2022 9th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob) | 978-1-6654-5849-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/BIOROB52689.2022.9925526

Experimental Evaluation Using Head Motion and Augmented Reality to Intuitively Control a Flexible Endoscope

Yoeko X. Mak[1], Maurits Zegel[2], Momen Abayazid[1], Massimo A. Mariani[2], and Stefano Stramigioli[1]

***Abstract*— Thoracoscopic procedures require an assistant to hold and control the camera while the surgeon performs the surgical task. This paper presents an approach in which surgeons can control camera orientation using their head movements, allowing them to steer a flexible endoscope without the need for a camera assistant during the operation. Additionally, an augmented reality headset has been integrated into the head movement control system to serve as a virtual display monitor capable of following the user's gaze. Experiments were conducted to assess the feasibility of the head-controlled approach compared to the manual control method by conducting camera-pointing experiments performed by clinicians and trained non-clinician participants at two difficulty levels. The results from the camera-pointing experiments have shown that the developed head-controlled endoscope has a statistically faster reaching time performance compared to manual use of the flexible endoscope in high difficulty index tasks with clinician participants (*p*=0.04), and in both lower and high difficulty index tasks with non-clinician participants (*p*=0.03). The head-controlled robotic endoscope approach enables surgeons to intuitively control the camera during an operation, while simultaneously performing other tasks using their hands, without sacrificing camera steering accuracy.**

I. INTRODUCTION

A crowded operating room environment and poor ergonomics during a minimally invasive surgery (MIS) operation can cause an increase in physical burden to operating room personnel. Several studies have stated that physical discomfort is frequently reported by clinicians who are involved routinely in laparoscopic or video-assisted thoracoscopic surgeries (VATS) [1]–[4]. Surgeons and camera assistants who control the surgical camera often work in a difficult operating position due to space limitations [5]. This often causes sustained upper body torsion and assymmetry in weight bearing, leading to discomfort and injury [6].

In addition to ergonomics, another bottleneck during an MIS procedure is the communication dependency between the surgeons and the first assistants who control the camera [7, 8]. This limitation can be overcome by robotizing the camera steering and letting the surgeon control the camera movement directly. This is commonly accomplished through the use of robotic systems in the operating room, such as the da Vinci Surgical System [9]. Despite the many benefits of the system, it is unattainable for many clinics and hospitals due to its very high cost and infrastructure requirements.

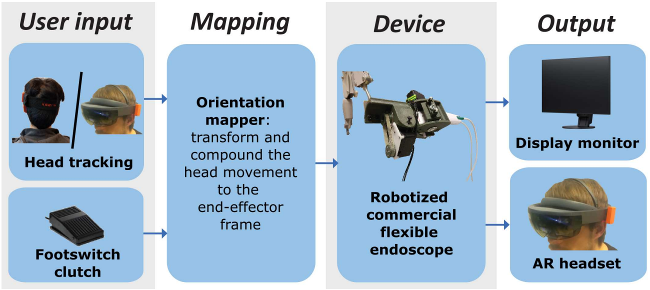


Fig. 1. An overall system flow diagram showing the implementation of the head-motion control approach.

Alternatively, various control input interfaces have been investigated to offload camera control from the first assistant, such as single-hand controlled [10], image guided [11, 12], eye-gaze controlled [13, 14], body-controlled [15], foot-controlled [16] and head-controlled surgical cameras [17]–[20]. Although the use of head motion for camera control has previously been demonstrated, the effectiveness of this method has not been investigated for flexible surgical cameras, more specifically, in combination with the use of a head-mounted display.

In recent years, the merits of using flexible endoscopes in place of rigid endoscopes for thoracoscopy have been discussed, such as single-port thoracic sympathectomy [21] and pleural biopsy procedures [22]. Flexible endoscopes are preferred due to their excellent maneuverability in accessing difficult-to-reach areas. However, it is relatively more difficult to control compared to its rigid counterpart and has a steep learning curve due to its non-intuitive steering mechanism [23, 24]. Robotizing the control mechanism and directly mapping the head motion to the movement of the endoscope tip can offer a more intuitive experience in steering a flexible endoscope. Using this method, users should be able to maneuver the camera as naturally as they would use their head to look around.

As the user's head is in motion while controlling the system, the utilization of a head mounted display (HMD) would enhance the ease of use as their gaze is not fixed toward a static display monitor (as in the case of VATS). Furthermore, the usage of augmented reality (AR) glasses has been found to improve ergonomics [25] and improve efficiency by augmenting the endoscopic image with additional information during VATS [26].

In this paper, we present the approach of using head movement to control a flexible endoscope system. A robotic control handle was designed and 3D printed to control the

[1] Y. X. Mak, M. Abayazid, and S. Stramigioli are with the Robotics and Mechatronics Laboratory, Technical Medical Centre at University of Twente, 7522 NH Enschede, the Netherlands (e-mail: y.x.mak@utwente.nl; m.abayazid@utwente.nl; s.stramigioli@utwente.nl).

[2] M. Zegel and M. A. Mariani are with the Department of Cardiothoracic Surgery, University Medical Center Groningen, Groningen, the Netherlands.

orientation of a commercially available endoscopic camera. Furthermore, the head-control approach is coupled with an AR headset, allowing users to see the image output without fixing their gaze on a stationary monitor. Additionally, the transparent headset enables users to be visually aware of their hands and external surroundings while operating the endoscopic system.

Participants with a medical background performed camera-pointing experiments to validate the proposed system. In the camera-pointing experiment, the operator steers the endoscope to visualize a certain landmark in the camera image, using both the head-controlled system and the conventional hand-controlled approach. The goal is to investigate the speed and accuracy of the head control interface compared to the conventional method of controlling the endoscope in the same setting. Furthermore, qualitative user experiences using manual and head-controlled methods were evaluated through questionnaires.

In essence, the novelty of this study lies in: (1) development of the head motion control method for a flexible endoscope and the integration of virtual display using an augmented reality headset, (2) mechanical design of a modular control handle to control a flexible endoscope in 2 degrees of freedom (DOF), and (3) a comprehensive validation study to systematically evaluate the accuracy of the head-controlled system and its user experience versus manual usage.

II. MATERIALS AND METHODS

A. *Head Orientation Tracking and Mapping*

Various measurement techniques can be used to track head orientation, for example, using optical tracking markers, distance sensors, electromagnetic (EM) based position sensors, inertial measurement unit (IMU), etc. In this study, a wireless IMU (Xsens Technologies B.V., the Netherlands) is used to track head motion (mounted on a headband). An IMU-based tracking method was selected for its advantage of not requiring a direct line of sight (or proximity) to a reference point, as in the case of optical or EM-based tracking.

Furthermore, the system is equipped with a foot-switch clutching mechanism that enables endoscope control when pressed. The clutching mechanism was implemented in order for the head control system to not restrain the surgeon's head movement and allow the surgeon to perform other clinical tasks when the foot-switch is not engaged.

The kinematic motion map consists of three components: (1) the mapping of orientation from the user's head motion to the endoscope tip motion, (2) an algorithm to continue the motion from the last saved orientation of the previous foot-switch session, and (3) the mapping to the servomotor's joint angles.

1) Orientation mapping from head to endoscope tip: The head orientation measurement from the IMU is given as the sensor orientation in reference to the magnetic north bearing of the earth. Therefore, a transformation between IMU frame to the user's head frame is given by:

$$\mathbf{q}_e^h = \mathbf{q}_{\text{IMU}}^h \odot \mathbf{q}_e^{\text{IMU}} \tag{1}$$

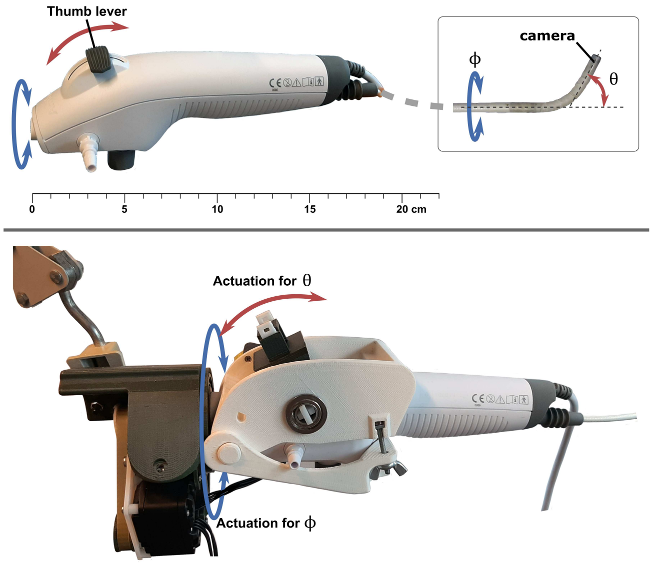


Fig. 2. A robotized mechanical gripper is developed to actuate the flexible endoscope camera in two degrees of freedom for bending and rotating the endoscopic tip. The disposable Ambu aScope 4 endoscope without the motorized control handle is shown in the top image.

where $\odot$ denotes a quaternion multiplication operator, $\mathbf{q}_{\text{IMU}}^h \in \text{SU}(2)$ is the transformation from IMU frame to the head frame, and $\mathbf{q}_e^{\text{IMU}} \in \text{SU}(2)$ is the orientation output from the IMU sensor.

2) Saving the orientation from the previous foot-switch session: The head orientation is mapped to the endoscope tip by a position-dependent mapping [27], where orientation of the head frame is mapped onto the endoscope tip frame, in contrast to velocity-dependent mapping (where angular velocities are used). Inspired by the usage method of a computer mouse, the endoscope tip should stay at the location where the foot-switch was previously released, and continue from that location when it is reengaged.

A foot-switch 'session' is defined as the duration from when the foot-switch is pressed until it is released. The output orientation can be calculated by accumulating the change in head orientation for each session. Therefore, the orientation of the current frame $\mathbf{q}(s,t)$ in session s at the time instance t can be calculated by:

$$\mathbf{q}(s,t) = \mathbf{q}_{h_{s,0}}^{h_{s,t}} \odot \mathbf{q}_{h_{s-1,0}}^{h_{s-1,\text{end}}} \odot \cdots \odot \mathbf{q}_{h_{1,0}}^{h_{1,\text{end}}} \,. \tag{2}$$

This calculation can be performed recursively by using the head orientation from the end of the previous session:

$$\mathbf{q}(s,t) = \mathbf{q}_{h_{s,0}}^{h_{s,t}} \odot \mathbf{q}(s-1, t_{\text{end}}) \,, \tag{3}$$

where t_{end} is the time at the end of the previous session.

3) Mapping to the joint space: Subsequently, the output orientation is mapped to the joint space: tip bending θ and rotation around the insertion tube axis ϕ (shown in Fig. 2). Afterward, these values are used to control the endoscope tip motion by a position control approach, processed on-board the servomotors. The joint rotation values are calculated by rotating a reference unit vector $\mathbf{r}_{\text{ref}} = \begin{bmatrix}1 & 0 & 0\end{bmatrix}^{\mathsf{T}}$ by the output orientation $\mathbf{q}$ (calculated in Eq. 3).

$$\begin{bmatrix}0\\ \mathbf{r}\end{bmatrix} = \mathbf{q} \odot \begin{bmatrix}0\\ \mathbf{r}_{\text{ref}}\end{bmatrix} \odot \mathbf{q}^* \,, \tag{4}$$

where $\odot$ represents the multiplication operation between two quaternions and $(\bullet)^*$ denotes the quaternion conjugate. Subsequently, the joint rotation values can be calculated by:

$$\phi = \text{atan2}\left(-r_y \,,\, r_z\right) \tag{5}$$

$$\theta = \text{atan2}\left(\sqrt{r_y^2 + r_z^2}\,,\, r_x\right) \tag{6}$$

where r_x, r_y, r_z are the three Cartesian components of $\mathbf{r}$.

B. Robotized Endoscope Control Handle

The developed head-motion control interface is implemented on a custom endoscope handle for a commercial disposable flexible endoscope, Ambu aScope 4 Regular (Ambu A/S, Denmark).

The developed robotic control handle mechanism is shown in Fig. 2. The commercial disposable flexible endoscope used in this project has a single DOF bending between −180° and 180°, which can be actuated using a thumb lever mechanism on the control handle. The robotic mechanism has been designed such that another DOF is added to the endoscopic camera motion by rotating the handle around the insertion tube axis. Two Herkulex DRS-0201 servomotors (Dongbu Robot, South Korea) were used in the control handle setup: one servomotor is used to rotate the handle around the insertion axis, and another servomotor is used to actuate the thumb lever mechanism to control the deflection of the endoscope tip.

To capture the analog image output from the Ambu endoscope and its monitor, a Terratec G1-USB frame grabber (TerraTec Electronic GmbH, Germany) is used to capture the endoscope image into digital format. The resulting output is a 25 frame per second (fps) video stream with 6-7 frame delay from the endoscope to the output screen.

The captured images are then processed to add visual cues to display target information for the validation experiments. OpenCV image processing library [28] was used to develop this image processing step.

C. Virtual Display using an Augmented Reality Headset

As an alternative to displaying the endoscopic image output on a display monitor, a HMD can be used. The main advantage of using HMD is that users do not have to focus on the display monitor when using the head control system. In this study, a Microsoft Hololens AR headset (Microsoft, USA) is used as HMD. One benefit of using such an AR headset is that the user is still able to see their hands and surroundings while wearing the HMD. When using Hololens as the visual output device, the head measurement IMU is attached to the left side of the headset, as shown in Fig. 3.

A Hololens Universal Windows Platform (UWP) application was developed based on the *ROS#* library [29]. This library will set up the communication between the UWP application running in the Hololens to the image processing module on a separate computer, in the form of JavaScript Object Notation (JSON) packets using a WebSocket protocol. The Hololens UWP application simply contains a virtual display plane that shows the processed endoscope output image, which stays in the center of the user's field of view while the head control system is running.

III. EXPERIMENTAL VALIDATION

User studies were conducted in which three modes of endoscope use were evaluated: the head-controlled method using a display monitor, the head-controlled method using Hololens, and the manual control method.

A. Evaluation Criteria

The performance metric is the time it takes for the user to point the center of the endoscope image to a target (shown in Fig. 3), defined as the reaching time. The accuracy requirement is prescribed to the task by changing the target diameter. A smaller target diameter corresponds to a more difficult task, which is characterized by a higher index of difficulty (ID).

Using a prescribed index of difficulty to define the target size, we aim to achieve a more controlled difficulty level for the camera-pointing experiment. The relation between time and accuracy is established in the context of a 2D camera-pointing experiment using Fitts' law [30]. The size of the target diameter w is determined by the starting distance to target d and ID, and the relation is given by:

$$\text{ID} = \log_2\left(\frac{d}{w} + 1\right) . \tag{7}$$

The camera-pointing experiments were conducted at two different difficulty settings, which correspond to the target size in the camera-pointing. ID of 2 and 3 were selected to simulate low- and high-accuracy use cases in practical use. The target outline was detected using a binary large object (blob) detector, from which the location of the target centroid can subsequently be calculated.

B. Experimental Scenarios

Reaching time is measured for each of the three usage modes:

a) Manual endoscope usage: For manual control (as shown in Fig. 4a), the endoscope handle can be held in either hand. With the thumb, the control lever can be moved. The control lever is used to flex and extend the tip of the endoscope in the vertical plane. Moving the control lever downward will make the tip bend anteriorly (flexion). Moving it upward will make the tip bend posteriorly (extension). The endoscope camera can be rotated by rotating the handle.

b) Head-controlled endoscope usage: For the head-controlled endoscope (as shown in Fig. 4b), a headband is placed on the participant's head, specifically with the IMU sensor on the left temple of the user. The movement of the image on the screen will correspond to the head movement: when the head moves up and down, the image moves up and down, and the same applies for the left and right motions.

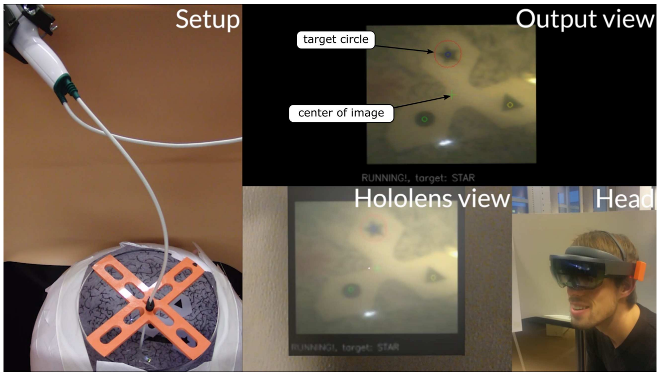
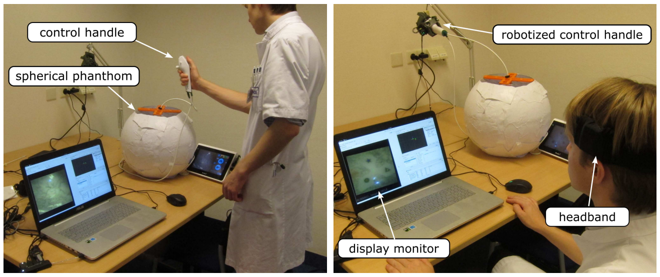


Fig. 3. The experimental setup for validating the feasibility of head-motion control. A spherical phantom which the visual targets are placed inside is shown in the left image. The image output to display monitor (top right image) shows the reticle at the center of the image and the target circle. The bottom center image shows the image output projected onto a virtual display in the Hololens field of view, as seen by the experiment participants. The bottom right image shows a user wearing the Hololens AR headset (with an attached IMU on the side).



(a) (b)

Fig. 4. Experimental setup of the user study (a) using manual control and (b) using the head control interface.

The camera motion is activated only when the foot-switch is pressed.

c) Head-controlled endoscope usage with Hololens AR headset: For head-controlled endoscope with Hololens (as shown in Fig. 3), the Hololens is adjusted to the participant's head and the IMU is placed on the left side of the headset, aligned with the orientation of the plane of the display glasses. The image output is presented to the user on a square image plane that covers most of the Hololens display. This image plane will stay directly on the user's sight, keeping its position steady in the center of the user's field of view. The method of operation for this system is identical to the normal head-controlled endoscope using the headband.

C. User Study Protocol

During the test, the endoscope camera was inserted into a spherical phantom and participants would see an overlaid image stream from the endoscope camera on the screen (Fig. 3). For head-controlled usage modes, the entire control handle mechanism was mounted on a balanced arm that is adjustable in position and able to keep the physical system steady while the device is in operation. The endoscope was fixed in the direction of the insertion depth, for both head-control usage and manual usage. The participants were tasked with moving the reticle (in the center of the image) inside the target circle that overlays the physical visual target. The target point is considered 'reached' when the center reticle is inside of the target circle for at least 0.4 seconds. This requirement is prescribed to force the participant to stop at the target location and not 'reaching' the target by moving quickly past the target location. One test session for each mode of use consists of reaching 20 targets consecutively.

The targets were placed $16.5\,\mathrm{cm}$ away from the endoscope camera on the inside of the spherical phantom, so that all targets were visible in the field of view of the camera and the experiment participants did not need to search for the location of the next target. The targets were positioned at $50°$ elevation angle from the vertical axis inside the phantom to avoid kinematic singularity when the endoscope robot is in a straight configuration. To counteract the muscle memory effect of repeating an identical task in the pointing experiments, two separate tracks were used. The overall shape of the tracks was both the same, however, the targeting order was reversed to counterclockwise for the second track.

TABLE I
EXPERIMENTAL SCHEDULE FOR EACH MODE OF ENDOSCOPE USE.

System	Trial type	Index of Difficulty	N	Track
Head-controlled	training	2	1×20	1
Head-controlled	test	2	1×20	1
Head-controlled	test	3	1×20	2
Manual usage	training	2	1×20	1
Manual usage	test	2	1×20	1
Manual usage	test	3	1×20	2
Hololens-scope	training	2	1×20	1
Hololens-scope	test	2	1×20	1
Hololens-scope	test	3	1×20	2

The complete schedule of the experiment for each participant is presented in Table I. The order of presentation for the three usage modes was alternated for each subsequent participant. The user study lasted about an hour (40 minutes when excluding the Hololens usage mode) per participant.

Training sessions were conducted before test sessions for each mode of use. The main purpose of these training sessions is to familiarize the subject participants with the experimental setup, rather than improving control proficiency, especially since all participants are trained in manual endoscope operation.

D. Participants

Camera-pointing experiments were carried out on two separate occasions with two different sets of participants. All participants were recruited on a voluntary basis and signed an informed consent form beforehand.

The first set of participants have intermediate to expert level in endoscopic skills (non-clinicians, trained using laparoscope and flexible endoscope). A total of eight trained non-clinician participants (6 males and 2 females, age between 23-33) participated in this experiment, 6 of which

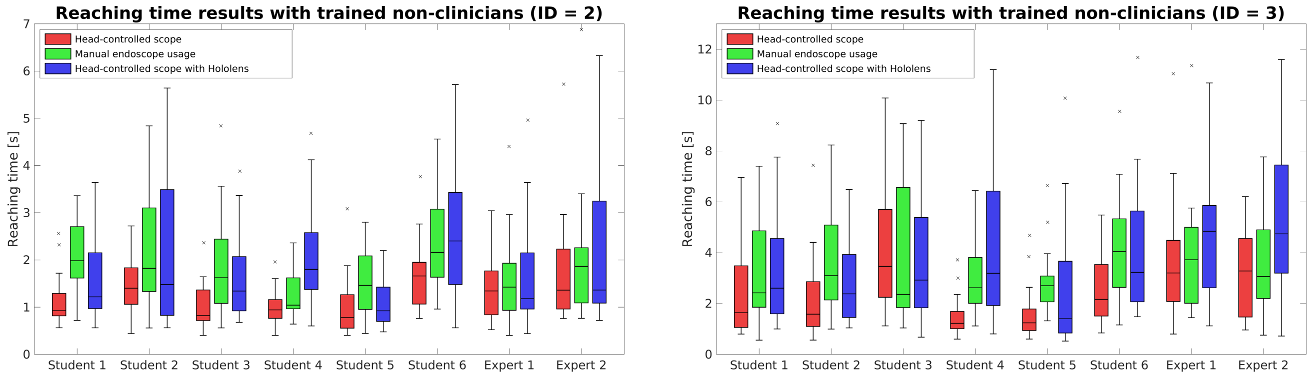


Fig. 5. Motion time results using head-controlled system (red), manual usage (green), and head-controlled system using Hololens (blue). Reaching time for the lower difficulty task (ID = 2) is shown in the left image, while results for the camera-pointing task with higher difficulty (ID = 3) is shown in the right image. Small circles (○) in the plot indicate outliers from the distribution.

are senior Technical Medicine[1] master students who had completed a training in clinical insertion or endoscopy, and 2 are lecturers of an endoscopic training course at the University of Twente.

The second set of participants consisted of four clinicians from the cardiothoracic department in University Medical Center Groningen (UMCG), three surgeons and one surgeon-in-training (males, age between 30-57). Due to the time constraints and the expected latency problem with head-controlled implementation using the Hololens headset, this usage mode was not included in the validation study with clinician participants.

E. Assessments

Statistical analyses were performed using the two-sample t-test to find the difference in reaching time means between head-controlled scope and manual usage. The statistical tests were conducted two times, for an ID of 2 and an ID of 3. Significance level $\alpha = 0.05$ was used for all statistical tests.

Additionally, participants were given questionnaires immediately after the experiments. Subject participants were asked to evaluate statements regarding their user experience with each respective usage method, using bipolar Likert-type five-point scales (‘Strongly disagree’, ‘Disagree’, ‘Neutral’, ‘Agree’, and ‘Strongly agree’). Furthermore, additional statements specific to the usage of the head-controlled system were asked to measure the reception of the proposed system among clinician and non-clinician participants.

IV. RESULTS AND DISCUSSION

A. Reaching time performance

The reaching time results for the non-clinician participants are presented in Fig. 5. Among the non-clinician participants ($N = 8$), at lower difficulty level (ID = 2) the two-sample t-test indicates significantly faster reaching times using the head-controlled scope with display monitor output ($\mu = 1.40\,\mathrm{s}$, $\sigma = 0.416\,\mathrm{s}$) compared to manual scope ($\mu = 1.90\,\mathrm{s}$, $\sigma = 0.413\,\mathrm{s}$), $p = 0.03$. Significantly shorter reaching times were also found at higher difficulty level (ID = 3) using the head-controlled scope with display monitor output ($\mu = 2.71\,\mathrm{s}$, $\sigma = 0.997\,\mathrm{s}$) over using manual scope ($\mu = 3.67\,\mathrm{s}$, $\sigma = 0.540\,\mathrm{s}$), $p = 0.03$. No significant differences were found in the mean reaching time between the use of the Hololens head control system and the manual use.

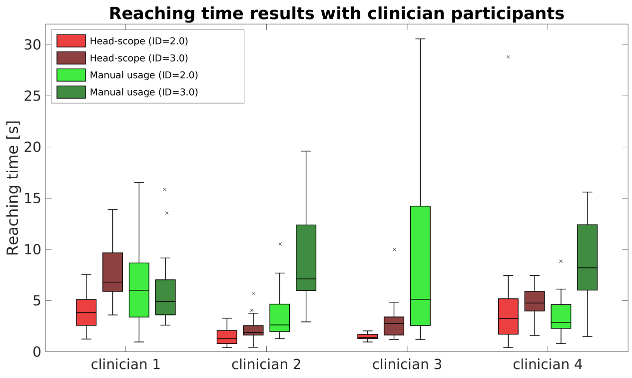


Fig. 6. Clinicians reaching time measurement in using the head-scope system with standard monitor display output (red) and manual scope (green). The small crosses indicate measurement points that are considered as outliers. Manual usage with higher difficulty level (ID = 3) for clinician 3 was not conducted due to his limited availability at the time.

The results of the camera-pointing experiment of the clinician participants are shown in Fig. 6. The results indicate a significant difference in reaching time performance between using the head-controlled ($\mu = 4.44\,\mathrm{s}$, $\sigma = 2.39\,\mathrm{s}$) and manual use ($\mu = 8.23\,\mathrm{s}$, $\sigma = 1.01\,\mathrm{s}$), $p = 0.04$ for clinician participants ($N = 3$) with higher difficulty ID = 3. The results of the experiments at the lower difficulty level (ID = 2) with clinician participants show no significant difference in reaching time between head control and manual use.

The results from the camera-pointing experiment conducted with non-clinician participants suggest that the difference in reaching time between head control and manual usage is larger with student participants compared to the more expert endoscopic course lecturers. This may indicate

---

[1] Technical Medicine is a Masters level program in which students learn to integrate advanced technologies within the medical sciences to improve patient care.

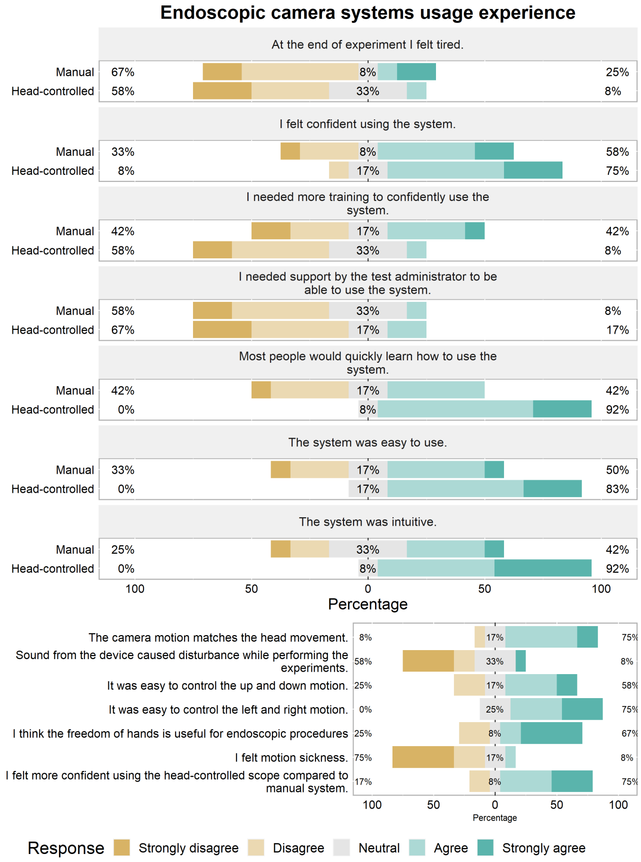


Fig. 7. Results of questionnaires in the user study. Usage experiences were compared between manual usage method and head-controlled method using display monitor output (top image). Participants were also asked to evaluate several statements which are specific to the usage of the head-controlled system (bottom image).

that a shorter learning time is required to use the head-controlled endoscope system, as opposed to manual use of the endoscope.

B. *Questionnaire results*

Questionnaires were given to the subject participants to assess their user experience in controlling the endoscope manually and using the head-controlled endoscope system. The questionnaire responses from all participants ($N = 12$) are presented in a Likert chart shown in Fig. 7.

The most notable differences in user responses between the two systems are regarding the confidence felt while using the system, the amount of training needed, and user friendliness of the respective endoscope systems. User responses indicate that participants felt more confident in using the head-controlled system compared to the manual flexible endoscope. Participants also thought that less training is needed with the head-scope and most people would quickly learn how to use the head-scope system. According to the user responses, the head-controlled system was easier to use and more intuitive compared to the manual usage method.

In terms of the amount of fatigue experienced by the user, the responses indicate that both systems were not fatiguing. However, forty minutes of total experiment time might not have been enough to induce fatigue. More research is required to compare the amount of fatigue caused by each endoscope system in a long-duration task to simulate long operating hours. Most participants also think that little instruction is needed to properly use both the manual and head-scope endoscope systems.

During the validation experiments, we observed that participants learned to ”reset” their head position to a more comfortable resting head position by utilizing the foot-switch clutch mechanism. We expect that operating the head-controlled system around the user’s natural resting position will lead to better ergonomics, especially when operation at large tip deflection is necessary, such as in the case of endoscope retroflexion.

C. *Limitations*

Although the reaching time performance of the head control interface for a flexible endoscope has been validated in this study, some limitations should be noted. The user study was performed in a synthetic environment that may not reflect the endoscope motion profile in practical cases.

The quality of the endoscope image, specifically the resolution (220×180 px) and the color contrast, was very limiting according to the qualitative evaluation by the clinicians. This substandard image quality adds to the difficulty of the camera-pointing experiment, especially in comparison to non single-use cameras (1920 × 1080 px resolution) which are routinely used by clinician participants.

During the development process, a limitation was discovered that a significant amount of latency and jitter was present (93 ± 118 ms later than the output of the display monitor) during the transmission of image data to the AR headset. This delay is caused by two main reasons: an image compression step must be performed before transmission (as the Hololens have limited processing power), and the headset can only be connected wirelessly, adding an extra encoding/decoding step, bandwidth limitation, and increasing jitter due to signal interference in the transmission process.

V. CONCLUSION

In this study, a head-motion controlled robotic endoscope has been realized on a flexible endoscope, in combination with an AR headset as an alternative display output.

The system developed in this study is a proof of concept that a head-motion controlled endoscope offers a possible alternative to control the camera in an MIS procedure. The results of the camera-pointing experiments show that the developed head-controlled endoscope has faster reaching time performance in the higher index of difficulty tasks with clinician participants ($p = 0.04$), and in both low and high difficulty tasks with non-clinician participants against the manual usage of the flexible endoscope ($p = 0.003$ for ID $= 2$ and $p = 0.03$ for ID $= 3$). On the basis of our experimental results and clinician feedback, the head-motion controlled endoscope is a promising approach that requires further investigation and development to reach clinical use.

ACKNOWLEDGMENT

The work in this paper was supported in part by ITEA under the 17021 IMPACT and 20044 ASSIST project.

REFERENCES

[1] C. C. J. Alleblas, A. M. de Man, L. van den Haak, M. E. Vierhout, F. W. Jansen, and T. E. Nieboer, “Prevalence of Musculoskeletal Disorders Among Surgeons Performing Minimally Invasive Surgery: A Systematic Review,” *Annals of Surgery*, vol. 266, no. 6, pp. 905–920, Dec. 2017.

[2] L. P. Aitchison, C. K. Cui, A. Arnold, E. Nesbitt-Hawes, and J. Abbott, “The ergonomics of laparoscopic surgery: a quantitative study of the time and motion of laparoscopic surgeons in live surgical environments,” *Surgical Endoscopy*, vol. 30, no. 11, pp. 5068–5076, Nov. 2016.

[3] P. R. Armijo, C.-K. Huang, R. High, M. Leon, K.-C. Siu, and D. Oleynikov, “Ergonomics of minimally invasive surgery: an analysis of muscle effort and fatigue in the operating room between laparoscopic and robotic surgery,” *Surgical Endoscopy*, vol. 33, no. 7, pp. 2323–2331, Jul. 2019.

[4] A. T. Schlussel and J. A. Maykel, “Ergonomics and Musculoskeletal Health of the Surgeon,” *Clinics in Colon and Rectal Surgery*, vol. 32, no. 6, pp. 424–434, Nov. 2019.

[5] C. D. P. van’t Hullenaar, P. Bos, and I. A. M. J. Broeders, “Ergonomic assessment of the first assistant during robot-assisted surgery,” *Journal of Robotic Surgery*, vol. 13, no. 2, pp. 283–288, Apr. 2019.

[6] A. Alhusuny, M. Cook, A. Khalil, and V. Johnston, “Visual symptoms, Neck/shoulder problems and associated factors among surgeons performing Minimally Invasive Surgeries (MIS): A comprehensive survey,” *International Archives of Occupational and Environmental Health*, vol. 94, no. 5, pp. 959–979, Jul. 2021.

[7] B. Penev, “The Experience of Robotic Camera Holder During Laparoscopic Radical Retropubic Prostatectomy,” *Urology and Nephrology*, vol. 1, pp. 92–95, Apr. 2019.

[8] N. A. Desai, A. L. Gubbels, and M. Hibner, “The Surgical Assistant in Robotic-Assisted Laparoscopy,” in *Textbook of Gynecologic Robotic Surgery*, A. El-Ghobashy, T. Ind, J. Persson, and J. F. Magrina, Eds. Springer International Publishing, 2018, pp. 235–238.

[9] J. Douissard, M. E. Hagen, and P. Morel, “The da Vinci Surgical System,” in *Bariatric Robotic Surgery: A Comprehensive Guide*, C. E. Domene, K. C. Kim, R. Vilallonga Puy, and P. Volpe, Eds. Springer International Publishing, 2019, pp. 13–27.

[10] S. Gillen, B. Pletzer, A. Heiligensetzer, P. Wolf, J. Kleeff, H. Feussner, and A. Fürst, “Solo-surgical laparoscopic cholecystectomy with a joystick-guided camera device: a case–control study,” *Surgical endoscopy*, vol. 28, no. 1, pp. 164–170, 2014.

[11] 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.

[12] S.-Y. Yen, H.-E. Huang, G.-S. Lien, C.-W. Liu, C.-F. Chu, W.-M. Huang, and F.-M. Suk, “Automatic lumen detection and magnetic alignment control for magnetic-assisted capsule colonoscope system optimization,” *Scientific reports*, vol. 11, no. 1, pp. 1–10, 2021.

[13] A. Sivananthan, A. Kogkas, B. Glover, A. Darzi, G. Mylonas, and N. Patel, “A novel gaze-controlled flexible robotized endoscope; preliminary trial and report,” *Surgical Endoscopy*, vol. 35, no. 8, pp. 4890–4899, Aug. 2021.

[14] S. G. Alletti, C. Rossitto, S. Cianci, E. Perrone, S. Pizzacalla, G. Monterossi, G. Vizzielli, S. Gidaro, and G. Scambia, “The Senhance™ surgical robotic system (“Senhance”) for total hysterectomy in obese patients: a pilot study,” *Journal of robotic surgery*, vol. 12, no. 2, pp. 229–234, 2018.

[15] A. M. Martinez, J. V. Gomez, R. O. Flores, and D. L. Espinoza, “Postural mechatronic assistant for laparoscopic solo surgery (pmass),” *Surgical endoscopy*, vol. 23, no. 3, p. 663, 2009.

[16] Y. Huang, W. Lai, L. Cao, E. Burdet, and S. J. Phee, “Design and Evaluation of a Foot-Controlled Robotic System for Endoscopic Surgery,” *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 2469–2476, Apr. 2021.

[17] J. N. Bucher, K. Bruewer, L. J. Dietz, N. Trebesius, J. Hidding, M. Wysocki, M. B. Schoenberg, J. Werner, and K. Karcz, “A Robotic Camera Holder Controlled by Head Movements: Exploring This New Robot-Surgeon Interface,” *Surgical Innovation*, vol. 27, no. 5, pp. 499–506, Oct. 2020.

[18] P. J. M. Wijsman, L. Molenaar, C. D. P. van‘t Hullenaar, B. S. T. van Vugt, W. A. Bleeker, W. A. Draaisma, and I. A. M. J. Broeders, “Ergonomics in handheld and robot-assisted camera control: a randomized controlled trial,” *Surgical Endoscopy*, vol. 33, no. 12, pp. 3919–3925, Dec. 2019.

[19] J. Sandoval, M. Laribi, J. Faure, C. Brèque, J. Richer, and S. Zeghloul, “Towards an Autonomous Robot-Assistant for Laparoscopy Using Exteroceptive Sensors: Feasibility Study and Implementation,” *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 6473–6480, Oct. 2021.

[20] J. M. Ali, K. Lam, and A. S. Coonar, “Robotic Camera Assistance: The Future of Laparoscopic and Thoracoscopic Surgery?” *Surgical Innovation*, vol. 25, no. 5, pp. 485–491, Oct. 2018.

[21] J.-B. Lin, M.-Q. Kang, J.-F. Chen, Q. Du, X. Li, F.-C. Lai, and Y.-R. Tu, “Transareolar single-port endoscopic thoracic sympathectomy with a flexible endoscope for primary palmar hyperhidrosis: a prospective randomized controlled trial,” *Annals of Translational Medicine*, vol. 8, no. 24, p. 1659, Dec. 2020.

[22] B. C. S. Hamilton and Y. B. Gesthalter, “Keep it Clean: Novel Use of Sterile Disposable Flexible Bronchoscopes for Pleuroscopy,” *Journal of Bronchology & Interventional Pulmonology*, vol. 29, no. 1, p. e2, Jan. 2022.

[23] B. Seeliger and L. L. Swanström, “Robotics in flexible endoscopy: current status and future prospects,” *Current Opinion in Gastroenterology*, vol. 36, no. 5, pp. 370–378, Sep. 2020.

[24] M. W. Gifari, H. Naghibi, S. Stramigioli, and M. Abayazid, “A review on recent advances in soft surgical robots for endoscopic applications,” *The international journal of medical robotics + computer assisted surgery: MRCAS*, vol. 15, no. 5, p. e2010, Oct. 2019.

[25] A. K. Lim, J. Ryu, H. M. Yoon, H. C. Yang, and S.-k. Kim, “Ergonomic effects of medical augmented reality glasses in video-assisted surgery,” *Surgical Endoscopy*, vol. 36, no. 2, pp. 988–998, Feb. 2022. [Online]. Available: https://doi.org/10.1007/s00464-021-08363-8

[26] 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, 2018.

[27] R. Reilink, G. de Bruin, M. Franken, M. A. Mariani, S. Misra, and S. Stramigioli, “Endoscopic camera control by head movements for thoracic surgery,” in *Proc. of 3rd IEEE RAS EMBS International Conference on Biomedical Robotics and Biomechatronics*, Sep. 2010, pp. 510–515.

[28] OpenCV, “Open source computer vision library,” https://github.com/opencv/opencv, 2015.

[29] D. Whitney, “UWP compatible ROS# library,” https://github.com/dwhit/ros-sharp, 2018.

[30] R. W. Soukoreff and I. S. MacKenzie, “Towards a standard for pointing device evaluation, perspectives on 27 years of fitts’ law research in hci,” *International journal of human-computer studies*, vol. 61, no. 6, pp. 751–789, 2004.