# Evaluating Human-Robot Interfaces for Maneuvering Surgical Laparoscopes using Robotic Scope Assistant Systems

SOFIA BASHA, MALEK ANBATAWI, NIHAL ABDURAHIMAN, and JHASKETAN PADHAN, Department of Surgery, Hamad Medical Corporation, Doha, Qatar
VICTOR M. BAEZ, Department of Electrical and Computer Engineering, University of Houston, Houston, Texas, USA
ABDULLA AL-ANSARI, Department of Surgery, Hamad Medical Corporation, Doha, Qatar
PANAGIOTIS TSIAMYRTZIS, Department of Mechanical Engineering, Politecnico di Milano, Milano, Italy
AARON T. BECKER, Department of Electrical and Computer Engineering, University of Houston, Houston, Texas, USA
NIKHIL V. NAVKAR, Department of Surgery, Hamad Medical Corporation, Doha, Qatar

Robotic scope assistant systems allow surgeons to adjust the operative field view during surgery by robotically maneuvering laparoscopes. A Human-Robot Interface (HRI) is used for issuing commands to these systems, with an interaction mode mapping these commands to laparoscope movements. Optimizing the HRI and interaction mode can streamline laparoscope positioning as well as reduce cognitive workload, helping the surgeon focus on the surgical procedure. Comparing and assessing various HRIs and interaction modes is essential for efficient laparoscope maneuvering. This study evaluates HRIs based on head-motion, eye-motion, hand-motion, and voice-input operating under three interaction modes (namely: discrete, continuous, and threshold). The participants performed a user study comparing different HRIs under two simulated surgical scenarios (one in a real environment and the other in a virtual environment). The results indicated that head and eye-based HRIs performed well in continuous interaction mode, while the voice-based interface suffered from a delay. Conversely, hand-based HRIs demonstrated superior performance in both scenarios across all evaluation parameters. The study provides a benchmark for the comparison of different HRIs and provides insights into the effectiveness, limitations, and potential advantages of different HRIs.

Research reported in this publication was supported by the Qatar Research Development and Innovation Council Academic Research Grant (ARG) award ARG01-0430-230047 and in part by National Priority Research Program (NPRP) award NPRP13S-0116-200084 from the Qatar National Research Fund (a member of the Qatar Foundation). All opinions, findings, conclusions, or recommendations expressed in this work are those of the authors and do not necessarily reflect the views of our sponsors.

Authors' Contact Information: Sofia Basha, Department of Surgery, Hamad Medical Corporation, Doha, Qatar; e-mail: sofiashafath@gmail.com; Malek Anbatawi, Department of Surgery, Hamad Medical Corporation, Doha, Qatar; e-mail: maleka@alumni.cmu.edu; Nihal Abdurahiman, Department of Surgery, Hamad Medical Corporation, Doha, Qatar; e-mail: nihalabdu1@gmail.com; Jhasketan Padhan, Department of Surgery, Hamad Medical Corporation, Doha, Qatar; e-mail: JPadhan@hamad.qa; Victor M. Baez, Department of Electrical and Computer Engineering, University of Houston, Houston, Texas, USA; e-mail: vjmontan@central.uh.edu; Abdulla Al-Ansari, Department of Surgery, Hamad Medical Corporation, Doha, Qatar; e-mail: aalansari1@hamad.qa; Panagiotis Tsiamyrtzis, Department of Mechanical Engineering, Politecnico di Milano, Milano, Italy; e-mail: panagiotis.tsiamyrtzis@polimi.it; Aaron T. Becker, Department of Electrical and Computer Engineering, University of Houston, Houston, Texas, USA; e-mail: atbecker@central.uh.edu; Nikhil V. Navkar (corresponding author), Department of Surgery, Hamad Medical Corporation, Doha, Qatar; e-mail: nnavkar@hamad.qa.


ACM 2573-9522/2025/5-ART52
https://doi.org/10.1145/3722121

CCS Concepts: • **Human-centered computing → Human computer interaction (HCI)**; **HCI design and evaluationmethods**; **User studies**;

Additional Key Words and Phrases: Human-robot Interface (HRI), Laparoscopic Surgery, Robotic Scope Assistant Systems

**ACM Reference format:**
Sofia Basha, Malek Anbatawi, Nihal Abdurahiman, Jhasketan Padhan, Victor M. Baez, Abdulla Al-Ansari, Panagiotis Tsiamyrtzis, Aaron T. Becker, and Nikhil V. Navkar. 2025. Evaluating Human-Robot Interfaces for Maneuvering Surgical Laparoscopes using Robotic Scope Assistant Systems. *ACM Trans. Hum.-Robot Interact.* 14, 3, Article 52 (May 2025), 20 pages.
https://doi.org/10.1145/3722121

## 1 Introduction

Minimally invasive laparoscopic procedures improve surgical outcomes by minimizing tissue damage, reducing post-operative pain and discomfort, and shortening hospital stays [1–3]. In these procedures, surgeons operate on tissue using laparoscopic instruments such as graspers, scissors, dissectors, needle holders, and other specialized tools. These instruments are inserted by creating small incisions (typically less than 1 cm in diameter) on the patient. This work focuses on *laparoscopes*, an elongated tubular surgical instrument comprised of a lens and light source at the distal end. It is used to visualize the operative field (i.e., the area where surgery is performed using laparoscopic instruments). The operative field view is displayed to the surgeon on a viewing screen. Since the surgeon uses both hands to manipulate laparoscopic instruments, a skilled assistant is required to steer the laparoscope. This assistant's task is to continuously provide necessary visualization of the operative field (Figure 1(a)).

During the laparoscopic surgery, a laparoscope is maneuvered by tilting, panning, and inserting/retracting it along an incision point on the abdominal wall (Figure 1(b)). The surgeon relies on an assistant to maneuver the laparoscope and provides a closeup yet unobstructed view of the operative field (Figure 1(c)). However, a surgeon's definition of an optimal view of the operative field may not always align with the assistant's definition. Different definitions can lead to bias and cooperation conflict [4]. Such miscommunication between the surgeon and the assistant may lead to complications and increase the duration of the surgical procedure [5–7]. Furthermore, assistants commonly experience fatigue, distractions, and natural hand tremors, which can lead to sudden shifts in the view on the screen and suboptimal visualization [7–9].

To address these challenges, robotic scope assistant systems have been developed. These systems enable the operating surgeon to directly control the laparoscope movements and adjust the operative field view during surgery. Robotic systems mitigate manual errors caused due to human fatigue during prolonged surgeries or misinterpretation of the surgeon's instructions for laparoscope maneuvering. The view of the operative field remains steady (and free from tremors) and can be interactively changed by issuing actuation commands to the robotic scope assistant system for maneuvering the laparoscope.

To effectively receive and interpret these actuation commands from the surgeon, an ergonomic **human-robot interface (HRI)** and an intuitive mode of interaction are required. An HRI acts as a modality to provide actuation commands to the robotic scope assistant system, whereas an interaction mode maps these actuation commands to movement of the laparoscope distal end. An ideal combination of HRI and interaction mode reduces the time required for positioning the laparoscope as well as minimizes the cognitive workload, allowing the surgeon to focus on the surgical procedure. Therefore, it is imperative to understand the functioning of different HRIs and

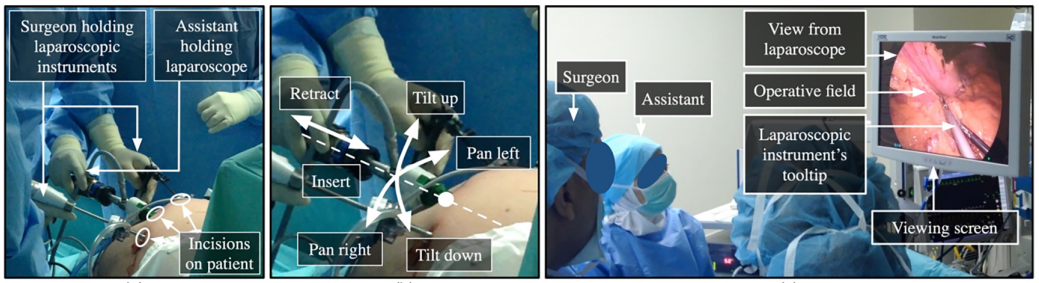


Fig. 1. (a) A typical laparoscopic surgery setup. (b) Maneuvering of the laparoscope by the assistant as per the surgeon's instruction to focus on the operative field. (c) The operative field view captured by the laparoscope is rendered on a viewing screen.

interaction modes to efficiently maneuver a laparoscope. The objective of this work is to compare and assess the performance of various HRIs under different interaction modes for maneuvering a laparoscope. In this work, a user study is conducted to assess the functionality of HRIs, each employing distinct input methods such as head-motion, hand-motion, eye-motion, and voice commands. In addition, three interaction modes are utilized, namely, continuous, threshold, and discrete. The novelty of this study lies in both the comprehensive performance analysis of multiple HRIs and the introduction of a continuous interaction mode specifically designed for head and eye-motion-based control of the laparoscope. In this study, laparoscope maneuvering tasks are simulated in virtual and real environments to assess the interaction of the operator with the robotic scope assistant system via the HRI and interaction mode.

## 2 Related Work

During a laparoscopic surgery, both hands of the operating surgeon are holding surgical instruments. This led to development of different HRIs based on the operating surgeon's foot motion [10, 11], hand-motion [12], head-motion [13, 14], or voice commands [11]. The HRI based on foot motion includes a foot pedal with multiple switches placed on the floor in front of the surgeon [10, 11]. Pressing multiple switches using the ankle and toe detects the motion of the foot during the surgery and causes specific movement of the laparoscope. The HRI based on hand-motion consists of a small hand-held miniature keypad with multiple buttons for laparoscope movements [12]. The keypad is attached to the rear-end of the laparoscopic instrument during the surgery. Pressing a button generates laparoscope motion in a particular direction. The HRI based on head-motion works under two interaction modes. In both modes, a gyro-sensor is strapped to a surgeon's head. When the head is moved in a particular way, the sensor detects the head-motion and registers a specific command for scope movement. In the first mode, when the surgeon presses a foot switch (or a hand-held button), the scope moves in that direction until the switch is pressed again. In the second mode, the surgeon's head behaves as a three-axis gimbal and the scope continues to move in the set direction while the head remains tilted. The HRI based on voice commands requires the surgeon to speak a word (such as "up," "down," "left," "right") into the microphone to set the required direction for scope movement [11]. Once the word is recognized and the direction is set, the surgeon presses a foot switch (or a hand-held button) to activate the selected motion. Table 1 summarizes commercial robotic scope assistance systems [15–23], their corresponding HRIs, and their application in various surgical scenarios.

Several studies evaluated robotic scope assistant systems for laparoscopic surgeries against human assistants. For instance, Lin et al. [24] found no significant differences in surgical outcomes

Table 1. HRIs Used in Commercial Robotic Scope Assistant Systems

| Robotic Scope Assistant System | Foot | Hand | Voice | Head | Surgical Scenarios |
|---|---|---|---|---|---|
| EMARO—Riverfield [15] | | | | ✓ | Inguinal hernia repair |
| FreeHand—HYB [16] | | | | ✓ | General, gynecology, urology, thoracic |
| SOLOASSIST—AKTORmed [17] | | ✓ | ✓ | | General, urology, gynecology, thoracic, cardiac |
| MTG-H100—HIWIN [18] | ✓ | | | | General, urology, gynecology, colon, rectal |
| RoboLens—SinaMed [19] | ✓ | | ✓ | | Cholecystectomy, ovarian cystectomy |
| ViKY—EndoControl [20][a] | | | ✓ | | Gynecology, abdominal, thoracoscopic |
| LAPMAN—MedSys [21][a] | | ✓ | | | Gynecology surgeries |
| Naviot—Hitachi [22][a] | | ✓ | | | Thoracoscopic, cholecystectomy |
| AESOP—Computer Motion [23][a] | ✓ | ✓ | ✓ | | Thoracic surgeries |

[a]Defunct commercial systems.

between robotic scope assistant systems and human assistant groups. Similarly, in colorectal cancer resections, robotic scope assistant systems demonstrated comparable performance to human assistants, with no significant differences in operating time or patient outcomes [25]. Another study by Kavoussi et al. [26] compared human-controlled laparoscope movement with robot-controlled movement. Despite the similar performance, robotic scope assistant systems were noted for offering greater precision, which is seen as a major advantage.

Further advancements in this area of research included studies exploring different methods of automating camera positioning for laparoscopic surgeries [27]. A recent review by Hamza et al. [28] and Nguyen et al. [29] summarized the different interfaces used in robotic scope assistant systems. To assess the effectiveness of different HRIs for maneuvering a laparoscope using a robotic scope assistance system, comparative studies were conducted [30–35]. Table 2 summarizes these studies, detailing the interfaces used in each study, the testing environment of the experiment, the metrics used for evaluation, and the outcome. While these studies provide insights into the working of HRIs, there were certain limitations. First, these previous studies compared at most two different HRIs. Most of the studies [30–33] compared HRIs based on foot motion (via a foot pedal) versus voice commands. In some of these studies ([30] and [32]), the outcome indicated a preference for foot motion over voice commands. This preference was due to misinterpretation/misrecognition while pronouncing the voice commands and corresponding lag in scope movement. Although foot motion was preferred over voice commands, eye-foot coordination was difficult as it required the operator to look down often to select the right pedal [34–36]. In another study [34], voice commands were preferred over hand-motion (via a keypad mounted on the surgical instrument). In the remaining studies ([31, 33], and [34]), both the HRIs were preferred. Thus, a user-study that compares multiple HRIs was lacking. Second, due to variations in testing environments and experimental tasks, the outcomes from these studies cannot be directly compared. Some studies utilized simulated surgical environments, others employed animal models, and a few were evaluated by surgeons performing live surgeries on humans. Also, the robotic scope assistant system used in these studies varied. Thus, a standardized environment to benchmark the HRIs was absent. Third, nearly all the studies measured total task completion time and error count as evaluation metrics. In the proposed study, we introduce additional metrics alongside those used in previous research. These metrics were designed specific to the task and provide more insights for evaluating interaction using an HRI.

While previous studies lay the groundwork for assessing various HRIs to maneuver laparoscopes, to the best of our knowledge, there has been no prior work done that comprehensively evaluates multiple HRIs (including head-motion) across a standard surgical use case and in conjunction with a single robotic scope assistant system. Such a user study could provide insights into the

Table 2. Previous Comparative Studies and Their Settings

| Study | Type of the HRI | | | | | Testing Environment/ Task | Evaluation Metrics | | | Outcome |
|---|---|---|---|---|---|---|---|---|---|---|
| | Foot | Hand | Voice | Head | Eye | | Duration | Accuracy | Others | |
| [30] | ✓ | | ✓ | | | Peg board transfer task, trajectory following, and LED board targeting | ✓ | ✓ | Questionnaire | Foot motion is better than voice |
| [31] | ✓ | | ✓ | | | Six training procedures on animal model | ✓ | ✓ | Setup and removal time | Both HRIs were preferred |
| [32] | ✓ | | ✓ | | | 30 laparoscopic procedures | | | Surgeon feedback | Foot motion is better than voice |
| [33] | ✓ | | ✓ | | | Targeting task on pelvic trainer | ✓ | ✓ | Durability, Operator-interface failure | Both HRIs were preferred |
| [34] | | ✓ | ✓ | | | Surgeries performed on cadavers and animals | | | Surgeon evaluation | Both HRIs were preferred |
| [35] | | ✓ | ✓ | | | Gynecological endoscopic surgeries | ✓ | | | Voice is better than hand |
| Proposed study | | ✓ | ✓ | ✓ | ✓ | Path following task and targeting task in simulated environment | ✓ | ✓ | Clutch count, error duration | Presented in results section |

effectiveness, limitations, and potential advantages of different HRIs. These insights could guide the development of efficient and intuitive HRIs for robotic scope assistant systems, eventually improving surgical outcomes. The development and testing of the five HRIs and three interaction modes, with the novel addition of a continuous mode, is the primary contribution of this work. Since such a comprehensive evaluation has not been done before, this article addresses the gap in the literature, offering a better understanding of HRIs performance in a standardized surgical setting.

In addition to laparoscopic procedures performed with rigid scopes, robotic scope assistant systems have also been developed for endoscopic procedures using flexible endoscopes. These include commercial systems such as the Avicenna Roboflex [37], which is controlled via a joystick, and the Endodrive system [38], which utilizes a footswitch, as well as research prototypes based on head-motion [39, 40], eye-gaze [41, 42], and joystick [43, 44]. Given the successful application of these HRIs in managing flexible endoscopes, adapting them for rigid laparoscope manipulation is a promising approach. Therefore, in this study, we incorporate these HRIs and systematically compare their performance in maneuvering a rigid laparoscope.

## 3 Methods

### 3.1 HRIs and Interaction Modes

Five different HRIs (as shown in Figure 2) were used to provide input commands for laparoscope maneuvering. Details of the HRIs are presented in Table 3. The head-motion was acquired using an optical tracking system (TrackIR5—Natural Point). The hand-motion was implemented using two distinct HRIs: a compact keypad (ExpressKeyTM—Wacom) and a 3D joystick (SpaceMouse—3DConnexion). The voice commands were incorporated using an off-the-shelf microphone. An interface (Eye Tracker—Tobii) based on eye-motion was also used. Foot motion was excluded from the study as it required a custom-made programmable foot pedal [30, 45] or commercial programmable foot switch (which are only available to surgical robotics companies [46, 47]). Since six distinct movements were necessary for laparoscope control, a foot pedal comprising of six individual switches or a mechanism capable of distinguishing six different foot movements was critical. However, such

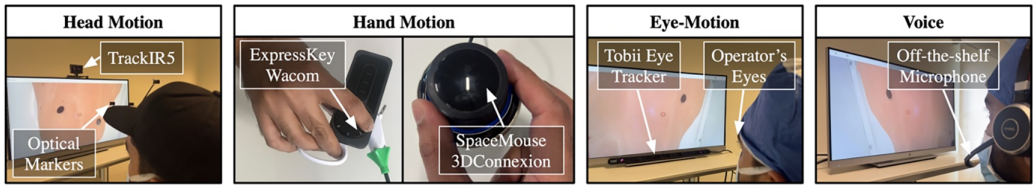


Fig. 2. HRIs used to provide input for maneuvering the laparoscope using the robotic scope assistant system.

Table 3. Devices Used for HRI using Different Interaction Mode

| Devices | C | D | T | Clutch Usage | Interaction of the Operator with the Device |
|---|---|---|---|---|---|
| **Head-motion** TrackIR5—Natural Point | ✓ | ✓ | ✓ | ✓ | A lightweight clip with optical markers is attached to a cap wore by the operator. An optical tracking system placed on the viewing screen tracks translations of operator's head and maps it to the motion of surgical laparoscope tip. |
| **Hand-Motion** ExpressKey Keypad—Wacom | | ✓ | | | A remote with keys is attached to a surgical instrument. Individual keystrokes (of six different keys) by operator's index finger are used to provide unidirectional motion commands to the surgical laparoscope tip. |
| **Hand-Motion** SpaceMouse—3DConnexion | | ✓ | ✓ | | Operator uses index finger and thumb to move the knob of the interface representing 3D joystick. The displacement of the knob from its initial position generates commands for motion of the surgical laparoscope tip. |
| **Eye-Motion** Tobii Eye Tracker | ✓ | ✓ | ✓ | ✓ | An infrared optical sensor (affixed at the bottom of the screen) is used to track operator's eye-motion. The translational motion of operator's eye pupils is mapped to the motion of surgical laparoscope tip. |
| **Voice** Off-the-shelf microphone | | ✓ | | ✓ | VOSK offline speech recognition API is used to identify operator's voice commands (Move-In, Move-Out, Move-Left, Move-Right, Move-Up, and Move-Down) and generate unidirectional discrete motion of the scope tip. |

C, continuous; D, discrete; T, threshold.

a device was not available commercially. To avoid inconsistent outcomes due to improper use of foot interface, foot pedal as an interface to issue actuation commands was not used. Instead, a foot pedal (Savant Elite2—Kinesis) in the form of a clutch was used. For head-motion and eye-motion, the laparoscope only moved if the clutch was pressed. This enabled the operator to ergonomically reposition their head without causing any laparoscope movements. For voice-based control, once a voice command was recognized, the clutch was used as a switch to start and stop the laparoscope movement.

The input from the HRIs was mapped to the movement of the laparoscope's distal end using three different interaction modes: continuous, discrete, and threshold. The continuous interaction mode was applicable to "Head-motion" and "Eye-motion." In this mode, the position of operator's head (or eye) was mapped to the position of the laparoscope tip. As the operator moved their head, the laparoscope tip followed, leading to multidirectional continuous (variable velocity) movements.

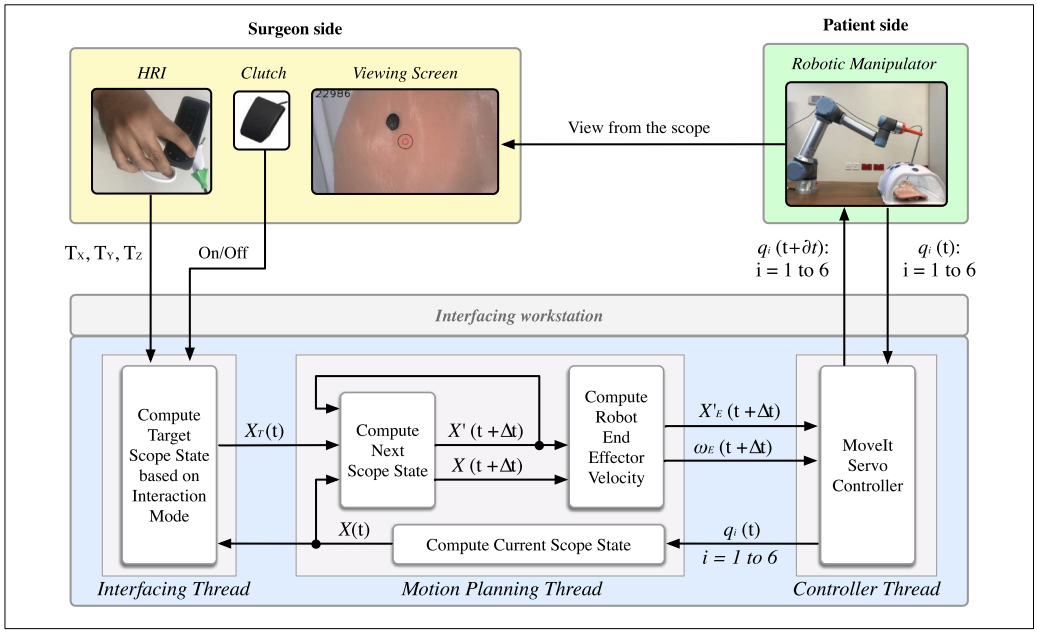


Fig. 3. Processing performed by the interfacing workstation to translate the user input acquired from the HRI into actuation commands for the robotic manipulator.

In discrete mode, the input provided via the HRI was translated into laparoscope tip motion along a discrete direction (i.e., +X, −X, +Y, −Y, +Z, or −Z) with a constant velocity. This enabled the operator to provide unidirectional commands (up, down, left, right, in, and out) one at a time. In threshold mode, the HRI acted as a three-axis gimbal. If the new position recorded by the HRI exceeded the current position by a threshold, the laparoscope tip started drifting in that direction at a constant velocity. The threshold interaction mode was not feasible for "Keypad" and "Voice." These HRIs and their interaction modes are presented in Table 3. The three interaction modes employed distinct operational mechanisms, each imposing varying levels of workload on the operator. While the continuous mode enabled precise movement through direct translation, it could lead to quicker physical exhaustion. The other two modes might alleviate this issue but at the cost of accuracy. These three modes were thus introduced to provide diverse study conditions and enhance the evaluation process.

### 3.2 Processing the Input Provided by HRI

The architecture of the robotic scope assistant system used in this study is presented Figure 3. An operator provided input using an HRI and a clutch. The clutch was used to toggle the on/off status of the system (the laparoscope was maneuvered only when the clutch was on). A zero-degree laparoscope (HD view, 1 cm diameter, and 33 cm length) used in surgical training was attached to a robotic manipulator (UR5e—Universal Robots) using a 3D-printed connector. The view acquired from the laparoscope was rendered to the operator on a viewing screen. The interfacing workstation receives the input (comprising of commands from the HRI, clutch status, and current state of the robotic manipulator), processes them to compute the next state for the robotic manipulator, and sends actuation commands to the robotic manipulator.

Three parallel running threads: the *Interfacing Thread*, the *Motion Planning Thread* running at $1/\Delta t$ Hz, and the *Controller Thread* running at $1/\partial t$ Hz are used by the interfacing workstation to actuate the robotic manipulator based on the input received from the HRI. The *Interfacing Thread* receives the current state of the laparoscope ($X(t)$), the relative translational motion of the HRI ($T_X, T_Y$ and $T_Z$), and the clutch state (On/Off) to compute the target laparoscope's tip position ($X_T(t)$) based on the interaction mode. The current position $X(t)$ of the laparoscope tip is computed based on joint positions $q_i(t) : i = 1$ to 6 of the robotic manipulator. To ensure a smooth continuous velocity profile as the laparoscope tip reaches the target laparoscope position for the next time instant $t + \Delta t$, the position $X(t + \Delta t)$ and velocity $X'(t + \Delta t)$ of the laparoscope tip are computed using a damped spring model (Appendix A.1). The linear velocity $X'_E(t + \Delta t)$ (Appendix A.2) and angular velocity $\omega_E(t + \Delta t)$ (Appendix A.3) of the robot's end-effector are computed by the *Motion Planning Thread* based on the laparoscope's tip position and velocity, such that a **remote center of motion (RCM)** is maintained at the incision point $X_I$. The *Motion Planning Thread* subsequently transmits these velocities to the *Controller Thread* that employs the MoveIt Servo controller (PickNik Robotics) to determine the joint positions $q_i(t + \partial t) : i = 1$ to 6 for the next time instant $t + \Delta t$, which is then sent to the robotic manipulator for activation. The threads were written in the C++ language and deployed in a ROS Melodic environment running on an Ubuntu 18.04 (Bionic Beaver) OS.

### 3.3 User Study

A user study was conducted with 16 subjects from the Department of Surgery at Hamad General Hospital, Qatar. The study was approved by the institutional review board ethical committee (Medical Research Center, Doha, Qatar, approval number MRC-01-20-087). The subjects were academic researchers experienced in developing surgical technologies and had no prior experience using the HRIs for laparoscope manipulation. To acquaint subjects with the HRIs and interaction modes, a preparatory session was conducted before the user study. The preparatory session lasted for approximately 30 minutes per subject, which included an initial 10-minute briefing on the scenarios and tasks, along with practice session using the proposed HRIs and interaction modes. The session ended when the subject demonstrated confidence in mapping the intended movements of the laparoscope based on the input provided using the HRIs to reflect the change in the operating field view displayed on the screen. After the practice session, it was ensured that the user exhibited comfort in operating the interfaces without any issues such as eye pain or fatigue. To minimize learning effect and bias during the user study, a simple randomization technique was used [48]. The subjects performed the study immediately after the practice session and without any breaks. Each subject performed the user study under two simulated scenarios, named Scenario-A and Scenario-B (shown in Figure 4). These scenarios were designed based on standard training curriculum for laparoscope navigation. Each subject also performed the study with all five proposed HRIs across different interaction modes. The preparatory session familiarized the users with the general task of both scenarios, and since each HRI-interaction mode combination was tested only once, any learning bias was minimized. The task duration for both scenarios was short, ranging from 2 to 4 minutes, and the proposed HRIs were designed to ensure user comfort, both of which helped prevent tiredness and its impact on the subjects' performance.

Scenario-A was performed in a virtual environment using the Gazebo simulation platform (Figure 4(a)) and assessed the ability of HRI to maneuver the laparoscope to visualize and follow a predefined path. This scenario was designed to simulate laparoscope movements that follow an instrument tip's trajectory [31, 49]. To ensure consistency across all trials, a predefined curved path was generated. Prior studies have demonstrated similar path-tracing tasks for assessing camera navigation skills [50–52]. A virtual track (~8.5 cm in length) was rendered in the Gazebo environment (Figure 4(c)).

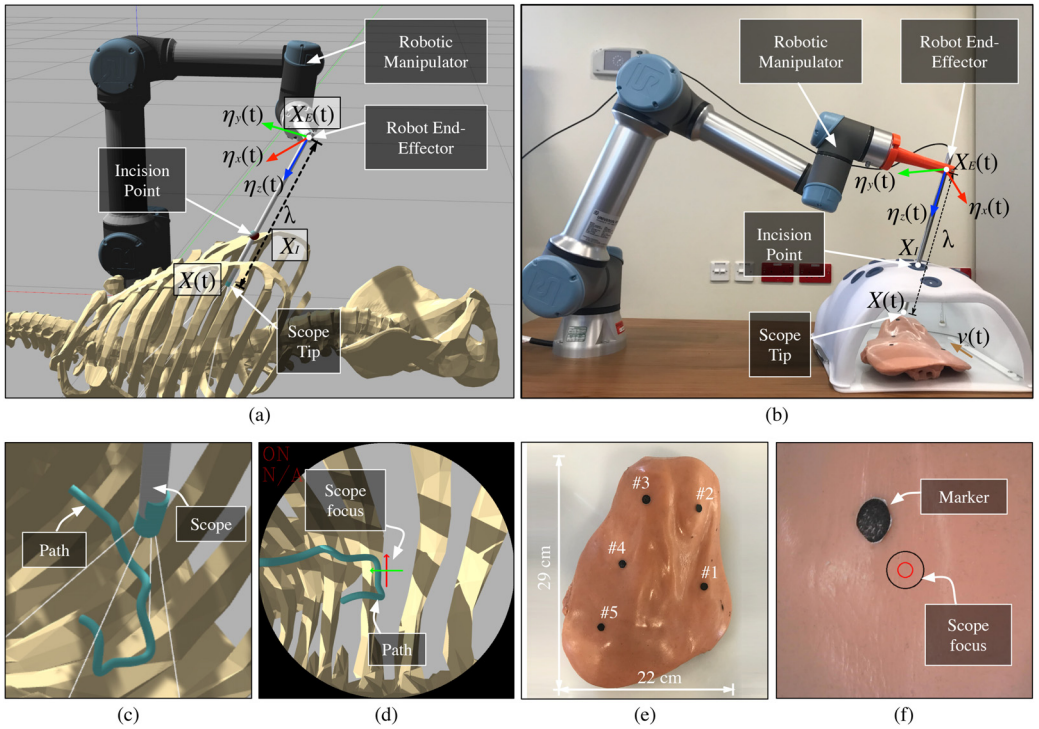


Fig. 4. (a) Setup for steering zero-degree laparoscope in Scenario-A. (b) Setup for steering zero-degree laparoscope in Scenario-B. (c) Operative field in Scenario-A. (d) View of operative field from the laparoscope in Scenario-A. (e) Operative field in Scenario-B. (f) View of operative field from the laparoscope in Scenario-B.

The same track was designed and tested in our previous work, allowing for the evaluation of camera navigation performance not only in the X and Y axes but also in the Z axis, thereby assessing depth [53]. A plus sign "+" was rendered at the center of the operative field view to show the focus of the laparoscope (Figure 4(d)). The task involved traversing the entire length of the track, while ensuring the laparoscope focus (black circle in Figure 4(f)) stays on the track.

Scenario-B was performed in the real world (Figure 4(b)) and assessed the ability of the HRI to maneuver the laparoscope to visualize specific target points on the operative field. The targeting tasks employed in training are directly applicable to real surgical procedures, particularly when identifying specific structures within the area of interest and maintaining a steady view. Some simulators, such as LapSim (Surgical Science), train operators to focus on gallstones and position them at the center of the view. Previous studies have utilized similar scenario to evaluate camera movement [40, 54–56]. Physical targets in the form of five circles (each 1 cm in diameter) were placed on a tissue phantom with 6–11 cm separation distances (Figure 4(e)). A circle was rendered at the center of the operative field view to show the focus of the laparoscope (Figure 4(f)). The task involved sequentially focusing on the targets from #1 to #5. The operator moved to the next target only if the previous target was within focus for at least 1 second.

The evaluation parameters utilized in this study are outlined in Table 4. While the duration to complete the tasks is a key parameter in comparative evaluation studies (as demonstrated in Table 2), other parameters were employed to quantify the errors made by the operator. Specifically, the clutch count reflects the number of instances the operator needed to pause movement to either

Table 4. Parameters Used for Evaluation

| Parameters | Parameters Recorded during Each Scenario | Scenario-A | Scenario-B |
|---|---|---|---|
| Duration | The total time taken by the subject to complete the task, in seconds | ✓ | ✓ |
| Clutch count | The number of times clutch was triggered by the subject during the task | ✓ | ✓ |
| Error duration | The total time for which the laparoscope was focused outside the path, in seconds | ✓ | |
| Mean distance | Summation of closest distance of the laparoscope viewing direction from the path with respect to time divided by the total time to complete the task, in millimeters. $$d_{Average} = \frac{\sum \Delta t_i d_i}{\sum \Delta t_i}$$ where $\Delta t_i$ is the change in time for a given instance $i$, and $d_i$ is the Euclidian distance between the laparoscope tip and the point on the path closest to the laparoscope tip. | ✓ | |
| Error count | The number of times the laparoscope focus overshoots a physical target | | ✓ |

realign the interface or adjust their head or eye focus. This count also indicates instances when movement was halted due to the limited range of motion available in different interfaces. A higher clutch count suggests a less effective HRI and interaction mode. Error duration directly represents the time during which the laparoscope's focus was unable to follow the designated path, while the mean distance indicates the extent to which the scope's focus deviated from the original trajectory. The error count was used to assess performance in scenario-B, similar to the error evaluation conducted in [33]. In real surgical settings, procedures with less duration are preferred, as well as minimizing errors is essential. The robotic scope assistant systems need to be intuitive and provide a seamless operational flow, ensuring that procedure times are not extended due to increased operator workload (resulting from excessive clutch usage), prolonged focus outside the area of interest (reflected in error duration), suboptimal operating field views (indicated by a larger mean distance), or jerky movements before settling on a region of interest (as indicated by an increased error count).

### 3.4 Data Analysis

During the user study, the state of the robotic scope assistant system and the HRI were logged. The recorded data were then processed to compute parameters for assessing the HRIs under different scenarios. These parameters are presented in Table 4. Under each scenario, the recorded parameters were analyzed using one-way ANOVA. Results of ANOVA showed significant differences among the HRIs (used under different interaction modes). Tukey's Honest Significant Differences was then used to find the pairwise significant differences. It was ensured that the homogeneity and linearity conditions of the tests were satisfied prior to obtaining the p-values and a confidence level of 0.05 was considered significant.

## 4 Results

In the results, the notation $HRI_{(Interaction\ Mode)}$ is used to denote a HRI working under single or multiple interaction modes. The subscripts "D," "T," and "C" are used to denote discrete, threshold, and continuous interaction modes, respectively. The recorded values of these parameters are presented in Table 5. Among these values, the five lowest measurements were observed for $head\text{-}motion_{(C)}$,

Table 5. Recorded Values of the Parameters from the User Study

| HRI$_{\text{(Interaction Mode)}}$ | Scenario-A: Duration (Seconds) | Scenario-A: Clutch Count (#) | Scenario-A: Error Duration (Seconds) | Scenario-A: Mean Distance (mm) | Scenario-B: Duration (Seconds) | Scenario-B: Clutch Count (#) | Scenario-B: Error Count (#) |
|---|---|---|---|---|---|---|---|
| Head-motion$_{(C)}$ | 176 ± 73 | **25 ± 15** | **9 ± 13** | **0.68 ± 0.36** | **183 ± 65** | **33 ± 12** | **3.4 ± 3.1** |
| Head-motion$_{(D)}$ | 153 ± 77 | 32 ± 15 | 22 ± 16 | 1.06 ± 0.45 | 251 ± 84 | 77 ± 27 | 6.4 ± 4.7 |
| Head-motion$_{(T)}$ | 199 ± 99 | 37 ± 21 | 35 ± 37 | 1.22 ± 0.68 | 275 ± 91 | 74 ± 21 | 6.8 ± 4.7 |
| Keypad$_{(D)}$ | **109 ± 25** | NA | **2 ± 2** | **0.54 ± 0.12** | **163 ± 51** | NA | **1.3 ± 1.7** |
| Space-mouse$_{(D)}$ | **120 ± 49** | NA | **2 ± 3** | **0.53 ± 0.13** | **157 ± 58** | NA | **1.2 ± 1.6** |
| Space-mouse$_{(T)}$ | **120 ± 37** | NA | **4 ± 6** | **0.58 ± 0.19** | **174 ± 68** | NA | **2.7 ± 4.0** |
| Eye-motion$_{(C)}$ | **120 ± 37** | **17 ± 7** | **1 ± 2** | **0.54 ± 0.13** | **180 ± 74** | **53 ± 24** | **2.8 ± 2.7** |
| Eye-motion$_{(D)}$ | 152 ± 66 | 29 ± 13 | 33 ± 41 | 1.38 ± 1.30 | 276 ± 124 | 83 ± 29 | 8.3 ± 5.6 |
| Eye-motion$_{(T)}$ | **150 ± 58** | 30 ± 16 | 19 ± 20 | 1.06 ± 0.49 | 275 ± 115 | 83 ± 27 | 7.9 ± 5.4 |
| Voice$_{(D)}$ | 203 ± 86 | 49 ± 45 | 30 ± 16 | 1.15 ± 0.50 | 237 ± 75 | 76 ± 43 | 4.8 ± 2.9 |

*keypad$_{(D)}$*, *space-mouse$_{(D,T)}$*, and *eye-motion$_{(C)}$*, with the exception of the duration for *head-motion$_{(C)}$*. Conversely, the five highest values were observed for *head-motion$_{(D,T)}$*, *eye-motion$_{(D,T)}$*, and *voice$_{(D)}$*, with the exception of the duration for *eye-motion$_{(T)}$*. The higher observed values are indicated in bold in Table 5.

The outcomes of the user study conducted in Scenario-A are illustrated in Figures 5, 6, and 7, and presented in supplementary materials. The duration (Figure 5) for *voice$_{(D)}$* was significantly higher as compared to *keypad$_{(D)}$* and *space-mouse$_{(D,T)}$*. Similarly, the duration for *head-motion$_{(C,T)}$* was higher as compared to *space-mouse$_{(T)}$*. Duration for *head-motion$_{(T)}$* was also higher as compared to *keypad$_{(D)}$*. The clutch counts were lower for *eye-motion$_{(C)}$* as compared to *head-motion$_{(T)}$*, while no significant differences were observed among the other pairs. The error duration (Figure 6) for *keypad$_{(D)}$*, *space-mouse$_{(D,T)}$*, and *eye-motion$_{(C)}$* was lower as compared to *eye-motion$_{(D)}$*, *voice$_{(D)}$*, and *head-motion$_{(T)}$*. Similarly, *head-motion$_{(C)}$* had lower error duration as compared to *head-motion$_{(T)}$*, *eye-motion$_{(D)}$*, and *voice$_{(D)}$*. The error duration for *keypad$_{(D)}$*, *space-mouse$_{(D)}$*, and *eye-motion$_{(C)}$* was lower as compared to *head-motion$_{(D)}$* and *eye-motion$_{(T)}$*. The mean distance (Figure 7) for *keypad$_{(D)}$*, *space-mouse$_{(D,T)}$*, and *eye-motion$_{(C)}$* was lower as compared to *eye-motion$_{(D,T)}$*, *head-motion$_{(D,T)}$*, and *voice$_{(D)}$*. Similarly, *head-motion$_{(C)}$* had lower mean distance as compared to *eye-motion$_{(D,T)}$*, *head-motion$_{(D,T)}$*, and *voice$_{(D)}$*.

The results of the user study conducted in Scenario-B are shown in Figures 8, 9, and 10, and presented in supplementary material. As depicted in Figure 8, the duration for *keypad$_{(D)}$*, *space-mouse$_{(D,T)}$*, and *eye-motion$_{(C)}$* was significantly shorter as compared to *eye-motion$_{(D,T)}$*. Also, the duration for *keypad$_{(D)}$* and *space-mouse$_{(D,T)}$* was less as compared to *head-motion$_{(T)}$*. Furthermore, the duration for *space-mouse$_{(D)}$* was significantly small compared to *head-motion$_{(D)}$* and *voice$_{(D)}$*. In case of clutch counts (Figure 9), *head-motion$_{(C)}$* showed a significant reduction compared to *head-motion$_{(D,T)}$*, *eye-motion$_{(C,D,T)}$*, and *voice$_{(D)}$*. Similarly, clutch counts for *eye-motion$_{(C)}$* were also lower as compared to eye-motion$_{(D,T)}$, *and head-motion$_{(D,T)}$*. The number of recorded errors (Figure 10) were higher for eye-*motion$_{(D,T)}$* as compared to *keypad$_{(D)}$*, *space-mouse$_{(D,T)}$*, and *eye-motion$_{(C)}$*. The number of recorded errors were also higher for *head-motion$_{(D,T)}$* and *voice$_{(D)}$* as compared to *keypad$_{(D)}$* and *space-mouse$_{(D)}$*.

## 5 Discussion

In both Scenario-A and Scenario-B, the duration to complete the task was lower when using hand-motion (i.e., *keypad$_{(D)}$* and *space-mouse$_{(D,T)}$*). This is because of the hand's dexterity in performing

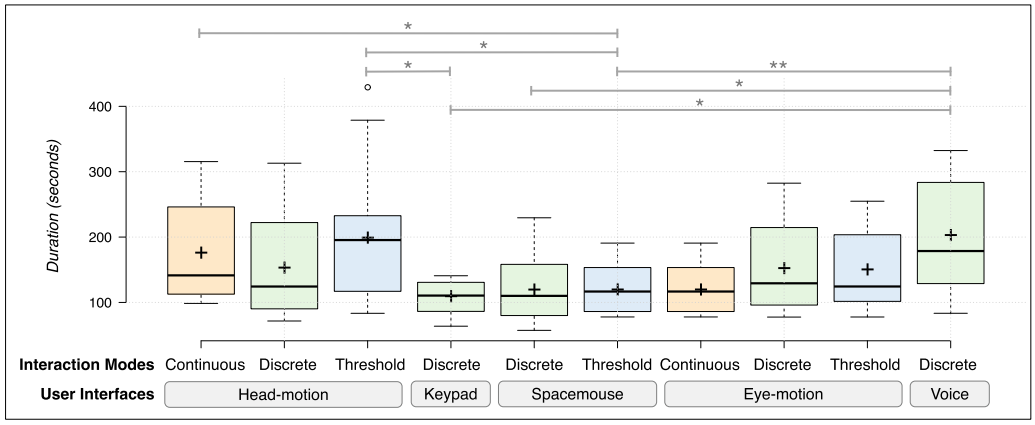


Fig. 5. Boxplots of the parameter "duration" under Scenario-A across different HRIs and interaction modes. "*" denotes a p-value between 0.05 and 0.01. "**" denotes a p-value between 0.01 and 0.001.

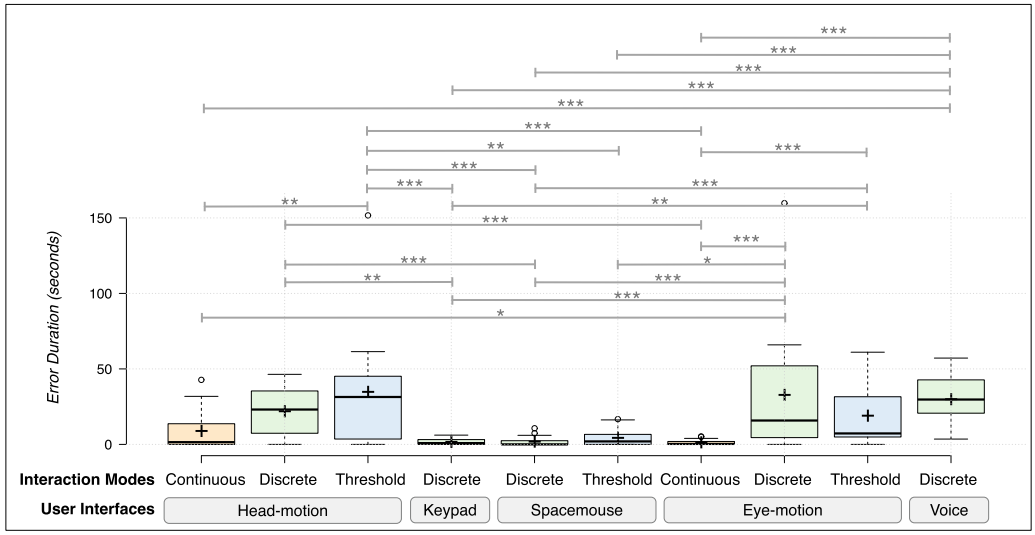


Fig. 6. Boxplots of the parameter "error duration" under Scenario-A across different HRIs and interaction modes. "*" denotes a p-value between 0.05 and 0.01. "**" denotes a p-value between 0.01 and 0.001. "***" denotes a p-value less than 0.001.

day-to-day activities, which enables the operator to naturally use different fingers to press multiple buttons or manipulate a 3D joystick. This ability allows the operator to swiftly transition between multiple actuation commands and issue them more rapidly. It was also noted that the eye-motion under continuous interaction mode (*eye-motion*$_{(C)}$) required relatively less duration to complete the tasks. In continuous interaction mode, there is one-to-one mapping between the operator's eye-motion and motion of the laparoscope's distal end, enabling the operator to easily switch directions and have smooth continuous motion to view the operating field.

The clutch acted as a trigger for activating or deactivating the system, enabling the surgeon to issue actuation commands for laparoscope movements only when required during the surgery. The count of clutches allows evaluation of different HRIs and interaction modes. The reliance on clutch

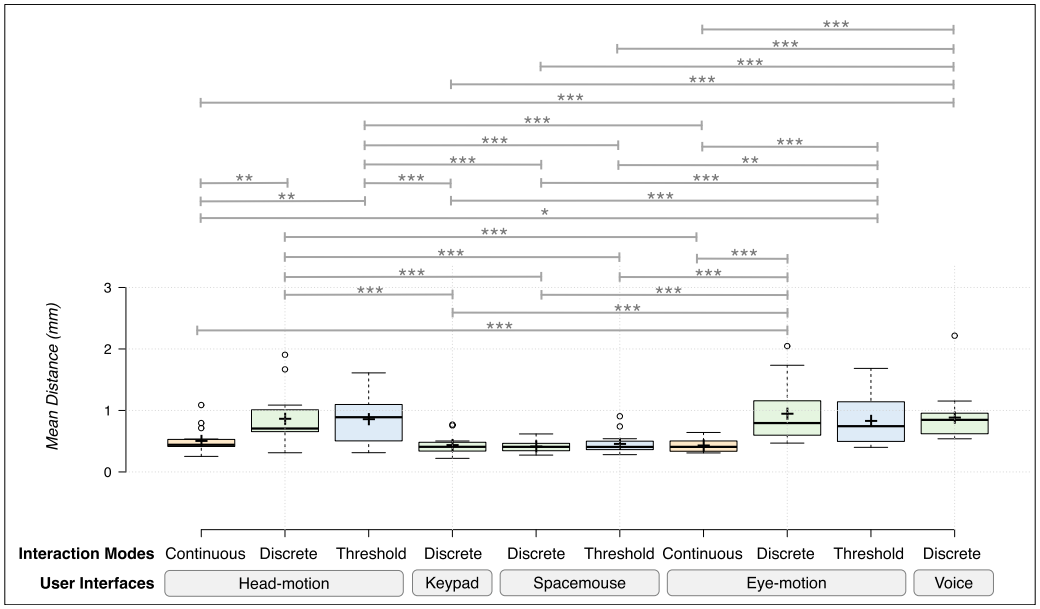


Fig. 7. Boxplots of the parameter "mean distance" under Scenario-A across different HRIs and interaction modes. "*" denotes a p-value between 0.05 and 0.01. "**" denotes a p-value between 0.01 and 0.001. "***" denotes a p-value less than 0.001.

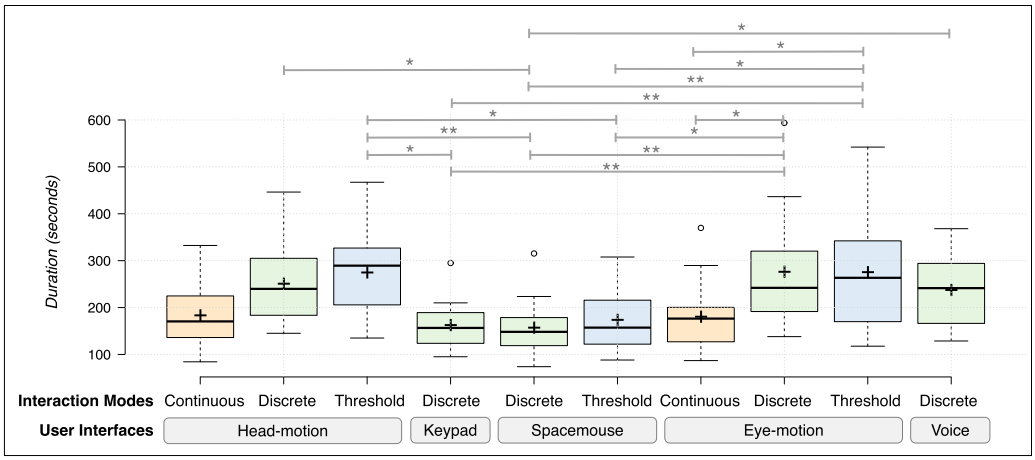


Fig. 8. Boxplots of the parameter "duration" under Scenario-B across different HRIs and interaction modes. "*" denotes a p-value between 0.05 and 0.01. "**" denotes a p-value between 0.01 and 0.001. "***" denotes a p-value less than 0.001.

to switch movement has been documented in previous study [57, 58] and reflects the frequency with which users need to make readjustments while performing the task. A higher number of clutches indicates a decreased ability to control movement using the primary interface alone. In real surgical setting, this metric reflects increased workload. In scenario-B, the clutch count for *head-motion*$_{(C)}$ and *eye-motion*$_{(C)}$ were lower as compared to *head-motion*$_{(D,T)}$ and *eye-motion*$_{(D,T)}$.

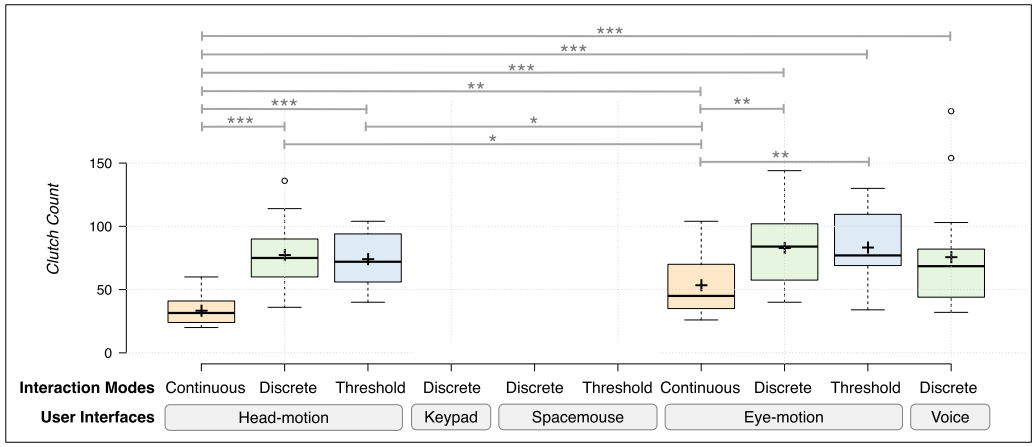


Fig. 9. Boxplots of the parameter "clutch count" under Scenario-B across different HRIs and interaction modes. "*" denotes a p-value between 0.05 and 0.01. "**" denotes a p-value between 0.01 and 0.001. "***" denotes a p-value less than 0.001.

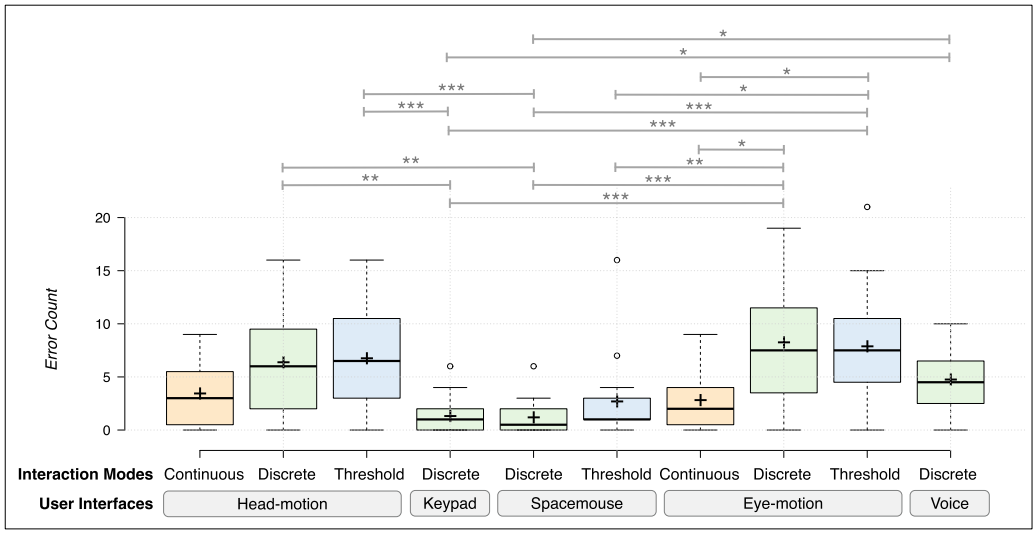


Fig. 10. Boxplots of the parameter "error count" under Scenario-B across different HRIs and interaction modes. "*" denotes a p-value between 0.05 and 0.01. "**" denotes a p-value between 0.01 and 0.001. "***" denotes a p-value less than 0.001.

This indicates that under continuous interaction mode, an operator can maneuver the laparoscope with fewer actuation commands than in discrete and threshold interaction modes. When using a keypad or space mouse, no clutch was employed and the operator was required to switch between manipulating the surgical instruments and pressing the keys on the keypad or moving the joystick. This limits the surgeon from simultaneously operating on the tissue using the surgical instruments and maneuvering the laparoscope using keypad or space mouse, inhibiting the workflow during the surgery.

The error duration and mean distance in Scenario-A and error count in Scenario-B were low for *keypad*$_{(D)}$, *space-mouse*$_{(D,T)}$, *head-motion*$_{(C)}$, and *eye-motion*$_{(C)}$. With these user interfaces and interaction modes, it was possible to perform the simulated surgical tasks both with reduced duration and fewer errors. On the contrary, *head-motion*$_{(D,T)}$, eye-*motion*$_{(D,T)}$, and *voice*$_{(D)}$ reported longer duration with higher errors to perform the simulated surgical tasks.

For discrete interaction mode, laparoscope maneuverability was constrained to one-axis (X, Y, or Z) at a constant speed, necessitating the issuance of multiple commands to navigate the designated path (Scenario-A). This resulted in prolonged task completion time and an increased use of the clutch, as users were compelled to employ a stepwise approach to advance forward, rather than having the flexibility to navigate along the curved path.

The overshooting of the laparoscope's focus (error count) when aligning with the targets in Scenario-B was prominent in both threshold and discrete interaction modes for head and eye-based motion. This could be due to three main factors. First, the user interface had to reach a preset threshold value to initiate the movement of the laparoscope's tip. Second, the laparoscope moved at a constant speed rather than being directly proportional to the distance moved by the user interface. Last, stopping the movement of the laparoscope required either switching off the clutch or returning the user interface to its original position. Alternatively, it was easier to perform precise movements in continuous mode as the user's interaction was directly translated into laparoscope movement and the motion could be paused easily by simply halting the head or eye movement.

In the case of *voice*$_{(D)}$, the system took time to detect and interpret the voice commands spoken by the operator, which subsequently led to lag in the laparoscope actuation. The poor performance of the voice mode can be attributed to factors such as voice clarity, accent, and background noise that made it difficult to detect the commands correctly. Additionally, the user often repeated the command due to the inherent delay in processing the previous command, which in turn led to increased mental demand for the user.

This study identified four limitations. First, foot-motion to issue actuation commands was not assessed due to the unavailability of a custom-made programmable foot pedal. These interfaces are designed to process foot motion under discrete interaction mode, where distinct pedals trigger specific laparoscope movements (without a clutch). As the surgeon is in a standing position, pressing the correct pedal is challenging. This may require the operator to shift focus from the viewing screen to the base of the operating table to select the appropriate pedal [34–36]. Also, repeatedly pressing multiple pedals may lead to operator fatigue. Second, the study was conducted using a zero-degree laparoscope. In laparoscopic surgery, a surgeon often uses angulated laparoscopes (with viewing direction angled at 30°, 45°, 60°, or 90°) or articulated laparoscopes (such as EndoCAMeleon—Karl Storz that offers adjustable viewing direction). Since the method of laparoscope movement was based on the laparoscope's distal end following a virtual target, the same approach can be adopted for angulated [14] and articulated laparoscopes [13, 53]. This would make the actuation commands (and corresponding laparoscope movements) independent of the angulation in the viewing direction. Thus, the same outcomes of the user interface can be extended to both angulated and articulated laparoscopes. Third, surgical tasks were carried out in a computer-generated environment (Scenario-A) and on an inanimate phantom (Scenario-B) instead of live tissue. This was necessary to ensure repeatability during the study consisting of multiple subjects. The choice of the user interface may alter when employed in a specific minimally invasive surgical procedure for visualizing the operative field. Last, this study did not consider the subjective evaluations of the operators, including factors such as cognitive workload and fatigue. Therefore, this work presents only a comparative benchmark based on the selected evaluation metrics.

## 6 Conclusion

This study evaluated five distinct HRIs and three interaction modes by performing simulated preclinical laparoscopic tasks. Head-motion and eye-motion-based HRIs performed better in the continuous interaction mode than the threshold and discrete interaction modes of these devices. Non-clutch-based devices, namely, keypad and space mouse, outperformed all other combinations for all the response variables in both scenarios. This could be due to the operator's hand dexterity in performing day-to-day activities as well as the absence of the complexity to control two devices (i.e., the primary HRI and additional foot pedal as a clutch). The voice interface had an overall moderate performance due to the delay in detecting, processing, and executing the actuation commands. The increased dexterity and fast response of hand movements and intuitive direct mapping of head/eye movement in continuous mode reflected promising performance of $keypad_{(d)}$, $space\text{-}mouse_{(d,t)}$, $head\text{-}motion_{(c)}$, and $eye\text{-}motion_{(c)}$ in the user-study. The continuous interaction modes for head and eye-motion for laparoscope control, developed as a novel aspect of this work, demonstrated superior performance compared to other interaction modes within the same interface. A comprehensive analysis of diverse user interfaces is presented; consequently, the selection of the HRI and interaction mode is contingent upon the proficiency and comfort level of the operating surgeon, as well as the nature of the minimally invasive procedure.

## References


[1] M. Siddaiah-Subramanya, K. W. Tiang, and M. Nyandowe. 2017. A new era of minimally invasive surgery: Progress and development of major technical innovations in general surgery over the last decade. *The Surgery Journal* 3, 4 (2017), e163–e166.

[2] M. Arnold, S. Elhage, L. Schiffern, B. Lauren Paton, S. W. Ross, B. D. Matthews, and C. E. Reinke. 2020. Use of minimally invasive surgery in emergency general surgery procedures. *Surgical Endoscopy* 34, 5 (2020), 2258–2265.

[3] A. Ng, N. Wang, and M. Tran. 2019. Minimally invasive surgery: Early concepts to gold standards. *British Journal of Hospital Medicine (London)* 80, 9 (Sep. 2019), 494–495. DOI: https://doi.org/10.12968/HMED.2019.80.9.494

[4] K. Omote, H. Feussner, A. Ungeheuer, K. Arbter, G.-Q. Wei, J. R. Siewert, and G. Hirzinger. 1999. Self-guided robotic camera control for laparoscopic surgery compared with human camera control. *The American Journal of Surgery* 177, 4 (Apr. 1999), 321–324.

[5] S. M. Ali, L. A. Reisner, B. King, A. Cao, G. Auner, M. Klein, and A. K. Pandya. 2008. Eye gaze tracking for endoscopic camera positioning: An application of a hardware/software interface developed to automate AESOP. *Studies in Health Technology and Informatics* 132 (2008), 4–7.

[6] B. M. Kraft, C. Jäger, K. Kraft, B. J. Leibl, and R. Bittner. 2004. The AESOP robot system in laparoscopic surgery: Increased risk or advantage for surgeon and patient?. *Surgical Endoscopy* 18, 8 (2004), 1216–1223. DOI: https://doi.org/10.1007/S00464-003-9200-Z

[7] P. B. Nebot, Y. Jain, K. Haylett, R. Stone, and R. Mccloy. 2003. Comparison of task performance of the camera-holder robots endoassist and AESOP. *Surgical Laparoscopy Endoscopy and Percutaneous Techniques* 13, 5 (2003), 334–338.

[8] S. Aiono, J. M. Gilbert, B. Soin, P. A. Finlay, and A. Gordan. 2002. Controlled trial of the introduction of a robotic camera assistant (endoassist) for laparoscopic cholecystectomy. *Surgical Endoscopy* 16, 9 (Sep. 2002), 1267–1270. DOI: https://doi.org/10.1007/S00464-001-9174-7

[9] N. Halín, P. Loula, and P. Aarnio. 2007. Experiences of using the endoassist-robot in surgery. *Studies in Health Technology and Informatics* 125 (2007), 161–163.

[10] J. Y. K. Chan, I. Leung, D. Navarro-Alarcon, W. Lin, P. Li, Dennis L. Y. Lee, Y.-H. Liu, and M. C. F. Tong. 2016. Foot-controlled robotic-enabled endoscope holder for endoscopic sinus surgery: A cadaveric feasibility study. *The Laryngoscope* 126, 3 (2016), 566–569.

[11] A. Mirbagheri, F. Farahmand, B. Ghanadi, K. Amini Khoiy, S. Porsa, M. J. Shamsollahi, M. H. Owlia, F. Karimian, and K. Toulabi. 2015. Operation and human clinical trials of robolens: An assistant robot for laparoscopic surgery. *Frontiers in Biomedical Technologies* 2, 3 (2015), 184–190.

[12] G. F. Buess, A. Arezzo, M. O. Schurr, F. Ulmer, H. Fisher, L. Gumb, T. Testa, and C. Nobman. 2000. A new remote-controlled endoscope positioning system for endoscopic solo surgery: The fips endoarm. *Surgical Endoscopy* 14, 4 (2000), 395–399.

[13] N. Abdurahiman, M. Khorasani, J. Padhan, V. M. Baez, A. Al-Ansari, P. Tsiamyrtzis, A. T. Becker, and N. V. Navkar. 2023. Scope actuation system for articulated laparoscopes. *Surgical Endoscopy* 37, 3 (2023), 2404–2413.

[14] N. Abdurahiman, J. Padhan, H. Zhao, S. Balakrishnan, A. Al-Ansari, and J. Abinahed. 2022. Human-computer interfacing for control of angulated scopes in robotic scope assistant systems. In *Proceedings of the 2022 International Symposium on Medical Robotics (ISMR). IEEE*, 1–7.
[15] EMARO pneumatic endoscope manipulator robot. Riverfield Inc. Retrieved 31 December 2023 from https://riverfieldinc.com/en/products/p01/
[16] Freehand. Freehand surgical. Retrieved 31 December 2023 from https://www.freehandsurgeon.com/
[17] Soloassist II. AKTORMED GMBH. Retrieved 31 December 2023 from https://aktormed.info/en/products/soloassist-ii
[18] HIWIN - MTG 100. Hiwin Technologies. Retrieved 31 December 2023 from https://www.hiwin.tw/products/me/mtg_h100.aspx
[19] Robolens: Laparoscopic surgery assistant robot (standalone model). Sina Robotics & Medical Innovators Co. Retrieved 31 December 2023 from https://sinamed.ir/robotic-tele-surgery/robolens-stand-alone-model/
[20] D. Gossot, W. Abid, and A. Seguin-Givelet. 2018. Motorized scope positioner for solo thoracoscopic surgery. *Video-Assisted Thoracic Surgery* 3 (2018), 47–47.
[21] R. Polet and J. Donnez. 2008. Using a laparoscope manipulator (Lapman) in laparoscopic gynecological surgery. *Surgical Technology International* 17 (2008), 187–191.
[22] K. Yamada and S. Kato. and 2008. Robot-assisted thoracoscopic lung resection aimed at solo surgery for primary lung cancer. *General Thoracic and Cardiovascular Surgery* 56, 6 (2008), 292–294. DOI: https://doi.org/10.1007/S11748-008-0240-0
[23] C.-A. O. Nathan, V. Chakradeo, K. Malhotra, H. D'agostino, and R. Patwardhan. 2006. The voice-controlled robotic assist scope holder AESOP for the endoscopic approach to the Sella. *Skull Base* 16, 3 (2006), 123–131.
[24] Y.-Y. Lin, M.-J. Hsieh, C.-Y. Wu, L.-Y. Yang, Y.-B. Pan, C.-F. Wu, D. Gonzalez-Rivas, and Y.-K. Chao. 2023. Comparison of active versus passive robotic-endoscope-holder-assisted unisurgeon uniportal thoracoscopic surgery in terms of surgical efficacy and patient safety. *Journal of Thoracic Disease* 15, 7 (2023), 3800–3810.
[25] Y. Ohmura, H. Suzuki, K. Kotani, and A. Teramoto. 2019. Comparative effectiveness of human scope assistant versus robotic scope holder in laparoscopic resection for colorectal cancer. *Surgical Endoscopy* 33, 7 (2019), 2206–2216.
[26] L. R. Kavoussi, R. G. Moore, J. B. Adams, and A. W. Partin. 1995. Comparison of robotic versus human laparoscopic camera control. *The Journal of Urology* 154, 6 (1995), 2134–2136.
[27] A. Pandya, L. Reisner, B. King, N. Lucas, A. Composto, M. Klein, and R. Ellis. 2014. A review of camera viewpoint automation in robotic and laparoscopic surgery. *Robotics* 3, 3 (2014), 310–329.
[28] H. Hamza, V. M. Baez, A. Al-Ansari, A. T. Becker, and N. V. Navkar. 2023. User interfaces for actuated scope maneuvering in surgical systems: A scoping review. *Surgical Endoscopy* 37, 6 (2023), 4193–4223.
[29] C. C. Nguyen, S. Wong, M. T. Thai, T. T. Hoang, P. T. Phan, J. Davies, L. Wu, D. Tsai, H.-P. Phan, N. H. Lovell, et al. 2023. Advanced user interfaces for teleoperated surgical robotic systems. *Advanced Sensor Research* 2, 4 (2023), 2200036.
[30] Y.-J. Yang, S. Udatha, D. Kulić, and E. Abdi. 2020. A novel foot interface versus voice for controlling a robotic endoscope holder. *In Proceedings of the 2020 8th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BIOROB). IEEE*, 272–279.
[31] A. A. Gumbs, F. Crovari, C. Vidal, P. Henri, and B. Gayet. 2007. Modified robotic lightweight endoscope (ViKY) validation in vivo in a porcine model. *Surgical Innovation* 14, 4 (2007), 261–264. DOI: https://doi.org/10.1177/1553350607310281
[32] A. Mirbagheri, F. Farahmand, A. Meghdari, and F. Karimian. 2011. Design and development of an effective low-cost robotic cameraman for laparoscopic surgery: Robolens. *Scientia Iranica* 18 (2011), 105–114.
[33] M. E. Allaf, S. V. Jackman, P. G. Schulam, J. A. Cadeddu, B. R. Lee, R. G. Moore, and L. R. Kavoussi. 1998. Laparoscopic visual field: Voice vs foot pedal interfaces for control of the AESOP robot. *Surgical Endoscopy* 12 (1998), 1415–1418.
[34] P. Berkelman, P. Cinquin, E. Boidard, J. Troccaz, C. Létoublon, and J.-A. Long. 2005. Development and testing of a compact endoscope manipulator for minimally invasive surgery. *Computer Aided Surgery* 10, 1 (2005), 1–13.
[35] L. Mettler, M. Ibrahim, and W. Jonat. 1998. One year of experience working with the aid of a robotic assistant (the voice-controlled optic holder AESOP) in gynaecological endoscopic surgery. *Human Reproduction (Oxford, England)* 13, 10 (1998), 2748–2750.
[36] M. Van Veelen, C. Snijders, E. van Leeuwen, R. Goossens, and G. Kazemier. 2003. Improvement of foot pedals used during surgery based on new ergonomic guidelines. *Surgical Endoscopy* 17, 7, 1086–1091.
[37] J. Rassweiler, M. Fiedler, N. Charalampogiannis, A. S. Kabakci, R. Saglam, and J.-T. Klein. 2018. Robot-assisted flexible ureteroscopy: An update. *Urolithiasis* 46, 1, 69–77.
[38] S. G. Lim. 2020. The development of robotic flexible endoscopic platforms. *International Journal of Gastrointestinal Intervention* 9, 1 (2020), 9–12. DOI: https://doi.org/10.18528/IJGII190022

[39] R. Reilink, G. D. Bruin, M. Franken, M. A. Mariani, S. Misra, and S. Stramigioli. 2010. Endoscopic camera control by head movements for thoracic surgery. *In Proceedings of the 2010 3RD IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics*, 510–515. DOI: https://doi.org/10.1109/BIOROB.2010.5627043
[40] Y. X. Mak, M. Zegel, M. Abayazid, M. A. Mariani, and S. Stramigioli. 2022. *Experimental evaluation using head motion and augmented reality to intuitively control a flexible endoscope*. IEEE, 1–7.
[41] A. Sivananthan, A. Kogkas, B. Glover, A. Darzi, G. Mylonas, and N. Patel. 2021. A novel gaze-controlled flexible robotized endoscope; preliminary trial and report. *Surgical Endoscopy* 35, 8 (2021), 4890–4899.
[42] A. Sivananthan, A. Kokgas, B. Glover, A. Darzi, G. Mylonas, and N. Patel. 2023. Eye controlled endoscopy-A benchtop trial of a novel robotic system. *Gastrointestinal Endoscopy* 97, 6 (2023), AB770–AB771.
[43] P. Allemann, L. Ott, M. Asakuma, N. Masson, S. Perretta, B. Dallemagne, D. Coumaros, M. De Mathelin, L. Soler, J. Marescaux, et al. 2009. Joystick interfaces are not suitable for robotized endoscope applied to notes. *Surgical Innovation* 16, 2 (2009), 111–116.
[44] S. Basha, M. Khorasani, N. Abdurahiman, J. Padhan, V. Baez, A. Al-Ansari, P. Tsiamyrtzis, A. T. Becker, and N. V. Navkar. 2024. A generic scope actuation system for flexible endoscopes. *Surgical Endoscopy* 38, 2 (2024), 1096–1105.
[45] E. Abdi, M. Bouri, J. Olivier, and H. Bleuler. 2016. Foot-controlled endoscope positioner for laparoscopy: Development of the master and slave interfaces. In *Proceedings of the 2016 4th International Conference on Robotics and Mechatronics (ICROM)*. IEEE, 111–115.
[46].] Steute Meditec - Foot Switch. Steute Technologies. Retrieved 31 December 2023 from https://www.steute-meditec.com/en/products/mfs-microscope-sw24le-med.html
[47] Y. Huang, E. Burdet, L. Cao, P. T. Phan, A. M. H. Tiong, P. Zheng, and S. J. Phee. 2019. Performance evaluation of a foot interface to operate a robot arm. *IEEE Robotics and Automation Letters* 4, 4 (2019), 3302–3309.
[48] K. P. Suresh. 2011. An overview of randomization techniques: An unbiased assessment of outcome in clinical research. *Journal of Human Reproductive Sciences* 4, 1 (2011), 8–11.
[49] D. Shabir, N. Abdurahiman, J. Padhan, M. Anbatawi, M. Trinh, S. Balakrishnan, A. Al-Ansari, E. Yaacoub, Z. Deng, A. Erbad, et al. 2022. Preliminary design and evaluation of a remote tele-mentoring system for minimally invasive surgery. *Surgical Endoscopy* 36, 5 (2022), 3663–3674.
[50] P. J. Roch, H. M. Rangnick, J. A. Brzoska, L. Benner, K.-F. Kowalewski, P. C. Müller, H. G. Kenngott, B.-P. Müller-Stich, and F. Nickel. 2018. Impact of visual–spatial ability on laparoscopic camera navigation training. *Surgical Endoscopy* 32, 3 (2018), 1174–1183.
[51] M. Graafland, K. Bok, H. W. Schreuder, and M. P. Schijven. 2014. A multicenter prospective cohort study on camera navigation training for key user groups in minimally invasive surgery. *Surgical Innovation* 21, 3 (2014), 312–319.
[52] J. R. Korndorffer, D. J. Hayes, J. B. Dunne, R. Sierra, C. L. Touchard, R. J. Markert, and D. J. Scott. 2005. Development and transferability of a cost-effective laparoscopic camera navigation simulator. *Surgical Endoscopy* 19, 2 (2005), 161–167.
[53] N. Abdurahiman, J. Padhan, M. Khorasani, H. Zhao, V. M. Baez, and A. Al-Ansari. 2023. Interfacing mechanism for actuated maneuvering of articulated laparoscopes using head motion. *In Proceedings of the 2023 IEEE International Conference on Mechatronics and Automation (ICMA). IEEE*, 2289–2296.
[54] V. Arikatla, S. Horvath, Y. Fu, L. Cavuoto, S. De, Steve S., and A. Enquobahrie. 2019. Development and face validation of a virtual camera navigation task trainer. *Surgical Endoscopy* 33, 6 (2019), 1927–1937.
[55] N. Y. Sherpa, A. E. Minawi, A. N. Askalany, and M. Abdalla. 2024. Comparative study to determine the proper sequence of simulation training, pelvic trainer versus virtual reality simulator: A pilot study. *Middle East Fertility Society Journal* 29, 1 (2024), 11.
[56] D. Shabir, M. Anbatawi, J. Padhan, S. Balakrishnan, A. Al-Ansari, J. Abinahed, P. Tsiamyrtzis, E. Yaacoub, A. Mohammed, Z. Deng, et al. 2022. Evaluation of user-interfaces for controlling movements of virtual minimally invasive surgical instruments. *The International Journal of Medical Robotics and Computer Assisted Surgery* 18, 5 (2022), E2414.
[57] N. Hashemi, M. B. S. Svendsen, F. Bjerrum, S. Rasmussen, M. G. Tolsgaard, and M. L. Friis. 2023. Acquisition and usage of robotic surgical data for machine learning analysis. *Surgical Endoscopy* 37, 8 (2023), 6588–6601.
[58] M. Khorasani, N. Abdurahiman, J. Padhan, H. Zhao, A. Al-Ansari, A. T. Becker, and N. Navkar. 2023. Preliminary design and evaluation of a generic surgical scope adapter. *The International Journal of Medical Robotics and Computer Assisted Surgery* 19, 2 (2023), e2475. Retrieved fro https://onlinelibrary.wiley.com/doi/10.1002/rcs.2475

# A Appendices

## A.1 Compute Next Scope State

The module labeled "Compute Next Scope State" in Figure 3 receives as input the target laparoscope tip position $X_T(t)$, the current laparoscope tip position $X(t)$ along with its current velocity $X'(t)$.

It then calculates the next laparoscope tip position $X(t+\Delta t)$ and velocity $X'(t+\Delta t)$ for the time instant $t+\Delta t$ as follows:

$$X(t+\Delta t) = X_T(t) + (X(t) - X_T(t))e^{-\frac{2}{t_s}\Delta t} + \left(X'(t) + \frac{2}{t_s}(X(t) - X_T(t))\right)\Delta t e^{-\frac{2}{t_s}\Delta t},$$
$$X'(t+\Delta t) = X'(t)e^{-\frac{2}{t_s}\Delta t} + \left(X'(t) + \frac{2}{t_s}(X(t) - X_T(t))\right)\frac{2}{t_s}\Delta t e^{-\frac{2}{t_s}\Delta t}.$$

$\Delta t$ is the timestep of every iteration of the "motion planning thread" and $t_s$ is the smoothness time.

### A.2 Compute Robot End-Effector Linear Velocity

The module labeled "Compute Robot End-Effector Velocity" depicted in Figure 3 receives the next laparoscope tip position $X(t+\Delta t)$ and velocity $X'(t+\Delta t)$ as input. It then calculates the linear velocity $X'_E(t+\Delta t)$ for the robot's end-effector, while maintaining a RCM at the incision point $X_I$. This ensures a consistent and smooth velocity profile for the laparoscope tip, avoiding abrupt changes. Even if the operator makes a sudden movement that significantly displaces the target laparoscope tip position $X_T(t)$, the system prevents the laparoscope from undergoing drastic movements to reach the desired target.

The position $X_E(t)$ of the end-effector at time instant $t$ is computed as

$$X_E(t) = X(t) - \lambda\frac{n_1(t)}{\|n_1(t)\|},$$
$$n_1(t) = X(t) - X_I,$$

where $\lambda$ is the laparoscope length from the laparoscope tip $X(t)$ to end-effector $X_E(t)$.

The time-derivative of end-effector position $X_E(t)$ provides the end-effector's linear velocity $X'_E(t)$:

$$X'_E(t) = X'(t) - \lambda\frac{(n_1(t)\cdot n_1(t))n'_1(t) - (n_1(t)\cdot n'_1(t))n_1(t)}{\|n_1(t)\|^3},$$
$$n'_1(t) = X'(t).$$

While manipulating the laparoscope, there is a possibility that the shaft of the laparoscope (represented as a line segment connecting the laparoscope tip $X(t)$ with the end-effector $X_E(t)$) may deviate from the incision point $X_I$ due to the accumulation of computational errors. To counteract this, a velocity element $X'_R(t)$ is introduced to the end-effector velocity $X'_E(t)$ and is computed as

$$X'_R(t) = \frac{n_3(t) - (n_3(t)\cdot n_1(t))n_1(t)}{\Delta t}X'_{Max},$$

where $n_3(t) = X_E(t) - X_I$ and $X'_{Max}$ is the maximum laparoscope tip speed.

### A.3 Compute Robot End-Effector Angular Velocity

The unit "Compute Robot End Effector Velocity" shown in Figure 3 takes input the next laparoscope tip's position $X(t+\Delta t)$ and velocity $X'(t+\Delta t)$, and computes the robot's end-effector angular velocity $\omega_E(t)$, while maintaining a RCM at an incision point $X_I$.

The skew symmetric matrix $S(\omega_E(t))$ is equated to the product of the derivative $R'_E(t)$ and the transpose $R^T_E(t)$ of the rotational matrix $R_E(t)$ representing a frame at the end-effector to compute

the angular velocity components $(\omega_x(t),\ \omega_y(t),\ \omega_z(t))$ of the end-effector:

$$S(\omega_E(t)) = \begin{bmatrix} 0 & -\omega_z(t) & \omega_y(t) \\ \omega_z(t) & 0 & -\omega_x(t) \\ -\omega_y(t) & \omega_x(t) & 0 \end{bmatrix} = R'_E(t)R_E^T(t),$$

$$R_E(t) = \left[\eta_x(t)\ \eta_y(t)\ \eta_z(t)\right],$$

$$R'_E(t) = \left[\eta'_x(t)\ \eta'_y(t)\ \eta'_z(t)\right],$$

where $\eta_x(t),\ \eta_y(t)$, and $\eta_z(t)$ are direction cosines representing X, Y, and Z axes of the frame at the end-effector. The direction cosine $\eta_z(t)$ and its derivative $\eta'_z(t)$ are computed as

$$\eta_z(t) = \frac{n_1(t)}{\|n_1(t)\|},$$

$$\eta'_z(t) = \frac{(n_1(t)\cdot n_1(t))n'_1(t) - (n_1(t)\cdot n'_1(t))n_1(t)}{\|n_1(t)\|^3}.$$

Similarly, the direction cosine $\eta_x(t)$ and its derivative $\eta'_x(t)$ are computed as

$$\eta_x(t) = \frac{n_2(t)}{\|n_2(t)\|},$$

$$\eta'_x(t) = \frac{(n_2(t)\cdot n_2(t))n'_2(t) - (n_2(t)\cdot n'_2(t))n_2(t)}{\|n_2(t)\|^3}.$$

where vector $n_2(t)$ is perpendicular to the plane defined by the laparoscope's horizon $v(t)$ and direction cosine $\eta_z(t)$ representing the direction of the laparoscope's shaft. The vector $n_2(t)$ and its derivative, denoted as $n'_2(t)$ are computed as

$$n_2(t) = v(t)\ \times\ \eta_z(t),$$

$$n'_2(t) = v(t)\ \times\ \eta'_z(t).$$

The direction cosine $\eta_y(t)$ and its derivative $\eta'_y(t)$ are computed based on $\eta_z(t)$ and $\eta_x(t)$ as

$$\eta_y(t) = \eta_z(t)\ \times\ \eta_x(t),$$

$$\eta'_y(t) = \eta_z(t) \times \eta'_x(t) + \eta'_z(t) \times \eta_x(t)\ .$$

Received 12 March 2024; revised 11 February 2025; accepted 20 February 2025