FlexiVision: Teleporting the Surgeon's Eyes via Robotic Flexible Endoscope and Head-Mounted Display

Long Qian*, Chengzhi Song*, Yiwei Jiang, Qi Luo, Xin Ma, Philip Waiyan Chiu, Zheng Li and Peter Kazanzides

***Abstract*— A flexible endoscope introduces more dexterity to the image capturing in endoscopic surgery. However, manual control or automatic control based on instrument tracking does not handle the misorientation between the endoscopic video and the surgeon. We propose an automatic flexible endoscope control method that tracks the surgeon's head with respect to the object in the surgical scene. The robotic flexible endoscope is actuated so that it captures the surgical scene from the same perspective as the surgeon. The surgeon wears a head-mounted display to observe the endoscopic video. The frustum of the flexible endoscope is rendered as an augmented reality overlay to provide surgical guidance. We developed the prototype, FlexiVision, integrating a 6-DOF robotic flexible endoscope based on the da Vinci Research Kit and Microsoft HoloLens. We evaluated the proposed automatic control method via a lesion observation task, and evaluated the AR surgical guidance in a lesion targeting task. The multi-user study results demonstrated that, for both tasks, FlexiVision significantly reduced the completion time (by 59% and 58%), number of errors (by 75% and 95%) and subjective task load level. With FlexiVision, the flexible endoscope could act as the surgeon's eyes teleported into the abdominal cavity of the patient.**

I. Introduction

In endoscopic surgery, an endoscope is inserted into the patient's body through a keyhole on the patient's skin or a natural orifice. Its advantages include minimized surgical trauma, accelerated recovery and reduction of hospital stay [1], [2]. Due to these benefits, endoscopy has been the gold standard for procedures such as cholecystectomy [3] and morbid obesity [4]. An annual survey in Japan showed that the number of endoscopic surgeries has been steadily increasing since its first use in the country [3].

However, an endoscopic procedure is not without drawbacks. In a typical endoscopy setup, the portion of the endoscope inside the patient's body is rigid and the surgeon holds the handle of the endoscope outside of the trocar. If the surgeon wishes to change the perspective of the endoscope to 'look around', the whole rigid endoscope is manipulated. As a result, the motion space occupied by the endoscope is large and it is possible to unintentionally hit other anatomy [5]. To alleviate this issue, researchers have proposed to use flexible endoscopes. Flexible endoscopes can alternate the perspective by re-configuring the flexible part of the endoscope, without motion of the entire apparatus.

*L. Qian and C. Song contributed equally and are considered joint first authors. Email: long.qian@jhu.edu and songchengzhiinhk@gmail.com

L. Qian, C. Song, Y. Jiang and P. Kazanzides are with the Laboratory for Computational Robotics and Sensing (LCSR), Johns Hopkins University, Maryland, United States.

C. Song, P.W. Chiu and Z. Li are with the Department of Surgery, The Chinese University of Hong Kong, HKSAR, China.

Q. Luo is with Pacific Lutheran University, Washington, United States

X. Ma is with Purdue University, Indiana, United States

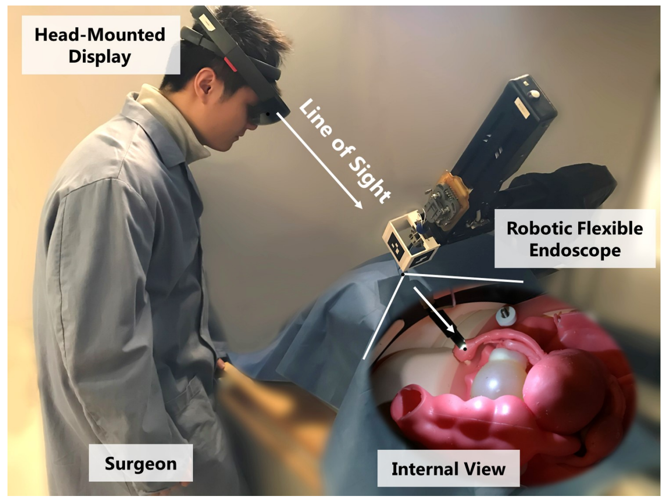


Fig. 1: FlexiVision is composed of a robotic flexible endoscope and a head-mounted display on the surgeon's head. With FlexiVision, the endoscope is actuated to match the surgeon's perspective, acting as the surgeon's eyes.

Another disadvantage is that the endoscope captures the 3D internal anatomy as a projection on 2D images [1]. The images are then visualized on cart or ceiling mounted monitors to guide the surgeon's operation. There may be a mismatch between the perspective of the endoscope and that of the surgeon, which limits the surgeon's hand-eye coordination for manipulating instruments [6]. Researchers have proposed to use augmented reality with a head-mounted display (HMD) to re-orient the displayed image to match the surgeon's perspective [7] or to provide in-situ visualization of the reconstructed 3D model of the anatomy [8], [9].

In this paper, we propose FlexiVision, where we integrate a robotic flexible endoscope with an optical see-through HMD for endoscopic surgery. With FlexiVision, the surgeon is able to observe the details of the anatomy with the HMD and, at the same time, autonomously control the pose of the flexible endoscope so that it matches his/her perspective. Fig. 1 illustrates the concept of FlexiVision. The major contributions of this paper include:

- an autonomous control method to actuate a robotic 6-DOF flexible endoscope to match the viewer's perspective.
- an AR visualization method for flexible endoscopy based on an optical see-through head-mounted display.

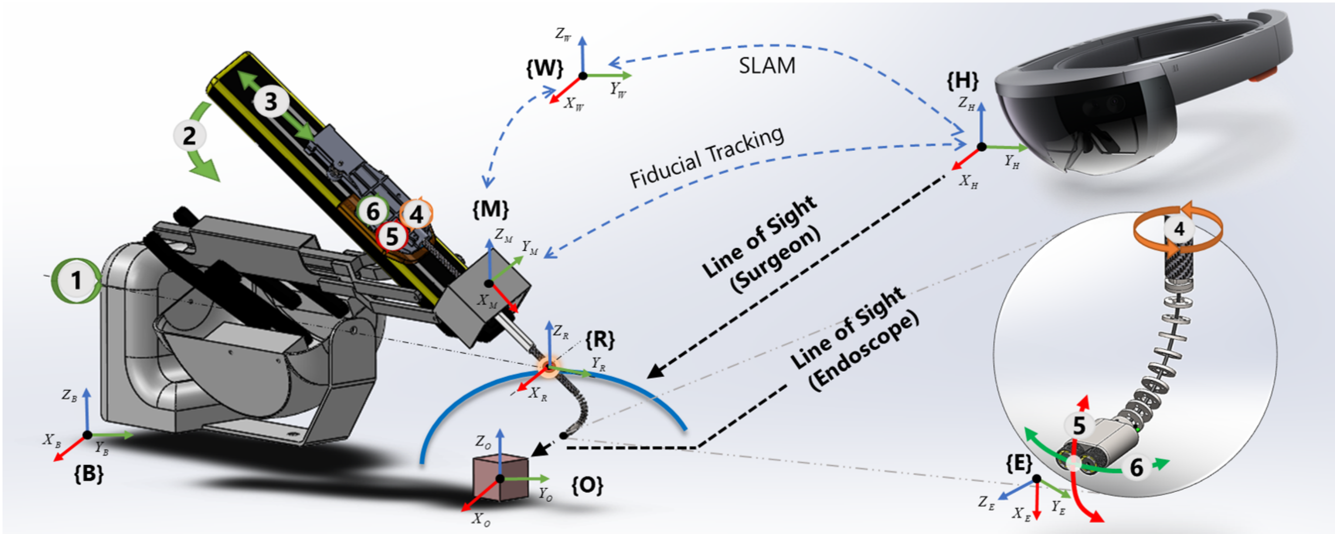


Fig. 2: System overview and components of FlexiVision

- evaluation of the above two contributions in a multi-user phantom study on a prototype of FlexiVision.

We introduce the related works in Sect. II and the system overview in Sect. III. We detail the proposed methods for autonomous flexible endoscope control and AR visualization in Sect. IV and Sect. V. We present our system setup and user study in Sect. VI and Sect. VII. The results of the experiments are presented and discussed in Sect. VIII.

II. Related Works

A. Control Methods of Flexible Endoscope

A flexible endoscope adds more degrees-of-freedom (DOF) to the system compared to rigid ones and therefore its control is not trivial. The control methods of flexible endoscopes can be generally categorized into manual control, remote control and autonomous control. The manually controlled flexible endoscope, such as the ENDOEYE FLEX 3D introduced by Olympus® [10], has the benefit of higher dexterity, but it requires a longer learning curve and will cause fatigue [11]. To reduce the fatigue and non-intuitive control of the flexible endoscope, robotic-actuated teleoperated flexible endoscopes have been developed [12], [13]. The teleoperation causes additional effort to switch between instrument operation and endoscope operation, which complicates the clinical workflow. To further release the surgeon from distraction beyond the operation itself, imaged-based autonomous control of a flexible endoscope was introduced by Song et al [14], [15], and Slawinski et al [16]. With the proposed autonomous flexible endoscope [17], surgeons can focus on the operation and use the instrument to guide the pose of flexible endoscope whenever needed. Nonetheless, the viewing perspective cannot be decided freely, thus causing misorientation during the operation and limiting the range of possible directions of observation.

B. AR Visualization for Endoscopy

With AR based on HMDs, the visualization of endoscopic video is not restricted to the designated monitors, which causes misorientation and poor hand-eye coordination [6]. Qian et al. proposed ARssist for robotic-assisted laparoscopic surgery, which has the 'frustum projection' option to display the endoscopic video through an HMD [7]. Specifically, the endoscopy is oriented correctly and perceived to be within the patient's body. An instrument manipulation task showed significant improvement of hand-eye coordination with ARssist [18]. Researchers have also proposed to display the reconstructed 3D anatomy with the correct physical coordinates, creating the effect of 'x-ray vision'. Fuchs et al. first introduced the concept in 1998, and implemented it using a structured light sensor and video see-through HMD [8]. Qian et al. recently proposed ARAMIS, which instead uses computer vision algorithms to reconstruct a real-time point cloud from stereo endoscopic video [9]. However, when the surgeon's perspective is significantly different from that of the endoscope, the point cloud reconstruction and visualization suffer from visual artifacts such as flickering and holes. In this paper, we use a combination of a basic heads-up display and frustum projection, to allow for observation of fine details or operation guidance depending on the specific surgical task, which will be presented in detail in Sect. V.

III. Overview of FlexiVision

The system architecture of FlexiVision is shown in Fig. 2. The surgeon, or the assistant, wears the HMD standing at the bedside. The coordinate system of the HMD is depicted as $\{H\}$. We use Microsoft HoloLens v1 as the HMD, which offers SLAM capabilities, so that the pose of the HMD within the world coordinate system $\{W\}$ is known ${}^{W}T_{H}$.

The flexible endoscope is mounted on a Patient-Side Manipulator (PSM) of the da Vinci Research Kit (dVRK) [19]. The robot base is $\{B\}$. The dVRK employs a mechanical Remote Center of Motion (RCM) at the cannula $\{R\}$. The rigid shaft of the endoscope has 4-DOF (pitch, yaw, insertion, roll), and the flexible tip of the endoscope has 2-DOF (orthogonal bending). The flexible tip is developed based on a tendon-driven continuum manipulator design. Therefore, the joint state of the flexible endoscope is represented by a 6×1 vector $\vec{q}$. The end effector frame of the flexible endoscope is denoted as $\{E\}$. The transformation between the end effector and the base is ${}^{B}T_{E}(\vec{q})$. We previously derived the forward and inverse kinematics of the flexible endoscope in [17].

A multi-surface fiducial marker $\{M\}$ is placed on the third link of the robotic arm, outside the cannula. With the joint status $\vec{q}$, the pose of the fiducial w.r.t. the robot base ${}^{B}T_{M}(\vec{q})$ and robot tip ${}^{M}T_{E}(\vec{q})$ can be calibrated. The HMD uses the front-facing camera to track the fiducial during runtime, as ${}^{H}T_{M}$. Therefore, the transformation between the HMD and the robot base is known: ${}^{B}T_{H} = {}^{B}T_{M} \cdot {}^{H}T_{M}{}^{-1}$. When the fiducial is not visible to the HMD, it uses SLAM to compensate for the motion [20].

We assume that there is an object of interest $\{O\}$ in the surgical scene, for example, a piece of tissue with a tumor. The 6-DOF flexible endoscope can be actuated to observe the tissue from different perspectives and the stereo video is wirelessly streamed to the HMD for AR visualization.

IV. Autonomous Flexible Endoscope Control Following the Surgeon's Perspective

We propose a novel autonomous flexible endoscope control method that tracks the surgeon's position and actuates the flexible endoscope so that it captures images from the same perspective as the surgeon.

Given the current transformation between the HMD frame and robot base frame ${}^{B}T_{H}$, and the position of the object of interest $\vec{t}_{BO}$, we would like to control the transformation between the end effector and the robot base:

$$ {}^{B}T_{E} = \left[\, R \mid \vec{t} \,\right] \tag{1} $$

such that the following constraints are satisfied:

1) The origins of the coordinate frame of object $\{O\}$, end effector $\{E\}$ and HMD $\{H\}$ are collinear.
2) The Y axis of end effector frame $\{E\}$ is horizontal, i.e. parallel with the $X_B - Y_B$ plane of the robot base frame.
3) The distance between the end effector and the object is r.
4) The RCM is fixed at frame $\{R\}$.

Constraint 1) guarantees that the viewing direction of the flexible endoscope is the same as that of the surgeon looking at the target anatomy. Since the flexible endoscope is stereoscopic, constraint 2) ensures that the stereo camera pair is horizontally separated, so that the stereo endoscopic video viewed by the surgeon provides good depth cues. Constraint 3) sets the flexible endoscope at a reasonable distance to the object of interest for observation according to the optical focal length. In our experiment, we set it to $5\,cm$. It is noticeable that the distance r is still adjustable during runtime based on the clinical requirements. Constraint 4) guarantees that there is no force exerted on the incision port, which is critical to the patient's safety. This constraint is mechanically satisfied via the dVRK robotic framework.

With the above constraints, we first calculate the translation $\vec{t}$ in Eq. 1 as follows. Without losing generality, we

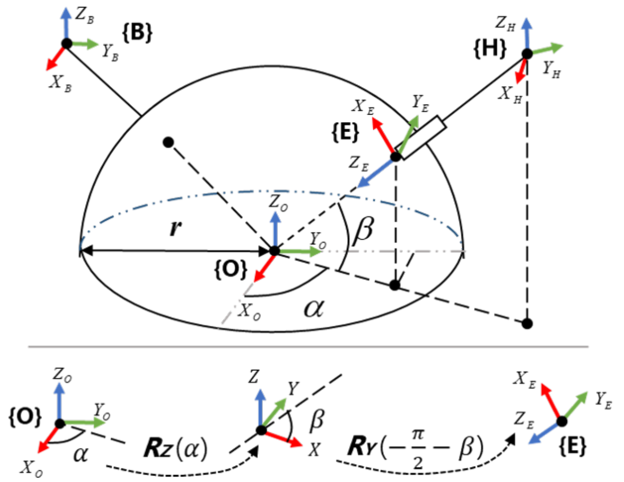


Fig. 3: The calculation of the desired end effector pose ${}^{B}T_{E}$

assume that there is only a linear translation (no rotation) between $\{B\}$ and $\{O\}$ because the object of interest is defined as a spatial point. We obtain the transformation between the object of interest $\{O\}$ and the HMD $\{H\}$ via ${}^{O}T_{H} = {}^{B}T_{O}{}^{-1} \cdot {}^{B}T_{H}$. Then, we extract the translation part as $\vec{t}_{OH} = (x_{OH}, y_{OH}, z_{OH})$. Since $\vec{t}$ is parallel to $\vec{t}_{OH}$ and has a fixed length of r, it can be calculated as:

$$ \vec{t} = \vec{t}_{BE} = \vec{t}_{BO} + \vec{t}_{OE}\ , \quad \vec{t}_{OE} = r \cdot \frac{\vec{t}_{OH}}{\|\vec{t}_{OH}\|} \tag{2} $$

where $\vec{t_{OE}}$ can be further written as (x_{OE}, y_{OE}, z_{OE}).

Secondly, we calculate R in Eq. 1, the rotation between the end effector frame $\{E\}$ and the robot base frame $\{B\}$. It is sufficient to calculate the rotation between $\{O\}$ and $\{E\}$ because we assume there is no rotation between $\{B\}$ and $\{O\}$. In order to satisfy constraint 2), we deliberately construct the rotation R using a special Euler angle XYZ representation without the third rotation about the X axis:

$$ R = R_z(\alpha) \cdot R_y(\beta) \tag{3} $$

As shown in Fig. 3, the $\{O\}$ is first rotated about the Z axis by α, and then rotated about the Y axis by $-\frac{\pi}{2} - \beta$. In this way, the resulting Y axis will always be horizontal. The horizon is defined by the X_B and Y_B axis of robot base frame. The rotation parameters α and β can be calculated by:

$$ \alpha = \arctan\left(\frac{y_{OE}}{x_{OE}}\right)\ , \quad \beta = \arctan\left(\frac{z_{OE}}{\sqrt{x_{OE}^2 + y_{OE}^2}}\right) \tag{4} $$

Combing Eq. 1 to Eq. 4, we have determined the desired transformation between the end effector and the robot base ${}^{B}T_{E}$, which is then controlled by the position-based control algorithm in the dVRK. At runtime, the desired pose of the end effector is constantly updated according to the HMD frame $\{H\}$. As a consequence, the end effector will gradually follow the perspective of the surgeon.

V. Visual Guidance with Augmented Reality

We propose to use AR to provide visual guidance for the surgeon when he or she is operating instruments in the surgical field. The AR guidance is comprised of two parts:

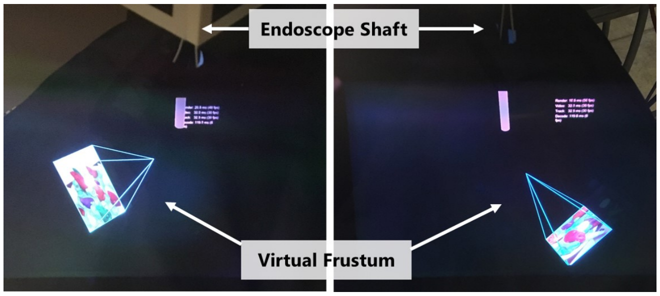


Fig. 4: AR visualization of the virtual frustum of the flexible endoscope. The rendering matches the physical position of the endoscope.

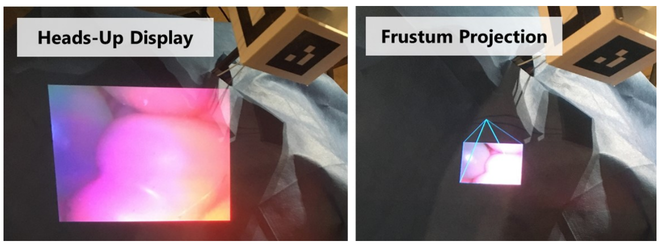


Fig. 5: AR visualization of the endoscopic video, as a heads-up display (left) or as frustum projection (right)

- the rendering of the frustum of the flexible endoscope
- the stereo endoscopic video visualized as a heads-up display or frustum projection.

A. Virtual Frustum

In a traditional endoscopy setup with a rigid endoscope, an experienced surgeon is able to figure out the orientation of the endoscope by observing the shaft outside the cannula. However, it is not possible to infer the orientation of the flexible endoscope tip by observing the shaft. It is also a challenge to understand the insertion depth. In this situation, it is difficult to introduce an instrument into the field-of-view of the flexible endoscope. Therefore, we propose to use AR to provide guidance, more specifically, to visualize the frustum of the flexible endoscope registered with the physical object (Fig. 4). The visualization requires the transformation between the HMD and the endoscope tip ${}^{H}T_{E}$, which can be calculated via ${}^{H}T_{E} = {}^{H}T_{M} \cdot {}^{M}T_{P}$, based on the fiducial tracking and the kinematics. The horizontal and vertical field-of-view of the endoscopic cameras is calibrated and used for rendering the virtual frustum.

B. Stereo Endoscopic Video

Two visualization options for the stereo endoscopy are available: heads-up display or frustum projection (Fig. 5). When the surgeon is observing the tissue, it is recommended to use the heads-up display visualization because the endoscopy is magnified and occupies the entire screen to allow easier observation. When the autonomous control in Sect. IV is activated, the perspective of the endoscope is aligned with that of the surgeon, as if the surgeon's eyes are teleported into the patient's body.

When the surgeon is operating or bringing in an instrument, the autonomous control is deactivated to avoid drift in the endoscopic video. In this case, it is recommended to use frustum projection to visualize the endoscopy because it incorporates the physical orientation of the flexible endoscope to restore the hand-eye coordination of the operator [18].

VI. Implementation

The flexible endoscope is mounted on the dVRK, which is powered by an Ubuntu 18.04 Desktop PC (Intel® Xeon (R) W-2145 CPU, 38.9 GB RAM). The stereo endoscopy (640×480) is available to the same desktop via two USB cameras. The two channels are retrieved, concatenated and wirelessly streamed to the HoloLens using FFmpeg[1]. The HoloLens application is developed using Unity[2], which decodes the received wireless video stream, tracks the fiducial marker using HoloLensARToolKit[3] [21], configures the robotic model using dVRK-XR[4] [22], and provides AR visualization. At runtime, the HoloLens sends the current head pose ${}^{B}T_{H}$ to the desktop via UDP protocol. The desktop then actuates the flexible endoscope accordingly. The AR application runs at $60\,Hz$. It is also important to mention that the perspective-based control is enabled and disabled with a foot-pedal switch.

VII. Evaluation

We evaluate FlexiVision with two tasks in the multi-user study, lesion observation and lesion targeting. In both tasks, we compare FlexiVision with a traditional setup. 21 users (mean age: 27.6, std: 7.5) from the Johns Hopkins University were recruited for the user study with IRB approval. Their familiarity with AR systems and minimally-invasive procedures is well distributed.

A. Lesion Observation

In endoscopic surgery, the surgeon needs to observe the tissue for diagnosis and planning. We set up a phantom-based lesion observation task. A 3D printed phantom with three colored faces is inserted into the phantom abdominal cavity (Fig. 6). On each of the three surfaces, there are a few triangles and circles. The users need to manipulate the flexible endoscope to look at all three surfaces, count the total number of triangles and circles and report to the researchers.

In the traditional setup, the users manually manipulate the rigid shaft of the endoscope (as currently done on the dVRK), and the flexible part is controlled via footpedals. The four footpedals can rotate the endoscope field-of-view to the left and right (yaw axis) and up and down (pitch axis), shown in Fig. 2. With FlexiVision, the footpedal is used to enable/disable the autonomous control. When enabled, the users can control the pose of the flexible endoscope via physically moving around the phantom. The heads-up display visualization of the endoscopic video is used.

[1]FFmpeg: https://www.ffmpeg.org/
[2]Unity: https://www.unity.com/
[3]HoloLensARToolKit: https://github.com/qian256/HoloLensARToolKit
[4]dVRK-XR: https://github.com/jhu-dvrk/dvrk-xr

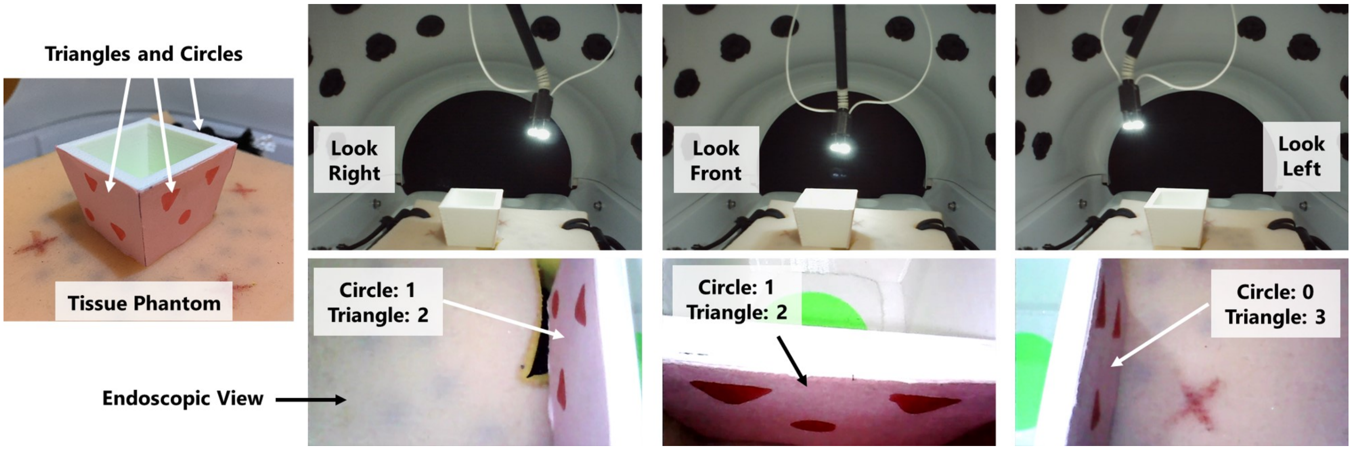


Fig. 6: The setup for the lesion observation task. The users need to manipulate the flexible endoscope to observe three surfaces of the phantom and count the number of the shapes. We compare FlexiVision to manual control using a footpedal.

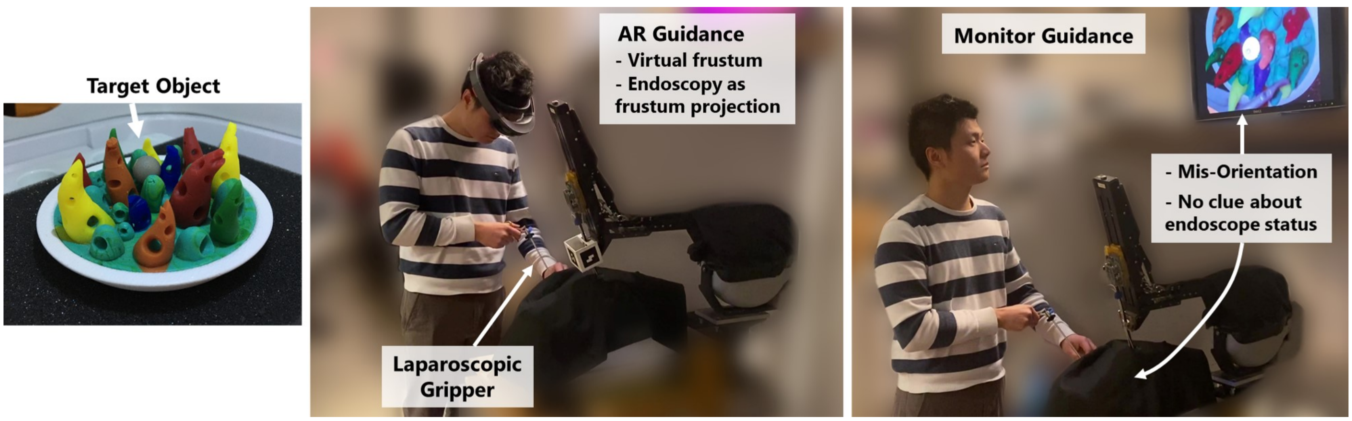


Fig. 7: The setup for the lesion targeting task. A phantom plate with a lesion target is placed inside the view of the flexible endoscope. The users manipulate a laparoscopic instrument to hit the target.

After a few training trials, each user performs the lesion observation task three times with the traditional setup and three times with FlexiVision. A few 3D printed phantoms with different numbers of shapes are randomized for these experiments to prevent the user from memorizing the exact phantom. We record the number of the observed shapes (triangles and circles), the time duration and task load index (NASA-TLX [23]) for each setup.

B. *Lesion Targeting*

Lesion targeting is a common task in endoscopy where the surgeon uses an instrument to reach a target in the surgical scene. Traditionally, the surgeon views the monitor which does not indicate the pose of the flexible tip. Therefore, the user needs to search for the field-of-view of the endoscope. With FlexiVision, the user is guided by the AR visualization, including the virtual frustum (Fig. 4) and the endoscopic video as frustum projection (Fig. 5-right), which helps them to locate the endoscope and its viewing direction.

After a few training trials, each user performs the lesion targeting task three times with conventional monitor guidance and three times with FlexiVision. The flexible endoscope is randomly configured for each experiment. We evaluate the time to hit the target, the number of accidental hits to the flexible endoscope, and the task load index (NASA-TLX).

VIII. Results and Discussion

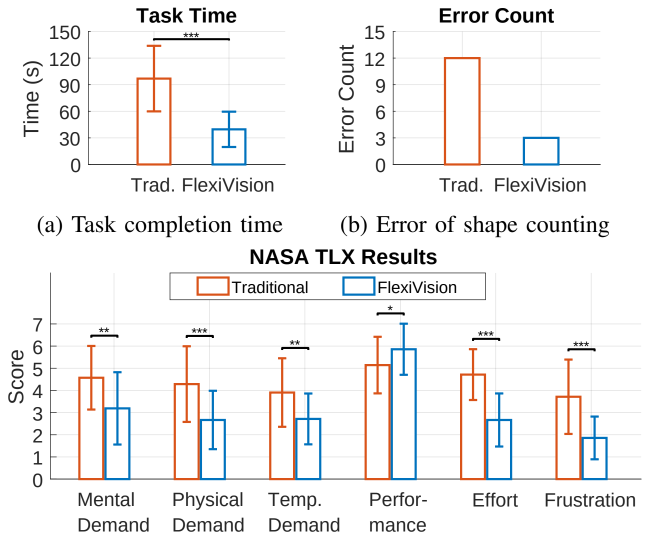
A. *Lesion Observation*

The evaluation results of the lesion observation task are shown in Fig. 8. In total, there are 63 trials (21 users $\times$ 3 repetitions) of the lesion observation task with the traditional manual flexible endoscope control and 63 trials with FlexiVision, where the flexible endoscope is controlled by the user's perspective.

The users spent in average $98.87 \pm 36.97s$ [5] to finish the observation task with the traditional setup. The time is significantly reduced to $39.63 \pm 19.92s$ with FlexiVision, $p = 3.97 \times 10^{-16}$ (via paired t-test). The reduction in time indicates that the user is able to manipulate the flexible endoscope to the desired pose much quicker via the autonomous control using head pose.

We recorded the results of the observation from each trial. The users mis-counted 12 times with the traditional manual control and 3 times with FlexiVision, which indicates that the pose to observe the surface was less optimal under manual

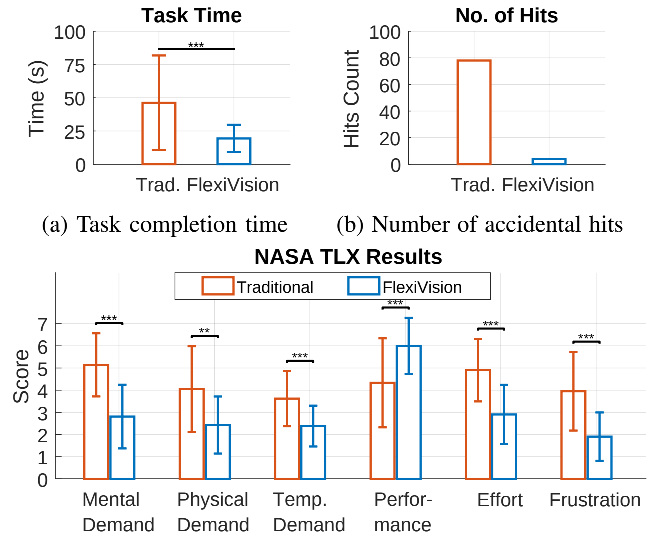
[5] Notation: mean $\pm$ standard deviation



(a) Task completion time (b) Error of shape counting

(c) Task load index of lesion observation

Fig. 8: Evaluation results for lesion observation task comparing traditional manual control and FlexiVision



(a) Task completion time (b) Number of accidental hits

(c) Task load index of lesion targeting

Fig. 9: Evaluation results for lesion targeting task comparing traditional monitor-based guidance and FlexiVision

control. The shapes are not seen clearly or are missed. FlexiVision provides better viewing perspective because the users can easily and interactively adjust it so that each shape is clearly visible. There are still mistakes with FlexiVision, which may be due to the difficulty in memorizing the counted numbers on each side. When the user spends too much time for the task, memory errors are more likely to occur.

The TLX results are shown in Fig. 8c. The score range is from 1 to 7 (1 is best). FlexiVision has shown improvement for all the categories. More specifically, the mental demand, physical demand, temporal demand, effort level, and frustration level are significantly decreased, and the self-assessed performance is significantly improved with FlexiVision.

In summary, FlexiVision is able to significantly reduce the task time, number of mistakes, and task load for anatomy observation.

B. Lesion Targeting

The evaluation results for the lesion targeting task are shown in Fig. 9. As in the lesion observation task, there are 63 trials with the traditional setup, where the user is guided by endoscopic video displayed on a monitor, and 63 trials with AR guidance of FlexiVision.

The task completion time is $46.20 \pm 35.61s$ with traditional monitor guidance and $19.38 \pm 10.29s$ with the AR-guidance of FlexiVision. The reduction in completion time is significant $p = 2.80 \times 10^{-8}$. The improved efficiency comes from two aspects. On one hand, with AR visualizing the frustum of the flexible endoscope, it is much easier to introduce the instrument into the field-of-view of the endoscope. With the traditional monitor, the users are not aware of the actual pose of the endoscope inside the non-transparent phantom. A manual random search is unavoidable, which also causes more accidental hits. On the other hand, once the instrument is inside the endoscopic video, the orientation offset provided by frustum projection visualization helps the user to manipulate the instrument in the correct direction to finally hit the target. However with the misorientation of the endoscopic video on the traditional monitor, the user has to first find out the correct manipulation direction.

With monitor guidance, there are 78 accidental hits to the flexible endoscope, whereas with FlexiVision there are only 4. As discussed above, the lack of guidance information in the conventional setup causes a random search process. Although the AR visualization does not overlay the virtual hand-held instrument, the user is still able to introduce the instrument into the field because the virtual frustum is registered with the physical space. During the training phase, we found that there is a learning curve to understand the AR interface, especially the frustum projection visualization. After a few trials, the users were able to perform the task relatively successfully.

As shown in Fig. 9c, the task load index demonstrated significant improvement using FlexiVision in all categories. Noticeably, even though the HMD has additional weight on the user's head, the users still found the traditional setup to be more physical demanding and more frustrating.

In summary, the evaluation results have demonstrated that surgical guidance provided by FlexiVision yields significantly shorter task completion time, much fewer number of accidental hits, and significantly smaller task load.

C. Future Work

In the future, we wish to evaluate FlexiVision in a more clinical setting, e.g., an ex-vivo phantom study. Currently, the object of interest $\{O\}$ that the flexible endoscope rotates around is defined in the robot coordinate system, and can be adjusted via voice command. We wish to support more intuitive methods to adjust the location of the object of interest, e.g., by integrating surgical scene tracking and task understanding to automatically determine $\{O\}$ [24], [25]. In the current implementation of FlexiVision, the scene depth observed from the endoscope is fixed as r (Fig. 3). This parameter could also be adjusted based on the surgeon's distance to the surgical site. Another future direction is to develop a hybrid automatic flexible endoscope control that combines instrument tracking, scene target tracking and surgeon perspective tracking to offer the best user experience.

IX. Conclusion

A flexible endoscope has the potential to improve endoscopic surgery by providing an increased visual field and smaller motion space. Traditionally, endoscope control is either done manually or via tracking of certain objects in the surgical scene, e.g., the instrument, and the endoscopic video is shown to the user via standard monitors. In this paper, we propose to actuate the flexible endoscope to align with the surgeon's viewing perspective and display the endoscopic video using augmented reality on a head-mounted display. We developed the FlexiVision prototype based on a 6-DOF flexible endoscope [15] and Microsoft HoloLens v1.

With FlexiVision, the surgeon is able to control the flexible endoscope by changing his/her own viewing perspective of the patient's anatomy. We evaluated it in a lesion observation task, comparing to traditional manual control. The results showed that the autonomous control method significantly reduced the time to observe the anatomy, the number of mistakes and the task load.

FlexiVision also offers an AR interface showing the current pose of the frustum and endoscopic video "inside" the patient's body. It offers guidance for introducing instruments to operate on the anatomy. We evaluate the efficacy of the AR visualization in a lesion targeting experiment. The results showed significant reduction in completion time, the number of mistakes and the subjective task load.

In summary, FlexiVision acts like the surgeon's eyes teleported into the patient's body, helping the surgeon to see and to operate more intuitively.

ACKNOWLEDGMENT

This work is supported by internal funds from the Johns Hopkins University, and grants from the Chinese University of Hong Kong (No. 14212316, No. 14207017 and No. 24204818). The da Vinci Research Kit (dVRK) is supported by NSF NRI-1637789.

References

[1] S. M. D. Sørensen, M. M. Savran, L. Konge, and F. Bjerrum, "Three-dimensional versus two-dimensional vision in laparoscopy: a systematic review," *Surgical Endoscopy*, vol. 30, no. 1, pp. 11–23, 2016.

[2] L. Medeiros, A. Stein, J. Fachel, R. Garry, and S. Furness, "Laparoscopy versus laparotomy for benign ovarian tumor: a systematic review and meta-analysis," *International Journal of Gynecologic Cancer*, vol. 18, no. 3, pp. 387–399, 2008.

[3] M. Inomata, H. Shiroshita, H. Uchida, T. Bandoh, S. Akira, S. Yamaguchi, Y. Kurokawa, Y. Seki, S. Eguchi, N. Wada, *et al.*, "Current status of endoscopic surgery in Japan: The 14th national survey of endoscopic surgery by the Japan Society for Endoscopic Surgery," *Asian Journal of Endoscopic Surgery*, 2020.

[4] F. Brody, "Minimally invasive surgery for morbid obesity," *Cleveland Clinic Journal of Medicine*, vol. 71, no. 4, pp. 289–302, 2004.

[5] M. Kukuk, "Modeling the internal and external constraints of a flexible endoscope for calculating its workspace: application in transbronchial needle aspiration guidance," in *Medical Imaging: Visualization, Image-Guided Procedures, and Display*, vol. 4681. SPIE, 2002, pp. 539–550.

[6] B. Wentink, "Eye-hand coordination in laparoscopy-an overview of experiments and supporting aids," *Minimally Invasive Therapy & Allied Technologies*, vol. 10, no. 3, pp. 155–162, 2001.

[7] L. Qian, A. Deguet, and P. Kazanzides, "ARssist: Augmented reality on a head-mounted display for the first assistant in robotic surgery," *Healthcare Technology Letters*, vol. 5, no. 5, pp. 194–200, Oct 2018.

[8] H. Fuchs, M. A. Livingston, R. Raskar, K. Keller, J. R. Crawford, P. Rademacher, S. H. Drake, A. A. Meyer, *et al.*, "Augmented reality visualization for laparoscopic surgery," in *Med. Image Comp. and Comp.-Assisted Interv. (MICCAI).* Springer, 1998, pp. 934–943.

[9] L. Qian, X. Zhang, A. Deguet, and P. Kazanzides, "ARAMIS: Augmented reality assistance for minimally invasive surgery using a head-mounted display," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI).* Springer, 2019, pp. 74–82.

[10] "Endoeye Flex 3D (LTF-190-10-3D)." [Online]. Available: https://medical.olympusamerica.com/products/laparoscopes/endoeye-flex-3d

[11] J. Perrone, C. Ames, Y. Yan, and J. Landman, "Evaluation of surgical performance with standard rigid and flexible-tip laparoscopes," *Surgical Endoscopy And Other Interventional Techniques*, vol. 19, no. 10, pp. 1325–1328, 2005.

[12] B. P. M. Yeung and T. Gourlay, "A technical review of flexible endoscopic multitasking platforms," *International Journal of Surgery*, vol. 10, no. 7, pp. 345–354, 2012.

[13] T. Iwasa, R. Nakadate, S. Onogi, Y. Okamoto, J. Arata, S. Oguri, H. Ogino, E. Ihara, K. Ohuchida, T. Akahoshi, *et al.*, "A new robotic-assisted flexible endoscope with single-hand control: endoscopic submucosal dissection in the ex vivo porcine stomach," *Surgical Endoscopy*, vol. 32, no. 7, pp. 3386–3392, 2018.

[14] C. Song, X. Ma, X. Xia, P. W. Y. Chiu, C. C. N. Chong, and Z. Li, "A robotic flexible endoscope with shared autonomy: a study of mockup cholecystectomy," *Surgical Endoscopy*, pp. 1–12, 2019.

[15] X. Ma, C. Song, P. W. Chiu, and Z. Li, "Autonomous flexible endoscope for minimally invasive surgery with enhanced safety," *IEEE Robotics and Automation Letters*, vol. 4, no. 3, pp. 2607–2613, 2019.

[16] P. R. Slawinski, A. Z. Taddese, K. B. Musto, K. L. Obstein, and P. Valdastri, "Autonomous retroflexion of a magnetic flexible endoscope," *IEEE Robotics and Auto. Letters*, vol. 2, no. 3, pp. 1352–1359, 2017.

[17] X. Ma, C. Song, P. W. Chiu, and Z. Li, "Visual servo of a 6-DOF robotic stereo flexible endoscope based on da Vinci Research Kit (dVRK) system," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 820–827, 2020.

[18] L. Qian, A. Deguet, Z. Wang, Y.-H. Liu, and P. Kazanzides, "Augmented reality assisted instrument insertion and tool manipulation for the first assistant in robotic surgery," in *2019 International Conference on Robotics and Automation (ICRA).* IEEE, 2019, pp. 5173–5179.

[19] P. Kazanzides, Z. Chen, A. Deguet, G. S. Fischer, R. H. Taylor, and S. P. DiMaio, "An open-source research kit for the da Vinci® surgical system," in *Intl. Conf. on Robotics and Automation (ICRA).* IEEE, 2014, pp. 6434–6439.

[20] J. Wang, L. Qian, E. Azimi, and P. Kazanzides, "Prioritization and static error compensation for multi-camera collaborative tracking in augmented reality," in *Virtual Reality.* IEEE, 2017, pp. 335–336.

[21] L. Qian, E. Azimi, P. Kazanzides, and N. Navab, "Comprehensive tracker based display calibration for holographic optical see-through head-mounted display," *arXiv preprint arXiv:1703.05834*, 2017.

[22] L. Qian, A. Deguet, and P. Kazanzides, "dVRK-XR: Mixed reality extension for da Vinci Research Kit," in *Hamlyn Symposium on Medical Robotics*, 2019.

[23] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (task load index): Results of empirical and theoretical research," in *Advances in Psychology.* Elsevier, 1988, vol. 52, pp. 139–183.

[24] B. Yang, W. Chen, Z. Wang, Y. Lu, J. Mao, H. Wang, and Y.-H. Liu, "Adaptive fov control of laparoscopes with programmable composed constraints," *IEEE Transactions on Medical Robotics and Bionics*, vol. 1, no. 4, pp. 206–217, 2019.

[25] A. Mariani, G. Colaci, T. Da Col, N. Sanna, E. Vendrame, A. Menciassi, and E. De Momi, "An experimental comparison towards autonomous camera navigation to optimize training in robot assisted surgery," *IEEE Robotics and Automation Letters*, 2020.