The Johns Hopkins University
Advanced Computer Integrated Surgery

Group 7

Hand-Gesture Interface for the Raven Robot

Paper Seminar on

"Implementation and Evaluation of a Gesture-Based Input Method in Robotic Surgery"

Authors:
Christoph Staub, Salman Can, and Alois Knoll
Robotics and Embedded Systems, Technische Universität München, D-85748 Garching, Germany;

Verena Nitsch, Ines Karl, and Berthold Färber
Human Factors Institute, Universität der Bundeswehr München, D-85577 Neubiberg, Germany

Presented:
IEEE Workshop on Haptic Audio Visual Environments and Games (HAVE), October 14-17, 2011

Seminar Presented by:
Kristine Sarnlertsophon
Project Partner: Alan Chancellor

Project Advisors:
Anton Deguet, Laboratory for Computational Sensing and Robotics, Johns Hopkins University
Kelleher Guerin, PhD Candidate, Department of Computer Science, Johns Hopkins University

CIS II Course Project Background

The Raven Surgical Robot was developed at the University of Washington in Seattle as a platform for researchers to experiment with surgical robots. It was developed with open-source software, and it uses two Phantom Omnis (haptic input devices) for input. We aim to increase its usability by implementing a hand-gesture interface for the Raven. The hand positions and orientations will be acquired using the 3Gear system, which uses two Kinect input devices in order to determine hand poses and positions. We will create interfaces to integrate the 3Gear software with existing libraries designed to compute data for surgical systems (CISST) and control the movement of robots (ROS), finally connecting these to the Raven Robot itself to not only allow for 3Gear hand-gesture input but also open the doors for further work on Raven Robot input devices built upon our library integrations.

Paper Selection

The paper presented in this seminar is: Staub, et. al., ""Implementation and Evaluation of a Gesture-Based Input Method in Robotic Surgery," presented at the IEEE Workshop on Haptic Audio Visual Environments and Games (HAVE) on October 14-17, 2011. This paper implements one type of gesture-based input for a surgery robot and explores its usability for commanding frequently used automated or semi-automated surgical actions. It presents a discussion on the motivation for gesture-based inputs in addition to the haptic or joystick-based inputs used in many current systems.

We aim to explore gesture-based input to control the robot arms themselves, although our ultimate goal is to integrate the software to ease future endeavors in integrating input devices for the Raven Robot and for surgical robots in general. Although our implementation of gesture-based input is vastly different from the one presented in the paper, its insights about surgical robot inputs and the consideration the authors have used in designing their experimental input mechanism are instructive in the design of our own system.

Summary of Problem

One advantage given by surgical robot systems over traditional surgeries is the ability to perform automated or semi-automated tasks, such as knot-tying or tissue-piercing. Surgeons must execute these tasks, along with other commands to the slave robot (such as measuring distance or retracting a 3rd robot arm) while keeping track of the main input device. Camera control is executed with foot pedals, voice commands, or tracking of the surgeons head or eyes; however, many automated tasks are still controlled using a menu system. In using the menu system, the surgeon must remove his hands from the input device and focus his gaze on the interactive menu, which interrupts the workflow of the surgery.

This paper attempts to improve commands for autonomous actions in the experimental Endoscopic Partial-Autonomous Robot (EndoPar) by integrating haptic gesture control. After implementing the controls, the authors evaluated its usability by comparing it to a menu input for speed, accuracy, and user experience as evaluated by the surgeons testing the system.

Key Result and Significance

The authors found that the gesture-based input for the routine commands was faster and more user-friendly than menu-based input, although more errors were made using gesture inputs (10.42% false gesture inputs vs. 5.21% false menu inputs. User-friendliness was evaluated based on the surgeons' opinions on the following established measures of user experience: pragmatic quality, attractiveness, hedonic quality-stimulation (the user feels the experience positively affected the development of knowledge and skills; the user felt more creative or engaged), and hedonic quality-identity (the product promotes one's self-worth; it feels more stylish and valuable).

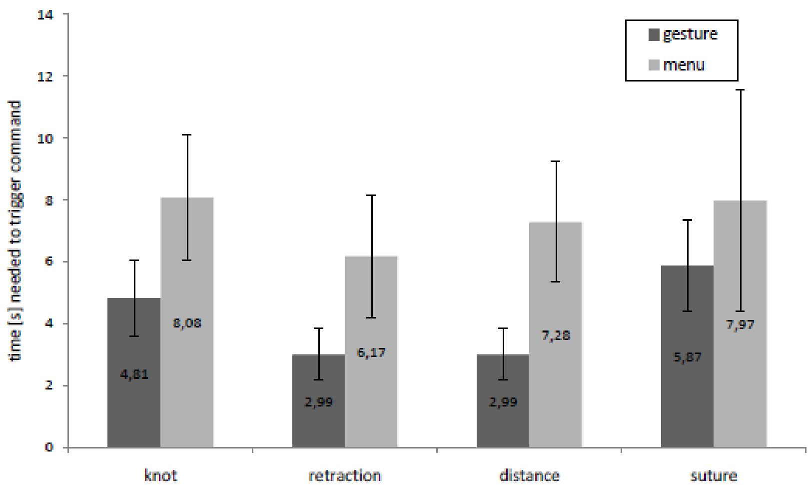

Fig. 7. Times needed to trigger an action: Gesture-based vs. menu.

The paper presents preliminary results showing the feasibility of gesture-based inputs for robot-assisted surgery. The particular implementation of gesture interpretation (discussed under Technical Background) can be further refined in future systems, and the results of the experiment are promising for improving the usability of surgical workstations.

Technical Background

The EndoPar system is an experimental robotic surgical system developed at the Technical University of Munich. Its four slave arms are mounted to the ceiling of the operating room and fitted with EndoWrist instruments (the same kind used in the Da Vinci), which are controlled by and provide force-feedback to two Phantom haptic displays.

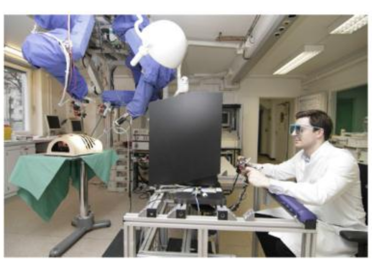

Fig. 1. **Hardware setup:** Ceiling mounted robots with surgical instruments

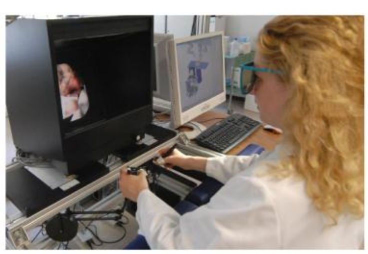

Fig. 2. **Master console** with Phantom™devices and 3D screen

The authors implemented a gesture-based input using the haptic input devices. In this case, surgeons would indicate different commands by gesturing with same two Phantom devices that control the main surgical instruments, thus circumventing the need for the surgeons to release the input device, visually and mentally focus on menu (presumably touch-based), and take hold of the haptic inputs to reposition them to continue with the surgery. The gestures were identified using hidden Markov models, using sampling points from 15 individual executions of each gesture (performed by a person not involved in the later gesture-execution phase of the experiment). A hidden Markov model describes the likelihood of a stochastic system to be in a particular state, based on the probabilities of observing the given output in each state and of transitioning from one state to another. In this experiment, the observed outputs were comprised of the following features for each sampling point: directional change of each instrument's trajectory, directional change of one instrument with respect to the second instrument, velocity of each instrument, distance between the two instruments, temporal change of distance between two instruments, and the state (open or closed) of each gripper.

Experiment

After developing the algorithm to interpret the surgeon's gestures, the authors conducted an experimental user study to compare gesture-based input to menu-based input. The authors conducted a preliminary study in order to choose the most intuitive (and thus most easily remembered and executed) gestures to represent surgical tasks: they asked 22 participants to perform two different gestures for nine pre-selected surgical functions, and chose the four most consistent and highly rated actions to conduct their experiment. These four functions were: knot-tying, suturing, distance measuring, and arm retraction.

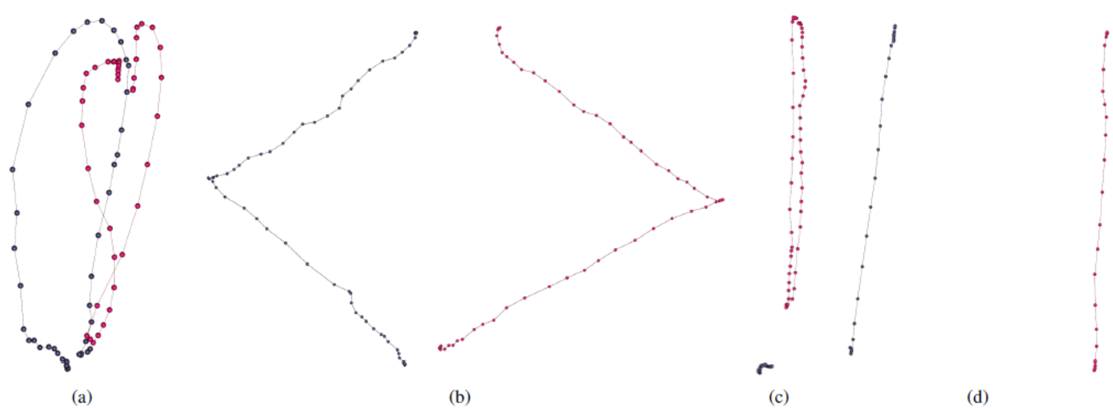

Fig. 6. **Trajectories of gesture instances:** The blue lines indicate the left instrument, the red ones show the right instrument. Fig. 6(a) shows an instance of the "knot-tying" gesture, Fig. 6(b) shows the gesture for "suturing", Fig. 6(c) shows the gesture that initializes the "distance measuring". The picture is rotated 90° counterclockwise to save space. The gesture depicted in Fig. 6(d) would initialize the retraction of the 3rd robot arm to supply the surgeon with new material (e.g., threads).

The main study evaluated the performance of 24 participants, half of whom had surgical experience, using both gesture input and menu input for the four surgical tasks. Prior to the experiment, participants were trained in the use of both input modalities; learning the gesture input took about 7 minutes, whereas learning the menu input took about 3 minutes. After being trained in a certain modality, the participants were tested by being asked to perform a certain action using the input they had just learned.

Input time (the time taken to activate a surgical action) and input success (the accuracy of triggering the correct action) were both measured. User experience was measured with the AttrakDiff, a questionnaire measuring four different aspects of user experience on a seven-point bipolar Likert-type scale: pragmatic quality, attractiveness, hedonic quality-stimulation, and hedonic quality-identity (these were discussed in the "Key Result and Significance" section).

Author's Discussions and Conclusions

This study is a preliminary one exploring the potential of haptic gesture input for surgical robots. While the authors are optimistic that results show gesture input to be faster and have a better user experience, they acknowledge that their experimental conditions and measurements were far from what would need to be conducted to implement a gesture-based interface in a clinical setting.

During the experiment, control of the main instruments was not decoupled when the users were performing their menu or gesture inputs. While this makes their experiment less realistic, implementing this feature would likely skew results even more in favor of gesture inputs, since more time would be taken in repositioning the main input devices to resemble that of the end effectors. Furthermore, while the authors did not explore the effects of training on their setup, it is likely that their participants are more experienced with menu input than with gesture input; hence, errors associated with gesture inputs could decrease with training. Finally, the authors acknowledge their user experience ratings as being biased towards novel and exciting technology.

Although further exploration is needed in comparing input modalities for different systems, and further development needed to implement a safe and effective haptic gesture-based input system, the authors conclude that haptic gesturing is a potential worthwhile addition to more traditional haptic inputs and offers a faster way to express surgical commands than menu inputs.

Evaluation of Paper

Positive points

- The paper was very detailed in explaining the motivation of the study. The authors gave a clear overview of current input devices for surgical robots and the needs and workflows of the surgeons operating them.
- The implementation of the gesture-based inputs was thoughtfully implemented, taking into consideration many factors which comprise a two-handed gesture and using gestures which are already intuitive to surgeons.
- The authors provided a thorough analysis of their experiment, explaining how different experimental conditions might have affected the results (e.g. the decrease in input time for haptic devices might have been more dramatic if participants had to reposition the haptic input devices to match the position of the end effectors after their surgical command) and acknowledging many areas of further study needed to support their conclusion that haptic gesturing is a feasible time-saving addition to current input modalities.

Negative points:

- I found the hidden Markov model to be poorly explained; the authors only explicitly stated that the features measured (such as distance between two instruments) were the observed outputs for each state, without clarifying what the hidden states were in the model and how they measured the probability of transition between states.
- While the authors outlined further studies needed to measure and improve the effectiveness of their particular gesture-based input system, they did not mention any other work or potential studies in other gesture-based input systems, which would be of great interest to many developing inputs for surgical robots (including me).
- They also only measured the four commands they found to be most intuitive to participants without addressing the need for training or thoughtful development for less intuitive commands.
- Commanding the robot to perform specific surgical actions seems a task well-suited to voice input. Using vocal commands, the surgeon would not have to take their hands away from the main input devices or move the input device from its current position. The authors acknowledged voice input in their introduction, but they did not address it as a possible alternative to their system.
- The user experience measures do not answer the problems the authors posed as their motivation for the study. The survey seemed to measure primarily how much the participants enjoyed using the system, without addressing issues such as cognitive burden, distraction from the operating situ, and integration into the surgical workflow.

My Conclusions

Although this paper is far from conclusive evidence for gesture-based input for surgical systems, its preliminary findings are hopeful that more intuitive input modalities can have a significant positive effect on the usability of these systems. Our project implements an (admittedly less thoughtful) system which uses the postures of a surgeon's hands to control the surgical robot; its main purpose is as a proof-of concept that we can join the various libraries we will integrate. This paper shows that the 3Gear input can be utilized to improve usability in ways we hadn't yet conceived, perhaps in combination with voice or other inputs to control surgical commands. In integrating the Raven Robot with the CISST and ROS libraries, we are making it simpler to experiment with combinations of multiple inputs. While the gesture-based inputs covered in this paper were not directly applicable to our gesture inputs, the paper's analysis of surgical robot input devices and its positive initial findings indicate that our work to integrate more input modalities is in the right direction.