Measuring Situation Awareness of Surgeons in Laparoscopic Training

Geoffrey Tien*
M. Stella Atkins†
School of Computing Science
Simon Fraser University

Bin Zheng‡
CESEI
The University of British Columbia

Colin Swindells§
The University of Victoria
Locarna Systems, Inc.

Abstract

The study of surgeons' eye movements is an innovative way of assessing skill and situation awareness, in that a comparison of eye movement strategies between expert surgeons and novices may show differences that can be used in training.

Our preliminary study compared eye movements of 4 experts and 4 novices performing a simulated gall bladder removal task on a dummy patient with an audible heartbeat and simulated vital signs displayed on a secondary monitor. We used a head-mounted Locarna PT-Mini eyetracker to record fixation locations during the operation.

The results showed that novices concentrated so hard on the surgical display that they were hardly able to look at the patient's vital signs, even when heart rate audibly changed during the procedure. In comparison, experts glanced occasionally at the vitals monitor, thus being able to observe the patient condition.

CR Categories: H.1.2 [Information Systems]: Models and Principles – User/Machine Systems—Human Factors; Human Information Processing; Software Psychology Design; I.4.8 [Computing Methodologies]: Image Processing and Computer Vision—Scene Analysis; Tracking

Keywords: eye tracking, surgery simulation, situation awareness

1 Introduction

1.1 Motivation

Minimally invasive surgery (MIS) is regularly used for many abdominal operations, where typically only two or three small incisions are required instead of one large wound, so recovery times are greatly reduced. To see inside the abdominal area, a small camera and light mounted at the end of a narrow tool (a laparoscope) is inserted into the patient through a small incision, and the resulting image is displayed on an overhead monitor which the surgeon views while manipulating the tools. The display monitor usually shows the ends of the tools within the internal tissue, as seen in Figure 1.

The laparoscope is usually controlled by another member of the surgical team, such as a resident [Eyal and Tendick 2001]. Laparoscopy is now used routinely for removing gall bladders (cholecystectomy) [Gelijns and Rosenberg 1995]. More recently, virtual reality (VR) simulations have been shown to improve surgical skills in an actual operating room (OR) setting [Grantcharov et al. 2004]. Our study attempts to describe differences between the eyegaze of experts and novices, for use in evaluating and potentially enhancing laparoscopic surgery training. The aim is to improve not only eye-hand coordination, but also to improve the situation awareness of the surgeon in the OR.

Figure 1: *Laparoscopic monitor on a simulated task showing tools and target tissues*

1.2 Eye movement differences in experts and non-experts

Tracking the eyegaze of experts and novices has revealed differences in several application domains: in pilots [Kasarskis et al. 2001], in radiology diagnosis tasks [Atkins et al. 2008], and in sports [Vickers 2007]. Kasarkis et al. found during landing that expert pilots fixated more frequently on the airspeed indicator and made fewer fixations on the altimeter whereas novices fixated more frequently on the altimeter. The experts' fixation behavior is learned by their knowledge that the airspeed indicator is more informative. In radiology diagnosis tasks, Atkins et al. showed that experts fixate and saccade more systematically than novices. In sports applications, Vickers shows that expert golfers fixate at one spot on the ball for several seconds, whereas novices tend to fixate on several places on the ball. In surgery, Law et al. [Law et al. 2004] showed differences between expert surgeons and novices performing a simple eye-hand co-ordination task useful for laparoscopy, where it was found that novices tend to look at the tool tip rather than at the target. Measures included fixation times on the tool vs. the target, the number of saccades between the tools and the target, and others. On a single surgical monitor these measures are related to the technical skills of the surgeon.

1.3 Situation awareness in the OR

Situation awareness (SA) is often defined as the mental representation of one's cognition on understanding the current surroundings and the ability to give correct responses based on judgment [End-

*e-mail: gtien@cs.sfu.ca
†stella@cs.sfu.ca
‡e-mail: bin@cesei.org
§e-mail: colin@locarna.com

ETRA 2010, Austin, TX, March 22 – 24, 2010.

sley and Kaber 1999]. Inadequate situation awareness has been shown to be the primary precursor to human error under stress [Endsley and Kaber 1999; Gaba et al. 1995]. To ensure the safety of patients undergoing surgical procedures, it is important for surgeons to possess a high level of mental judgement ability in the OR [Aggarwal et al. 2008; Stefanidis et al. 2007; Zheng et al. 2009]. Explicitly, a competent surgeon should have the ability to pay attention to life-threatening signs during their performance on any complex surgical procedure [Stefanidis et al. 2007; Zheng et al. 2009]. In particular, Stefanidis et al. monitored performance on a visual-spatial multi-task demanding visual attention at different locations, indicating vision as an important component of SA. However, the SA of a surgeon is often measured subjectively by pre-designed assessment forms [Carswell et al. 2005; Berguer et al. 2003]. Results are inconsistent and lack reliability [Gaba et al. 1998; Harrison et al. 2006]. We proposed in this study to use eye-tracking information of surgeons as a probe to measure their SA ability during a simulated laparoscopic cholecystectomy.

2 Methodology

2.1 Experimental apparatus

The experimental apparatus shown in Figure 2 consisted of four components. A PC-based SurgicalSim VR laparoscopic training simulator provided the tactile and main visual interface. A surgical training dummy was placed beside the simulator to provide a more immersive experience, and another PC-based Emergency Care Simulator connected to the dummy provided a simulated anaesthetic display of heart rate, blood pressure, and blood oxygen saturation, in addition to controlling the physical motion of the dummy's breathing. Eye movements were recorded using a lightweight head-mounted eyegaze tracker produced by Locarna Systems. The SurgicalSim VR and Emergency Care Simulator systems are designed by Medical Education Technologies International.

The inputs to the simulator include two long tools inserted into "trocars" in a panel to simulate the full range of motion and actions of a real laparoscopic grasper and electrocautery hook, and a foot pedal. The simulated laparoscopic view was displayed on a 17" LCD monitor located at about eye level at a distance of approximately 180 cm, and depression of the foot pedal produced an audible hum. The anaesthetic display was placed above the dummy at the same height as the main display, and produced audible beeps synchronized to the simulated patient's heartbeat.

Figure 2: *Experimental setup with display monitors, training dummy, and head-mounted eyetracker*

2.2 Description of participants

Eight participants were recruited from the Centre of Excellence for Surgical Education and Innovation (CESEI) at Vancouver General Hospital. Four are surgical residents or practicing surgeons and were categorized as experts for this study, four CESEI staff and research fellows were categorized as novices. Three of the participants wear eyeglasses and one of the experts is female. The average age of the novices is 32.5 and the experts 36.5. No one had any prior experience using a head-mounted eyetracker.

2.3 Procedure

Participants were briefed on the nature of the study, and completed a background survey to measure their experience performing laparoscopic surgery. The task was to use the laparoscopic simulator to perform a cholecystectomy while being considerate of the time taken and patient safety.

Two experimental conditions were presented as different patient histories. The first condition involved a relatively low-risk patient - the anaesthetic simulator produced relatively stable vital signs and breathing rate. The second condition involved a higher risk patient, for whom the anaesthetic simulator cycled through stable - abnormal - recovery vital signs at roughly one-minute intervals. Participants were required to perform a single trial on each condition in a counterbalanced order following a practice trial.

For each condition, participants were first calibrated on the Locarna eyetracker, and given a written description of the patient history. They then performed one trial, and completed the NASA Task Load Index survey to measure the difficulty of the task.

2.4 Data collection and analysis

The Locarna eyetracker consists of two synchronized cameras recording video at 30 Hz which are sent to a customized notebook computer. One camera faces forward relative to the wearer's head and captures the scene at 720 $\times$ 480 resolution. The other camera is aimed towards the wearer's pupil and records at 352 $\times$ 240 resolution. This eye tracking setup, involving a wearable Locarna eye tracker to study participant viewing between a computer screen and nearby physical objects on different viewing planes, is similar to previous work such as Swindells et al. [Swindells et al. 2009].

The recorded video streams are post-processed using Locarna's Pictus software to produce the eyegaze coordinates and fixation locations. The fixation parameters used for this study are minimum duration of 3 video frames (100 ms) with a maximum radius of 40 pixels relative to the video frame recorded from the scene camera. Accuracy was verified to be within 1° by applying 1° error bound circles, using Locarna's Pictus software, and checking the participant's eye gaze towards nine known reference markers that spanned the scene camera's field of view. Due to the nature of calculating fixations relative to the scene camera, these are not strictly traditional fixations as stationary objects recorded by the scene camera appear to move as the head rotates. However, our participants seldom made large, rapid head motions that would cause the Locarna software to miss important details with the chosen parameters. Moving fixations were handled using velocity-based algorithms similar to those described by Salvucci and Goldberg [Salvucci and Goldberg 2000].

Fixations produced from the Locarna Pictus software were annotated as gazing on the main laparoscopic display, the vitals display, or elsewhere. Consecutive fixations on the same item were then collapsed into a single extended dwell. Any change in which item is being dwelled upon is classified as a saccade to the new item.

		P1			P2	
	Total time	Time on lap screen (% total)	Time on vitals screen (% total)	Total time	Time on lap screen (%total)	Time on vitals screen (% total)
Novice 1	238.4*	222.5 (93.3)	0	388.1	365.6 (94.2)	0.2 (< 0.1)
2	123.8	121.8 (98.4)	0	179.3	168.9 (94.2)	0
3	190.4	188.6 (99.1)	0	144.4	143.3 (99.2)	0
4	122.3	188.8 (97.1)	0	167.2	162.5 (97.2)	0.8 (0.5)
Expert 1	274.2	261.8 (95.5)	1.2 (0.4)	177.2	163.7 (92.4)	4.8 (2.7)
2	84.9	78.5 (92.4)	1.5 (1.8)	98.4	86.0 (87.4)	2.4 (2.4)
3	380.3	348.8 (91.7)	14.4 (3.8)	330.8	311.2 (94.1)	7.0 (2.1)
4	135.5	132.7 (97.9)	0	196.9	190.5 (96.8)	0

Table 1: *Task times and time spent dwelling on each display per participant. Times are reported in seconds, with percentage of the total time in parentheses. An asterisk indicates the task was not completed.*

3 Results

The total task time and cumulative dwell durations on each of the main and secondary displays for each participant are illustrated in Table 1. P1 represents the first condition with the stable patient and P2 is the operation on the risky patient.

The number of saccades to the vitals display is presented in Table 2, showing that the experts tended to glance at the vitals screen more often than the novices, for both patient conditions. Only two novices looked at the secondary monitor, and only for the second patient when they heard an irregular heartbeat. However, expert #4 never looked at the vitals screen. Figure 3 shows a histogram of durations of the fixations reported in Table 2 on the vitals screen, for both trials.

Furthermore, expert #1 noted following the operation on the patient with the unstable condition: "There was some arrhythmia, but the patient was fine, nothing dangerous." This expert also mentioned the simulator is only about 30% accurate compared to a real operation – while the simulator is visually realistic, it lacks haptic feedback and important events such as accidental perforation of the gall bladder, which would be a serious incident should it occur during an actual operation.

	P1	P2
Novice 1	0	1
2	0	0
3	0	0
4	0	1
Expert 1	2	6
2	4	6
3	14	9
4	0	0

Table 2: *Saccades to vitals display per subject for each patient.*

The average times for the novice and expert groups to perform their first and second trials are reported in Table 3 and the corresponding average unweighted NASA Task Load Index (TLX) scores are shown in Figure 4.

	First trial (s)	Second trial (s)
Novice	193.8*	194.6
Expert	237.5	182.1

Table 3: *Group average trial times for first and second trials performed. * note one novice did not complete the task*

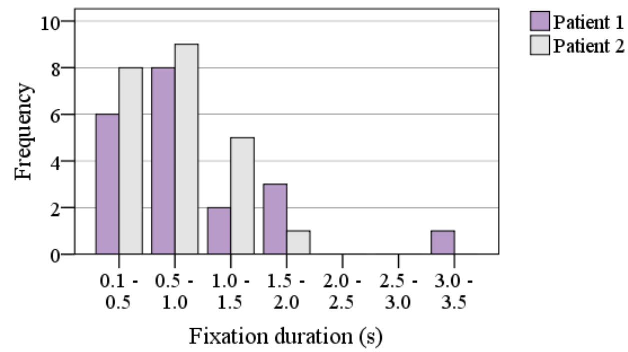

Figure 3: *Histogram of fixation durations for all participants*

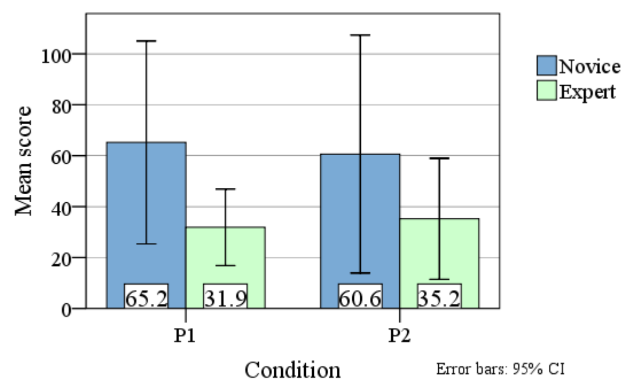

Figure 4: *Mean TLX scores with 95% confidence interval for both patient conditions*

4 Discussion

Table 1 shows that for all participants, over 90% of the elapsed trial times were captured in fixations on the two monitors. This implies our chosen parameters for fixation detection were reasonable. Furthermore, the most time fixated on the vitals screen was by expert #3 at about 3% of the total trial time. Figure 3 shows that the participants usually looked at the monitor for 1 – 2 seconds with just one fixation at 3.5 seconds. This shows that only a short time is needed to obtain relevant information from the vitals monitor. Although seeing does not always equal attention, in this situation experts did indeed gain useful information from looking at the vitals monitor, as indicated by the comments of expert #1.

Table 3 shows that participants in both groups completed the task faster on their second trial (noting that one novice did not complete), most likely due to learning. More practice trials may have been beneficial for the participants to become more familiar with operating the simulation apparatus.

Figure 4 shows that not surprisingly, novices found the task to be more challenging than the experts, although only expert #2 had any prior experience with the simulator. This is likely the reason for expert #2's short trial times and wide variance in general.

The discovery of limitations such as those noted by expert #1 may have led the experts to perform the task more quickly and less gently on their second trial. Regarding the quick trial times, performing the operation very quickly as expert #2 did could be dangerous for the patient, as frequent cautery causes heat and smoke to build up in the body cavity without allowing sufficient time to dissipate.

Furthermore, knowledge of the experimental study and environment may have introduced a different mindset in the experts compared to a live operation. This may have been the case for expert #4 who never looked at the patient's vitals but may likely have done so had it been an actual operation. It should be noted he was the most junior and relatively inexperienced of the experts. In contrast, expert #3, the most senior expert, was very careful and made frequent saccades to the secondary screen even for the stable patient.

5 Conclusion

In this preliminary study, expert surgeons were observed to check the patient's vitals more frequently throughout the duration of a simulated minimally-invasive gall bladder operation. While glances at the vitals monitor cannot always be interpreted to mean expert surgeons have a high level of awareness of the patient condition, the comments of one of our experts indicates that useful patient information was acquired during glances at the vitals monitor. These early results show promise for gaining valuable insights into the development of situation awareness skills over the course of surgical training.

A more rigorous study involving a larger number of experts and more repeated trials will be necessary to observe statistical validity. In the future, a longitudinal evaluation of the clinical awareness of a selected cohort of surgeons in training could be conducted, along with a more thorough post-operative assessment of patient condition and self-evaluation from the participants.

Acknowledgements

The authors would like to thank the Canadian Natural Sciences and Engineering Research Council for funding the eyegaze tracker and CESEI for use of their facilities.

References

AGGARWAL, R., LEONG, J., LEFF, D., WARREN, O., YANG, G.-Z., AND DARZI, A. 2008. New technologies for the surgical curriculum. *World Journal of Surgery 32*, 2 (February), 213–216.

ATKINS, M. S., NICOLAOU, M., AND YANG, G.-Z. 2008. Eye monitoring applications in medicine. In *Passive eye monitoring: Algorithms, applications and experiments*, Springer-Verlag, R. I. Hammoud, Ed., 323–346.

BERGUER, R., CHEN, J., AND SMITH, W. D. 2003. A comparison of the physical effort required for laparoscopic and open surgical techniques. *Archives of Surgery 138*, 967–970.

CARSWELL, C. M., CLARKE, D., AND SEALES, W. B. 2005. Assessing mental workload during laparoscopic surgery. *Surgical Innovation 12*, 1, 80–90.

ENDSLEY, M. R., AND KABER, D. B. 1999. Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics 42*, 3, 462–492.

EYAL, R., AND TENDICK, F. 2001. Spatial ability and learning the use of an angled laparoscope in a virtual environment. In *Proceedings of Medicine Meets Virtual Reality (MMVR)*, IOS Press, 146–153.

GABA, D. M., HOWARD, S. K., AND SMALL, S. D. 1995. Situation awareness in anesthesiology. *Human Factors 37*, 1, 20–31.

GABA, D. M., HOWARD, S. K., FLANAGAN, B., SMITH, B. E., FISH, K. J., AND BOTNEY, R. 1998. Assessment of clinical performance during simulated crises using both technical and behavioral ratings. *Anesthesiology 89*, 1 (July), 8–18.

GELIJNS, A. C., AND ROSENBERG, N. 1995. From the scalpel to the scope: Endoscopic innovations in gastroenterology, gynecology, and surgery. In *Sources of Medical Technology: Universities and Industry*, National Academy Press, N. Rosenberg, A.Gelijns, and H. Dawkins, Eds.

GRANTCHAROV, T. P., KRISTIANSEN, V. B., BENDIX, J., BARDRAM, L., ROSENBERG, J., AND FUNCH-JENSEN, P. 2004. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *British Journal of Surgery 91*, 146–150.

HARRISON, T. K., MANSER, T., HOWARD, S. K., AND GABA, D. M. 2006. Use of cognitive aids in a simulated anesthetic crisis. *Anesthesia & Analgesia 103*, 551–556.

KASARSKIS, P., STEHWIEN, J., HICKOX, J., ARETZ, A., AND WICKENS, C. 2001. Comparison of expert and novice scan behaviors during VFR flight. In *Proceedings of the 11th International Symposium on Aviation Psychology*.

LAW, B., ATKINS, M. S., KIRKPATRICK, A. E., AND LOMAX, A. J. 2004. Eye gaze patterns differentiate novice and experts in a virtual laparoscopic surgery training environment. In *ETRA '04: Proceedings of the 2004 symposium on Eye tracking research & applications*, ACM, New York, NY, USA, 41–48.

SALVUCCI, D. D., AND GOLDBERG, J. H. 2000. Identifying fixations and saccades in eye-tracking protocols. In *ETRA '00: Proceedings of the 2000 symposium on Eye tracking research & applications*, ACM, New York, NY, USA, 71–78.

STEFANIDIS, D., SCERBO, M., KORNDORFFER, J. J., AND SCOTT, D. 2007. Redefining simulator proficiency using automaticity theory. *The American Journal of Surgery 193*, 4, 502–506.

SWINDELLS, C., TORY, M., AND DREEZER, R. 2009. Comparing parameter manipulation with mouse, pen, and slider user interfaces. In *Proceedings of EuroVis, Computer Graphics Forum*, Blackwell Publishing, vol. 29, 919–926.

VICKERS, J. N. 2007. *Perception, cognition, and decision training: The quiet eye in action*. Human Kinetics publishers.

ZHENG, B., CASSERA, M. A., MARTINEC, D. V., SPAUN, G. O., AND SWANSTRÖM, L. L. 2009. Measuring mental workload during the performance of advanced laparoscopic tasks. *Surgical endoscopy online* (May).