

Using an Endoscopic Solo Surgery Simulator for Quantitative Evaluation of Human-Machine Interface in Robotic Camera Positioning Systems

Atsushi Nishikawa¹, Daiji Negoro¹, Haruhiko Kakutani¹, Fumio Miyazaki¹, Mitsugu Sekimoto², Masayoshi Yasui², Shuji Takiguchi², and Morito Monden²

¹ Department of Systems and Human Science,
Graduate School of Engineering Science, Osaka University
1-3 Machikaneyama-cho, Toyonaka City, Osaka 560-8531, Japan

² Department of Surgery and Clinical Oncology,
Osaka University Graduate School of Medicine
2-2 Yamadaoka, Suita City, Osaka 565-0871, Japan

Abstract. An endoscopic solo surgery simulator was designed to quantitatively evaluate human-machine interface in robotic camera positioning systems. Our simulator can assess not only the quantitative efficiency of laparoscopic cameraworks but also the influence of cameraworks upon the accuracy of surgical actions. Two human-machine interfaces: a face motion navigation system and a voice activated system were developed and compared. As a result, the face control interface was more efficient in cameraworks than the voice control, even under a stress to control the instruments. However, it was also found that the face motion may have bad influence on precise surgical actions.

1 Introduction

In laparoscopic surgery, the vision of the operating surgeon depends on a camera assistant responsible for guiding the laparoscope. Recently, several robotic camera positioning systems have been devised towards the realization of “solo surgery” in which the operating surgeon directly controls all of the interventional procedure including the laparoscope guiding task. In such robotic systems, the human(surgeon)-machine(robot) interface is of paramount importance because it is the means by which the surgeon communicates with and controls the robotic camera assistant.

These days comparative experimental studies evaluating the performance of laparoscope positioning systems with different human-machine interfaces are reported[1,2,3,4,5,6,7]. These studies can be classified into two types.

- In [1,2,3,4], typical laparoscopic cameraworks such as pan/tilt/zoom movements and their combinations were objectively and quantitatively assessed. However, no instrument operation was included in the assessment tasks.
- In [5,6,7], a standard intervention task with the regular use of instruments such as laparoscopic cholecystectomy was performed/simulated on a phantom model or a pig(in vivo). However, the task procedure time, the number of camera control commands, and the subjective judgement such as rating on the comfort were only used as criteria for assessment.

The purpose of this project is not only to comparatively evaluate the quantitative efficiency of laparoscopic cameraworks under a lot of stress to control surgical instruments such as forceps and foot pedal, but also to investigate the influence of cameraworks through the human-machine interface upon the accuracy of instrument operation. To date, there has been no research to accomplish this.

We have developed two human-machine interfaces for controlling a laparoscope positioner: a face-motion navigation system and a voice-activated system. These two “real” human-machine interfaces are directly connected with our computer-based assessment system, “Endoscopic Solo Surgery Simulator (ESSS)”, that simulates a laparoscope, the patient’s body and tissues, surgical instruments such as left/right forceps and a camera-holding robot manipulator. We implemented three basic intervention tasks on the simulator for the comparative experiment. These tasks include the model of motions in laparoscopic surgery such as removing the gallbladder. We recorded a lot of data such as the time to complete the task, the camera moving time, the frequency of the surgeon’s commands, the number of errors of the instrument operation, the total moving distance of the tip of the virtual laparoscope, and the total moving distance of the tip of the virtual instruments, then we analyzed them.

2 Overview of Laparoscope Control Interface

First of all, we briefly explain our laparoscope control interface. The details of the system can be found in [8].

2.1 Face Control Interface(FACe MOUSe)

One of our laparoscope control interface, FACe MOUSe system, generates the robot manipulator control commands based on the operating surgeon’s face motion. It mainly consists of a CCD camera placed just over the TV monitor, and a Linux PC with a video capturing board. The core system in the PC can detect and track the surgeon’s face features in real-time from a sequence of video images captured through the CCD camera. The system estimates the position and the pose of the surgeon’s face in real-time from the image processing result and then recognizes the surgeon’s face gestures.

2.2 Voice Control Interface

The other interface forms a voice-controlled laparoscope positioning system, which is based on the commercial voice recognition engine, ViaVoice(IBM Corporation). It mainly consists of a microphone worn by the surgeon and a Windows PC with a USB device, and generates the manipulator control commands by interpreting the surgeon’s simple Japanese words inputted through the USB device from the microphone.

3 Endoscopic Solo Surgery Simulator

In order to compare the above two human-machine interfaces for positioning the laparoscope, we developed a computer-based assessment system, the Endoscopic Solo Surgery Simulator(ESSS), which mainly consists of an instrument control interface and a main computer. The hardware configuration of our assessment system is shown in Fig. 1(a).

3.1 Instrument Control Interface (VLI)

The Virtual Laparoscopic Interface (VLI) (Immersion Corporation) was utilized as the instrument simulation device. The VLI tracks the motion of a pair of laparoscopic surgical instruments simultaneously, each moving in 5 degrees of freedom (a 1st and 2nd D.O.F. for pivoting about its insertion point, i.e., pitch and yaw motions, a 3rd D.O.F. for insertion/retraction, a 4th D.O.F. for spinning about its insertion axis, i.e., roll motion, and the rest for open-close motion of the instrument handle). The latency is less than 1 ms, resulting in sampling rates of 1000Hz. The VLI also monitors the on/off signals of a foot switch in real-time.

3.2 Main Computer

A Windows PC with AMD Athlon 1.4 GHz processor was used as a main computer. It generates three-dimensional computer graphics (3DCG) using the OpenGL graphics library. All of the software modules on the PC were developed by using Visual C++ version 6.0 (Microsoft Corporation). The PC has two standard RS-232 serial ports for respectively connecting to the laparoscope/instrument control interfaces, and also has a graphic card.

The software system configuration is shown in Fig. 1(b). The system has a laparoscopic camera model, a robot manipulator kinematics model, a patient body model, and surgical instrument(forceps) shape/structure models. Analyzing the inputs from the laparoscope/instrument control interfaces based on these models, the simulator in real-time calculates:

1. the position and orientation of the laparoscopic camera,
2. the position and orientation of the two surgical instruments, and the open-close state of their handle,
3. the distance between the tip of the laparoscope/instrument and the center of gravity of the target object(corresponding to the virtual tissue), and
4. the angle between the direction from the tip of the laparoscope/instrument to the centroid of the target point and the direction of the longitudinal axis of the laparoscope/instrument.

Based on the above informations, the simulator decides the state of the target object such as “*appropriately captured by the laparoscope (can be touched/grasped with the instrument)*”, “*in touch with the instrument*”, “*grasped with the instrument*”, “*sticked with the instrument(indicating an error)*”, “*attached to the*

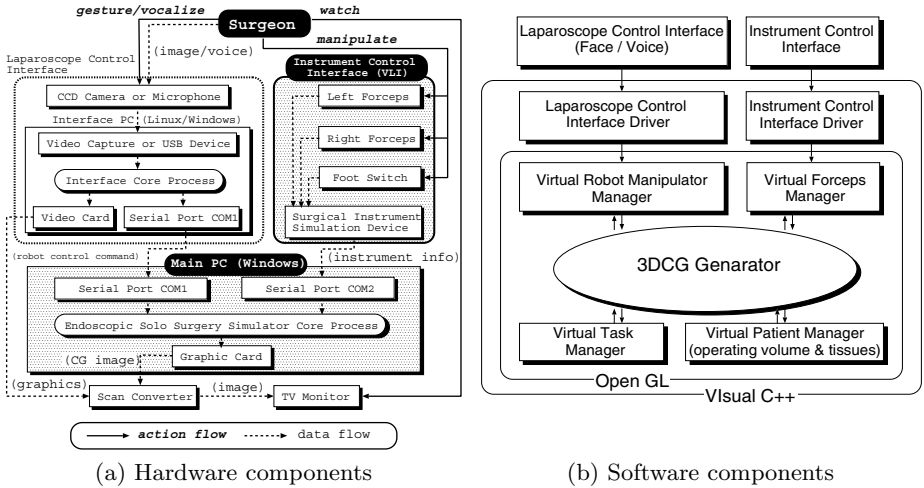


Fig. 1. System configuration.

instrument”, “*can be detached from the instrument*”, and so on. As a result, the system can manage the timing of cameraworks and the number of errors (For details, see the next chapter). Finally, these calculation results are sent to the graphics software module (3DCG generator) and the virtual laparoscopic image is generated using the OpenGL library. The state of the target object is represented by its color, as a feedback information for the operator.

4 Intervention Tasks

Design of the study involved creation of a laparoscopic intervention task that accurately simulated in vivo laparoscopic surgical conditions. We chose the laparoscopic cholecystectomy (gallbladder removal) as the model operation because it is one of the standard laparoscopic interventions and nowadays performed frequently. We made the camerawork analysis of an in vivo laparoscopic cholecystectomy that had been performed[8] using our laparoscope positioning system. Finally, we worked out the following three tasks, in which the instrument operation part was designed based on that of the commercial computer-based laparoscopic trainer, MIST VR[9].

4.1 Task Setting

Fig. 2 shows comparison between the screenshots of the three basic tasks performed in an in vivo laparoscopic cholecystectomy and the corresponding images generated by our simulator. Now we briefly explain these tasks.

Task 1. It was designed to simulate basic maneuvers in the process of “retrieval of gallbladder”. Notice that the target sphere and the target box in Fig. 2(a)

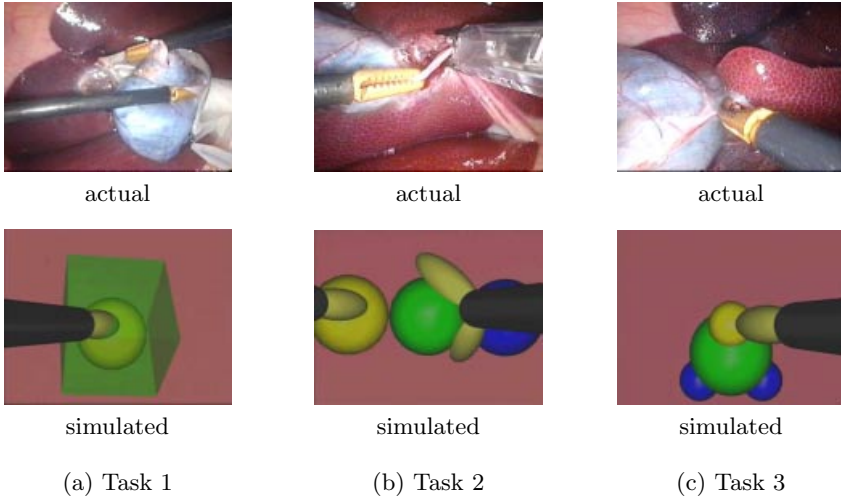


Fig. 2. Screenshots of basic tasks. (a), Task 1: retrieval of gallbladder. (b), Task 2: dissection and ligation of cystic duct and vessel. (c), Task 3: dissection of gallbladder from liver. Upper part: real laparoscopic images taken during an actual cholecystectomy operation. Lower part: virtual laparoscopic images simulated by the ESSS system.

correspond to the gallbladder dissected from liver and the disposal bag respectively.

Task 2. It simulates a part of basic operations in the process of “dissection and ligation of cystic duct and vessel”, especially “clipping of gallbladder duct”. In this task, the target object, corresponding to the cystic duct, is composed of four spheres, as shown in Fig. 2(b).

Task 3. It was designed to simulate basic operations in the time segment of “dissection of gallbladder from liver”. During this process, the surgeon applies a foot pedal-controlled cutting tool (diathermy) to the dissection spot in the gallbladder bed and burns the spot off by contact for a few seconds using a foot switch control. In Fig. 2(c), the main sphere and the small sub-spheres correspond to the gallbladder and the dissection/bleeding spots on the gallbladder surface respectively.

4.2 Task Management

In this section, we summarize the major points of task management in the ESSS.

Error scoring system— If at any time during these processes the instrument tip goes inside the target object, the color of the target object changes to red indicating an error which is recorded. In this case, contact with the instrument tip must be restored before the task can be continued. The judgement of errors

is made based on the radius of the target sphere (r_{target}) and the distance between the instrument tip and the centroid of the target, say d_{tool} . If $d_{tool} \leq C_{error} \cdot r_{target}$ (where C_{error} is a predefined positive constant less than 1), the system will consider that the target is stucked with the instrument.

Camerawork induction mechanism— Let d_{scope} be the distance between the tip of the laparoscope and the center of gravity of the target object. Also let θ_{scope} be the angle between the direction of the longitudinal axis of the laparoscope and the direction from the tip of the scope to the centroid of the target. In the ESSS, the target under consideration can be touched/grasped if and only if it is captured with appropriate magnification in the video center, that is, $d_{scope} < T_d$ and $\theta_{scope} < T_\theta$ (where T_d and T_θ indicates predetermined thresholds). This inevitably induces regular cameraworks without exception. The operator can tell whether the above condition concerning d_{scope} and θ_{scope} is satisfied or not by confirming the color of the target (e.g., the light blue indicates the target is appropriately captured.).

Log file generation— During the task, the frame number, the coordinate of the tip point of the laparoscope, the coordinate of the tip point of the left and right surgical tools, errors if any, and the type of camera control commands are sequentially logged on a text file. After the completion, task beginning time and termination time are also logged on it.

5 Experiment

In this study, the face motion-based laparoscope control interface[8] was evaluated and compared to a voice control interface. All the tasks(Task 1-3) on the ESSS system were applied to this experiment. Camerawork time, the number of camerawork commands, camera moving distance, instrument moving distance, the number of errors, and task completion time were analyzed from the log files generated during each task execution. The task repetition number was set to two. All of the tasks involved four subjects. One of them was familiar with the face controlled interface but the rest had used the system several times before.

Experimental results are illustrated in Fig. 3. In these figures, the error bar with “×” indicates data for the face tracking based system while the “+” marked one indicates the result by voice control. Each error bar illustrates the mean(“×” or “+” position) and minimum/maximum values. Although there was no significant difference between the interfaces for the task completion time (Fig. 3(f)), we got two interesting results:

- For all tasks, both the camerawork time and the laparoscope moving distance were shorter for the face motion-based interface compared to the voice control interface (Fig. 3(a)(b)). Furthermore, the number of camerawork commands in task 3 using the face control was smaller than the number with voice control interface, although there was no significant difference between the interfaces for that in task 1 and task 2 (Fig. 3(c)). In Task 3, the

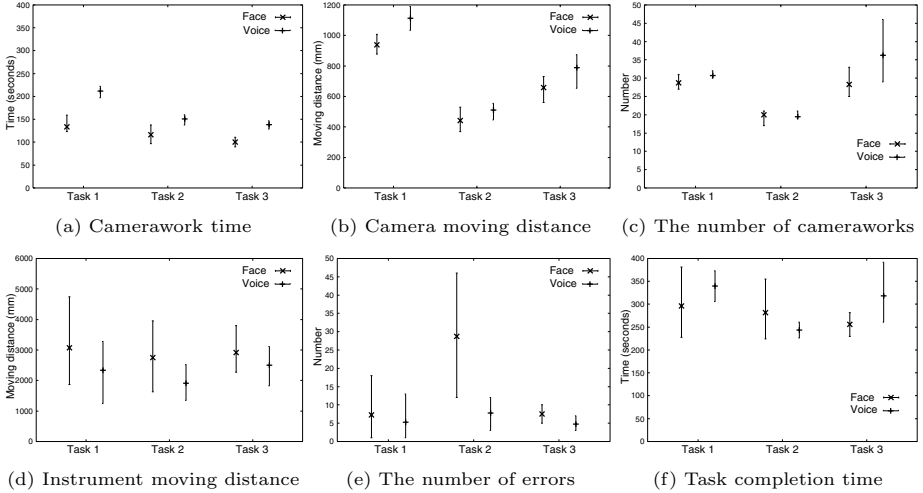


Fig. 3. Results of comparison of the two human-machine interfaces. (a), Camerawork time. (b), Laparoscope moving distance, (c), The number of camerawork commands. (d), Instrument moving distance (mean of the values for the left/right instruments). (e), The number of errors. (f), Task completion time. In these figures, the error bar with “x” indicates data for the face tracking based system while the “+” marked one indicates the result by voice control. Each error bar illustrates the mean (“x” or “+” position) and minimum/maximum values.

subject had to position the laparoscope in high accuracy for removing small-sized subtargets. These results demonstrate the quantitative efficiency of laparoscopic cameraworks through face/voice commands.

- For all tasks, the instrument moving distance was longer for the face interface compared to the voice interface (Fig. 3(d)). Especially, the number of errors in Task 2 was significantly greater for the face motion-based system compared to voice control system (Fig. 3(e)). In Task 2, the subject had to control the laparoscope and the instrument simultaneously and carefully. These results demonstrate the influence of face motion/vocalization upon the instrument operation.

6 Conclusion

Two human-machine interfaces for controlling the laparoscope were comparatively evaluated using a solo surgery simulator. The face control interface was more efficient in laparoscopic cameraworks than the voice control, even under a lot of stress to control the instruments. This is because the face motion can easily and flexibly represent not only the omni-direction of scope motion but also its motion velocity. As shown in the experimental results, however, in our current model it may have a bad influence on complex and precise surgical actions (may

interrupt the flow of the procedure) especially in case that the laparoscope and the instrument should be simultaneously controlled. To solve this problem we are studying an improved model for guiding the laparoscope based on the visual tracking of both the surgeon's face and surgical instruments[10].

References

1. Ballester, P., Jain, Y., Haylett, K. R., McCloy, R.F.: Comparison of Task Performance of Robotic Camera Holders EndoAssist and Aesop. Computer Assisted Radiology and Surgery, Proc. of the 15th Intl. Congress and Exhibition, Elsevier Science (2001) 1071–1074
2. Yavuz, Y., Ystgaard, B., Skogvoll, E., Marvik, R.: A Comparative Experimental Study Evaluating the Performance of Surgical Robots Aesop and Endosista. Surgical Laparoscopy, Endoscopy & Percutaneous Techniques **10** (3) (2000) 163–167
3. Allaf, M.E., Jackman, S.V., Schulam, P.G., Cadeddu, J.A., Lee, B.R., Moore, R.G., Kavoussi, L.R.: Laparoscopic Visual Field. Voice vs Foot Pedal Interfaces for Control of the AESOP Robot. Surg Endosc **12** (12) (1998) 1415–1418
4. Kobayashi, E., Nakamura, R., Masamune, K., Sakuma, I., Dohi, T., Hashimoto, D.: Evaluation Study of A Man-machine Interface for A Laparoscope Manipulator. Proc. of the 8th Ann Meeting of Japan Soc Comput Aided Surg (1999) 121–122
5. Arezzo, A., Ulmer, F., Weiss, O., Schurr, M.O., Hamad, M., Buess, G.F.: Experimental Trial on Solo Surgery for Minimally Invasive therapy. Comparison of Different Systems in A Phantom Model. Surg Endosc **14** (10) (2000) 955–959
6. Buess, G.F., Arezzo, A., Schurr, M.O., Ulmer, F., Fisher, H., Gumb, L., Testa, T., Nobman, C.: A New Remote-Controlled Endoscope Positioning System for Endoscopic Solo Surgery. The FIPS Endoarm. Surg Endosc **14** (4) (2000) 395–399
7. Kobayashi, E., Masamune, K., Sakuma, I., Dohi, T., Shinohara, K., Hashimoto, D.: Safe and Simple Man-machine Interface for a Laparoscope Manipulator System. Comparison of Command Input Methods. J. Japan Soc Comput Aided Surg **3** (2) (2001) 71–78
8. Nishikawa, A., Hosoi, T., Koara, K., Negoro, D., Hikita, A., Asano, S., Miyazaki, F., Sekimoto, M., Miyake, Y., Yasui, M., Monden, M.: Real-Time Visual Tracking of the Surgeon's Face for Laparoscopic Surgery. Medical Image Computing and Computer-Assisted Intervention, Proc. of 4th Intl. Conference, Springer-Verlag (2001) 9–16
9. Wilson, M.S., Middlebrook, A., Sutton, C., Stone, R., McCloy, R.F.: MIST VR: A Virtual Reality Trainer for Laparoscopic Surgery Assesses Performance. Ann R Coll Surg Engl. **79** (6) (1997) 403–404
10. Nishikawa, A., Asano, S., Miyazaki, F., Sekimoto, M., Yasui, M., Monden, M.: A Laparoscope Positioning System based on the Real-Time Visual Tracking of the Surgeon's Face and Surgical Instruments. Computer Assisted Radiology and Surgery, Proc. of the 16th Intl. Congress and Exhibition, Springer-Verlag (2002)