

## Multimodal and Mixed Control of Robotic Endoscopes

Ignacio Avellino, Gilles Bailly, Mario Aricò, Guillaume Morel, Geoffroy  
Canlorbe

► **To cite this version:**

Ignacio Avellino, Gilles Bailly, Mario Aricò, Guillaume Morel, Geoffroy Canlorbe. Multimodal and Mixed Control of Robotic Endoscopes. CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2020), Apr 2020, Honolulu, United States. 10.1145/3313831.3376795 . hal-02475088

**HAL Id: hal-02475088**

**<https://hal.archives-ouvertes.fr/hal-02475088>**

Submitted on 11 Feb 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

# Multimodal and Mixed Control of Robotic Endoscopes

Ignacio Avellino<sup>1,2</sup> Gilles Bailly<sup>1,2</sup> Mario Aricò<sup>1,2</sup> Guillaume Morel<sup>1,2</sup> Geoffroy Canlorbe<sup>3,4,1</sup>

<sup>1</sup>Sorbonne Université | <sup>2</sup>CNRS, ISIR, Paris, France

<sup>3</sup>Department of Obstetrics and Gynaecology, Pitié-Salpêtrière University Hospital, AP-HP

<sup>4</sup>INSERM UMRS938

## ABSTRACT

Bedside robotic endoscopes render surgeons autonomous from assistants, potentially improving surgical outcome and decreasing costs. Why then have they not been widely adopted? We take a step back and first characterize classic (non-robotic) endoscope use through observations, literature and a domain expert interview. We review the literature on bedside robotic endoscopes and find that existing controls, individually, do not have the power to support both intended and appropriated endoscope uses. We thus explore combining controls to support this diversity of uses. Through an iterative cycle, we design and implement a multimodal and mixed-initiative technique that combines two user controls and one system control. Our evaluations confirm that individual controls do not satisfy the diversity of endoscope uses, and also that our technique indeed does so. Our work highlights the relevance of HCI research in the medical domain through robotic systems.

## Author Keywords

Robotic Endoscope Manipulator; Minimally Invasive Surgery; Mixed-Initiative Interfaces

## CCS Concepts

•Human-centered computing → Interactive systems and tools;

## INTRODUCTION

In Minimally Invasive Surgery (MIS), surgeons insert an endoscope (camera) and instruments into a patient's abdomen through small incisions to operate. The surgeon typically manipulates the instruments while relying on an assistant to manipulate the endoscope. However, delegating control (1) increases variability of the surgical outcome, as communication is error prone [50] and the assistant can create an unsteady image or unintentionally rotate the view [17], and (2) it occupies the assistant with a repetitive task, adding cost to surgery. Our goal is to offload assistants from endoscope manipulation.

Telemanipulated surgical robots, such as the *daVinci* [21], free the assistant as the surgeon controls the endoscope from a distance. However, recent studies question the move towards

telemanipulated robots as they have unintended impacts on perceptual senses and team work [43, 4, 10].

In this paper, we focus on *bedside*<sup>1</sup> robotic endoscopes, which are manipulated at the bedside, next to the patient. Commercial bedside robotic endoscopes have implemented *user* controls through the hands [19], feet [46], voice [20] or the head [16], and research has also explored *system* control relying on image processing [30, 62, 58, 42, 28], fully delegating control to the robotic system. Surprisingly, bedside robotic endoscopes have not been adopted. We thus first take a step back to better understand classic endoscope use and their low adoption rate.

We derive a list of 13 tasks in classic (non-robotic) laparoscopic surgery that reflect both intended and appropriated endoscope uses. While reviewing bedside robotic endoscopes, we find that their design does not take into account this wide variety of uses. We thus turn our research towards controlling bedside robotic endoscopes with the goal of supporting the specific needs and constraints of such identified uses.

Through an iterative cycle, we design a novel interaction technique. First, we select a set of five controls based on their interactive properties and previous empirical findings. Second, we conduct a formative study in the form of video prototypes and retain three controls based on surgeon's perceived advantages and limitations. The result is a mixed-initiative multimodal interaction technique as it combines user and system controls. Finally, we implement this interaction technique and evaluate it in a summative study. We show that our technique indeed supports the different endoscope uses we initially identified.

Our contributions are (1) a better understanding of classic endoscope use, (2) the design of a novel multimodal mixed-initiative interaction technique for bedside robotic endoscopes, and 3) its implementation, including a novel tool tracking algorithm based on robotics inverse kinematics. Altogether, our findings highlight the subtle but complex role of endoscopes in surgical tasks, they shed new light on the relevance of existing bedside robotic endoscopes and demonstrate potential benefits of multimodal and mixed-initiative interaction techniques.

## BACKGROUND

**Minimally Invasive Surgery (MIS)** eliminates the need for large incisions as in open surgery. Surgeons insert tubular devices (trocars) through small incisions on the patient's skin,

---

CHI 2020, April 25–30, 2020, Honolulu, HI, USA  
© 2020 Association for Computing Machinery.

This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2020), April 25–30, 2020, Honolulu, HI, USA*. <http://dx.doi.org/10.1145/3313831.3376795>.

---

<sup>1</sup>We coin this term to describe robots that keep surgeons at the bedside, independently of how they are manipulated. We distinguish from the *daVinci*, which is teleoperated and moves the surgeon from the bedside.

which restricts air flow when introducing and removing instruments. They use elongated mechanical instruments to operate and an endoscope to see inside the body through a 2D monitor. Using small incisions greatly improves the clinical outcome: it lowers morbidity of surgical site infections [12], it lowers estimated blood loss, and it shortens hospital stay and recovery [11, 26, 61]. MIS has become the routine procedure for interventions as cholecystectomy and appendectomy [37, 55].

An assistant manipulates the endoscope while the surgeon operates using both hands, which can introduce variability to the surgical outcome among surgeries of the same type. First, manual endoscope manipulation creates an unsteady image and unintentional rotations of the view, which can degrade performance, lead to errors and increase surgery time [17]. Second, communication between surgeons and assistants is error prone, as often the latter misinterpret instructions on where to move the view [50]. Third, manipulating the endoscope hampers assistants from performing other tasks such as suction, exposition or specimen retrieval [49]. Thus, giving control to the surgeon has great potential to improve surgery.

**Robotic-Assisted MIS** is when a robot assists surgeons. Today, there are mostly *telemanipulated* surgical robots, with the *daVinci* (Intuitive Surgical, Mountain View, CA, USA) [21] dominating the market with over six million procedures performed in 2018 [24]. *Telemanipulation* however (1) separates the surgeon from the patient and surgical team, eliminating haptic feedback and reducing visual feedback which impacts communication [4] and (2) increases costs. Moreover, the clinical benefits with respect to classic MIS have yet to be proven [52].

In this work, we focus exclusively on *bedside* robotic endoscopes, where the surgeon is next to the patient and surgical team when controlling the endoscope. Keeping the surgical team in proximity does not incur the impacts of teleoperation on perceptual senses and communication [43, 4, 10]. Although commercial bedside robotic endoscopes have implemented hand [19], feet [46], voice [20] or even head [16] control, adoption rates have been low. We thus take a step back to understand this problem and propose alternative designs.



Figure 1. Classic Laparoscopic Surgery. The surgeon holds instruments and an assistant the endoscope and an instrument. The displays shows the live video feed of the endoscope. (source: wikipedia)

## RESEARCH METHODOLOGY

**Part I: Identifying The Problem.** We gain understanding of classic (non-robotic) endoscope use through identifying and describing two types of surgical tasks: those in which the endoscope is used for viewing (intended use), and those for which the endoscope is appropriated as a tool (beyond intended use). We then review existing bedside robotic endoscopes, we find that they propose only one control at a time (e.g. the voice or the feet), and, more importantly, that their evaluations do not focus on understanding the tradeoffs of different controls for different tasks. Our conclusion is that there is a fundamental problem in the design of bedside robotic endoscopes: they do not consider the diversity of endoscope uses.

**Part II: Exploring, Designing and Building Interaction.** We design a novel interaction technique for bedside robotic endoscopes that takes into account the variety of tasks identified in our analysis. It is informed by the interactive properties of individual control modalities and previous empirical findings of existing bedside robotic endoscopes. We iteratively explore the combination of controls in two studies. In a first study we use video prototypes to explore the *individual* benefits of five controls (four by the user and one by the system) within our defined tasks. In a second study, we implement three controls and explore their *combination* on the same set of tasks.

## CHARACTERIZING ENDOSCOPE USE: TASKS ANALYSIS

We start our exploration by taking a step back from the technology and analyzing the current (i.e. non-robotic) endoscope uses to characterize it. We use observations from a previous study in HCI [4] and hierarchical task decompositions in the literature for different surgical procedures, including: the Nissen fundoplication procedure [31], laparoscopic radical prostatectomy [1] and cholecystectomy [8]. We incorporate a gynecology surgeon with 10 years of laparoscopic surgery experience into our team to provide feedback. We separate the tasks into two: surgical and endoscope appropriation.

### Intended Use: Surgical Tasks

These tasks correspond to surgical gestures where the endoscope is used to see the actions performed by the tools.

1. *Object Insertion*: when an object such as a trocar, an instrument, a plastic bag (endobag) or a needle is inserted into the patient. The endoscope needs to be focused on the insertion point so that surgeons can make sure the object does not damage any part of the body. Surgeons perform a quick motion of backing the endoscope and turning towards the insertion point until finding it.
2. *Exploration*: when surgeons insert the endoscope into the patient, they scrutinize the body checking the anatomy, e.g. for abnormalities or traces of previous surgeries. Surgeons make large and fast movements with the endoscope.
3. *Focused Work*: when surgeons work in a focused area, for instance for dissecting or cutting. The endoscope should be steady and centered on the instruments.
4. *Anastomosis*: when surgeons do sutures to connect two adjacent channels, such as blood vessels or parts of the intestine. The endoscope should focus on the knot and also the needle when pulling the thread to tighten the knot.

5. *Exposition*: when surgeons push a structure, such as an organ or an artery, with an instrument and make room for the endoscope to pass and show what is behind it. The endoscope is moved inwards behind the structure.
6. *Organ Mobilization*: when instruments are used to move organs and reveal a zone of interest. The endoscope needs to center on the instruments pushing the organ for safety but also on the zone of interest.
7. *Specimen Retrieval*: when a structure is inserted into a plastic bag after it is removed. There is little room for the endoscope, so it is pulled back into the trocar and should show the bag, the specimen, and the tools.
8. *Specimen Extraction*: when the specimen is extracted. The endoscope follows the tool that is grabbing the specimen (or a bag containing it) as it is removed from the body.

### Beyond Intended Use: Endoscope Appropriation

We were surprised to identify instances where the endoscope is used beyond its intended use of observing the body. While the previous tasks are grounded on existing literature, we identified four tasks from observations and from interviewing the expert surgeon part of our team, where the endoscope is used in an accessory way to support surgical tasks.

9. *Endoscope Cleaning*: when surgeons rub the endoscope against tissue inside the body. Debris such as blood often dirty the endoscope, which normally should be removed and cleaned by rubbing it against a sterile gauze. To avoid this interruption and to save time, surgeons oftentimes rub off the endoscope tip against human tissue to remove debris.
10. *Endoscope Dissection*: when the camera shaft is used to break through tissue. The tools normally used for dissection are bipolar forceps and monopolar scissors, yet oftentimes surgeons turn to the endoscope shaft to dissect tissue by using it as a lever. They apply force to the distal end which is multiplied at the proximal end due to the trocar acting as a pivot, a phenomenon known as the *fulcrum effect* [40].
11. *Transillumination*: consists in glowing light through tissue to differentiate inner structures, as each structure reflects light in a particular way. Surgeons shine the endoscope light through the skin to identify blood vessels and other structures that should be avoided when piercing a trocar.
12. *Endoscope Exposition*: exposition is when surgeons push a structure to the side to make room for operating. To expose, an assistant normally uses the grasper tool to hold an organ and advances the endoscope so that the surgeon can advance both surgical tools and operate. Surgeons can also use the endoscope shaft to push an organ sideways while at the same time sliding behind a structure.

We acknowledge this list has limitations: surgical tasks greatly vary according to the specialty and specific procedure. Although the observed appropriation uses were consistent among the different surgeons (from study [4]), this list is not exhaustive as there might be other uses yet to be identified. We now review existing bedside robotic endoscopes and their controls.

## ROBOTIC BEDSIDE ENDOSCOPES

We distinguish the modalities used by existing bedside robotic endoscopes where surgeons execute control from those where surgeons delegate control to the system.

### Surgeons Execute Control

Bedside robotic endoscopes have proposed user controls that solicit the surgeon's head, feet, voice and hands.

#### Single Modality

*Head*. *EndoAssist* (Armstrong Healthcare Ltd.) [16, 14, 15], previously called *EndoSista*, is a commercial robotic endoscope that proposed control using the head. The surgeon wears a headband over the surgical hat and a receiver unit on the belt. The system measures the head movement relative to the sensor, so that the surgeons can move the endoscope to the left (resp. right/up/down) by turning their head to the left (resp. right/up/down). A foot pedal must be pressed for the system to perform movements in order to avoid false activation. *Free-Hand* (Prosururgics Ltd.) [53], the current successor, uses the same headband but here the system continuously shows the action to execute on a small display, then the surgeon pushes the foot pedal to confirm its execution.

Previous research has proposed two techniques without commercialization. The first one is *FAce MOUSE* [38], where surgeons perform a counterclockwise roll motion back and forth with the head to activate pan mode explicitly, then pitching and yawing the head controls the camera, and a new counterclockwise roll motion exits the mode. Likewise, with a clockwise roll motion the surgeon activates zoom in/out through pitching/yawing. The second one implements a parallax effect [56] which is always activated. The surgeon moves to the sides to see beyond the edges of the screen, increasing the amount of spatial information and thus increasing certainty, speed and accuracy of movement.

*Feet*. Three commercial robots propose foot control through a pad with 4 buttons for moving up, down, left and right, plus two buttons for moving in and out: the *AESOP* (Computer Motion) on version 1000 [46] (the first commercial robotic endoscope), on version 2000 [2] and on version 3000 [36]; *ViKY* (Vision Kontrol endoscopY, ViKY Ste Endocontrol-Medical SAS 38000, Grenoble, France) [20]; and *RoboLens* [34]. Another approach is to use foot orientation, which has been explored for sinus surgery [9]. Here, stepping on the ball of the feet, the heel or the sides control the direction of movement, which is only possible when the surgeon is sitting.

*Voice*. *ViKY* [20] implements a voice control interface, where the surgeon says "Viky" and then either "right", "left", "up", "down", "in" or "out". At the same time, the surgeon must push a foot pedal to confirm the action and avoid false activation. *RoboLens* [34] implements a similar interface: while the surgeon is pressing the foot switch, the system listens to the microphone and interprets commands. The system repeats the interpreted command over headphones, so that the surgeon can release the pedal in case of misinterpretation.

*AESOP 2000* [2] also interprets voice commands to control the direction of the movement. The system provides two modes, a first one where it moves the endoscope until the surgeon instructs “stop”, and, a second one where the system makes a small movement per command spoken (e.g. “left”). To avoid false activation the system attains a more accurate recognition by requiring a voice-training stage instead of using pedals, consisting of repeating 23 words six times.

*Hand.* Two commercial robots propose control through a joystick for 2D motion (left–right, up–down) and buttons for moving inwards and outwards: *SOLOASSIST* [19, 29] (AKTORmed GmbH) and *AESOP 1000* [46]. *SOLOASSIST* provides a joystick attached to the instrument, whereas *AESOP*, provides a cabled remote which the surgeon needs to grab with one hand, letting go of the surgical instrument. Regarding research, both 2D [54] (a ThinkPad TrackPoint joystick attached to the instrument) and 3D joysticks (finger through ring [7]) have been proposed.

### Multimodality

We define multimodality as having multiple modalities available while executing a task. We discuss systems that propose more than one modality and the studies that compare them.

*Voice and Feet.* Mirbagheri et al. [34] evaluated *RoboLens* to compare voice vs. feet, first measuring technical aspects (accuracy, range of motion) and then preference after 30 laparoscopic operations. Surgeons preferred the foot pedal as the voice command had a large delay in the execution. Still, they remarked that the pedal is not ideal as they need to look down to find the correct one among the other pedals for activating electrical tools.

In a study with six porcine models [20] using *ViKY*, half performed with voice control and half with feet, surgeons preferred pedals because voice commands were sometimes misinterpreted by the system (12 times in average, range = 9–21).

Allaf et al. [2] evaluated voice vs. feet on the *AESOP 2000* using a task that required surgeons to move the endoscope and focus on targets (the endoscope had a cross hair that had to be matched with the target). Foot was faster and learned quicker, command misinterpretation was higher with voice.

*Feet and Hand.* The *AESOP* [46] is one example, but we did not find studies comparing these interfaces.

*Feet, Hand and Voice.* One study [33] compared the *AESOP 2000* in 50 operations, half using the hand and feet control and half voice. This is the only study we found that reports on two available modalities at once, although, there are no findings reported on how hand and voice interacted, their individual or combined benefits or the challenges in combining them. Also, it is not clear if surgeons used them interchangeably or exclusively, even if both were available. Authors report that feet is preferred to hand, as the application of the handpiece limits the surgeon to using two arms. However, authors observe that for feet surgeons have to look down to find the pedal and this decreases coordination. Lastly, voice appeared to be faster.

### Surgeons Delegate Control

Here, the system controls the endoscope. The success depends on the capacity of automatically estimating the viewpoint that surgeons need. Existing systems have adopted one of two approaches: tracking instruments (object) or taking into account the activity (semantics) to compute the viewpoint.

Tracking objects relies on visual servoing, the process of tracking features in a video by comparing their current and desired position and then moving the robotic endoscope to reduce this distance [5]. Implementations of this approach include: computer vision using classification on the *AESOP* [30], monochrome markers for instruments [62], color markers [58, 42, 28] and later more complex methods, such as model-free efficient second-order minimization (ESM) tracking algorithm based on histogram matching [6]. All these methods have the limitation of relying on a clear view of the scene for processing the image, although frequently in surgery, blood, internal structures or smoke from cauterization occlude the endoscope lens.

Tracking using semantics relies on extracting knowledge from the surgery. One approach is using information from previous interventions to determine the optimal zoom ratio [39]. Another to simulate a cognitive approach with long-term memory of interventions and surgical facts, procedural memory of learned behaviors and an episodic memory that stores user experiences; and a short-term memory where the surgical state and focus of attention are kept [45].

While these two approaches are promising, a major limitation is that they do not take into account surgeons’ intentions, for example, when a surgeon wants to closely examine a structure or show a zone with the purpose of teaching an assistant.

### Conclusion

While several studies evaluated bedside robotic endoscopes, they have focuses on showing that certain controls perform better than an assistant. It still remains unclear what are the impacts of the evaluated control properties on interaction. Moreover, these evaluations are performed in actual surgeries (in-vivo or porcine model) without a clear formalization of different tasks. This is appropriate for ecological validity, although it does not let us understand the individual benefits of a control for a specific task, i.e. why a control works or not in a particular situation.

We have previously shown the variety of tasks imposing different requirements (e.g. haptic feedback) and constraints (e.g. the need to use both hands) on endoscope use. It appears that existing bedside robotic endoscopes have not taken this into consideration enough in their design and evaluation. We argue this is one of the reasons behind the low adoption of bedside robotic endoscopes, as it limits the surgical tasks they can effectively support. We believe that one solution is to provide different user controls and combine them with a system controls that automatically moves the endoscope.

## DESIGNING AN INTERACTION TECHNIQUE

We aim at designing a novel interaction technique for bedside robotic endoscopes that supports the diversity of endoscope uses we identified. We study the potential benefits of combining user and system controls as we believe providing surgeons with different types of controls is critical to support our identified endoscope uses. However, this design space is large: it is unclear which controls to consider and whether they can easily be combined altogether. We now present the reduction process to select the final controls.

### Excluded Controls

We exclude *voice* based on empirical findings and theory. Previous studies show voice is error prone, both because of system command misinterpretation [20, 36] and because having to repeat commands distracts the surgeon [2]. Moreover, voice remains more appropriate for discrete control (e.g. “*move left .. move left .. move left*”) rather than continuous control. We exclude *feet* based on previous empirical evidence showing that when surgeons use bedside robotic endoscopes that rely on pedals, they look down to find the correct one among those for electric tools, which draws their concentration away from the surgery [34]. Our team’s expert surgeon confirms this from his own experience using pedals.

Finally, in a brainstorm with our team’s expert surgeon, we discard different body parts (e.g. elbow, knees, hips, tongue) that could be used for control as they affect the comfort and/or have limited degrees of freedom. For instance, in surgery, elbows are kept close to the body purposely to avoid fatigue.

### Retained User Controls

#### Head

Despite the negative results in commercial bedside robotic endoscopes, we believe this modality has potential as HCI has done advanced explorations and implementations [60, 51, 22]. The head provides not only a high level of immersion [57] but also a high level of efficiency when navigating in 3D environments while the hands are used for a primary task [25]. Current commercial bedside robotic endoscopes require an explicit activation with the feet as already highlighted, and they only provide discrete control which impairs both performance and comfort. Moreover, studies do not provide information about the control–display ratio which is critical for such controls.

**1. POSTURE TRACKING:** we retain the head for control and offer more freedom by tracking the whole posture to control the endoscope’s left-to-right directional movement.

#### Hand

We also retain *hand* control despite the fact that surgeons already manipulate surgical instruments, as it provides full haptic feedback, which is generally not possible with other modalities.

**2. HAND MANIPULATION:** first, we keep the classical approach where surgeons grab and move the endoscope with one of their hands, as they are already familiar.

**3. JOYSTICK:** second, we also consider control with a joystick mounted on to the side of one instrument. One commercial (*SOLOASSIST*) and two academic works [54, 7] implement a tool-mounted joystick. These do not suggest inadequacy for endoscope control, and given it provides eyes-free control in other contexts such as text entry [59], we retain it to study its individual benefits and integration with other techniques.

**4. FORCE SENSING:** finally, we propose a novel control with the hand, which conceptually consists in directly using the instrument as a joystick. Surgeons first lock the instrument by pushing a pedal, then, they apply forces to it which the system detects and uses as input to control the endoscope: when the surgeon tries to push inwards, the system moves the endoscope in, same for out or sideways motion. We believe this control is a good fit for tasks where surgeons cannot move their tools (e.g. doing a knot during an anastomosis) but they need to see a place (e.g. the entry point when inserting a scissor to cut the suture).

### Retained System Controls

Automatic camera control has been implemented in virtual environments either by tracking *objects* of interest such as a character’s head [27] or by using *semantics*, such as the narrative of a story [18]. These two approaches have been explored for endoscopes, but never compared to or integrated with user controls.

**5. TOOL TRACKING:** we implement automatic control through tool tracking. One limitation of existing solutions is that they rely on computer vision, thus, we leverage our existing setup where a robot holds the tool, and compute the instrument position based on inverse kinematics, not the endoscope image. We thus avoid problems from occlusion. The endoscope automatically centers the view on the primary tool tip as it moves across the surgical site. To avoid permanent movement of the view as the surgeon operates, we divide the screen into two regions: the *center* where tool movement does not trigger a corrective action of the endoscope and the screen *edge*, defining a spatial activation threshold beyond which a compensatory movement is activated to recenter the instrument tip.

We retain five controls as our goal is to combine them to support surgeons with an appropriate control for each task. A key aspect of this combination is that it is a mixed-initiative [3] approach: it mixes different surgeon controls (head and hand) with system controls (tool tracking). In mixed control, the user and the system cooperate to accomplish a task, thus, in our context, surgeons can choose to execute or delegate control according to the task at hand. This is promising as delegating control offloads surgeons while executing control lets surgeons deal with complex, unusual situations.

We now conduct a video-based evaluation to refine our choice of individual controls and, in a second step, investigate how they can be combined.

## STUDY 1: EXPLORATION OF ENDOSCOPE CONTROLS

We investigate the advantages and limitations of the retained controls by evaluating the endoscope uses we identified using a set of criteria developed with the team’s expert surgeon and semi-structured interviews. To increase transparency and reproducibility, we preregistered this study and provide access to the protocol, questionnaire and video prototype<sup>2</sup>.

### Participants

We recruited 6 laparoscopic surgery surgeons (Table 1).

	Experience (years)	Gender	Institution	Specialty
<i>P1</i>	12	M	H1	Gynecology
<i>P2</i>	11	F	H1	Gynecology
<i>P3</i>	7.5	M	H1	Gynecology
<i>P4</i>	15	M	H2	Digestive Surgery
<i>P5</i>	31	M	H2	Digestive Surgery
<i>P6</i>	20	M	H3	Pediatric Surgery

Table 1. Study 1 participants

### Method and Data Collection

We video-recorded an actor enacting each control using the robotic arms described in System Implementation, although a confederate moved the endoscope from a computer. This *wizard-of-oz* method lets us identify the relevance of controls early-on. We acknowledge it has limitations as watching video of a control is different from testing it hands-on. Moreover, it might lead to discarding some controls too quickly, thus more work is necessary to completely discard the excluded controls.

Participants first watched the user controls: HAND MANIPULATION, POSTURE TRACKING, JOYSTICK, FORCE SENSING evaluating them in a 7-point likert scale using the following criteria: stability, movement fluidity, movement economy, movement amplitude (magnitude of distance), precision, mental load and intuitiveness. Following, they watched the system control video, TOOL TRACKING, answering to the same questionnaire. In a semi-structured interview, participants evaluated the controls for our surgical tasks, followed by an open interview. Interviews were audio recorded, quotes are translated into English.

### Data Analysis

We present a visualization of the questionnaire data (Figure 2) as this is not a controlled experiment and the number of participants is low for statistical analysis. For the semi-structured questions and open interviews, one of the principal investigators listened to the interviews and took notes. We then performed a qualitative analysis to answer our research questions, underpinned by a bounded relativist ontological stance and constructivist epistemological stance [35]. We took an inductive approach, creating meaning from specific examples in the interviews. The investigator that performed the analysis had a background in HCI and is heavily influenced by the view that computer systems, and especially robots, should not be designed with the goal of replacing human activity but to enhance it in a partnership.

<sup>2</sup><https://osf.io/w6ud3/>

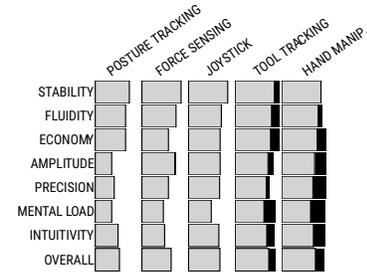


Figure 2. Tabular visualization [44] summarizing rates for the criteria. Bars indicate mean scores. The black fillings indicate the portion above the mean value for the criteria. Controls are sorted by Overall score. Mental Load scores are inverted to simplify readability.

### Results

The best ranked controls were HAND MANIPULATION and TOOL TRACKING, with scores above the average (Figure 2).

HAND MANIPULATION was perceived as having the lowest mental load and being the most intuitive. It was noted as the best for all the surgical tasks that require ample and free movements: *Exploration*, *Object Insertion* and *Specimen Extraction*. Also, it is the only one adequate for the tasks where the endoscope is appropriated: *Endoscope Cleaning*, *Endoscope Dissection*, *Transillumination* and *Endoscope Exposition*. However, participants noted that it requires one hand which interrupts the surgical task, and *P6* noted that when surgeons stand on one side of the patient, they cannot reach the other side to grab and manually move the endoscope, making this control cumbersome for some types of surgery.

POSTURE TRACKING received mixed comments. Although surgeons tend to lean their bodies sideways when trying to see beyond the screen edge, this movement is considered “*parasite*” (*P3*) and should be avoided. *P1* saw a great potential as this control does not solicit the hands, contrary to all other controls, provided that it works flawlessly. The biggest fear shared across all participants is false activation: the system reacting to posture changes when turning the head to talk to colleagues or picking up a tool from the table. Participants also showed concerns about imposed posture constraints, as posture plays an important role: to avoid fatigue surgeons keep elbows low and shoulders straight (*P1*, *P4*); to rest, surgeons change position and move their head to the sides. *P2* mentioned the control might be cognitively demanding. Overall, participants noted that this control might work for tasks with low movement amplitude where the tools stay on the field of view, such as *Focused Work* and *Anastomosis*.

JOYSTICK was appreciated by all surgeons as they can manipulate the endoscope without letting go of the tools, and it was perceived as appropriate for narrow movements (*P4*) Still, it was perceived as slow and that it could affect the ergonomics during tool manipulation as it solicits one finger (*P1*, *P2*, *P4*, *P5*). Lastly *P4* had used the *SOLOASSIST* [19] before with a joystick interface and acknowledged that, although in theory it does not pose problems, in practice the joystick direction is inverted when the tool is rotated, possibly leading to confusion.

FORCE SENSING was perceived similar to the JOYSTICK, in general surgeons saw the potential but were worried about it being imprecise (*P1*), leading to high mental load (*P2*), and being slow. Moreover, the problem of how to activate this control still remains an issue, with participants suggesting using a pedal or a button. Still, two surgeons saw potential in the control (*P3*, *P4*), especially *P4* as one can move to one place (e.g. instrument insertion) while doing a task that requires the tool to be still (e.g. suture).

Both JOYSTICK and FORCE SENSING were perceived adequate for all tasks with low amplitude movement, both when tools stay in the view (*Focused Work* and *Anastomosis*) and when they might be outside (*Exposition*, *Organ Mobilization*, *Specimen Retrieval* and *Specimen Extraction*).

TOOL TRACKING was every participant's favorite control. *P1* and *P3* thought it was "*seducing*", *P5* that it was "*very intuitive*" and "*very interesting*", *P6* that "*this is something very interesting [...] it's exactly what we need*". Every participant acknowledged that the control does not interrupt surgery, and that it would work during most of the surgical tasks. *P1* was concerned with practical issues such as how to select which tool to follow or how the system would behave when following two tools. All were concerned about how to use this control to see places where there are no tools. We thus explored during interviews ideas on how to enable and disable the control. One possibility that came up was displaying a dashed line on the screen to separate the center from the edge regions, such that tracking is activated when the tool crosses the line but not when it passes between lines. However, as *P3* put it, it is hard to focus on two different layers of information. Another possibility is having the system follow only the dominant tool. This control was perceived adequate for all surgical task except for *Exploration* and *Object Insertion* since the task does not necessarily involve handling instruments.

### Conclusion and Final Design

We retained HAND MANIPULATION as it is adequate for tasks that require high movement amplitude and/or haptic feedback. Both JOYSTICK and FORCE SENSING are adequate for low amplitude movement, we chose to retain JOYSTICK as it does not require explicit activation. We initially valued POSTURE TRACKING as it is hands-free and enables natural camera control, however, this study showed that we underestimated its impact on surgeons' posture. We thus discarded this control. Finally, we retained TOOL TRACKING as it was extremely well received and it lowers the cognitive load as surgeons can focus on operating without thinking about controlling the endoscope.

Our final design is a multimodal and mixed-initiative technique. It combines TOOL TRACKING (system control), which offloads the surgeon when working in a focused area, with HAND MANIPULATION and JOYSTICK (user controls) which lets surgeons do both ample and narrow movements.

## SYSTEM IMPLEMENTATION

We implement a first prototype of our multimodal mixed-initiative technique. We perform a preliminary evaluation with 6 surgeons, mostly the same ones as Study 1, to calibrate the technique and improve usability. This led us to use a smaller joystick than originally planned, to fine tune the speed of movement in tool tracking, and to include tracking in depth, moving the endoscope inwards and outwards as surgeons move the tools in these directions.

### Setup

The setup (Figure 3-a) has two main components: a robotic instrument manipulator and a robotic endoscope manipulator. We first describe the technology for both robotic arms.

**Robotic Arms.** These are modified Haption Virtuose 3D<sup>3</sup> robots: compact, lightweight and characterized by six rotational joints that allow full motion across the surgical workspace. One important feature is their high transparency. Contrary to commercial bedside robotic endoscopes, such as the *daVinci* [21], *AESOP* and *ViKY* that exhibit *stiff* responses to external actions, these robotic arms are characterized by a highly *compliant* behavior: the cable-driven actuation significantly lowers the frictional phenomena at the joint level, conferring a high transparency, ease of hand-manipulation and safety. We implemented by software (1) tool weight compensation, to make tool-holding more transparent to surgeons, and (2) a lock function that after 3 seconds locks the arm into a precise position (status shown by LEDs mounted on the arms). The surgeon can unlock the system by gently pulling to overcome a defined force threshold of 2N.

**Robotic instrument.** Surgical instruments can be mounted onto the robotic arm (Figure 3-1) by using a magnetic clipper. This arm can sense the cartesian tool tip position in real-time (1000Hz), which is a key feature for TOOL TRACKING.

**Robotic endoscope.** An endoscope is attached to the robotic arm (Figure 3-2) through a customized metallic support, ensuring the correct orientation of the optical axis. It has a FullHD (1980 × 1080) resolution and a refresh rate of 50FPS, its video feed is displayed on the tablet screen (Figure 3-3). The lock function ensures a stable view.

**Calibration.** It is essential that the two robots know (1) their relative position and (2) the relative position of the entry points (trocars) for precise interaction, especially TOOL TRACKING. A calibration procedure is required in the pre-operative phase to estimate the 3D pose of the robotic endoscope with respect to the robotic instrument. Then, our system is capable of automatically estimating the position of the trocars by computing the best intersection of a set of straight lines that represent the instrument position at different times [13].

**Controls Implementation.** TOOL TRACKING relies on the robots' kinematics to compute the tool tip position in the 3D reference frame of the endoscope robot. We then project the tool tip position onto the video feed, using the endoscope calibration parameters, which is divided into three regions:

<sup>3</sup><https://www.haption.com/en/products-en/virtuose-3d-en.html>

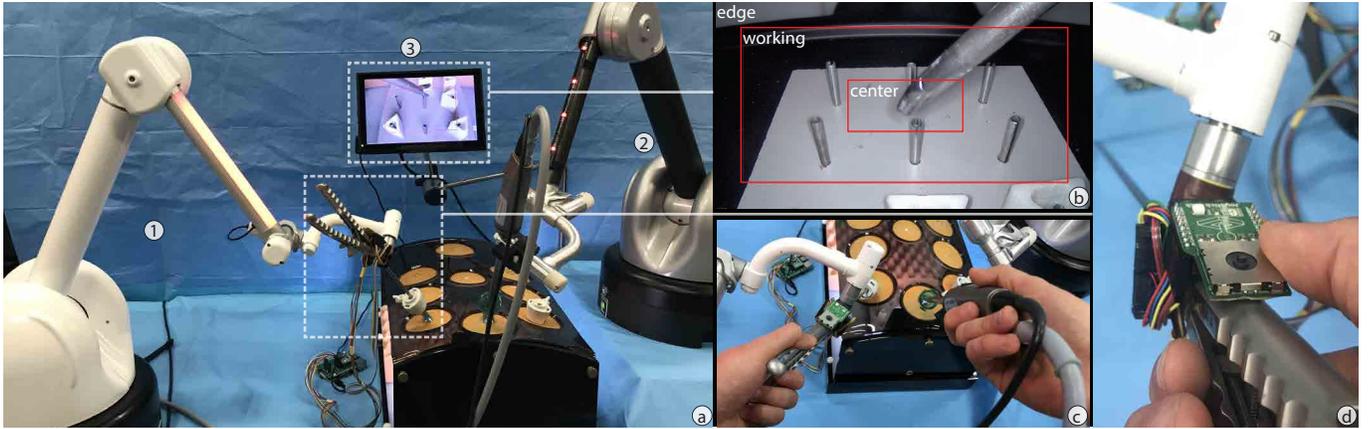


Figure 3. (a) System. (1) A robotic instrument that can capture the tool tip position, (2) a robotic endoscope controlled by the surgeon and the system and (3) a screen displaying the endoscope video feed. Controls: (b) TOOL TRACKING, (c) HAND MANIPULATION and (d) JOYSTICK.

*center*, *working* and *edge* (Figure 3-b). The endoscope does not move while the tool is inside *center* or *working*; when it enters the *edge*, the endoscope centers the tip into *center*. Our approach has great advantages over visual-based tracking as it is precise and robust to visual occlusion, a common problem in surgery (e.g. by blood, organs or smoke). The first prototype implemented only lateral tracking of the endoscope pose. However, the pilot study showed that tracking in depth was necessary for tasks as *Focused Work* and *Anastomosis*, so we added tracking according to the tool’s depth of insertion. Finally, our teams’ expert surgeon calibrated movement speed and the size of the areas for optimal performance.

JOYSTICK relies on a finger-controlled joystick clipped at the base of the instrument handle, within finger reach (Figure 3-d). In the first prototype, we used a *KY-023* push joystick<sup>4</sup>. The pilot study showed that it was not easy to manipulate due to its size. Therefore, we switched to the more compact *MIKROE-1506* push joystick<sup>5</sup>. We read the output using a Raspberry PI and pass it on through a web socket to our main application. Left, right, up and down movements are mapped to the same directions in the endoscope tip; if the joystick is pushed, the zoom control is activated: within a 1-second window, the up/down input will be mapped to in/out, controlling the depth of the endoscope.

HAND MANIPULATION is enabled when the surgeon pulls and unlocks the endoscope, enabling movement without restrictions while the robot only provides gravity compensation.

### Control Switching

Our technique seamlessly integrates system and user controls. All controls are always enabled. Tracking becomes active when the surgeon moves the tool tip inside the *center* (Figure 3-a), and inactive when another control is currently used. At any time, the surgeon can switch to HAND MANIPULATION by grabbing the endoscope and pulling it (Figure 3-c), or to JOYSTICK by operating it (Figure 3-d).

<sup>4</sup>[http://sensorkit.en.joy-it.net/index.php?title=KY-023\\_Joystick\\_module\\_\(XY-Axis\)](http://sensorkit.en.joy-it.net/index.php?title=KY-023_Joystick_module_(XY-Axis))

<sup>5</sup><https://www.mikroe.com/joystick-click>

## STUDY 2: TECHNIQUE AND PROTOTYPE EVALUATION

We evaluate our interaction technique, which combines two user controls and one system control. To increase transparency and reproducibility, we preregistered this study and provide access to the protocol and questionnaire<sup>6</sup>.

### Participants and Apparatus

We recruited 6 laparoscopic surgery surgeons, all different from Study 1 (Table 2). We used a laparotrainer, the standard equipment for evaluating surgical skills [48], which has proactive transfer to real surgery [47].

	Experience (years)	Gender	Institution	Specialty
<i>P1</i>	3	M	H1	Pediatric Surgery
<i>P2</i>	10	F	H2	Gynecology
<i>P3</i>	14	F	H2	Gynecology
<i>P4</i>	6	F	H3	Gynecology
<i>P5</i>	14	M	H2	Gynecology
<i>P6</i>	25	M	H2	Gynecology

Table 2. Study 2 participants

### Method and Data Collection

Participants tested HAND MANIPULATION, JOYSTICK and TOOL TRACKING individually first, then their combination. We chose six tasks: four surgical, where two require ample movements (*Exploration*, *Object Insertion*), two require narrow movements (*Focused Work*, *Anastomosis*); and two where the endoscope is appropriated (*Endoscope Cleaning* and *Endoscope Exposition*). Participants performed these tasks:

1. *Exploration*: showing the four corners of the laparotrainer;
2. *Object Insertion*: inserting a tool under eye control, thus first visualizing the trocar and then inserting the tool;
3. *Endoscope Cleaning*: rubbing the endoscope onto a surface;
4. *Focused Work*: transferring 4 pegs in a  $3 \times 4$  peg transfer board using two tools, transferring the peg between them;
5. *Anastomosis*: suturing two structures; and,
6. *Endoscope Exposition*: pushing a structure with the endoscope to see underneath it.

<sup>6</sup><https://osf.io/mx6gn>

The user evaluation consisted of two parts.

- *Part I: All Tasks with All Controls.* Participants perform each task using each control randomly to get acquainted with them and to perceive their trade-offs. Then, they evaluate if each control is appropriate for each task using a 7-point scale.
- *Part II: Series of Tasks.* Participants perform all tasks in the following sequence: *Exploration, Object Insertion, Endoscope Cleaning, Focused Work, Anastomosis* and *Endoscope Exposition*. This sequence was defined with our team’s expert surgeon to best reflect task ordering during surgical procedures. We observe if participants choose the control rated as the highest in each task. After finishing, we ask questions in an open interview focusing on whether they chose their preferred control and why (or why not), problems switching controls and aspects to improve.

We video-recorded participants in both parts.

### Data Analysis

For likert-scale questions (Part I) we analyze the data visually as this is not a controlled experiment and the number of participants is low for a statistical analysis. One of the researchers watched the recorded videos and listened to the interview audios using the same approach as in Study 1.

### Results

First and foremost, this study showed that the implementation was robust, it worked for our selected tasks and participants believed it could be used for real surgery. Participants were surprised by the implementation of the different controls, especially TOOL TRACKING. Feedback focused on high-level (e.g. control combination) rather than low-level (e.g. size of the joystick) interaction considerations, as it was the case on the pilot study during the implementation phase. We now present results for each control both from the questionnaire (Figure 4) and interviews. Then, we discuss their combination.

#### Part I: Individual Controls

HAND MANIPULATION was rated as the best for all surgical tasks, which was expected since this is how surgeons work today, with the exception of *Focused Work* for which the automatic control was rated best. We want to highlight that for the two endoscope appropriation tasks, this control received the highest rating and it was the only one with which participants could actually perform *Endoscope Exposition*. As four participants highlighted, it is necessary to have the full haptic feedback that only manipulation through the hand provides to perform the task correctly.

JOYSTICK received a high rating overall except for *Endoscope Exposition*, as it requires feeling the tissue. We were surprised to see the high rating for *Endoscope Cleaning*, and that most participants could do this task, although in our case they did not have to pay attention to how much pressure they apply as this was not human tissue. Despite the high rating, five participants thought the joystick was not intuitive and had difficulties finding the best way to use it—moving to the wrong side or overshooting. Only *P5* mastered it from the start. Still, they thought they could use it in a real surgery by adapting their

gestures, and highlighted that it would not be an impediment as they adapt their gestures very often, such as when working with a new assistant. *P1, P2* and *P4* further added that they could master the joystick with practice.

Regarding our implementation, three participants thought the joystick was sensible, and *P4* actually touched it by accident several times resulting in unintended movements.

TOOL TRACKING was highly rated for all surgical tasks and low for the endoscope appropriation tasks. Participants mentioned this is because the tasks require haptic feedback which is only present when grabbing the endoscope.

Although every participant was impressed with this control, it became clear that it fails when the point of interest is not the tip of the instrument. The most notable example is when making sutures (*Anastomosis*) and pulling the string to tighten the knot: it is important to see the knot but here the endoscope moved to focus on the tool performing the pulling action, leaving the knot outside of the view. Again, participants remarked that they can use this control by adapting their gesture, in this example by making repeated small pulling motions instead of a long pull, which they already do when suturing in small areas.

In the specific case where the point of interest is not the tip of the tracked instrument, rather the other instrument, participants suggested that the system could switch the tool being tracked. Indeed, one limitation of our setup is that only one instrument is attached to a robot. Thus, the strength of our method from using kinematics and not image processing in this case is a limitation, as we can only track instruments attached to a robotic arm.

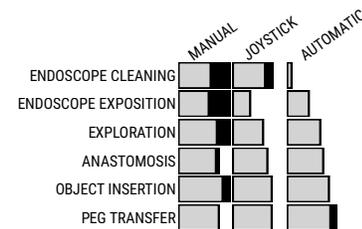


Figure 4. Tabular visualization [44] summarizing scores for each criteria. Bars indicate mean scores. The black fillings indicate the portion above the mean value for the criteria.

#### Part II: Combination

We observed that, generally, participants switched to their preferred control in each task. When asked explicitly if the previous task influenced their choice, all answered no. We believe that the cost of switching is lower to the cost of performing a task with a sub-optimal control. The mixed-initiative was perceived as “*seducing*” (*P6*), as participants saw the benefit of system and user control for each task. Our choice of an implicit switch from user to system control by placing the tip at the center, and from system to user control by grabbing the endoscope or using the joystick was seen as natural and intuitive.

Regarding the activation of TOOL TRACKING, four participants would have liked to have a way to explicitly deactivate the control. After having found a good plane of view for *Focused Work* or *Anastomosis*, participants did not want to risk the system moving the endoscope. They were hesitant of their movements and unsure of how to avoid the system taking over.

Similarly, all participants acknowledged that while they want the system to follow them sideways, they want to have control over the zoom level. Our method keeps the tool tip at a set distance from the endoscope tip, so once participants found a comfortable distance and started working on a task, the automatic tracking would take over and set the distance to a defined value. Participants suggested that placing the endoscope using HAND MANIPULATION or JOYSTICK should set the distance for when TOOL TRACKING takes over.

We explored in the interview how the joystick could be used to activate and deactivate TOOL TRACKING by clicking it (as a button), and to use one of the directions (up–down) to control the zoom. This suggestion was well received by four participants, and we were surprised that two (*P3*, *P6*) made this suggestion before we mentioned it.

## IMPLICATIONS FOR DESIGN

### The Importance of Real Tasks

Previous robotic endoscope research has not taken into account the variety of situated endoscope uses. Our evaluations confirmed that different tasks impose different constraints on the surgeon (e.g. must use two hands) and on the type of movement needed (e.g. high amplitude movements). The design and evaluation of future systems should consider this variety of uses. Moreover, specific surgeries may include different task which we have not yet considered, thus, we also encourage their further exploration.

### Combining Controls

Combining controls has been studied in HCI for interacting with systems in the surgery room, although there is more ground to cover to better understand interaction in this critical-context environment. Mentis et al. [32] studied voice and hand gesture control for a system that shows preoperative images, and conclude that there is a case for redundancy as their benefits are circumstantial. O’Hara et al. [41] highlights that control techniques should be designed to work within operating rooms, as individual and also collaborative practices vary. Our findings show this is also the case for endoscopes: existing controls alone do not cover the wide space of tasks. Combining controls creates a synergy where surgeons can take advantage of their individual benefits and chose according to the task at hand and their personal preference. Future robotic endoscopes should also explore the combination of different controls.

### Automatic Control

Automatic tool tracking showed great potential in our studies as it offloads the surgeon, to the point where they take control for granted. Still, vision-based approaches studied so far are greatly limited as image occlusion is common during surgery. We introduced a viable alternative: taking advantage

of robotics inverse kinematics. Future robotic endoscopes can also benefit from this approach, and even explore the combination with visual tracking. Moreover, our evaluation revealed that automatic control is not enough and has limitations: although implicit activation based on instrument position is adequate and provides a smooth transition from user controls, it is important to provide additional explicit deactivation for cases where surgeons want to keep a particularly good point of view or in case of emergencies.

### Domain Application for HCI

Given the high rate of adoption of robotic systems in the operating room, research in HCI and robotics must keep up with this trend to improve interaction. The *daVinci* robot went from 652,000 surgeries in 2015 [23] to over six million by the end of 2018 [24]. But it is only the beginning of a long exploration of robots in the operating room. We encourage more research in HCI that can be applied for the development of robotic systems in the surgical domain.

Conversely, HCI research can also learn from the surgical domain. In the case of camera control for instance, research has highly evaluated head control, however, for surgeons, this control is not well adapted. We hope our work inspires further work on virtual camera control, to include the case of surgeons and endoscope control.

## ACKNOWLEDGMENTS

This research was supported by the Labex CAMI (project ANR-11-LABX-0004) operated by the ANR as part of the program “*Investissement d’Avenir*”. We thank all the participants of this study for their valuable time.

## REFERENCES

- [1] Latif Al-Hakim, Ming Wang, Jiaquan Xiao, Dennis Gyomber, and Shomik Sengupta. 2019. Hierarchical task analysis for identification of interrelationships between ergonomic, external disruption, and internal disruption in complex laparoscopic procedures. *Surgical Endoscopy* Article not assigned to an issue (Jan. 2019), 1–15. DOI : <http://dx.doi.org/10.1007/s00464-018-06656-z>
- [2] 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. *Surgical Endoscopy* 12, 12 (Dec. 1998), 1415–1418. DOI : <http://dx.doi.org/10.1007/s004649900871>
- [3] J. E. Allen, C. I. Guinn, and E. Horvtz. 1999. Mixed-initiative interaction. *IEEE Intelligent Systems and their Applications* 14, 5 (Sept. 1999), 14–23. DOI : <http://dx.doi.org/10.1109/5254.796083>
- [4] Ignacio Avellino, Gilles Bailly, Geoffroy Canlorbe, Jérémie Belghti, Guillaume Morel, and Marie-Aude Vitrani. 2019. Impacts of Telemanipulation in Robotic Assisted Surgery. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, 583:1–583:15. DOI : <http://dx.doi.org/10.1145/3290605.3300813>

- [5] Andreas Bihlmaier. 2016. Endoscope Robots and Automated Camera Guidance. In *Learning Dynamic Spatial Relations: The Case of a Knowledge-based Endoscopic Camera Guidance Robot*, Andreas Bihlmaier (Ed.). Springer Fachmedien Wiesbaden, Wiesbaden, 23–102. DOI: [http://dx.doi.org/10.1007/978-3-658-14914-7\\_2](http://dx.doi.org/10.1007/978-3-658-14914-7_2)
- [6] F. Bourger, C. Doignon, P. Zanne, and M. de Mathelin. 2007. A Model-free Vision-based Robot Control for Minimally Invasive Surgery using ESM Tracking and Pixels Color Selection. In *Proceedings 2007 IEEE International Conference on Robotics and Automation*. IEEE, New York, NY, USA, 3579–3584. DOI: <http://dx.doi.org/10.1109/ROBOT.2007.364026>
- [7] 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 (April 2000), 395–399. DOI: <http://dx.doi.org/10.1007/s004640020066>
- [8] C. G. Cao, C. L. MacKenzie, J. A. Ibbotson, L. J. Turner, N. P. Blair, and A. G. Nagy. 1999. Hierarchical decomposition of laparoscopic procedures. *Studies in Health Technology and Informatics* 62 (1999), 83–89.
- [9] Chan Jason Y. K., Leung Iris, Navarro-Alarcon David, Lin Weiyang, Li Peng, Lee Dennis L. Y., Liu Yun-hui, and Tong Michael C. F. 2015. Foot-controlled robotic-enabled endoscope holder for endoscopic sinus surgery: A cadaveric feasibility study. *The Laryngoscope* 126, 3 (Sept. 2015), 566–569. DOI: <http://dx.doi.org/10.1002/lary.25634>
- [10] Amy Cheatle, Hannah Pelikan, Malte Jung, and Steven Jackson. 2019. Sensing (Co)Operations: Articulation and Compensation in the Robotic Operating Room. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW (Nov. 2019), 225:1–225:26. DOI: <http://dx.doi.org/10.1145/3359327>
- [11] Hanyu Chen, Bin Ma, Peng Gao, Hongchi Wang, Yongxi Song, Linhao Tong, Peiwen Li, and Zhenning Wang. 2017. Laparoscopic intersphincteric resection versus an open approach for low rectal cancer: a meta-analysis. *World Journal of Surgical Oncology* 15, 1 (Dec. 2017), 229. DOI: <http://dx.doi.org/10.1186/s12957-017-1304-3>
- [12] M. W. Dobson, D. Geisler, V. Fazio, F. Remzi, T. Hull, and J. Vogel. 2011. Minimally invasive surgical wound infections: laparoscopic surgery decreases morbidity of surgical site infections and decreases the cost of wound care. *Colorectal Disease* 13, 7 (2011), 811–815. DOI: <http://dx.doi.org/10.1111/j.1463-1318.2010.02302.x>
- [13] Lin Dong and Guillaume Morel. 2016. Robust trocar detection and localization during robot-assisted endoscopic surgery. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, New York, NY, USA, 4109–4114. DOI: <http://dx.doi.org/10.1109/ICRA.2016.7487602>
- [14] Patrick Finlay. 1996. Clinical experience with a goniometric head-controlled laparoscope manipulator. (1996).
- [15] Patrick Finlay. 2001. A robotic camera holder for laparoscopy. In *Proceedings of 10th International Conference on Advanced Robotics. Workshop 2 on Medical Robotics, 2001*. IEEE, New York, NY, USA, 129–132.
- [16] P. A. Finlay and M. H. Ornstein. 1995. Controlling the movement of a surgical laparoscope. *IEEE Engineering in Medicine and Biology Magazine* 14, 3 (May 1995), 289–291. DOI: <http://dx.doi.org/10.1109/51.391775>
- [17] Anthony G. Gallagher, Musallam Al-Akash, Neal E. Seymour, and Richard M. Satava. 2008. An ergonomic analysis of the effects of camera rotation on laparoscopic performance. *Surgical Endoscopy* 23, 12 (Dec. 2008), 2684. DOI: <http://dx.doi.org/10.1007/s00464-008-0261-x>
- [18] Quentin Galvane, Rémi Ronfard, Marc Christie, and Nicolas Szilas. 2014. Narrative-driven Camera Control for Cinematic Replay of Computer Games. In *Proceedings of the Seventh International Conference on Motion in Games (MIG '14)*. ACM, New York, NY, USA, 109–117. DOI: <http://dx.doi.org/10.1145/2668064.2668104>
- [19] Robert Geiger and Jürgen Scherr. 2009. Surgery assistance system for guiding a surgical instrument. U.S. Patent 8433389B2. (20 March 2009). <https://patents.google.com/patent/US8433389B2/en> Filed March 20, 2009.
- [20] Andrew A. Gumbs, Fernando Crovari, Clement Vidal, Patrick Henri, and Brice Gayet. 2007. Modified Robotic Lightweight Endoscope (ViKY) Validation In Vivo in a Porcine Model. *Surgical Innovation* 14, 4 (Dec. 2007), 261–264. DOI: <http://dx.doi.org/10.1177/1553350607310281>
- [21] G. S. Guthart and J. K. Salisbury. 2000. The Intuitive(TM) telesurgery system: overview and application. In *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No.00CH37065)*, Vol. 1. IEEE, San Francisco, CA, USA, 618–621 vol.1. DOI: <http://dx.doi.org/10.1109/ROBOT.2000.844121>
- [22] Chris Harrison and Anind K. Dey. 2008. Lean and Zoom: Proximity-aware User Interface and Content Magnification. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*. ACM, New York, NY, USA, 507–510. DOI: <http://dx.doi.org/10.1145/1357054.1357135>
- [23] Surgical Intuitive. 2016. Intuitive Surgical Announces Preliminary Fourth Quarter and Full Year 2015 Results. (Jan. 2016). <https://isrg.intuitive.com/news-releases/news-release-details/intuitive-surgical-announces-preliminary-fourth-quarter-and-1>

- [24] Surgical Intuitive. 2019. Intuitive Surgical Announces Preliminary Fourth Quarter and Full Year 2018 Results. (Jan. 2019). <https://isrg.intuitive.com/news-releases/news-release-details/intuitive-surgical-announces-preliminary-fourth-quarter-and-4>
- [25] Thibaut Jacob, Gilles Bailly, Eric Lecolinet, Géry Casiez, and Marc Teyssier. 2016. Desktop Orbital Camera Motions Using Rotational Head Movements. In *In proceedings of SUP'16, the ACM Symposium on Spatial User Interaction*. Tokyo, Japan, 139–148. DOI : <http://dx.doi.org/10.1145/2983310.2985758>
- [26] Thomas Jaschinski, Christoph G. Mosch, Michaela Eikermann, Edmund Am Neugebauer, and Stefan Sauerland. 2018. Laparoscopic versus open surgery for suspected appendicitis. *The Cochrane Database of Systematic Reviews* 11 (2018), CD001546. DOI : <http://dx.doi.org/10.1002/14651858.CD001546.pub4>
- [27] Sing Bing Kang. 1999. Hands-free interface to a virtual reality environment using head tracking. U.S. Patent 6009210A. (Dec. 1999). <https://patents.google.com/patent/US6009210A/en> Filled March 05, 1997.
- [28] Brady W. King, Luke A. Reisner, Abhilash K. Pandya, Anthony M. Composto, R. Darin Ellis, and Michael D. Klein. 2013. Towards an Autonomous Robot for Camera Control During Laparoscopic Surgery. *Journal of Laparoendoscopic & Advanced Surgical Techniques* 23, 12 (Nov. 2013), 1027–1030. DOI : <http://dx.doi.org/10.1089/lap.2013.0304>
- [29] Julia Kristin, Robert Geiger, Peter Kraus, and Thomas Klenzner. 2015. Assessment of the endoscopic range of motion for head and neck surgery using the SOLOASSIST endoscope holder. *The International Journal of Medical Robotics and Computer Assisted Surgery* 11, 4 (2015), 418–423. DOI : <http://dx.doi.org/10.1002/rcs.1643>
- [30] Cheolwhan Lee, Yuan-Fang Wang, D.R. Uecker, and Yulun Wang. 1994. Image analysis for automated tracking in robot-assisted endoscopic surgery. In *Proceedings of 12th International Conference on Pattern Recognition*, Vol. 1. IEEE, New York, NY, USA, 88–92 vol.1. DOI : <http://dx.doi.org/10.1109/ICPR.1994.576232>
- [31] L. MacKenzie, J. A. Ibbotson, C. G. L. Cao, and A. J. Lomax. 2001. Hierarchical decomposition of laparoscopic surgery: a human factors approach to investigating the operating room environment. *Minimally Invasive Therapy & Allied Technologies* 10, 3 (Jan. 2001), 121–127. DOI : <http://dx.doi.org/10.1080/136457001753192222>
- [32] Helena M. Mentis, Kenton O'Hara, Gerardo Gonzalez, Abigail Sellen, Robert Corish, Antonio Criminisi, Rikin Trivedi, and Pierre Theodore. 2015. Voice or Gesture in the Operating Room. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '15)*. ACM, New York, NY, USA, 773–780. DOI : <http://dx.doi.org/10.1145/2702613.2702963>
- [33] 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* 13, 10 (Oct. 1998), 2748–2750. DOI : <http://dx.doi.org/10.1093/humrep/13.10.2748>
- [34] 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, 1 (Feb. 2011), 105–114. DOI : <http://dx.doi.org/10.1016/j.scient.2011.03.012>
- [35] Katie Moon and Deborah Blackman. 2014. A Guide to Understanding Social Science Research for Natural Scientists. *Conservation Biology* 28, 5 (2014), 1167–1177. DOI : <http://dx.doi.org/10.1111/cobi.12326>
- [36] Pedro Ballester Nebot, Yatin Jain, Kevin Haylett, Robert Stone, and Rory McCloy. 2003. Comparison of Task Performance of the Camera-Holder Robots EndoAssist and Aesop. *Surgical Laparoscopy Endoscopy & Percutaneous Techniques* 13, 5 (Oct. 2003), 334. [https://journals.lww.com/surgical-laparoscopy/Abstract/2003/10000/Comparison\\_of\\_Task\\_Performance\\_of\\_the.10.aspx](https://journals.lww.com/surgical-laparoscopy/Abstract/2003/10000/Comparison_of_Task_Performance_of_the.10.aspx)
- [37] E. Neugebauer, H. Troidl, C. K. Kum, E. Eypasch, M. Miserez, and A. Paul. 1995. The E.A.E.S. Consensus development conferences on laparoscopic cholecystectomy, appendectomy, and hernia repair. *Surgical Endoscopy* 9, 5 (May 1995), 550–563. DOI : <http://dx.doi.org/10.1007/BF00206852>
- [38] A. Nishikawa, T. Hosoi, K. Koara, D. Negoro, A. Hikita, S. Asano, H. Kakutani, F. Miyazaki, M. Sekimoto, M. Yasui, Y. Miyake, S. Takiguchi, and M. Monden. 2003. FAcE MOUSe: A novel human-machine interface for controlling the position of a laparoscope. *IEEE Transactions on Robotics and Automation* 19, 5 (Oct. 2003), 825–841. DOI : <http://dx.doi.org/10.1109/TRA.2003.817093>
- [39] Atsushi Nishikawa, Hiroaki Nakagoe, Kazuhiro Taniguchi, Yasuo Yamada, Mitsugu Sekimoto, Shuji Takiguchi, Morito Monden, and Fumio Miyazaki. 2008. How Does the Camera Assistant Decide the Zooming Ratio of Laparoscopic Images? Analysis and Implementation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008 (MICCAI '08)*, Dimitris Metaxas, Leon Axel, Gabor Fichtinger, and Gábor Székely (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 611–618. DOI : [http://dx.doi.org/10.1007/978-3-540-85990-1\\_73](http://dx.doi.org/10.1007/978-3-540-85990-1_73)
- [40] Ilana Nisky, Felix Huang, Amit Milstein, Carla M. Pugh, Ferdinando A. Mussa-ivaldi, and Amir Karniel. 2012. Perception of Stiffness in Laparoscopy – the Fulcrum Effect. *Studies in health technology and informatics* 173 (2012), 313–319. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4102265/>

- [41] Kenton O'Hara, Gerardo Gonzalez, Abigail Sellen, Graeme Penney, Andreas Varnavas, Helena Mentis, Antonio Criminisi, Robert Corish, Mark Rouncefield, Neville Dastur, and Tom Carrell. 2014. Touchless Interaction in Surgery. *Commun. ACM* 57, 1 (Jan. 2014), 70–77. DOI : <http://dx.doi.org/10.1145/2541883.2541899>
- [42] T. Osa, C. Staub, and A. Knoll. 2010. Framework of automatic robot surgery system using Visual servoing. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, New York, NY, USA, 1837–1842. DOI : <http://dx.doi.org/10.1109/IRoS.2010.5650301>
- [43] Hannah R. M. Pelikan, Amy Cheatle, Malte F. Jung, and Steven J. Jackson. 2018. Operating at a Distance - How a Teleoperated Surgical Robot Reconfigures Teamwork in the Operating Room. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW (Nov. 2018), 138:1–138:28. DOI : <http://dx.doi.org/10.1145/3274407>
- [44] C. Perin, P. Dragicevic, and J. Fekete. 2014. Revisiting Bertin Matrices: New Interactions for Crafting Tabular Visualizations. *IEEE Transactions on Visualization and Computer Graphics* 20, 12 (Dec. 2014), 2082–2091. DOI : <http://dx.doi.org/10.1109/TVCG.2014.2346279>
- [45] I. Rivas-Blanco, B. Estebanez, M. Cuevas-Rodriguez, E. Bauzano, and V. F. Muñoz. 2014. Towards a cognitive camera robotic assistant. In *5th IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics*. IEEE, New York, NY, USA, 739–744. DOI : <http://dx.doi.org/10.1109/BIOROB.2014.6913866>
- [46] J. M. Sackier and Y. Wang. 1994. Robotically assisted laparoscopic surgery. *Surgical Endoscopy* 8, 1 (Jan. 1994), 63–66. DOI : <http://dx.doi.org/10.1007/BF02909496>
- [47] Daniel J. Scott, Patricia C. Bergen, Robert V. Rege, Royce Laycock, Seifu T. Tesfay, R. James Valentine, David M. Euhus, D. Rohan Jeyarajah, William M. Thompson, and Daniel B. Jones. 2000. Laparoscopic training on bench models: better and more cost effective than operating room experience?1. *Journal of the American College of Surgeons* 191, 3 (Sept. 2000), 272–283. DOI : [http://dx.doi.org/10.1016/S1072-7515\(00\)00339-2](http://dx.doi.org/10.1016/S1072-7515(00)00339-2)
- [48] Daniel J. Scott, William N. Young, Seifu T. Tesfay, William H. Frawley, Robert V. Rege, and Daniel B. Jones. 2001. Laparoscopic skills training. *The American Journal of Surgery* 182, 2 (Aug. 2001), 137–142. DOI : [http://dx.doi.org/10.1016/S0002-9610\(01\)00669-9](http://dx.doi.org/10.1016/S0002-9610(01)00669-9)
- [49] Anna Serur, Rebecca Rhee, and Matthew M. Philp. 2015. The Role of a Bedside Assistant. In *Robotic Approaches to Colorectal Surgery*, Howard Ross, Sang Lee, Bradley J. Champagne, Alessio Pigazzi, and David E. Rivadeneira (Eds.). Springer International Publishing, Cham, 111–117. DOI : [http://dx.doi.org/10.1007/978-3-319-09120-4\\_9](http://dx.doi.org/10.1007/978-3-319-09120-4_9)
- [50] J. Bryan Sexton, Martin A. Makary, Anthony R. Tersigni, David Pryor, Ann Hendrich, Eric J. Thomas, Christine G. Holzmueller, Andrew P. Knight, Yun Wu, and Peter J. Pronovost. 2006. Teamwork in the operating room: frontline perspectives among hospitals and operating room personnel. *Anesthesiology* 105, 5 (Nov. 2006), 877–884.
- [51] Torben Sko and Henry J. Gardner. 2009. Head Tracking in First-Person Games: Interaction Using a Web-Camera. In *Proceedings of the 12th IFIP TC 13 International Conference on Human-Computer Interaction: Part I (INTERACT '09)*. Springer-Verlag, Berlin, Heidelberg, 342–355. DOI : [http://dx.doi.org/10.1007/978-3-642-03655-2\\_38](http://dx.doi.org/10.1007/978-3-642-03655-2_38)
- [52] Karem Slim and Michel Canis. 2017. La chirurgie robotique. Peut-on (doit-on) nager à contre-courant ? *La Presse Médicale* 46, 6 (June 2017), 557 – 560. DOI : <http://dx.doi.org/10.1016/j.lpm.2017.04.004>
- [53] Jens-Uwe Stolzenburg, Toni Franz, Panagiotis Kallidonis, Do Minh, Anja Dietel, James Hicks, Martin Nicolaus, Abdulrahman Al-Aown, and Evangelos Liatsikos. 2011. Comparison of the FreeHand® robotic camera holder with human assistants during endoscopic extraperitoneal radical prostatectomy. *BJU International* 107, 6 (2011), 970–974. DOI : <http://dx.doi.org/10.1111/j.1464-410X.2010.09656.x>
- [54] R. H. Taylor, J. Funda, B. Eldridge, S. Gomory, K. Gruben, D. LaRose, M. Talamini, L. Kavoussi, and J. Anderson. 1995. A telerobotic assistant for laparoscopic surgery. *IEEE Engineering in Medicine and Biology Magazine* 14, 3 (May 1995), 279–288. DOI : <http://dx.doi.org/10.1109/51.391776>
- [55] Charlotte Tsui, Rachel Klein, and Matthew Garabrant. 2013. Minimally invasive surgery: national trends in adoption and future directions for hospital strategy. *Surgical Endoscopy* 27, 7 (July 2013), 2253–2257. DOI : <http://dx.doi.org/10.1007/s00464-013-2973-9>
- [56] F. A. Voorhorst, D. W. Meijer, and C. J. Overbeeke. 1999. Head-controlled laparoscopy: experiment, prototype, and preliminary results. *Journal of Laparoendoscopic & Advanced Surgical Techniques. Part A* 9, 5 (Oct. 1999), 379–388. DOI : <http://dx.doi.org/10.1089/lap.1999.9.379>
- [57] Colin Ware, Kevin Arthur, and Kellogg S. Booth. 1993. Fish Tank Virtual Reality. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems (CHI '93)*. ACM, New York, NY, USA, 37–42. DOI : <http://dx.doi.org/10.1145/169059.169066>
- [58] Guo-Qing Wei, K. Arbter, and G. Hirzinger. 1997. Real-time visual servoing for laparoscopic surgery. Controlling robot motion with color image segmentation. *IEEE Engineering in Medicine and Biology Magazine* 16, 1 (Jan. 1997), 40–45. DOI : <http://dx.doi.org/10.1109/51.566151>

- [59] Jacob O. Wobbrock, Brad A. Myers, AungHtet Htet, and Edmund F. LoPresti. 2003. Text entry from power wheelchairs: edgewrite for joysticks and touchpads. *ACM SIGACCESS Accessibility and Computing* (Sept. 2003). <https://dl.acm.org/doi/abs/10.1145/1029014.1028650>
- [60] Kota Yamaguchi, Takashi Komuro, and Masatoshi Ishikawa. 2009. Ptz Control with Head Tracking for Video Chat. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems (CHI EA '09)*. ACM, New York, NY, USA, 3919–3924. DOI: <http://dx.doi.org/10.1145/1520340.1520594>
- [61] Piaopiao Ye, Na Zhao, Jing Shu, Heping Shen, Yanpeng Wang, Lifeng Chen, and Xiaojian Yan. 2019. Laparoscopy versus open surgery for adnexal masses in pregnancy: a meta-analytic review. *Archives of Gynecology and Obstetrics* 299, 3 (March 2019), 625–634. DOI: <http://dx.doi.org/10.1007/s00404-018-05039-y>
- [62] Xiaoli Zhang and Shahram Payandeh. 2002. Application of visual tracking for robot-assisted laparoscopic surgery. *Journal of Robotic Systems* 19, 7 (2002), 315–328. DOI: <http://dx.doi.org/10.1002/rob.10043>