frame-rates (12). The fundamental advantage of programmable GPUs is their ability to execute highly parallelized per-vertex and per-pixel user routines (shaders). Our implementation uses pixel shaders to cast rays through the volumetric dataset in a ray-per-pixel fashion.

For stereoscopic viewing, two 7800GT (nVidia Corp, Santa Clara, CA) consumer graphics cards render the ultrasound volume in parallel. Left- and right-eye views are separately generated by rendering the 3DUS volume from two viewpoints, mimicking a left and right eye. The separation of these two views is set such that the ultrasound volume appears 0.5 m from the subject's eyes. This distance corresponds to the actual distance between the subject's eyes and their hands on the control console, providing a natural interaction with the robot controls.

An overview of the experimental set-up is shown in Figure 2. The ultrasound volumes are produced by a SONOS 7500 3DUS machine (Philips Medical Systems, Andover, MA). The 3DUS volumes, typically $128 \times 48 \times 204$ voxels, are created at 25 Hz and sent over a 1 Gb TCP/IP network to a personal computer running the rendering algorithm. As the data is received from the ultrasound machine, it is loaded to both GPUs through a PCI-Express bus. Each GPU renders the volume and produces an image from either the left- or the

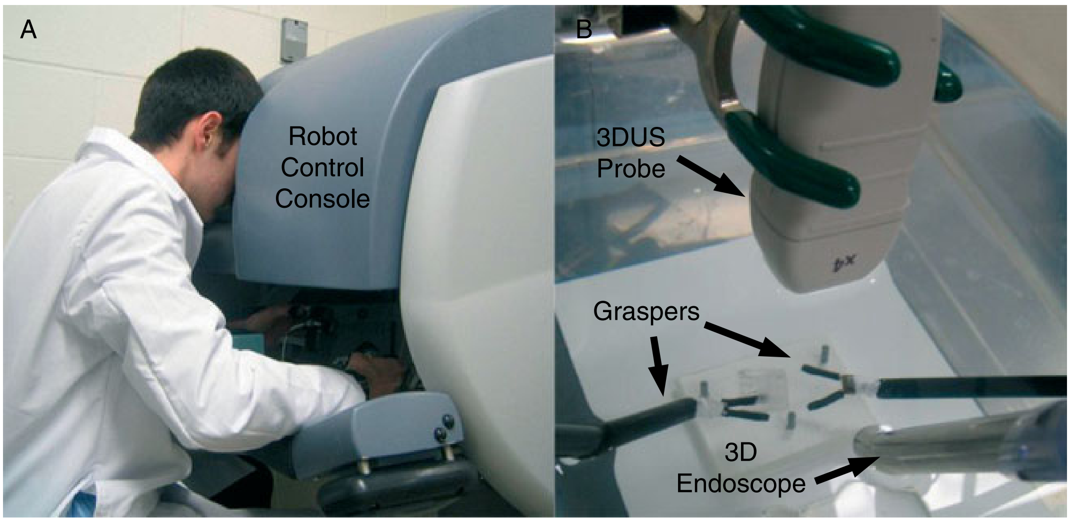

Figure 1. (A) The stereo 3D display is shown on a binocular display on the robot control console. User motions at the controls are mapped to two surgical graspers shown in (B)

DOI: 10.1002/rcs

right-eye viewpoint. This image is then passed to the binocular display of the surgical robot through an S-Video connection. As a result, surgeons use the stereo-rendered ultrasound data for guiding a surgical procedure as they control the robot from the console. The console also contains all of the controls necessary for the surgeon to control the movements of the robot during surgery.

The surgical robot has a pair of seven degrees-of-freedom manipulators with 10 mm laproscopic instruments mounted on each manipulator (Figure 1). The surgeon's movements at the master console are mapped to two surgical graspers controlled by the robot. The workspace of the robot consisted of a 45 cm × 60 cm × 15 cm tank of water over which the robot manipulators, the ultrasound probe and a 10 mm, 0° stereo-endoscope were mounted. Due to the high acoustic reflectance of metal, surgical graspers such as those used in this study appear distorted and incomplete in ultrasound. To minimize distortion and improve their appearance, the metal graspers were coated with electrical heat-shrink tubing and Teflon tape to cover their highly reflective surfaces. An absorptive nickel-impregnated rubber mat was also placed in the workspace to reduce ultrasound reflections from the bottom of the tank.

Evaluation of surgical performance

To quantitatively evaluate surgical performance, 16 test subjects performed 'pegboard' and 'rope pass' tasks using each vision system. These two tasks were selected from among a number of laparoscopic training tasks shown to correlate well with laparoscopic surgical skill (13–15). The tasks were chosen for their emphasis on depth perception as well as their suitability to the limiting factors of the ultrasound's imaging characteristics and field of view. The test subjects were required to complete these tasks guided by both 2D-displayed and stereo-displayed 3DUS as well as by a stereo surgical endoscope which provided a benchmark for comparison.

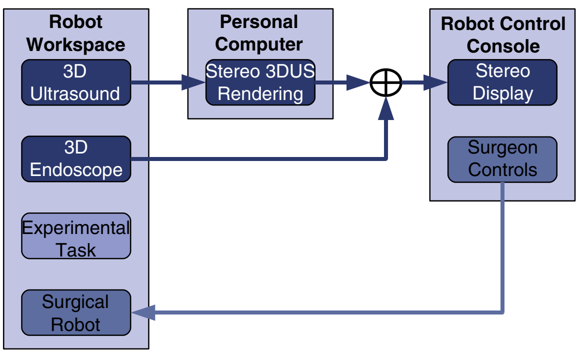

Figure 2. Flow of information and control commands in experimental set-up. The 3DUS data is streamed to a personal computer where it is rendered from a left- and right-eye view. The stereo display of the robot console displays either the left- and right-eye view of the 3D endoscope or the rendered 3DUS. The surgeon sits at the control console and controls the movements of the surgical robot while viewing the stereo display

1. *Pegboard*. In this task, the test subject picked up a plastic collar sitting around one peg, passed the collar between manipulators and placed the collar around a second peg (Figure 3A). To minimize the effects of any learned muscle motion, the initial and destination pegs for the collar were rotated during every trial. The pegs were arranged in a triangular pattern on a 5.0 cm × 6.5 cm acrylic base. The acrylic collar used was 1.3 cm × 1.3 cm × 1.2 cm with a 1.2 cm hole. This relatively large hole made the collar easier to grasp with the surgical manipulators. This emphasizes the amount of time required to manoeuvre the manipulators into the correct position to grip the collar, the task portion most affected by depth perception. The test subject could use the left and right manipulators in either order to move the collar, depending on their own preferences. One trial of the pegboard task consisted of repeating the basic grasp, pass and replace motion once under each vision system; subjects therefore performed nine trials, three for each of the three vision systems.
2. *Rope pass*. The rope pass task consisted of passing a knotted rope from the left to the right manipulator (Figure 3B). A 0.5 cm diameter nylon rope with five knots on 3.0 cm centres was used. The test subject started each trial by gripping the rope to the left of the first knot with his/her left manipulator. The rope was passed one knot at a time by gripping the rope with the right manipulator to the right of the knot currently held by the left manipulator and then moving the left manipulator to the left of the next knot. One trial of the rope pass task consisted of passing the rope five knots under one vision system, starting from the same knot under each vision system. Nine trials were performed, three trials for each of the three vision systems.

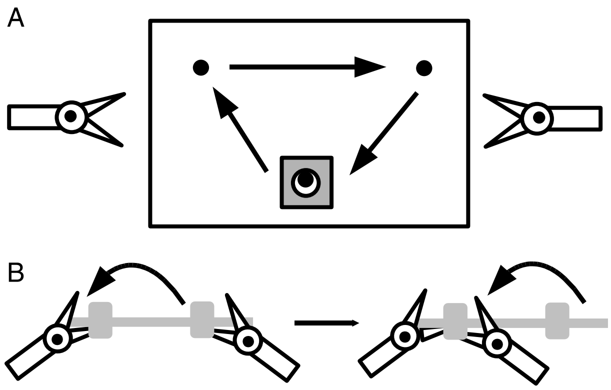

Figure 3. (A) Diagram of the pegboard task. Subjects were instructed to pick up the collar from one peg, pass the collar to the opposite grasper, and place the collar on the next peg. (B) Diagram of the rope pass task. In this task, subjects pass the rope from grasper to grasper. Subjects were instructed to first grab the rope just to the left of the next knot, then grab just to the right of the same knot with the right grasper. In this fashion they moved along the rope until they passed five knots

Procedure

The 3DUS vision system was evaluated by three groups of test subjects. The first group consisted of eight graduate students with no prior surgical experience. These subjects' relevant previous experience ranged from research involving ultrasound imaging and surgical robots to no familiarity with the systems being tested. The second group consisted of four medical doctors with 1–10 years of surgical experience but without experience with surgical robotics. The final group consisted of four surgeons with 7–10 years of surgical experience and at least 1 year of experience with surgical robotics.

Before performing the actual trials, the test subjects were required to complete a short practice programme. This practice session was intended to bring the subjects to a standard level of ability and to limit learning effects during trials. Practice typically took 15 min and was divided into two sections. In the first, the test subject worked under the guidance of the stereoscopic endoscope and manipulated the rope and collar that would be used in the two tasks. This familiarized the subject with the controls of the robot and the objects used in the tasks. The practice session continued until subjects were able to demonstrate proficiency by picking up and passing the collar between graspers five times and passing the rope five knots. In the second practice section, the subjects completed trials of each task with each vision system to become familiar with guiding the robot under the trial conditions. These final practice trials were run like the actual data-collecting trials and provided the test subject with an opportunity to become comfortable with the tasks and testing procedure.

Following the practice programme, the test subjects completed the actual tests (Figure 4). The tests consisted of nine trials of the pegboard task followed by nine trials of the rope pass task, for a total of three trials for each task–vision system combination. During the actual tests, the order of the vision systems used for each trial was counter balanced in order to further remove any effects due to learning.

At the end of the session, the subjects completed a short survey about their experiences with the three vision systems. They were asked to rate the ease of use of each vision system on a scale of 1–5. In addition, the survey asked the subjects to rate their confidence and level of mental fatigue while using the three vision systems.

Analysis

Trial time and number of errors were used to measure task performance. In the pegboard task, errors were defined as dropping the collar during the transfer between the hands, or incorrectly placing the collar. In the rope pass, each time the subject dropped the rope was considered one error. Statistical analysis was used to determine if there is a significant effect of the stereo display

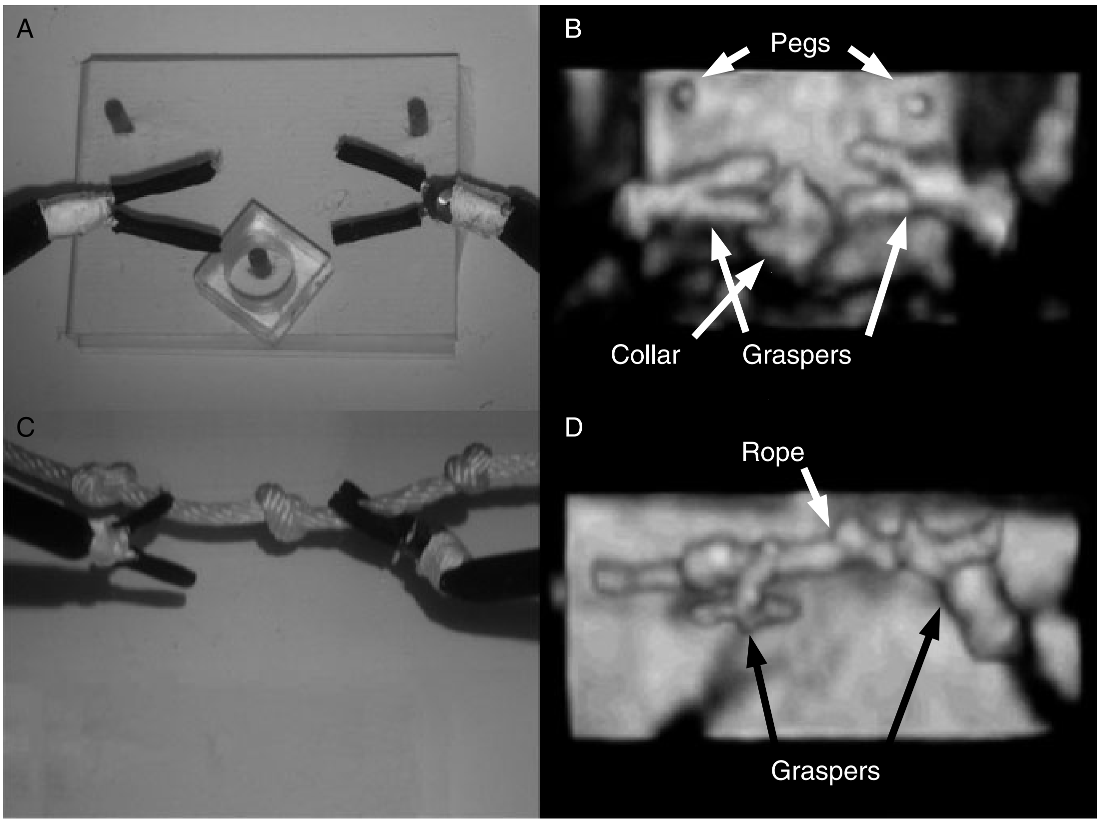

Figure 4. Endoscope image (A) and corresponding 3DUS image (B) of the pegboard task. The endoscope image (C) and the 3DUS image (D) of the rope pass

DOI: 10.1002/rcs

on subjects' performance. Comparisons were done to determine the effects of stereo-displayed 3DUS on mean and variance of the performance metrics. An F-test was used to determine if the variances of the distributions were significant ($p < 0.05$) and a Student's t-test to determine any significant difference in mean ($p < 0.05$) (16).

Results

The results of the study show that stereo-displayed 3DUS improves subjects' ability to complete simulated surgical tasks with a surgical robot. As shown in Figure 5, task completion times for both the rope pass and pegboard decreased when stereo-displayed 3DUS was used compared to traditionally-displayed 3DUS. Across all subjects the completion times for the pegboard task were 10 ± 1 s (mean $\pm$ standard error) with the stereo endoscope, 36 ± 8 s with the stereo 3DUS, and 56 ± 18 s with normal 3DUS. For the rope pass, the completion times were 16 ± 1 s with the stereo endoscope, 40 ± 4 s with the stereo 3DUS display, and 51 ± 7 s with the normal 3DUS display. Statistical analysis was done to investigate the effect of the stereo display on the standard deviation and mean on completion times. An F-test demonstrated that the standard deviation of completion times for both rope pass and pegboard were significantly decreased with the use of a stereo display with 3DUS ($p < 0.05$). Mean times were also significantly decreased ($p < 0.05$), by 35% for the pegboard and 20% for the rope pass.

The vision system also had an effect on the number of errors that the users committed while performing the tasks (Figure 6). Subjects experienced almost no errors while using endoscopic guidance. In the case of the stereo-displayed 3DUS, the average number of errors increased to 0.41 ± 0.16 errors/trial. However, this error rate was not as high as with 2D-displayed 3DUS, where subjects either dropped the collar or rope at a rate of 0.81 ± 0.19 times/trial, a 100% increase over stereo-displayed 3DUS. A Student's t-test confirms that using a stereo display significantly decreases the number of errors during the trials ($p < 0.05$).

Analysis of the survey responses showed subjects' preference for stereo-displayed 3DUS over 2D-displayed 3DUS (Figure 7). When asked to rate the three vision systems on a five-point scale from very difficult (1) to very easy (5), subjects rated the stereo endoscope, stereo 3DUS, and 3DUS as 5.0, 3.5, and 1.8, respectively. Similar responses were given when subjects were asked to rate

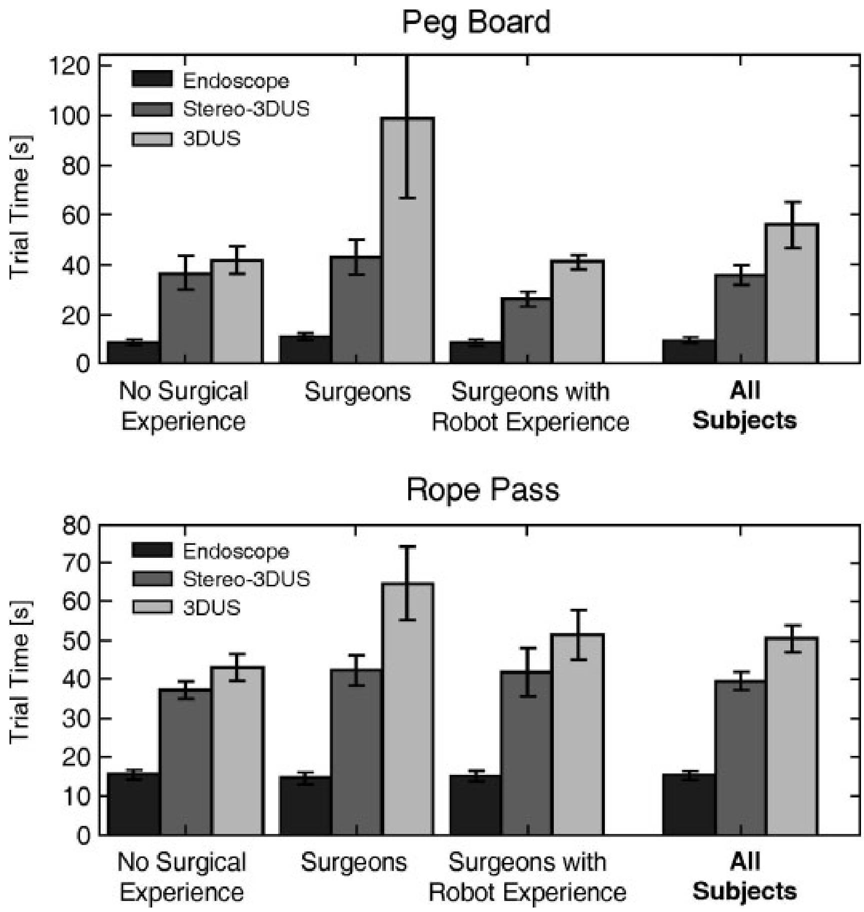

Figure 5. Across all experience levels and both tasks, the trial times decreased with 3DUS when used with a stereo display. The data are shown for each of the three subject groups: subjects with no surgical experience; surgeons; and surgeons with surgical robotic experience. The combined data for all subjects are also shown. Error bars indicate standard error

DOI: 10.1002/rcs

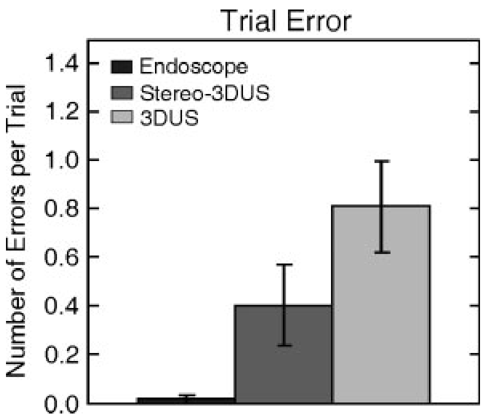

Figure 6. With 3DUS, the use of a stereo display reduced the average number errors per trial. Error bars indicate standard error

their confidence with each vision system when completing the tasks from not confident (1) to very confident (5). Subjects rated the 3D endoscope at 5.0, stereo-displayed 3DUS at 3.5 and 2D-displayed 3DUS at 2.1. Finally, when asked to rate the level of mental fatigue they experienced on a scale from very fatiguing (1) to not at all fatiguing (5), they rated the stereo endoscope, stereo-displayed 3DUS and 3DUS at 4.9, 2.8 and 1.5, respectively.

Discussion

This study showed that stereo-displayed 3DUS improves surgical performance as compared to the normal 2D-displayed 3DUS. The study also demonstrates that surgeons can complete dexterous tasks with a surgical robot under 3DUS guidance. Stereo-displayed 3DUS presents an attractive enhancement of 3DUS by helping improve a surgeon's ability to interpret the noisy ultrasound images. This is especially useful in procedures where endoscopes are not feasible, such as intracardiac and fetal surgery.

To study the effects of depth information with 3DUS-guided procedures, two manipulation tasks were used to evaluate surgical proficiency. These tasks compared the ability to conduct manipulations with a surgical robot using 3DUS with and without a stereo display. These tasks, however, did not fully model the complete surgical environment or fully explore the set of movements or tasks a surgeon performs during a procedure. In addition, the tasks were conducted inside a water tank instead of within a dynamic *in vivo* environment. To mitigate these differences, the tasks used were carefully selected from laparascopic training tasks that are known to correlate with surgical skill (13–15). As a result, differences in performance with these tasks in a water tank should be representative of differences seen in actual surgical procedures.

The performance improvement seen with the stereo display is consistent with results found with 3D endoscopes in laparoscopic (7–9) and robotic (10) surgery. These studies have demonstrated the improvement of surgical performance when comparing normal 2D endoscopes with 3D endoscopes. Widespread use of 3D endoscopes has not come about, due to lower image quality and larger size compared to traditional endoscopes. Furthermore, surgeons are adept at using depth cues, such as foreshortening, occlusion and shading, that infer depth in endoscopic images. In the case of 3DUS, it is a different story. There is no image quality degradation or need for

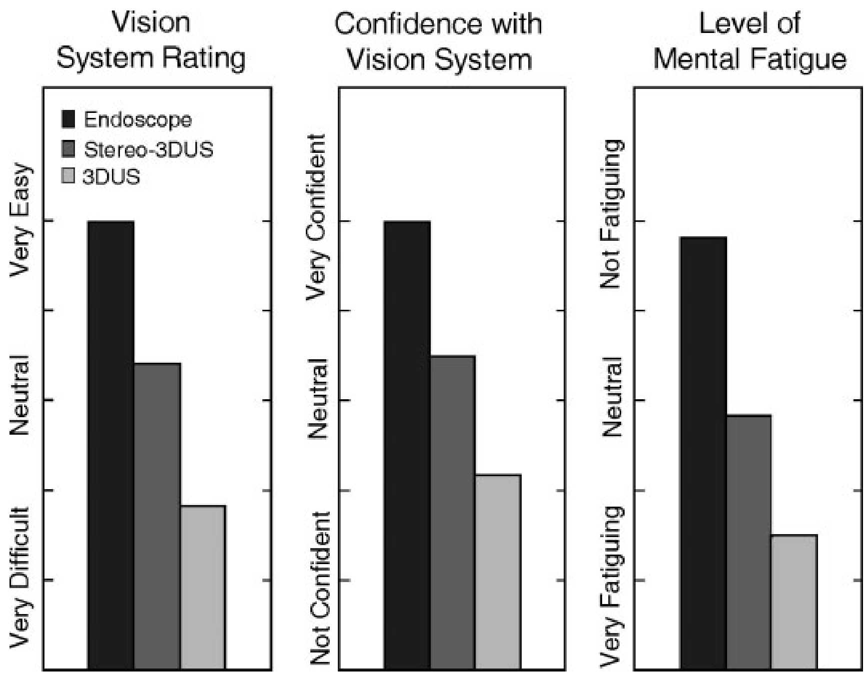

Figure 7. Stereo-displayed 3DUS was subjectively rated easier as user, more confidence-inspiring and less fatiguing than mono-displayed 3DUS

DOI: 10.1002/rcs

a larger ultrasound probe. The stereo display uses the same probe and 3DUS volume for both stereo and non-stereo display. In addition, natural depth cues are not present in ultrasound images, making it difficult to infer three dimensions. As a result, there are few foreseeable disadvantages to adoption of stereo 3DUS displays.

The advantages of using a stereo display with 3DUS are highlighted in this paper. Specifically, incidences of errors, task completion times and repeatability were significantly improved with the stereo-displayed 3DUS. Analysis of trial times showed that subjects were faster and more consistent when using stereo-displayed 3DUS; the mean and variance of the completion time significantly decreased for both tasks in this study. While fundamentally the subjects were performing the same movements to complete the tasks regardless of vision system, the effect of the vision system was most apparent when things went wrong. With the 3D endoscope, a subject's understanding of his/her movements and the environment is near perfect and errors are quickly corrected. With 3DUS, both stereo and non-stereo, if everything is going smoothly, users were able to complete the tasks without trouble. However, when errors in trajectory or understanding of the environment occur, they are not immediately comprehended by the user. These errors are more quickly recognized and corrected under stereo 3DUS than mono 3DUS. The effect is highlighted by the significant difference in the standard deviation of trial time between stereo 3DUS and 3DUS. In the worst case, subjects dropped the rope or misplaced the collar. Such errors occurred 100% more often with the 2D-displayed 3DUS.

In addition to the objective results, subjects also expressed a subjective preference to using the stereo-displayed 3D ultrasound over normal 3DUS. Subjects felt more confident in their movements and experienced a lower level of mental fatigue with the stereo-display. Subjects preferred the 3D endoscope in all aspects over 3DUS, but using stereo is a more feasible alternative to normal 3DUS in situations where endoscopic guidance is not possible.

Future work will focus on improving the appearance of the ultrasound volume. The volume renderer algorithm used in this study was purposely designed without any additional image processing. Additional enhancements to the ultrasound image were excluded in order to isolate the effects of the stereo display. However, a promising next step is to investigate various image-processing techniques that would improve a surgeon's ability to interpret ultrasound images. Techniques such as edge enhancement, intensity correction (17) and direct modification of the opacity transfer function used by the volume renderer (18) will be explored. In addition to improving the rendering, future work will include animal trials to verify the benefits of stereo-displayed 3DUS in realistic surgical scenarios.

The results of this experiment demonstrate the utility of using stereo-displayed 3DUS for improving surgical performance. As 3DUS-guided procedures become more prevalent, these results suggest that clinicians should seriously consider the addition of a stereo display. As we have shown, the adoption of a stereo display lowers error rates, increases speed and improves consistency, three traits very important for surgical procedures.

Acknowledgements

This work is supported by the US National Institutes of Health under Grant No. NIH R01 HL073647-01. The authors would also like to thank all the surgeons and graduate students who volunteered to participate in this study.

References

1. Cannon JW, Stoll JA, Salgo IS, *et al*. Real-time three-dimensional ultrasound for guiding surgical tasks. *Comput Aided Surg* 2003; **8**: 82–90.
2. Suematsu Y, Marx GR, Triedman JK, *et al*. Three-dimensional echocardiography-guided atrial septectomy: an experimental study. *J Thorac Cardiovasc Surg* 2004; **128**: 53–59.
3. Suematsu Y, Martinez JF, Wolf BK, *et al*. Three-dimensional echo-guided beating heart surgery without cardiopulmonary bypass: atrial septal defect closure in a swine model. *J Thorac Cardiovasc Surg* 2005; **130**: 1348–1357.
4. Murkin JM, Boyd WD, Ganapathy S, Adams SJ, Peterson RC. Beating heart surgery: why expect less central nervous system morbidity?. *Ann Thorac Surg* 1999; **68**: 1498–1501.
5. Zeitlhofer J, Asenbaum S, Spiss C, *et al*. Central nervous system function after cardiopulmonary bypass. *Eur Heart J* 1993; **14**: 885–890.
6. Bellinger DC, Wypij D, Kuban KC, *et al*. Developmental and neurological status of children at 4 years of age after heart surgery with hypothermic circulatory arrest or low-flow cardiopulmonary bypass. *Circulation* 1999; **100**: 526–532.
7. Taffinder N, Smith SG, Huber J, Russell RC, Darzi A. The effect of a second-generation 3D endoscope on the laparoscopic precision of novices and experienced surgeons. *Surg Endosc* 1999; **13**: 1087–1092.
8. Huber J, Stringer N, Davies I, Field D. Only stereo information improves performance in surgical tasks. *Proc Int Soc Opt Eng* 2004; **5372**: 463–470.
9. Durrani A, Preminger G. Three-dimensional video imaging for endoscopic surgery. *Comput Biol Med* 1995; **25**: 237–247.
10. Falk V, Mintz D, Grunenfelder J, Fann JI, Burdon TA. Influence of three-dimensional vision on surgical telemanipulator performance. *Surg Endosc* 2001; **15**: 1282–1288.
11. Novotny PM, Kettler DT, Jordan P, *et al*. Stereo display of 3D ultrasound images for surgical robot guidance. IEEE International Conference of the Engineering in Medicine and Biology Society, New York, NY, 2006.
12. Kruger J, Westermann R. Acceleration techniques for GPU-based volume rendering. *IEEE Visualization* 2003; 287–292.
13. Scott DJ, Bergen PC, Rege RV, *et al*. Laparoscopic training on bench models: better and more cost effective than operating room experience?. *J Am Coll Surg* 2000; **191**: 272–283.
14. Derossis AM, Fried GM, Abrahamowicz M, *et al*. Development of a model for training and evaluation of laparoscopic skills. *Am J Surg* 1998; **175**: 482–487.
15. Cotin S, Stylopoulos N, Ottensmeyer M, *et al*. Metrics for laparoscopic skills trainers: the weakest link! Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2002, 5th International Conference, Tokyo, Japan, 2002. Proceedings, Part I, 2002; 35–43.
16. Press WH, Flannery BP, Teukolsky SA, Vetterling WT. *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press: Cambridge, 2002.
17. Xiao G, Brady M, Noble J, Zhang Y. Segmentation of ultrasound b-mode images with intensity inhomogeneity correction. *IEEE Trans Med Imaging* 2002; **21**: 48–57.
18. Honigmann D, Ruisz J, Haider C *et al*. Adaptive design of a global opacity transfer function for direct volume rendering of ultrasound data. *IEEE Visualization* 2003; 489–496.

DOI: 10.1002/rcs