International Journal of Computer Assisted Radiology and Surgery
https://doi.org/10.1007/s11548-020-02176-1

ORIGINAL ARTICLE

A multi-camera, multi-view system for training and skill assessment for robot-assisted surgery

Alaa Eldin Abdelaal[1] · Apeksha Avinash[1] · Megha Kalia[1] · Gregory D. Hager[2] · Septimiu E. Salcudean[1]

Received: 16 November 2019 / Accepted: 21 April 2020

Abstract

Purpose This paper introduces the concept of using an additional intracorporeal camera for the specific goal of training and skill assessment and explores the benefits of such an approach. This additional camera can provide an additional view of the surgical scene, and we hypothesize that this additional view would improve surgical training and skill assessment in robot-assisted surgery.

Methods We developed a multi-camera, multi-view system, and we conducted two user studies ($N = 12$) to evaluate its effectiveness for training and skill assessment. In the training user study, subjects were divided into two groups: a single-view group and a dual-view group. The skill assessment study was a within-subject study, in which every subject was shown single- and dual view recorded videos of a surgical training task, and the goal was to count the number of errors committed in each video.

Results The results show the effectiveness of using an additional intracorporeal camera view for training and skill assessment. The benefits of this view are modest for skill assessment as it improves the assessment accuracy by approximately 9%. For training, the additional camera view is clearly more effective. Indeed, the dual-view group is 57% more accurate than the single-view group in a retention test. In addition, the dual-view group is 35% more accurate and 25% faster than the single-view group in a transfer test.

Conclusion A multi-camera, multi-view system has the potential to significantly improve training and moderately improve skill assessment in robot-assisted surgery. One application of our work is to include an additional camera view in existing virtual reality surgical training simulators to realize its benefits in training. The views from the additional intracorporeal camera can also be used to improve on existing surgical skill assessment criteria used in training systems for robot-assisted surgery.

Keywords Surgical training · Surgical skill assessment · Multi-camera system · Surgical robotics · Robot-assisted surgery · Minimally invasive surgery

Introduction

Video-based methods have been successfully used for training and skill assessment in many domains such as sports [30] and education [22]. The main philosophy in such methods is to monitor, then evaluate the professional activity, with the goal of modifying it to improve future performance in what is called monitor-evaluate-modify cycle. Previous research in psychology has shown the benefits of using this cycle for training and coaching [5].

In surgery, video-based methods have been used as part of the training process for novice surgeons. For example, Causer et al. [6] used videos of the eye gaze behavior of both experts and trainees as part of their proposed training protocol to help novices acquire the gaze patterns of experts. Jowett et al. [12] show the effectiveness of computer-based video training in a one-handed knot tying task in open surgery setting. More recently, Soucisse et al. [24] show that video-based training is a time-efficient method for teaching novices intestinal anastomosis. In minimally invasive surgery, Singh et al. [23] show that video-based coaching improves the qual-

✉ Alaa Eldin Abdelaal
aabdelaal@ece.ubc.ca

1 Electrical and Computer Engineering Department, University of British Columbia, Vancouver V6T 1Z4, Canada

2 Department of Computer Science, Johns Hopkins University, Baltimore, MD 21218, USA

ity of the trainees' performance on a virtual reality (VR) simulator for laparoscopic surgery. In real-world inanimate tasks using the da Vinci surgical robot, recorded motion data, along with videos of experts' performances, have been shown to be effective as well [1]. All of the above works use videos that show only a single view of the surgical scene.

Multi-camera systems have also been used for surgical training and skill assessment. In [29], Wang et al. show that the use of multiple cameras is significantly better than using only one for teaching residents hip surgery in an open surgery setting. Islam et al. [10] propose a multi-camera web-based skill assessment system for conventional laparoscopic surgery. Their system takes one view of the surgical scene and a view from each of the trainee's hands as inputs. Then, they extract features from these views and train a machine learning classifier to classify the surgical skill. Moreover, Goldenberg et al. [7] develop the OR Black Box system which collects several types of data in the operating room such as video, audio and physiological parameters in a synchronized way. The videos are collected from multiple cameras including one intracorporeal camera only in addition to other cameras recording other footage in the operating room. The collected data are then analyzed to identify intraoperative errors and assess the surgical skill [13].

In the context of robot-assisted surgery, several groups designed additional cameras that can be used inside the patient's body in addition to the original endoscope. Velasquez et al. [28] designed a tele-manipulated 2D imaging probe that can be mounted on any of the existing tools of the da Vinci surgical system. The probe is controlled by a custom-made pedal system that does not require any modifications to the da Vinci system's hardware. Our group also designed a pick-up stereoscopic camera probe with special mating interface for easy repeatable grasping by one of the da Vinci system's tools [4]. This camera probe can be inserted through a conventional trocar, requiring an additional trocar into the patient, or, alternatively, adjacent to one of the existing trocars. In the latter case, the incision is made by the trocar, which is then pulled out, the camera probe is inserted, and then the trocar is re-inserted. The tissue around the incision is enlarged by one camera cable diameter. In either case, once the camera is in the body, it can be grasped, due to its specially designed mounting surface [20], by the most commonly used da Vinci instrument, the ProGrasp forceps. The camera can be positioned by one of the da Vinci's arms and moved as required, just as easily as re-positioning the da Vinci endoscope. It is also important to note that the idea of providing additional views from additional cameras has been tested in vivo in laparoscopic cholecystectomy as in [15]. All these cameras, however, were proposed to overcome occlusions in real surgeries, and they were not used in surgical training or surgical skill assessment.

In contrast to all the above work, this paper explores the use of two cameras and two views inside the human body for surgical training and skill assessment in robot-assisted surgery. This is unlike the system in [10], which had only one view of the surgical scene (from the laparoscope) and the remaining views were for the hands of the user (from cameras outside the patient). The system proposed in [29] used eight cameras outside the patient's body and was designed for an open surgery setting; therefore, it is not applicable in robot-assisted surgery. Furthermore, the system was designed for training only; its use for skill assessment was not considered in this work.

From a human–computer interaction perspective, while a second view provides additional information, it also introduces additional workload on the user because of the need to switch focus back and forth between two different sources of information that could degrade his/her overall performance. This motivates the need to design, implement and quantitatively evaluate our system in surgical training and skill assessment. In particular, the contributions of this paper are as follows:

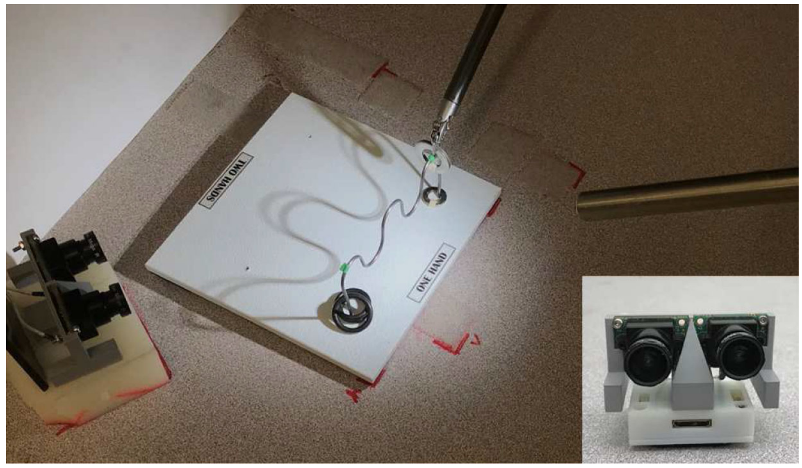
- We design and present a multi-camera, multi-view system that can show multiple views simultaneously at the surgical console in robot-assisted surgery. Our system makes use of two cameras: the first one is the surgical robot's endoscope and the other is a miniature camera that can be inserted inside the patient's abdomen, similarly to the concepts proposed in [4] and [28].
- We propose a novel use of our system in the context of surgical training and skill assessment in robot-assisted surgery.
- We evaluate the system in these two applications by conducting two user studies comparing it with the current case of using a single view of the surgical scene.

Materials and methods

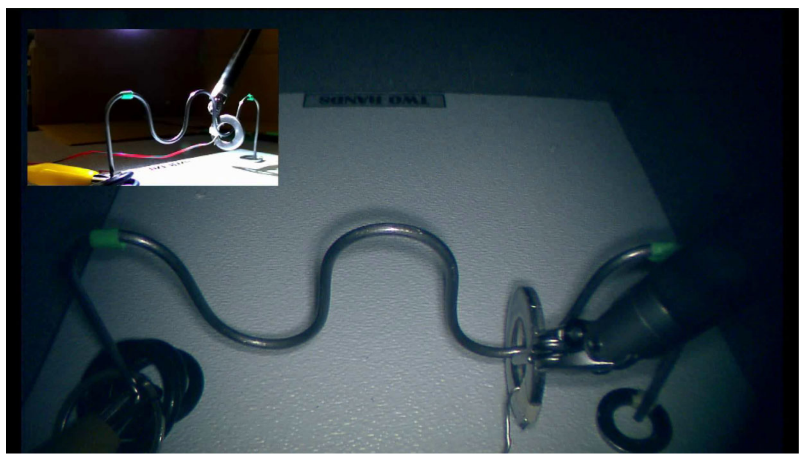
Multi-camera system

The training method used in this work is based on providing visual feedback to trainees after each training trial. This visual feedback is in the form of playing back the trainee's performance multiple times for him/her after each training trial to use it to improve future performance. These videos are recorded from two different perspectives using two cameras to show more useful information to the trainees. The training is conducted on a da Vinci S surgical system, and the videos are shown on its console to provide a more immersive training environment. We note that our training method differs from [29] where trainees watched a video of an expert

Fig. 1 The two-camera setup used in this work along with the views of the task from the two cameras



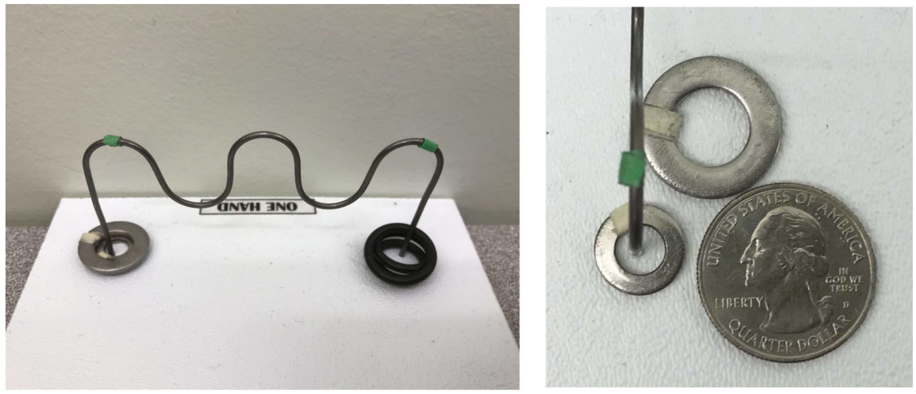
(a) The two-camera setup used in this work showing the da Vinci endoscope on the left and the additional camera on the right. A closer view of the additional camera used in the user studies is shown in the bottom right corner.



(b) The two views used for training in this paper. The main one is from the robot's endoscope and the secondary one, shown in the corner, is from the additional camera.

performance and then performed the task with no further viewing or feedback.

The two cameras used for our training system are the da Vinci S endoscope and the Leopard Imaging LI-OV580-STEREO cameras as shown in Fig. 1a. While the ultimate goal of this work is to use a miniature camera that can be inserted in the human body (as in our previous work in [4]), we use larger cameras in this work as they provide better imaging quality than the smaller ones available on the market. We chose these cameras to reduce the impact of the resolution limitations present in the early prototypes of the cameras in [4] or [28]. The BlackMagic DeckLink card is used to grab frames from the da Vinci endoscope and OpenCV is used to grab frames from the Leopard Imaging cameras. The video frames are recorded in synchronized fashion using a Pearl Epiphan recording device [16]. Using our system, the user can use the endoscope view during the task execution as usual. Either a single view (from the endoscope alone) or a combined view (from the endoscope and the Leopard cameras) is played back to the user for training. The BlackMagic DeckLink card is again utilized to playback the recorded videos onto the da Vinci console. OpenGL is used to draw the smaller Leopard Imaging camera images onto the endoscopic view when showing two views on the surgical console. All code is written in C++ with Visual Studio 2015 and OpenCV 3.4.



(a) The ring and rail task.

(b) The two metal rings used in the user study.

Fig. 2 The task used in the user studies

Task

We used the "ring and rail" task in our studies, which is part of a validated training curriculum in laparoscopic surgery [21]. This task is also one of the validated training tasks used in the ROSS simulator for robot-assisted surgery training [25]. *Face, construct, content and concurrent validity studies have been conducted showing that good performance in this task is correlated with good performance in the operating room* [2,11,17]. The task has also been validated for multiple disciplines in surgery including urology, gynecology and general surgery [3]. The original task is to pick up rubber rings with the surgical tool and move through a curved rail/wire. This task measures manual dexterity, fine motor skills, spatial awareness and hand-eye coordination of trainees. We used the one-handed version of this task as shown in Fig. 2a. Moreover, we made some modifications to it to suit our needs. We used metal rings instead of the original rubber ones. In addition to measuring the speed of task execution, we were also interested in measuring the precision of the movements as well. That is why we asked our trainees to avoid touching the rail with the ring and/or surgical tools. This is consistent with the fact that the speed and accuracy of the task execution are both important aspects used in the above curricula to measure the trainees performance. Therefore, our choice of the performance metrics is well justified.

The ring and rail task is a good candidate for multi-camera, multi-view systems. With the requirement of moving precisely without touching the rail, two cameras can be helpful. One of the cameras shows the face view of the rail itself in a position as recommended in the original ring and rail task as in [21], and the second camera shows its side view as seen in Fig. 1b. The first view is used to monitor the movement of the ring along the path between the start and end locations, and the second view helps in making sure that this movement is precise, i.e., the ring is not touching the rail.

To generate ground truth data, an electric circuit was constructed. One end of the circuit was connected to one end of the metallic rail of the ring and rail task. The other end of the circuit was connected to the metal ring. If either the ring or the metallic tooltip holding it touched the rail, the circuit would be closed and an LED would light up. We positioned an ordinary webcam to monitor this LED. The video feed from this webcam was also recorded with the Epiphan device to maintain synchronization with the video feeds. A MATLAB script was used to monitor the number of times, and the total duration the LED was switched on. This recorded the number of times the subject made contact with the rail (either with the ring or the tooltip; both were considered as errors) and the contact time.

User studies

We conducted two user studies ($N = 12$), one to evaluate the proposed system in training and the other to evaluate it in skill assessment. All our subjects were right-handed university students with no or little experience with robot-assisted surgery. The training study was a between-subjects study for which we had two six-subject groups: the control group, which used the traditional single-view system, and the experimental group, which used the dual-view system. To facilitate the narrative, in the rest of this paper, we refer to the former group as the single-view group and to the latter as the dual-view group. The skill assessment study was a within-subject study, for which all the 12 subjects were exposed to two conditions. This type of user studies measures the effect of changing the study conditions on each user [19].

The first user study ($N = 12$) was to determine the effect of using two cameras during surgical training using the da Vinci S surgical system. After briefing the subjects about the study and experimental setup, subjects were offered five minutes to familiarize themselves with the da Vinci system by performing a peg transfer task [18]. We then showed them a video of the ring and rail task to familiarize them with it. After that, subjects were randomly assigned to either the single view or dual-view group. The study then had three phases. The first was the baseline phase when subjects performed the task once before going through any training at all. The second phase involved the actual training: subjects were asked to carry out the task nine times. We chose nine training trials because this is the maximum number of trials needed for novices to reach expert's proficiency level on the ring and rail task according to many performance metrics [8]. Regardless of their group, subjects were shown the endoscopic view only when they performed their training trials. During each training trial, we recorded a video of the task from our two cameras. After each training trial, each subject was shown the video of his/her performance on the da Vinci surgical console. The single-view group members were shown their performance from the endoscopic camera only, while the dual-view group members were shown their performance from the two cameras as a picture-in-picture view, having the endoscopic view as the primary one as shown in Fig. 1b. Showing these videos gave the subjects the opportunity to reflect on their performance and identify how to improve it. The third and last phase of the user study was the evaluation phase. In this phase, subjects were asked to perform the same training task one more time without any feedback at all. We refer to this task as the large ring evaluation task or task 1. After that, they were asked to carry out a more difficult version of the task using a smaller ring than the one used during the baseline and training phases as shown in Fig. 2b. We refer to this task as the small ring task or task 2. The goal of this last evaluation task was to measure if the acquired skills during the training are transferable to other more difficult and unseen cases. The use of different sizes of the ring in the ring and rail task is already part of the validated training curricula as in [21].

The second user study ($N = 12$) focused on the effect of using a dual-view system when assessing surgical skill. Unlike the first study, this one was a within subject study: each subject was exposed to the two conditions (single vs. dual view). Each subject watched six videos. The first three videos were either single-view videos or dual-view ones. The last three videos were from the complementary condition, i.e., if the first three videos were single-view videos, then the rest were dual-view ones and vice versa. We employed counter-balancing to account for any biases meaning that half of the subjects watched the single-view videos first and the other half watched the dual-view videos first. The three single-view videos showed the endoscopic view only, and the remaining three showed the views from the endoscope and additional camera as a picture-in-picture similar to the corresponding case in the first user study. The first two videos of each condition were used to train subjects how to process the visual information provided to them to count the number of errors. Hence, we were only interested in the subjects' performance in the last video in each condition.

Performance metrics

In the training user study, we used the completion time and error time as our performance metrics. The error time is defined by the sum of all the times the ring or the surgical tool touches the rail. The lower the error time the better. We used the error circuit described in "Task" section to compute this metric. In the assessment study, the metric was the absolute difference between the ground truth and the estimated number of errors by the subjects. The lower the value the better. We used the same error circuit to compute the ground truth.

Results

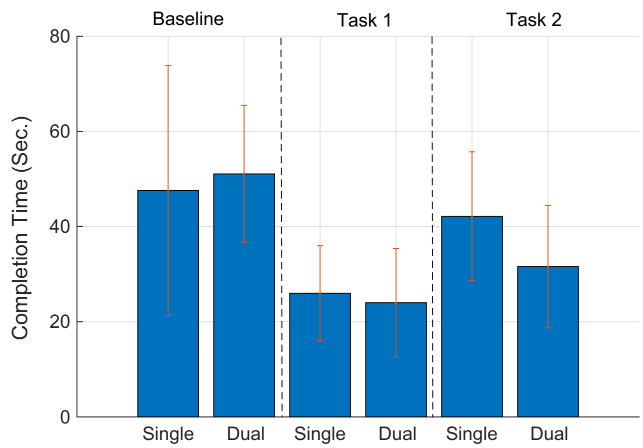
Overall, the results of the two user studies show that using two views is better than using the endoscopic view alone in all the metrics and cases for both training and skill assessment. In addition, for the training study, results show improvement in performance after training regardless of the training method used.

Figure 3a shows the results of the training study in terms of completion time. In task 1 (the large ring evaluation task), the dual-view group performance (23.98 ± 11.44 s) is slightly better than the single-view group performance (26.01 ± 9.95 s), by approximately 8% on average. The dual camera group show the best improvement compared with their baseline performance by taking 53% less time on average in task 1 (the large ring evaluation task) compared with an improvement of approximately 45% on average for the single camera view group. In task 2 (evaluation task with a smaller ring), the dual-view group performance (31.58 ± 12.88 s) is better than the single-view group (42.19 ± 13.52 s) by approximately 25% on average.

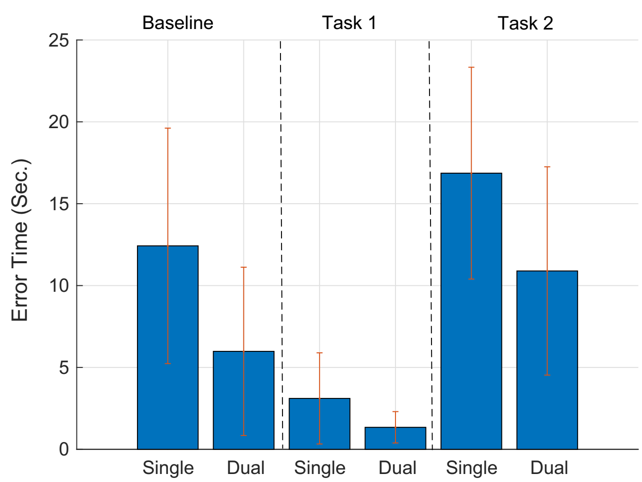
In terms of the error time measure, as shown in Fig. 3b, the dual-view group performance (1.34 ± 0.96 s) is better than the single-view group (3.11 ± 2.79 s) by approximately 57% in task 1 after training (the large ring evaluation task). Compared with their baseline performances, the two groups show comparable improvements, with the dual-view group's improvement (approximately 78%) being slightly better than the single-view group's improvement (approximately 75%). In task 2, (evaluation task with a smaller ring), the dual-view group performance (10.89 ± 6.36 s) is better than the single-view group (16.86 ± 6.47 s) by approximately 35% on average.

We use the NASA Task Load Index (or NASA TLX) assessment tool [9] to subjectively evaluate the use of two cameras for training compared with using only one. Subjects are asked to rate their training method on a 21-point

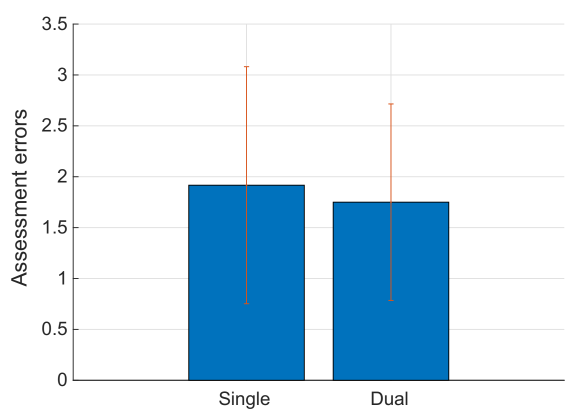
Fig. 3 The results of the two user studies



(a) Training study: The completion time for the single and dual view groups during their baseline task, the large ring evaluation task after training (task 1) and the testing task with smaller ring (task 2).



(b) Training study: The error time for the single and dual view groups during their baseline task, the large ring evaluation task after training (task 1) and the testing task with smaller ring (task 2).



(c) Skill assessment study: The assessment errors in the single and dual view conditions.

scale in terms of six measures. The lower the ratings, the better. The only exception is the performance measure, for which a higher rating is better. Subjects in the dual-view group give a higher rating to their training method than the single-view group's ratings in all but one of the measures of the NASA TLX tool. The only exception is the temporal demand measure. The three important measures in our case are the mental demand, effort and performance measures. Subjects in the dual-view group rate their method better than the other group's ratings by approximately 23% and 26% in mental demand and effort, respectively. The two groups give comparable rating to their corresponding method for the performance measure of the NASA TLX, with the single-view group's rating being slightly better by approximately 8%.

In the assessment study, all subjects were exposed to single- and dual-view videos. Their results are comparable in the two cases, with their assessment scores being slightly more accurate (by approximately 9%) in the dual-view case than their scores in the single-view one as shown in Fig. 3c.

The standard deviation values (shown as red lines in Fig. 3) in task 1 and task 2 in the training results along with the skill assessment results are comparable in the dual- and single-view groups/conditions and the major differences are in the average values. This means that while the differences in the individual performances in the two groups are comparable, the performance of the dual-view group as a whole is better than the single-view group.

We perform hypothesis testing using a t test on the results of the two studies. For the training study, there is no statistical significance between the baseline performance of the two groups according to the completion time and error time. In addition, we perform a t test on the results before and after training for each of the two groups in terms of our performance metrics. For the dual-view group, we find that there is statistical significance in terms of completion time ($p < 0.01$) and marginal statistical significance ($p = 0.0549$) in terms of error time. For the single-view group, there is only statistical significance in terms of error time ($p < 0.05$). For the skill assessment study, there is no statistical significance between the two conditions.

Discussion

This paper makes the case for using an additional intracorporeal camera for training and skill assessment in robot-assisted surgery. The results of the two user studies we designed and conducted show that using this additional camera along with the endoscopic camera of the surgical robot improve both training and skill assessment. While the improvement in skill assessment is modest, the use of the two cameras shows the best potential in training as evident by the large improvements after training in terms of completion time and accuracy. Using two cameras for training also leads to better performance on unseen and more difficult tasks compared with using the endoscope alone, which shows the transferability of this training method. Moreover, subjective evaluation shows that subjects of the dual-view group achieve these improvements with less mental demand and effort compared with the single-view group.

The single-view group rated their own performance better than the dual-view group in the NASA TLX performance measure. Their performance according to the objective measures (completion and error times) was, however, worse than the dual-view group. It is possible that having been exposed to two different views, the dual-view group could utilize the additional view to identify errors that are hard to see in the single endoscopic view. With this information in hand, it is likely that they grasped the difficulty of the task better than the single-view group and hence had a harsher judgment of their own performance. In contrast, oblivious to errors made that are hidden from the single endoscopic view, the single-view group rated their own performance highly.

In the training user study, we designed our training intervention so that all trainees regardless of their group (single or dual view) would carry out the task while viewing the endoscopic view alone as it is the most common case in actual surgery. The dual-view group watched a video of their own performance that shows both the endoscopic and additional camera views as a picture in picture only after each training trial. Moreover, in the transfer test with the smaller ring, all subjects were shown the endoscope view only and results show that the dual-view group was better according to all the performance metrics.

The feedback from the dual-view group shows that they used the additional camera view to perform root-cause analysis for their committed errors to improve their performance. For instance, subjects highlighted that the playback of their performances helped them figure out that some of the errors they made at the end of the rail were because of the way they hold the surgical tool at the very beginning of it. This is a sign of the potential of using our system for root-cause analysis even if it is not intentionally designed to do so.

We chose $N = 12$ for the number of subjects in this study because in a preliminary study, the results indicated that $N = 12$ would provide us with statistical significance in all the comparisons performed based on an a priori power analysis. It is possible that the subjects in the preliminary study ($N = 4$) were better than average. A larger sample size can lead to more statistically significant results.

For the skill assessment study, the "number of errors" metric was used as a proxy for how a user would handle tissue in real surgery. This number can represent, for example, the number of undesired collisions between the tools and tissue. We argue that having an additional camera would help in identifying such collisions better than using the endoscope alone.

In real surgery, surgical skill assessment questionnaires ask the assessors to evaluate such collisions. For example, the Objective Structured Assessment of Technical Skills (OSATS) tool has two relevant criteria namely "respect for tissue" and "handling of endovascular material" [26] and our multi-camera system can enable the assessor to either identify violations that are not visible in the endoscopic view or increase his/her confidence in the ratings in general. A good analogy for this is the use of Video-Assisted Referees (VAR) technology in soccer/football [27] that helps referees in the critical decisions in football matches.

The actual use of our system for skill assessment in real surgery in the absence of the ground truth is the same as the normal video-based skill assessment methods using tools such as the OSATS. More than one assessor will be involved in the process, and the final score will be the combination/average of each assessor's scores. More assessors can be included if there are significant differences in the original assessors' scores.

Another potential future use of our multi-camera system is in computer-based methods for skill assessment without human intervention using computer vision and machine learning methods. The additional view provided by our system can help in overcoming occlusions in the endoscopic view and may help improve the skill assessment accuracy of such systems.

Conclusion and future work

In this paper, we introduced the use of an additional intracorporeal camera for training and skill assessment in robot-assisted surgery. We developed a multi-camera, multi-view system that is capable of showing two views (or more) of the surgical scene simultaneously to the user. We used this system to record users' performance on a surgical training task from two perspectives. We conducted two user studies ($N = 12$) to evaluate the effect of the proposed system on training and skill assessment compared with the traditional case of using only a single view. Our results showed that a two-view system can improve both training and skill assessment with its application in training showing the most promise for improving both the speed and accuracy of the task execution after training. Our results also showed that skills acquired from training with the proposed system are more transferable in a transfer test than those acquired from training with the traditional single-view system. Subjective evaluation showed that subjects trained with the proposed system needed less mental demand and effort to achieve their performances.

We believe that the results of this work open up many avenues for future work in terms of system design and evaluation. For instance, conducting user studies with larger number of subjects, preferably surgical residents, is needed to strengthen our results. It would be interesting to study and quantify the effect of having an additional view on the cognitive load of trainees and assessors. In addition, it is important to study where to place the additional camera view on the surgical console to maximize its benefits. Another interesting direction is to study how frequently in our setting, trainees and assessors use the information of the additional view. Eye gaze tracking can provide cues to answer this question and help improve the interface design. Moreover, studying other effects such as the camera position, camera focus and zoom level on the results is another interesting venue of research.

There are many potential practical implications of our system. For example, our results can be useful for improving current virtual reality simulators for surgical training by adding a simulated additional view to the main one that is currently being used. Moreover, another interesting application is the inclusion of an additional intracorporeal camera in systems such as the OR Black Box system [7] and studying its effect in training, error analysis and skill assessment. Furthermore, the existence of an additional camera for training and skill assessment can lead to modifications in some of the assessment tools currently being used such as the Objective Structured Assessment of Technical Skills (OSATS) [14]. These modifications can employ the additional visual information available by the additional camera to improve the assessment criteria.

Acknowledgements This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (Discovery Grant), in part by the Canada Foundation for Innovation (infrastructure and operating funds), in part by Intuitive Surgical (equipment donation), in part by the C.A. Laszlo Chair in Biomedical Engineering held by Prof. Salcudean, and in part by the Vanier Canada Graduate Scholarship held by Alaa Eldin Abdelaal. We would like also to thank Keith Tsang for his assistance in this work.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the University of British Columbia Research Ethics Board (Study Number H17-03100) and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Abdelaal AE, Sakr M, Avinash A, Mohammed SK, Bajwa AK, Sahni M, Hor S, Fels S, Salcudean SE (2018) Play me back: a unified training platform for robotic and laparoscopic surgery. IEEE Robot Autom Lett 4(2):554–561
2. Aghazadeh MA, Mercado MA, Pan MM, Miles BJ, Goh AC (2016) Performance of robotic simulated skills tasks is positively associated with clinical robotic surgical performance. BJU Int 118(3):475–481
3. Alzahrani T, Haddad R, Alkhayal A, Delisle J, Drudi L, Gotlieb W, Fraser S, Bergman S, Bladou F, Andonian S, Anidjar M (2013) Validation of the da vinci surgical skill simulator across three surgical disciplines: a pilot study. Can Urol Assoc J 7(7–8):E520
4. Avinash A, Abdelaal AE, Mathur P, Salcudean SE (2019) A pickup stereoscopic camera with visual-motor aligned control for the da Vinci surgical system: a preliminary study. Int J Comput Assist Radiol Surg 14:1197–1206
5. Carver CS, Scheier MF (2001) On the self-regulation of behavior. Cambridge University Press, Cambridge
6. Causer J, Vickers JN, Snelgrove R, Arsenault G, Harvey A (2014) Performing under pressure: quiet eye training improves surgical knot-tying performance. Surgery 156(5):1089–1096
7. Goldenberg MG, Jung J, Grantcharov TP (2017) Using data to enhance performance and improve quality and safety in surgery. JAMA Surg 152(10):972–973
8. Hardon SF, Horeman T, Bonjer HJ, Meijerink WJ (2018) Force-based learning curve tracking in fundamental laparoscopic skills training. Surg Endosc 32(8):3609–3621
9. Hart SG, Staveland LE (1988) Development of NASA-TLX (task load index): results of empirical and theoretical research. In: Advances in psychology, vol 52. Elsevier, pp 139–183
10. Islam G, Kahol K, Li B, Smith M, Patel VL (2016) Affordable, web-based surgical skill training and evaluation tool. J Biomed Inform 59:102–114
11. Jarc AM, Curet M (2015) Face, content, and construct validity of four, inanimate training exercises using the da vinci® si surgical system configured with single-site instrumentation. Surg Endosc 29(8):2298–2304
12. Jowett N, LeBlanc V, Xeroulis G, MacRae H, Dubrowski A (2007) Surgical skill acquisition with self-directed practice using computer-based video training. Am J Surg 193(2):237–242
13. Jung JJ, Jüni P, Lebovic G, Grantcharov T (2018) First-year analysis of the operating room black box study. Ann Surg 271(1):122–127. https://doi.org/10.1097/SLA.0000000000002863
14. Martin J, Regehr G, Reznick R, Macrae H, Murnaghan J, Hutchison C, Brown M (1997) Objective structured assessment of technical skill (osats) for surgical residents. Br J Surg 84(2):273–278
15. Oleynikov D, Rentschler M, Hadzialic A, Dumpert J, Platt S, Farritor S (2005) Miniature robots can assist in laparoscopic cholecystectomy. Surg Endosc 4(19):473–476
16. Pearl Epiphan Specification (2017). https://www.epiphan.com/products/pearl/tech-specs/. Accessed 15 Nov 2019
17. Perrenot C, Perez M, Tran N, Jehl JP, Felblinger J, Bresler L, Hubert J (2012) The virtual reality simulator dv-trainer® is a valid assessment tool for robotic surgical skills. Surg Endosc 26(9):2587–2593
18. Peters JH, Fried GM, Swanstrom LL, Soper NJ, Sillin LF, Schirmer B, Hoffman K (2004) Sages FLS committee: development and validation of a comprehensive program of education and assessment of the basic fundamentals of laparoscopic surgery. Surgery 135(1):21–27
19. Salkind NJ (2010) Encyclopedia of research design, vol 1. Sage, Thousand Oaks
20. Schneider C, Nguan C, Rohling R, Salcudean S (2015) Tracked pick-up ultrasound for robot-assisted minimally invasive surgery. IEEE Trans Biomed Eng 63(2):260–268
21. Schreuder H, Van Den Berg C, Hazebroek E, Verheijen R, Schijven M (2011) Laparoscopic skills training using inexpensive box trainers: which exercises to choose when constructing a validated training course. BJOG Int J Obstet Gynaecol 118(13):1576–1584
22. Sherin M, van Es E (2005) Using video to support teachers ability to notice classroom interactions. J Technol Teach Educ 13(3):475–491
23. Singh P, Aggarwal R, Tahir M, Pucher PH, Darzi A (2015) A randomized controlled study to evaluate the role of video-based coaching in training laparoscopic skills. Ann Surg 261(5):862–869
24. Soucisse ML, Boulva K, Sideris L, Drolet P, Morin M, Dubé P (2017) Video coaching as an efficient teaching method for surgical residents—a randomized controlled trial. J Surg Educ 74(2):365–371
25. The Robotics Surgical Simulator (ROSS). http://www.simulatedsurgicals.com/ross2.html. Accessed 15 Nov 2019
26. Van Herzeele I, Aggarwal R, Malik I, Gaines P, Hamady M, Darzi A, Cheshire N, Vermassen F (2009) European Virtual Reality Endovascular Research Team EVEResT: validation of video-based skill assessment in carotid artery stenting. Eur J Vasc Endovasc Surg 38(1):1–9
27. Video Assisted Referees (VAR). https://football-technology.fifa.com/en/media-tiles/video-assistant-referee-var. Accessed 15 Nov 2019
28. Velasquez CA, Navkar NV, Alsaied A, Balakrishnan S, Abinahed J, Al-Ansari AA, Yoon WJ (2016) Preliminary design of an actuated imaging probe for generation of additional visual cues in a robotic surgery. Surg Endosc 30(6):2641–2648
29. Wang H, Sugand K, Newman S, Jones G, Cobb J, Auvinet E (2019) Are multiple views superior to a single view when teaching hip surgery? A single-blinded randomized controlled trial of technical skill acquisition. PLoS ONE 14(1):e0209904
30. Wilson BD (2008) Development in video technology for coaching. Sports Technol 1(1):34–40