

Robotic Surgery Remote Mentoring via AR with 3D Scene Streaming and Hand Interaction

Yonghao Long*, Chengkun Li* and Qi Dou

Dept. of Computer Science and Engineering, The Chinese University of Hong Kong

ARTICLE HISTORY

Compiled September 13, 2022

ABSTRACT

With the growing popularity of robotic surgery, education becomes increasingly important and urgently needed for the sake of patient safety. However, experienced surgeons have limited accessibility due to their busy clinical schedule or working in a distant city, thus can hardly provide sufficient education resources for novices. Remote mentoring, as an effective way, can help solve this problem, but traditional methods are limited to plain text, audio, or 2D video, which are not intuitive nor vivid. Augmented reality (AR), a thriving technique being widely used for various education scenarios, is promising to offer new possibilities of visual experience and interactive teaching. In this paper, we propose a novel AR-based robotic surgery remote mentoring system with efficient 3D scene visualization and natural 3D hand interaction. Using a head-mounted display (i.e., HoloLens), the mentor can remotely monitor the procedure streamed from the trainee's operation side. The mentor can also provide feedback directly with hand gestures, which is in-turn transmitted to the trainee and viewed in surgical console as guidance. We comprehensively validate the system on both real surgery stereo videos and ex-vivo scenarios of common robotic training tasks (i.e., peg-transfer and suturing). Promising results are demonstrated regarding the fidelity of streamed scene visualization, the accuracy of feedback with hand interaction, and the low-latency of each component in the entire remote mentoring system. This work showcases the feasibility of leveraging AR technology for reliable, flexible and low-cost solutions to robotic surgical education, and holds great potential for clinical applications.

KEYWORDS

Robotic Surgery; Augmented Reality; Remote Mentoring

1. Introduction

Robotic surgery has been revolutionized in past decades (Navab (2021); D'Ettoire et al. (2021)), greatly enhancing the surgeon's perception and operational ability in minimally invasive interventions. With dramatic rise of adoption of robotic surgery, education is important and requires urgent enhancement (Chen et al. (2021); Long et al. (2022)) for teaching the next generation of surgeons how to use robots to conduct a surgery step-by-step, instead of using traditional laparoscopic paradigms. Unfortunately, the novices always have limited access to experienced surgeons who have intensive clinical schedule or work in a distant city, making the surgical education cur-

*Authors contributed equally to this work.

Email: {yhlong,ckli,qdou}@cse.cuhk.edu.hk

rently challenging and expensive (Collins et al. (2021); Erridge et al. (2019); Huang et al. (2019)).

Remote mentoring, i.e., experts monitor surgical progress and provide guidance for trainees from a different geographical location, is envisaged as a promising way to reduce cost and facilitate surgical education. Hinata et al. (2014) employ video-audio communication and transfer handwritten notes with a touch tablet for remote mentoring. Their results present similar surgical outcomes (e.g., operating time, complication, early continence status, positive margin rate) between the remote and onsite mentoring groups. Mendez et al. (2005) propose to transmit video, audio communication and surgical instructions with electronic stylus and pad, for a 400km long-distance mentoring in neurosurgery. Experiments validate its safety and trustworthiness. However, common shortcomings of these methods are that, the mentoring and monitoring information are conveyed through text, voice, image and 2D video, which are not intuitive nor vivid for perceiving rich circumstances of surgery and provide comprehensive feedback by mentors.

Recently, augmented reality (AR) brings new possibility to remote mentoring along with many benefits (Yoon et al. (2018)). AR can provide an immersive way of displaying the 3D information (e.g., surgical scene and tool movement) and allow easier and natural interactions (Kovoor et al. (2021); Qian et al. (2019a); Wang et al. (2022)). Some initial investigations exist with encouraging progress. For examples, Jarc et al. (2017) overlay ghost tools for guidance within the view of trainee’s console; the mentor sees the surgical scene with a 3D display and glass, and controls the ghost tools through an extra handle device. Shabir et al. (2021) propose a robotic surgery remote mentoring system, which uses haptic devices to control virtual tools, but with a high latency observed even in a local area network. Though much benefits have been demonstrated, existing solutions still cannot achieve high-fidelity streaming of 3D scenes via AR, which constrains both efficiency and quality of remote mentoring. In addition, the requirement of extra interaction devices would unavoidably lead to inconvenience of education. In this paper, to address current limitations, we design a novel robotic surgery remote mentoring system relying on AR with 3D surgical scene streaming and hand interaction. Specifically, we first design an efficient streaming scheme to transmit the 3D surgical scene from the operation side to mentor side, using color and disparity images to reduce transmission cost. Then, we allow the mentor to use a portable head-mounted display to perceive 3D surgical scene in point cloud representation and give feedback guidance with hand gestures conveniently. Last, the instructions are returned to the trainee side and can be viewed in the console for straight-forward guidance. The system is validated on both real surgery video data and common ex-vivo scenarios of robotic training tasks, demonstrating promising results in terms of fidelity of streamed scene visualization, accuracy of hand interaction, and low-latency of transmission. To the best of our knowledge, this is the first completely AR-based remote mentoring system with only a head-mounted display for robotic surgical education. This new type of reliable, flexible and low-cost solution has great potential to be used in clinical practice.

2. Methods and Materials

The overview of our AR-based robotic surgery remote mentoring system is shown in Fig. 1. It consists of four components: operation side, surgical scene streaming, mentor side with hand interaction, guidance feedback for closed-loop education.

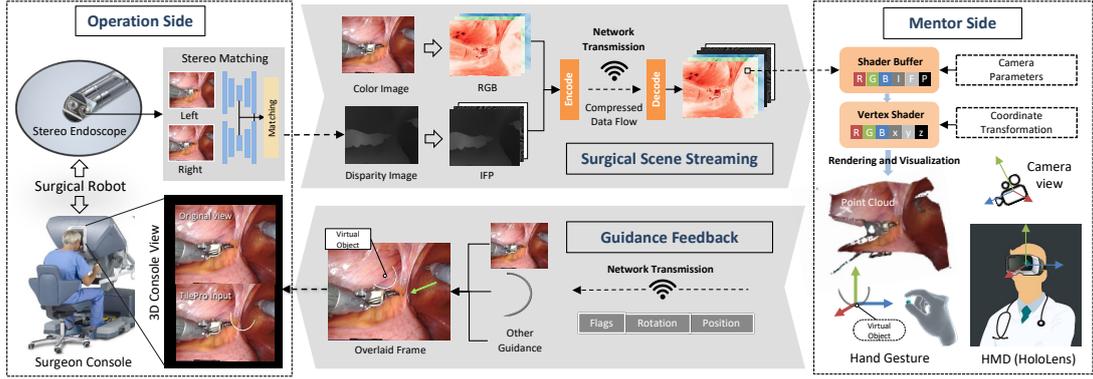


Figure 1. The overview of our proposed novel system for robotic surgery remote mentoring via augmented reality (AR) with 3D scene streaming and hand interaction in HoloLens.

Briefly speaking, in the operation side, a trainee conducts training tasks or a surgery at the console, which shows stereo surgical video, as well as the mentor’s guidance information displayed in AR via TilePro toolkit (Intuitive Surgical Inc). For surgical scene streaming, the RGB-D information of 3D scene is efficiently transmitted via high-speed internet from the operation side to the remote mentor side. Then, the mentor (i.e., an expert surgeon) uses a head-mounted AR glass to see the surgical scene in real-time and accordingly provide guidance with hand gestures. Finally, the mentor’s inputs with virtual objects are streamed back and visualized in trainee’s console as a closed-loop education system.

2.1. Efficient Surgical Scene Representation and Streaming

To represent the surgical scene, we opt for RGB-D as 3D representation in operation side, owing to the advantages that it can accurately describe relatively complex objects with finite points. It can also be easily obtained without heavy reconstruction computation compared to mesh, as well as can be straightforwardly and efficiently projected to point clouds for 3D visualization (Long et al. (2021); Wei et al. (2022)). Here, we first apply the HITNet (Tankovich et al. (2021)), a real-time deep learning method which can extract the high-level multi-scale context features from a stereo image pair (I_L, I_R) to calculate the correspondence for each pixel, and yield the disparity map D of the left image. These RGB and disparity images are leveraged for transmission, in order to save memory and keep up the streaming speed for fast system responses.

Different from previous work, we design a new compression operation to reduce the burden of wireless communication in-between mentor-operation sides. Qian et al. (2019b) suggests to transmit flattened RGB-D with three-bytes RGB color and one-byte disparity value. However, it would sacrifice quality due to compact representation which compresses a float type disparity value to one-byte type. Apart from this, the flatten array representation will also cause heavy network communication pressure as the frame resolution increasing. For instance, in the case of 640×512 resolution, it would cost $640 \times 512 \times (3 + 1) \times 8 \text{ bits} = 10 \text{ Mb}$ for every frame. Instead, our calculated disparity map (which is a float single channel image) is separated to two channels, with one denoting the integer part (I channel) of the disparity map, and the other for the fractional part (F channel). These two channels are then combined with an all-zero placeholder P channel to generate a 3-channel image to fulfill the format

requirement of the efficient “JPEG” compression algorithm, whose information loss and computational time will not exhibit exponential growth with larger data. Finally, the IFP image and corresponding RGB frame are both encoded as “JPEG” format and then combined to be our surgical scene streaming data flow. In this way, we can reduce the transmission data size to around 8 times smaller than previous work (Qian et al. (2019b)), i.e. 1.25 Mb for the resolution of 640×512 . The data flow is then sent to mentor side using the internet socket TCP/IP. Similarly, we do “JPEG” decoding in remote side to obtain RGB and IFP images for further processing.

2.2. HMD Visualization and Hand-based Natural Interaction

As mentor side is an expert who requires to setup and use the system as convenient as possible, we leverage optical see-through HMD to fulfill the requirement. Specifically, we first convert the IFP channels to disparity image and transfer it to a point cloud with the following projection equation:

$$z = f \cdot b \cdot \frac{1}{D(u, v)}, \quad x = z \cdot (u - c_x) \cdot \frac{1}{f}, \quad y = z \cdot (u - c_y) \cdot \frac{1}{f}, \quad (1)$$

where c_x, c_y are the camera centroid, f is focal length, and b is the baseline length of the stereo camera, which can all be obtained by camera calibration. With Eq.(1), the disparity vector $\mathbf{q} \in \mathbb{R}^3 = (u, v, D(u, v))^T$ is projected to a spacial point $\mathbf{p} \in \mathbb{R}^3 = (x, y, z)^T$. Given the color information from its corresponding RGB image I_L , we can obtain the point cloud representation of the surgical scene as $\mathbf{p} \in \mathbb{R}^6 = (x, y, z, r, g, b)^T$. Then, we leverage renderer’s shader to render and visualize the point cloud of the surgical scene. For correctly rendering the point cloud frame in the HMD view, it is important to perform coordinate transformations. In this regard, we propose the following scheme as:

$$\dot{\mathbf{p}}_H = \mathbf{K}_W^H \cdot \mathbf{K}_O^W \cdot \mathbf{K}_C^O \cdot \dot{\mathbf{p}}, \quad \mathbf{K} = \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{pmatrix}, \quad \dot{\mathbf{p}} = (\mathbf{p}^T | 1)^T, \quad (2)$$

where \mathbf{p}_H and \mathbf{p} are respectively the points in the HMD and camera coordinate frame, the dot notation means their corresponding homogenous vectors. \mathbf{K} is transformation matrix where \mathbf{R} and \mathbf{t} denote rotation and translation shift within the euclidean group $SE3 := \{\mathbf{R}, \mathbf{t} | \mathbf{R} \in \mathbb{SO}_3, \mathbf{t} \in \mathbb{R}^3\}$. For \mathbf{K} , We use different subscripts to indicate different coordinate system, $\{O\}$ denotes the rendering centroid which is defined by box boundary of the first point cloud frame, $\{C\}$ is the camera coordinate frame, $\{W\}$ is the real world coordinate system, $\{H\}$ is the HMD coordinate frame. For instance, \mathbf{K}_W^H indicates the transformation matrix from the real world to HMD coordinate system. In our case, \mathbf{K}_C^O is fixed and \mathbf{K}_W^H can be automatically obtained from HMD. Users can flexibly change the location and scale of the point cloud scene by \mathbf{K}_O^W for interaction.

Importantly, we design all the mentor’s interactions to be conducted purely by human hand which is highly user-friendly. Specifically, the cameras on HMD will capture multi-view images, compute the depth information, and reconstruct the environment. Then we employ built-in computer vision algorithms to analyze the spatial and pose information to recognize hand gestures, and locate the hand position w.r.t. the visu-

alized virtual scene to achieve a hand-based interaction. For example, we design to raise the hand for waking up the setting menu, move finger in the air for pointing to certain critical anatomy in the scene. Moreover, we create assisting virtual objects (e.g., needle) with AR to facilitate necessary demonstration provided by the mentor. We design CAD models of the virtual objects and import them into HMD. With our system, the mentor can manipulate the virtual object purely using a hand, just like naturally moving an object in the real-world. For example, the mentor can manipulate a virtual needle with a pinch to rotate and move, and draw lines as instructions of suturing trajectories. These interactions can be recorded in real-time as the mentor’s feedback.

2.3. Closed-loop Feedback for Guidance in Surgeon Console

To feedback the guidance information of the mentor, we record the hand interactions and virtual objects, and transmit them back to the operation side. Instead of directly sending rendered image with overlaid guidance information, we again design a light-weight data transfer approach to boost the efficiency. Specifically, we only send well-formatted location information to operation side, which can be the position of the virtual surgical tool, position of the draw trajectory, etc. The format of the feedback data is as $\{m, i, \alpha_C, \beta_C, \gamma_C, x_C, y_C, z_C\}$, where m is a flag to denote whether this information is available, i is used to decide which kinds of information is contained (such as different kinds of virtual object), the $(\alpha_C, \beta_C, \gamma_C)$ means the yaw, pitch, and roll angle in the camera coordinate frame respectively, the (x_C, y_C, z_C) represents the position of specified object. The feedback message is also sent to operating side using the TCP/IP Socket. After receiving the feedback information, operation side will decode the message and overlay related guidance information on the stereo video, and finally input to TilePro in the surgical robot console as augmented view to be perceived.

In this way, we build our system as an effective close-loop education paradigm. The mentor of expert surgeon can observe remote surgical procedures by simply wearing an AR glass. By manipulating virtual objects or indicating instructions with hands, the mentor can easily provide supportive advice to trainees in real-time. This new AR-based remote mentoring solution is distinct from traditional ones that use 2D video or haptic device. We have advantages of immersive visual experience, natural interaction, high-speed and low-cost, as well as great potential for richer information to be added on top of the augmented environment.

3. Experiments

3.1. System Setup and Experiment Design

For system setup, we use an in-house *da Vinci Si* surgical robot as the operation side, which is retired but generally similar to the robot used in clinic. Its stereo video is streamed to a computer equipping NVIDIA TITAN RTX GPU for computing 3D representation of the surgical scene. The mentor side uses a *Microsoft HoloLens 2*, the prevalent portable AR device, to visualize the 3D surgical scene and provide hand interaction with Unity toolkit. To realize remote surgical scene streaming and guidance feedback, we set up a local network communication environment using a low-cost WIFI router (TP-LINK AX3000). And the computer and HoloLens 2 are connected to the

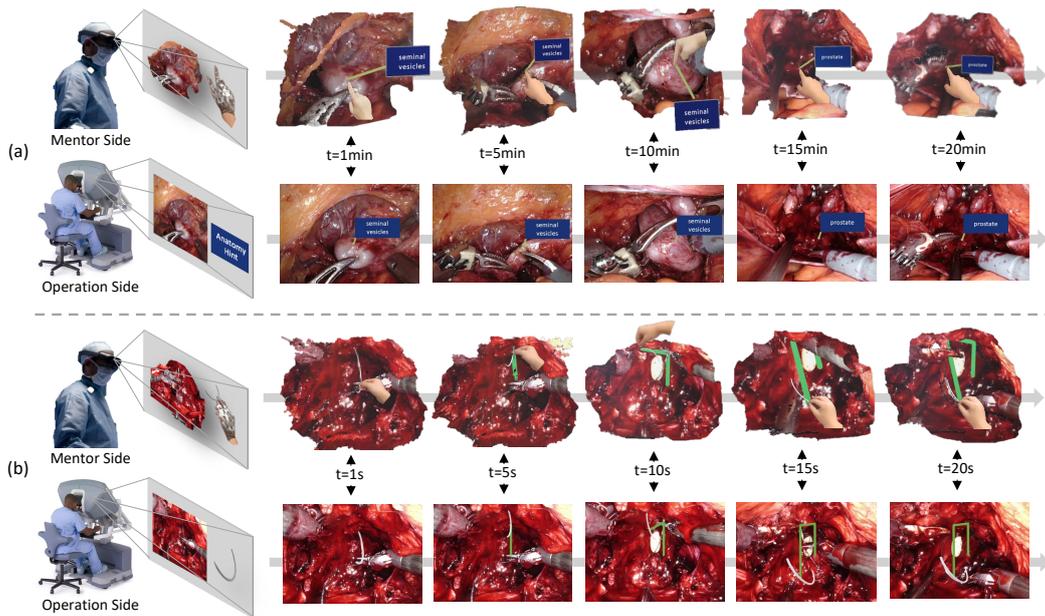


Figure 2. Case studies on real surgery data of prostatectomy. (a) Mentor uses the hand to point key anatomies in surgical scene which can be shown in the operation console for the trainee. (b) Mentor manipulates a virtual needle and draw suturing trajectories (green line) which can be demonstrated to the trainee. More details are in Sec. 3.2.

router using cable and wireless network respectively.

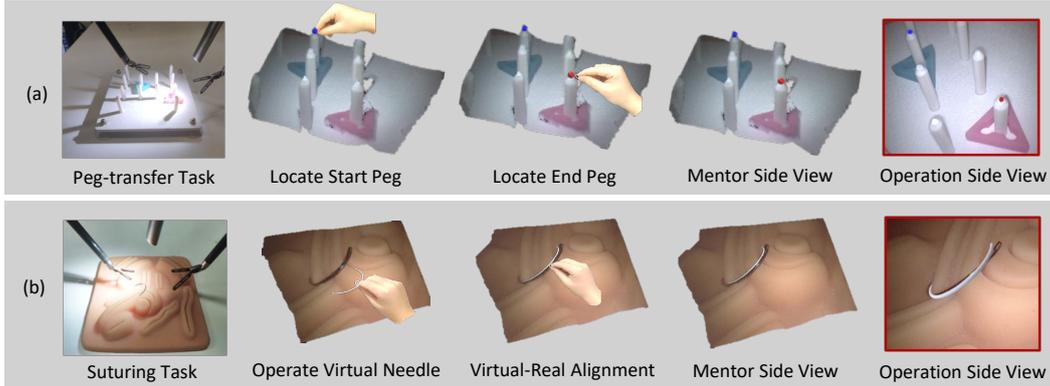
For experiment design, we extensively validate our proposed system on both real surgery stereo videos and ex-vivo scenarios of common robotic training tasks. For real surgery, we use two typical cases of DaVinci robotic surgery of prostatectomy procedure, with stereo videos captured in the resolution of 1280×1024 at $60fps$. To reduce computation cost of disparity map with HITNet (Tankovich et al. (2021)), we first downsample the resolution to 640×512 . We play the video in its original speed for simulation of the real clinical scenario with our system setup. The HoloLens can render the scene at the range of $25 \sim 30fps$ for the mentor side. For ex-vivo scenarios with *da Vinci Si*, we capture stereo videos with the same resolution and frame rate as the real surgery setting. We implement two scenarios, i.e., peg-transfer and suturing, which are common tasks for surgical skill training (Joseph et al. (2010)). In peg-transfer, the trainee picks up the block from one peg and moves it to another target peg. In suturing, the trainee manipulates the needle to suture the phantom wound. We use these ex-vivo experiments for quantitative evaluations of the interaction accuracy and remote mentoring latency of our system.

3.2. Feasibility Study on Real Surgical Scenes

To demonstrate the application scenario of our remote mentoring system in the real clinical environment, we show two case studies with the prostatectomy procedure. As depicted in Fig. 2(a), the surgeon is conducting the step of “dissection of the seminal vesicles and rectum” in prostatectomy. Inexperienced surgeons can meet difficulties to precisely identify the positions of key anatomies for dissection, due to fuzzy blood occlusion and poor lighting condition. With our system, the expert can monitor the

Table 1. The accuracy results for two ex-vivo scenarios.

Scenario	Peg-transfer	Suturing
Re-projection error (pixel)	7.35 ± 2.1	5.12 ± 1.47

**Figure 3.** Two scenarios for testing the accuracy, with peg-transfer evaluating pointing accuracy and suturing evaluating the virtual needle interaction accuracy.

procedure for a minute-long duration, and point to the key anatomies in the air with hand from time to time. These hints are feedback to the operator’s console for guidance. Similarly, in Fig. 2(b), the surgical step is “van velthoven anastomosis and rocco stitch”, in which suturing is demanding. The surgeon needs to incorporate as much of the muscular structural support behind the urethra as possible, therefore has to handle multiple suturing targets and complex needle trajectories which always require expertise. Using our system, the expert can provide help with a hand holding a virtual needle to indicate the next suturing target and demonstrate suturing trajectory. Such feedback together with the virtual needle can again be transmitted back to operation console for helping trainees. These results on real surgical data present the fidelity of streamed surgical scenes and flexibility of hand interactions, thus demonstrating promising feasibility for application in clinical practice.

3.3. Quantitative Evaluation on Ex-vivo Scenarios

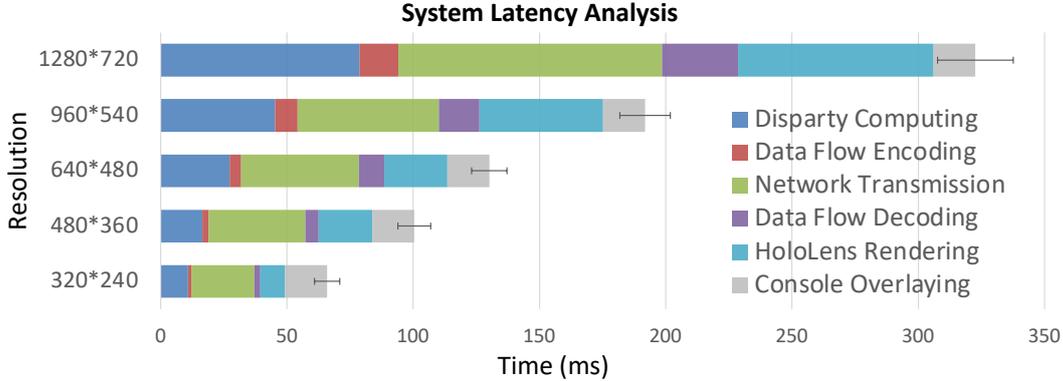
3.3.1. Accuracy evaluation.

To evaluate how accurate the hand-based interactions can achieve in this system, we conduct following experiments. Specifically, we evaluate the pointing accuracy using peg-transfer scenario, where mentor points to the position of starting peg and ending peg with hand in HoloLens, and then feedback the position to surgeon’s console in operation side, as shown in Fig. 3(a). We first manually locate 2D location of the starting and ending peg tips as ground truth. Then we get the hand-indicated 3D position from HoloLens and re-project it to 2D location on console view. We then calculate the error between re-projected 2D location and ground truth. In this way, we can obtain the close-loop pointing error in operation side. We conduct 30 trials and average the final results. As shown in Tab. 1, the position error is 7.35 pixel in average, which is hard to be perceived by human eyes and also within acceptable range compared with HoloLens calibration error (Qian et al. (2017)), thus precise enough for pointing guidance purpose.

Moreover, we evaluate the interaction accuracy of hand manipulation of virtual-object using suturing scenario, where the mentor side is asked to move a virtual-

Table 2. The update frame rate in HoloLens under different transmission resolutions.

Resolution	1280×720	960×540	640×480	480×360	320×240
FPS (Hz)	10.43	21.35	32.68	41.13	53.42

**Figure 4.** The latency analysis of each component in our proposed system.

with hand and align it to the real needle seen in the HoloLens, as shown in Fig. 3(b). Similarly, we manually locate 2D location of needle as ground truth, then record the “aligned” pose of virtual needle and re-project it to 2D console view as “aligned” 2D location. We calculate the error between ‘aligned’ 2D location and ground truth. In this way, we can obtain the close-loop virtual-object based interaction error in operation side. We also conduct 30 trials and average the results. As shown in Tab. 1, 2D re-projection error is 5.12 pixel in average, also consistent to HoloLens calibration error (Qian et al. (2017)), thus precise enough for virtual-object based interaction of surgical mentoring.

3.3.2. Latency evaluation.

To thoroughly evaluate the overall latency, we test our system under different common-used video resolutions from 320×240 to 1280×720 . Specifically, Tab. 2 reports the HoloLens update frame rate in mentor side, which can achieve satisfying real-time update speed (32.68fps) with standard definition resolution (640×480). Even when given high definition resolution (1280×720), the HoloLens can still achieve acceptable update rate (10.43fps) to fulfill the low requirement of mentor side for basic visual-hand interaction. We also evaluate the end-to-end closed-loop latency for operation side and distinguish different components’ latency, as shown in Fig 4. The total latency ranges from $65.9 \pm 5.0\text{ms}$ to $322.6 \pm 15.1\text{ms}$ with different resolution with even the highest latency is within the acceptable latency (330ms) in terms of surgeon’s perception of safety according to prior studies (Lacy et al. (2019); Marescaux et al. (2001)), thus feasible and efficient for remote mentoring. We further discuss the latency of each component. Relative constant time for console overlaying is observed, as GPU can visualize and overlay guidance in paralleled. And the disparity computing and HoloLens rendering time decrease dramatically when given lower resolution, showing the high efficiency of our system.

4. Conclusion and Future Work

In conclusion, we present the first completely AR-based remote mentoring system with a head-mounted display for robotic surgical education. Experiments and the achieved promising results have demonstrated its potential as a new reliable, flexible and low-cost solution for next-generation remote mentoring. For future work, we will extend the system to allow cooperative two-hands interaction, in order to facilitate mentoring for more complex procedures. Moreover, we will conduct systematic user study with our surgeons, who have already expressed interest and potential usefulness for their clinical practice. The future experiments shall include comparison against non AR remote mentoring systems, and investigation on surgeons' user experiences of using the proposed system, to determine the degree of benefit introducing AR to this clinical problem that would bring.

5. Acknowledgements.

This project was supported by CUHK Shun Hing Institute of Advanced Engineering (project MMT-p5-20), Hong Kong RGC TRS Project No.T42-409/18-R, and InnoHK Multi-Scale Medical Robotics Center.

References

- Chen I, Alan H, Ghazi A, Sridhar A, Stoyanov D, Slack M, Kelly JD, Collins JW. 2021. Evolving robotic surgery training and improving patient safety, with the integration of novel technologies. *World Journal of Urology*. 39(8):2883–2893.
- Collins JW, Ma R, Beaulieu Y, Hung AJ. 2021. Telementoring for minimally invasive surgery. In: *Digital surgery*. Springer; p. 361–378.
- D’Ettorre C, Mariani A, Stilli A, Valdastrì P, Deguet A, Kazanzides P, Taylor RH, Fischer GS, DiMaio SP, Menciassi A, et al. 2021. Accelerating surgical robotics research: Reviewing 10 years of research with the dvrk. *arXiv preprint arXiv:210409869*.
- Erridge S, Yeung DK, Patel HR, Purkayastha S. 2019. Telementoring of surgeons: a systematic review. *Surgical innovation*. 26(1):95–111.
- Hinata N, Miyake H, Kurahashi T, Ando M, Furukawa J, Ishimura T, Tanaka K, Fujisawa M. 2014. Novel telementoring system for robot-assisted radical prostatectomy: impact on the learning curve. *Urology*. 83(5):1088–1092.
- Huang EY, Knight S, Guetter CR, Davis CH, Moller M, Slama E, Crandall M. 2019. Telemedicine and telementoring in the surgical specialties: a narrative review. *The American Journal of Surgery*. 218(4):760–766.
- Jarc AM, Stanley AA, Clifford T, Gill IS, Hung AJ. 2017. Proctors exploit three-dimensional ghost tools during clinical-like training scenarios: a preliminary study. *World journal of urology*. 35(6):957–965.
- Joseph RA, Goh AC, Cuevas SP, Donovan MA, Kauffman MG, Salas NA, Miles B, Bass BL, Dunkin BJ. 2010. “chopstick” surgery: a novel technique improves surgeon performance and eliminates arm collision in robotic single-incision laparoscopic surgery. *Surgical endoscopy*. 24(6):1331–1335.
- Kovoor JG, Gupta AK, Gladman MA. 2021. Validity and effectiveness of augmented reality in surgical education: A systematic review. *Surgery*.
- Lacy A, Bravo R, Otero-Piñero A, Pena R, De Lacy F, Menchaca R, Balibrea J. 2019. 5g-assisted telementored surgery. *British Journal of Surgery*. 106(12):1576–1579.
- Long Y, Cao J, Deguet A, Taylor RH, Dou Q. 2022. Integrating artificial intelligence and aug-

- mented reality in robotic surgery: An initial dvrc study using a surgical education scenario. arXiv preprint arXiv:220100383.
- Long Y, Li Z, Yee CH, Ng CF, Taylor RH, Unberath M, Dou Q. 2021. E-dssr: efficient dynamic surgical scene reconstruction with transformer-based stereoscopic depth perception. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer. p. 415–425.
- Marescaux J, Leroy J, Gagner M, Rubino F, Mutter D, Vix M, Butner SE, Smith MK. 2001. Transatlantic robot-assisted telesurgery. *Nature*. 413(6854):379–380.
- Mendez I, Hill R, Clarke D, Kolyvas G, Walling S. 2005. Robotic long-distance telementoring in neurosurgery. *Neurosurgery*. 56(3):434–440.
- Navab N. 2021. Robotic imaging, machine learning and augmented reality for computer assisted interventions. In: *Medical Imaging 2021: Imaging Informatics for Healthcare, Research, and Applications*; vol. 11601. International Society for Optics and Photonics. p. 1160103.
- Qian L, Azimi E, Kazanzides P, Navab N. 2017. Comprehensive tracker based display calibration for holographic optical see-through head-mounted display. arXiv preprint arXiv:170305834.
- Qian L, Wu JY, DiMaio SP, Navab N, Kazanzides P. 2019a. A review of augmented reality in robotic-assisted surgery. *IEEE Transactions on Medical Robotics and Bionics*. 2(1):1–16.
- Qian L, Zhang X, Deguet A, Kazanzides P. 2019b. Aramis: Augmented reality assistance for minimally invasive surgery using a head-mounted display. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. p. 74–82.
- Shabir D, Abdurahiman N, Padhan J, Trinh M, Balakrishnan S, Kurer M, Ali O, Al-Ansari A, Yaacoub E, Deng Z, et al. 2021. Towards development of a tele-mentoring framework for minimally invasive surgeries. *The International Journal of Medical Robotics and Computer Assisted Surgery*. 17(5):e2305.
- Tankovich V, Hane C, Zhang Y, Kowdle A, Fanello S, Bouaziz S. 2021. Hitnet: Hierarchical iterative tile refinement network for real-time stereo matching. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. p. 14362–14372.
- Wang Y, Long Y, Fan SH, Dou Q. 2022. Neural rendering for stereo 3d reconstruction of deformable tissues in robotic surgery. arXiv preprint arXiv:220615255.
- Wei R, Li B, Mo H, Lu B, Long Y, Yang B, Dou Q, Liu Y, Sun D. 2022. Stereo dense scene reconstruction and accurate localization for learning-based navigation of laparoscope in minimally invasive surgery. *IEEE Transactions on Biomedical Engineering*.
- Yoon JW, Chen RE, Kim EJ, Akinduro OO, Kerezoudis P, Han PK, Si P, Freeman WD, Diaz RJ, Komotar RJ, et al. 2018. Augmented reality for the surgeon: systematic review. *The International Journal of Medical Robotics and Computer Assisted Surgery*. 14(4):e1914.