Multimodal Training by Demonstration for Robot-Assisted Surgery

Alaa Eldin Abdelaal
Electrical and Computer Engineering Dept., University of British Columbia
Vancouver, BC, Canada
aabdelaal@ece.ubc.ca

Gregory D. Hager
Dept. of Computer Science, Johns Hopkins University
Baltimore, MD, USA
hager@cs.jhu.edu

Septimiu E. Salcudean
Electrical and Computer Engineering Dept., University of British Columbia
Vancouver, BC, Canada
tims@ece.ubc.ca

ABSTRACT

Improving surgical training has the potential to reduce medical errors and consequently to save many lives. We briefly present our efforts to improve this training for robot-assisted surgery. In particular, we explore how data collected from expert demonstrations can enhance the training efficiency for novices. Thus far, our results show that combining hand-over-hand training based on experts' motion data with trial and error training can improve the training outcomes in robotic and conventional laparoscopic surgery settings. We briefly describe our current efforts for exploring how gaze-based training methods, based on experts' eye gaze data, can improve the training outcomes as well.

KEYWORDS

robot-assisted surgery, surgical training, training by demonstration, gaze tracking

ACM Reference Format:
Alaa Eldin Abdelaal, Gregory D. Hager, and Septimiu E. Salcudean. 2020. Multimodal Training by Demonstration for Robot-Assisted Surgery. In *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (HRI '20 Companion), March 23–26, 2020, Cambridge, United Kingdom.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3371382.3377448

1 INTRODUCTION

Medical errors are the third cause of death in the United States right after cancer and heart diseases and well before stroke and car accidents. According to [13], approximately 250,000 people die every year because of medical errors in the United States only. Recent regulations in medical facilities, however, limit the amount of working time available for medical residents, leaving very little time for them to practice their clinical skills [16]. Furthermore, the traditional method of teaching surgery - "See one, Do One, Teach One" - is clearly out of date. In addition to the limited training time and the need for more efficient training methods, other factors contribute to medical errors. For example, the high cognitive load of the surgeons, as well as fatigue, can affect their performance, leading to medical errors [26][8][20]. Thus improving surgeons' training could have a significant societal impact.

Our research targets surgeons' training for robot-assisted surgery. *We hypothesize that eye gaze data in addition to motion and kinematic data collected from expert demonstrations can improve the training efficiency for surgical residents in robot-assisted surgeries.* Thus, our research has the following objectives: (i) To explore the potential benefits of the collected data in facilitating the acquisition of surgical skills; (ii) To investigate the implications of these benefits on the design of surgeon-robot interfaces; and (iii) To develop new training algorithms, based on recorded training and surgery data to allow experienced mentors to remotely transfer their skills to their trainees.

2 RELATED WORK

2.1 Training in Robotic Surgery

Several modalities and interfaces have been proposed for robot-assisted surgical training. Visual interfaces have been extensively explored, mostly in the form of virtual reality simulators [15] [19] such as the Robotics Surgical Simulator (ROSS) [17] and the dV-Trainer [7]. The combination of visual and auditory interfaces has also been tested for surgical training in a knot tying task as in [3]. Haptic interfaces that guide the trainees' hands have also been explored in different ways as in [18], [9] and [5].

2.2 Eye Gaze for Surgical Training

Gaze-based training techniques have been tested in surgical training [14]. This line of research is motivated by the success of these techniques in sports training as in basketball [25], soccer [27] and golf [24]. For example, Causer *et al.* [4] propose the use of a technique called quiet eye training in a one-handed knot tying task in open surgery setting. Quiet eye refers to the final fixation period before starting the actual movement to perform the task. Because expert surgeons tend to have longer fixation periods before moving their hands, the main idea is to train novice surgeons to follow the same strategy. Vine *et al.* [23] present software that is used to guide novice surgeons during their training by showing the gaze of experts which the trainees should follow.

3 WORK ACCOMPLISHED

To achieve the above goals, we developed a system to collect and play back motion and video data from the da Vinci surgical robot [6], which is by far the most widely used surgical robotics system worldwide with an installed base of over 5,000 robots. The da Vinci system is a master-slave system where the master contains the surgeon's console which in turn contains a display system, a user interface, two Master Tool Manipulators (MTMs), and four foot-pedals. The

HRI '20 Companion, March 23–26, 2020, Cambridge, United Kingdom

ACM ISBN 978-1-4503-7057-8/20/03.
https://doi.org/10.1145/3371382.3377448

slave part of the system is called the patient-side cart which has four patient-side manipulators (PSMs), three of which can be used for teleoperation of surgical tools and the fourth arm holds the endoscopic camera of the da Vinci robot. The surgeon can then control the PSMs at the patient-side cart using the MTMs at the surgeon's console.

The developed system consisted of two main components: recording and playback. *In the recording component*, the joint angles of both the MTMs and PSMs were acquired. Forward kinematics was then used to derive the end effector's positions and orientations of all the MTMs and PSMs. Furthermore, the recording component was used to acquire the stereovision feed from the endoscopic camera and synchronize it with the collected motion data. This component was activated when an expert performed a surgical training task. *In the playback component*, the PSMs or MTMs were automatically moved based on the recorded motion data. This was done using the da Vinci Research Kit (dVRK) [11] and Robot Operating System (ROS). Moreover, the recorded videos were also played back in synchronization with the motion data. This component was activated during the training process.

We used the collected data for (i) robot-assisted surgery training; and (ii) standard laparoscopic surgery training. For the former, we played back the recorded motion data of the MTM and asked trainees to hold it passively while it was moving. The recorded videos were played back on the surgical console. For the latter, we played back the motion data of the PSM and we added a handle to it to make it easy for trainees to hold it. The recorded videos were played back on a computer monitor in front of the trainees.

We conducted two user studies (N = 21 for each) to evaluate the effectiveness of the proposed system and training approach for standard laparoscopic and robotic surgery settings. In these user studies, we compared three training approaches: *discovery training* (i.e., trial and error training), *hand-over-hand training* (i.e., holding the robot arm passively while it played back the recorded motion of an expert) and *a combination of the two*. For each study, participants were divided into three groups, one for each training approach. We measured the baseline performance of the trainees before training and we measured it again after training in retention and transfer tests based on completion time, number of committed errors and the multiplication of the completion time and errors. For the surgical robotics training study, the retention and transfer tests were carried out on the da Vinci robot. For the standard laparoscopic surgery training study, the two tests were conducted using a standard laparoscopic surgery tool and a training box.

The results of the two user studies show three main findings. First, combining discovery training with hand-over-hand training leads to higher accuracy (as measured by the number of committed errors) than using any of the two training approaches alone. Second, this combined training approach enables trainees to balance their performance speed and accuracy in a much better and statistically significant way (as measured by the multiplication of completion time and number of committed errors) than any of the two approaches alone. Third, our results show for the first time that training on a surgical robotics platform leads to statistically significant improvements of the performance of trainees when tested on a standard laparoscopic (non-robotic) surgery setting. The detailed description of the user studies and results can be found in [1].

4 ONGOING AND REMAINING WORK

Our next step is adding gaze tracking to our training system and collecting data from both expert and novice surgeons. We will use the eye gaze tracker in [12, 21] to test the feasibility of using several training techniques from the literature that use eye gaze such as quiet eye training [4]. We updated our recording and play back system to acquire raw eye gaze data in synchronization with all the data the system can collect. We can now overlay the estimated gaze positions on the collected videos using our system. Moreover, we designed new versions of this gaze tracker's hardware that are compatible with newer models of the da Vinci system. Furthermore, we are currently working on improving the gaze estimation accuracy of this tracker. The best binocular accuracy we achieved so far is $1.32° \pm 0.55°$. We are also exploring the use of machine learning methods to facilitate the tracker's calibration process.

Besides, we are currently exploring the quiet eye phenomenon in multiple surgical specialties. In particular, we are exploring this phenomenon in an eye gaze tracking data set for sinus surgery, which was used in [2]. We will also work with our clinical collaborators to explore this phenomenon in urology. This work will help define candidate tasks to apply quiet eye in surgical training.

In addition, we have the Institutional Board Ethics approval to deploy our data collection system at the Vancouver General Hospital and the surgical robotics training facility at our university to collect motion and eye gaze data during surgeries as well as during residents' surgical training sessions. The collected data will be used with machine learning algorithms to find the important features that distinguish experts and novices. This will be the basis for developing novel training methods that will focus on these features to improve the training efficiency.

We are going to conduct user studies with novices to evaluate the effectiveness of the proposed training methods. For evaluation, we will use general performance metrics such as accuracy and completion time. We also plan to use task-specific performance metrics (e.g., the knot tying assessment criteria proposed in [22]). We will also test the robustness of these training methods to high levels of pressure and anxiety.

Methodologically, we will use techniques such as Hidden Markov models (HMM) and Gaussian Mixture Model (GMM) and results from motor skill learning literature to improve our training framework. HMM and GMM will be used to find general patterns and relations between several demonstrations of different surgeons carrying out the same task. This will help us understand the underlying principles of surgical skills. We will also use established motor skill learning models such as Fitts and Posner's model [10] to design the steps of our training procedures to facilitate the acquisition of the motor skills. We will also build upon our previous work on exploring additional benefits of the data collected from robotic surgeries in standard laparoscopic ones.

ACKNOWLEDGMENTS

This work is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canada Foundation for Innovation, Intuitive Surgical, the C.A. Laszlo Chair in Biomedical Engineering held by Dr. Septimiu E. Salcudean and the Vanier Canada Graduate Scholarship held by Alaa Eldin Abdelaal.

REFERENCES

[1] Alaa Eldin Abdelaal, Maram Sakr, Apeksha Avinash, Shahed K Mohammed, Armaan Kaur Bajwa, Mohakta Sahni, Soheil Hor, Sidney Fels, and Septimiu E Salcudean. 2018. Play Me Back: A Unified Training Platform for Robotic and Laparoscopic Surgery. *IEEE Robotics and Automation Letters* 4, 2 (2018), 554–561.

[2] Narges Ahmidi, Gregory D Hager, Lisa Ishii, Gabor Fichtinger, Gary L Gallia, and Masaru Ishii. 2010. Surgical task and skill classification from eye tracking and tool motion in minimally invasive surgery. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 295–302.

[3] Mohammed J Al Fayyadh, Ramy A Hassan, Zachary K Tran, Jason W Kempenich, Leonid Bunegin, Daniel L Dent, and Ross E Willis. 2017. Immediate Auditory Feedback is Superior to Other Types of Feedback for Basic Surgical Skills Acquisition. *Journal of surgical education* 74, 6 (2017), e55–e61.

[4] Joe Causer, Adrian Harvey, Ryan Snelgrove, Gina Arsenault, and Joan N Vickers. 2014. Quiet eye training improves surgical knot tying more than traditional technical training: A randomized controlled study. *The American Journal of Surgery* 208, 2 (2014), 171–177.

[5] Margaret M Coad, Allison M Okamura, Sherry Wren, Yoav Mintz, Thomas S Lendvay, Anthony M Jarc, and Ilana Nisky. 2017. Training in divergent and convergent force fields during 6-DOF teleoperation with a robot-assisted surgical system. In *Proc. of IEEE World Haptics Conference (WHC’17)*. 195–200.

[6] Jonathan Douissard, Monika E Hagen, and P Morel. 2019. The da Vinci Surgical System. In *Bariatric Robotic Surgery*. Springer, 13–27.

[7] dVTrainer [n.d.]. *The dV-Trainer*. Retrieved January 14, 2020 from http://www.mimicsimulation.com/products/dv-trainer/

[8] Brian J Eastridge, Elizabeth C Hamilton, Grant E O’Keefe, Robert V Rege, Rawson J Valentine, Daniel J Jones, Seifu Tesfay, and Erwin R Thal. 2003. Effect of sleep deprivation on the performance of simulated laparoscopic surgical skill. *The American Journal of Surgery* 186, 2 (2003), 169 – 174. https://doi.org/10.1016/S0002-9610(03)00183-1

[9] Nima Enayati, Allison M. Okamura, Andrea Mariani, Edoardo Pellegrini, Margaret M. Coad, Giancarlo Ferrigno, and Elena De Momi. 2018. Robotic Assistance-as-Needed for Enhanced Visuomotor Learning in Surgical Robotics Training: An Experimental Study. In *Proc. of IEEE International Conference on Robotics and Automation (ICRA’18)*. 6631–6636.

[10] Paul M Fitts and Michael I Posner. 1967. *Human performance*. Brooks/Cole Publishing Company.

[11] Peter Kazanzides, Zihan Chen, Anton Deguet, Gregory S Fischer, Russell H Taylor, and Simon P DiMaio. 2014. An open-source research kit for the da Vinci® Surgical System. In *Proc. of IEEE International Conference on Robotics and Automation (ICRA’14)*. 6434–6439.

[12] Zhaoshuo Li, Irene Tong, Leo Metcalf, Craig Hennessey, and Septimiu E. Salcudean. 2018. Free Head Movement Eye Gaze Contingent Ultrasound Interfaces for the da Vinci Surgical System. *IEEE Robotics and Automation Letters* 3, 3 (July 2018), 2137–2143. https://doi.org/10.1109/LRA.2018.2809512

[13] Martin A Makary and Michael Daniel. 2016. Medical error-the third leading cause of death in the US. *The British Medical Jouranl (BMJ)* 353 (2016), i2139.

[14] Nabeel Merali, Darmarajah Veeramootoo, and Sukhpal Singh. 2019. Eye-Tracking Technology in Surgical Training. *Journal of Investigative Surgery* 32, 7 (2019), 587–593.

[15] Andrea Moglia, Vincenzo Ferrari, Luca Morelli, Mauro Ferrari, Franco Mosca, and Alfred Cuschieri. 2016. A systematic review of virtual reality simulators for robot-assisted surgery. *European urology* 69, 6 (2016), 1065–1080.

[16] Reena Pattani, Peter E Wu, and Irfan A Dhalla. 2014. Resident duty hours in Canada: past, present and future. *Canadian Medical Association Journal* 186, 10 (2014), 761–765.

[17] ROSS [n.d.]. *The Robotics Surgical Simulator (ROSS)*. Retrieved January 14, 2020 from http://www.simulatedsurgicals.com/ross2.html

[18] Mahya Shahbazi, S Farokh Atashzar, and Rajni V Patel. 2013. A dual-user teleoperated system with virtual fixtures for robotic surgical training. In *Proc. IEEE International Conference on Robotics and Automation (ICRA’13)*. 3639–3644.

[19] Roger Smith, Mireille Truong, and Manuela Perez. 2015. Comparative analysis of the functionality of simulators of the da Vinci surgical robot. *Surgical endoscopy* 29, 4 (2015), 972–983.

[20] NJ Taffinder, IC McManus, Y Gul, RCG Russell, and A Darzi. 1998. Effect of sleep deprivation on surgeons’ dexterity on laparoscopy simulator. *The lancet* 352, 9135 (1998), 1191.

[21] Irene Tong, Omid Mohareri, Samuel Tatasurya, Craig Hennessey, and Septimiu Salcudean. 2015. A retrofit eye gaze tracker for the da Vinci and its integration in task execution using the da Vinci Research Kit. In *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS’15)*. 2043–2050.

[22] MG Tytherleigh, TS Bhatti, RM Watkins, and DC Wilkins. 2001. The assessment of surgical skills and a simple knot-tying exercise. *Annals of the Royal College of Surgeons of England* 83, 1 (2001), 69.

[23] Samuel J Vine, Rich SW Masters, John S McGrath, Elizabeth Bright, and Mark R Wilson. 2012. Cheating experience: Guiding novices to adopt the gaze strategies of experts expedites the learning of technical laparoscopic skills. *Surgery* 152, 1 (2012), 32–40.

[24] Samuel J Vine, Lee J Moore, and Mark R Wilson. 2011. Quiet eye training facilitates competitive putting performance in elite golfers. *Frontiers in psychology* 2 (2011).

[25] Samuel J Vine and Mark R Wilson. 2011. The influence of quiet eye training and pressure on attention and visuo-motor control. *Acta psychologica* 136, 3 (2011), 340–346.

[26] Johanna I Westbrook, Magdalena Z Raban, Scott R Walter, and Heather Douglas. 2018. Task errors by emergency physicians are associated with interruptions, multitasking, fatigue and working memory capacity: a prospective, direct observation study. *BMJ Qual Saf* 27, 8 (2018), 655–663.

[27] Greg Wood and Mark R Wilson. 2011. Quiet-eye training for soccer penalty kicks. *Cognitive Processing* 12, 3 (2011), 257–266.