UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

MATHEUS DIAS NEGRÃO

Characterizing Affordances using AR for Laparoscopic Surgeries and Challenges of Planning a Mixed Reality Remote Collaboration Application

Thesis presented in partial fulfillment of the requirements for the degree of Master of Computer Science

Advisor:
Anderson Maciel

Porto Alegre
January 2024

CIP — CATALOGING-IN-PUBLICATION

Dias Negrão, Matheus

Characterizing Affordances using AR for Laparoscopic Surgeries and Challenges of Planning a Mixed Reality Remote Collaboration Application / Matheus Dias Negrão. – Porto Alegre: PPGC da UFRGS, 2024.

94 f.: il.

Thesis (Master) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação, Porto Alegre, BR–RS, 2024. Advisor:

Anderson Maciel.

1. Augmented Reality. 2. Laparoscopy. 3. Surgical Training. 4. Remote Collaboration. I. Anderson Maciel,

. II. Título.

"Eu insisto, persisto, não mando recado...
Eu tenho algo a dizer, não vou ficar calado."

— Sabotage

ACKNOWLEDGEMENTS

Em primeiro lugar, eu não poderia deixar de agradecer a minha mãe, por todo suporte e crédito depositado em mim ao longo da minha vida inteira, em especial minha vida acadêmica, uma realidade de conquistas distantes, mas que graças a ela eu pude alcançar e continuar lutando para chegar ainda mais longe. Eu agradeço também ao meu pai, que sempre acreditou que a educação poderia mudar a minha vida e me apoiou em continuar estudando. Agradeço especialmente a minha noiva Zaira, por todo amor e companheirismo durante os períodos mais difíceis, felizes e também durante as minhas maiores conquistas.

Eu gostaria de agradecer ao meu orientador, Anderson Maciel, por todo apoio, dedicação e confiança durante essa jornada, sempre presente compartilhando seus conhecimentos. As professoras Carla Freitas e Luciana Nedel, pelo acolhimento no VISLab, os valiosos conselhos e todo conhecimento compartilhado. Agradeço aos amigos que fiz no Lab, pelo incentivo, dicas e ajudas no dia-a-dia, além das melhores conversas e almoços no campus. Especialmente ao Elizer com o seu vasto conhecimento em computação gráfica e ao Jorge com seus conselhos acadêmicos.

ABSTRACT

The advantages of Minimally Invasive Abdominal Surgery (MIAS) over traditional open surgery include shorter recovery time, less blood loss, and less postoperative pain and complications. However, laparoscopic surgeries are more complex to perform and learn, with various difficulties that surgeons face during learning and procedures. AR and VR devices in surgical settings have become increasingly suitable, offering hands-free, non-intrusive, and portable solutions to help guide and identify the points of interest during the operation, complementing verbal explanations and reducing the complexity of learning.

This dissertation presents a comprehensive overview of the literature and related work exploring remote collaboration, augmented reality (AR), virtual reality (VR), and mixed reality (MR) for health environments, specifically in laparoscopic surgeries. We describe the primary perspective for planning and developing the XR application, its modes of utilization, and the challenges faced and turned into research questions, focusing on the AR interface and its standalone mode of utilization in laparoscopic surgeries intraoperatively. Further, we present the final developed AR interface, the hardware, and interaction designs.

Finally, it details a user experiment to evaluate hand and head-gaze affordances using AR for annotation during laparoscopy surgeries and similar settings. The main study aims to understand the impact and improvements that head-gaze average filtering and the scale method can bring to perform annotations in the laparoscopy video feed in a virtual monitor positioned straight in front of the user. The user experiment was performed in a between-subject protocol with 32 volunteers from the Institute of Informatics. The study found that the users are confident about their performance and demonstrated low physical and temporal demand, and the head-gaze interaction with HL1 is very distributed. The conclusion presents an overview of the user study's main findings, data analysis, and the interface's limitations. A brief explanation of the future experiment is also proposed to evaluate surgeons' interface use while performing physical tasks.
Keywords: Augmented Reality. Laparoscopy. Surgical Training. Remote Collaboration.

Caracterizando recursos usando AR para cirurgias laparoscópicas e desafios de planejar um aplicativo de colaboração remota de realidade mista

RESUMO

As vantagens da Cirurgia Abdominal Minimamente Invasiva (MIAS) em relação à cirurgia aberta tradicional incluem menor tempo de recuperação, menor perda de sangue e menos dor e complicações pós-operatórias. Porém, as cirurgias laparoscópicas são mais complexas de realizar e aprender, com diversas dificuldades que os cirurgiões enfrentam durante o aprendizado e os procedimentos. Os dispositivos de AR e VR em ambientes cirúrgicos têm se tornado cada vez mais comums, oferecendo soluções sem o uso ads mãos, não intrusivas e portáteis para ajudar a orientar e identificar os pontos de interesse durante a operação, complementando as explicações verbais e reduzindo a complexidade do aprendizado.

Esta dissertação apresenta uma visão abrangente da literatura e trabalhos relacionados explorando colaboração remota, realidade aumentada (AR), realidade virtual (VR) e realidade mista (MR) para ambientes de saúde, especificamente em cirurgias laparoscópicas. Descrevemos a perspectiva primária para o planejamento e desenvolvimento da aplicação XR, seus modos de utilização e os desafios enfrentados e transformados em questões de pesquisa, com foco na interface AR e seu modo autônomo de utilização em cirurgias laparoscópicas intraoperatórias. Além disso, apresentamos a interface AR final desenvolvida, o hardware e os designs de interação.

Por fim, é detalhado um experimento de usuário para avaliar as possibilidades de interação em AR usando a direção da cabeça e um anel an mão para anotação durante cirurgias de laparoscopia e configurações semelhantes. O estudo principal tem como objetivo compreender o impacto e as melhorias que o método de filtragem por média do mapeamento da direção da cabeça, e o método de escala do mapeamento podem trazer para realizar anotações no feed de vídeo de laparoscopia em um monitor virtual posicionado diretamente na frente do usuário. O experimento do usuário foi realizado em um protocolo entre sujeitos com 32 voluntários do Instituto de Informática. O estudo constatou que os usuários foram confiantes em seu desempenho e demonstraram baixa demanda física e temporal, e a interação com a direção da cabeça com HL1 é muito distribuída, embora apresente uma diferença significante entre os eixos dos posiconamentos. A conclusão apresenta uma visão geral das principais conclusões do estudo do usuário, análise de dados e limitações da interface. Uma breve explicação do experimento a seguir também é proposta para avaliar o uso da interface dos cirurgiões durante a execução de tarefas físicas.

Palavras-chave: Realidade Aumentada. Laparoscopia. Treinamento Cirurgico. Colaboração Remota.

LIST OF ABBREVIATIONS AND ACRONYMS

AR	Agumented Reality
CT	Computed Tomography
FOV	Field of View
HCI	Human-Centered Interaction
HCPA	Clinical Hospital of Porto Alegre
HL1	HoloLens Gen-1
HMD	Head-Mounted Display
MIAS	Minimally Invasive Abdominal Surgery
MR	Mixed Reality
MRI	Magnetic Resonance Imaging
NASA	National Aeronautics and Space Administration
OR	Operation Room
TLX	Task Load Index
UI	User Interface
VR	Virtual Reality
XR	Extended Reality

LIST OF FIGURES

Figure 1.1 Laparoscopic instruments: Right: Laparoscope camera. Left: Different types of surgery instruments. 19

Figure 2.1 Representation of the Reality-Virtuality Continuum. 26
Figure 2.2 Simulation of a pediatrician appointment. The doctor is conducting the information collection and anamnesis (left). The physician performs cardiac auscultation on the patient (center) as part of the physical exam. As one of the rules of biosafety, she is washing her hands in the sink (right). 28
Figure 2.3 Left: 1st STAR Version: remote collaboration using tablet for both remote and local surgeons. Center and Right: 2nd STAR Version using an AR HMD (HoloLens Gen-1). 29
Figure 2.4 An example of a simulated view of what the operator wearing the HoloLens sees. 32

Figure 3.1 A sketch for initial augmented reality interface. 35
Figure 3.2 A sketch representing the complete system. Left: VR interface for the remote surgeon. Right: AR interface for a local surgeon. 36
Figure 3.3 First Version With Two Video Panels 39
Figure 3.4 Head-Gaze representation and the mapped reticle for interaction. 41

Figure 4.1 The ring device with two buttons. 46
Figure 4.2 Initial 3D printed handcrafted foot pedals. 47
Figure 4.3 Commercially manufactured wireless foot pedals (we use only two of three inputs). 47
Figure 4.4 The interface panels through the HL1 view. 48
Figure 4.5 A macro view of the panels in the interface. 48
Figure 4.6 The arrow annotation and its manipulation points. The line manipulation is equal to the arrow. 49
Figure 4.7 The circle annotation and the manipulation dots for positioning and resizing. 50
Figure 4.8 Left: The video panel with annotations. Right: Floating Marker Selector Panel. 53
Figure 4.9 Left: the grid-view in the image panel. Right: Full view of a picture in the image panel. 53

Figure 5.1 Laparoscopic Physical Tasks. Left: Thread Passing. Right: Peg Transfer 55
Figure 5.2 Problem examples for the two tasks. Left: Peg transfer task, where the red circles are positioned above the rings, except the green one, and the yellow is a reference (In this picture angle, no one ring color is visible). Right: Passing thread task, the yellow arrow is a reference, and the arrows connect the points that the thread passes through. 57
Figure 5.3 Dots and Circles Detection 60
Figure 5.4 Nasa Task Load Index for each Task 64
Figure 5.5 Time Groups. Left: Peg Transfer. Right: Thread Passing. 65
Figure 5.6 Head-Gaze Methods Peg Transfer Time Variation. 66
Figure 5.7 Head-Gaze Methods Thread Passing Time Variation. 66

Figure 5.8 Circle center and arrow end positioning errors per axis. Left: circle positions are more scattered vertically. Right: arrow end positions are distributed around 2 radii from the center with a larger vertical spread than horizontal. 67
Figure 5.9 Circle center and arrow end positioning errors. 68
Figure 5.10 Center Positioning Error Filters. 69
Figure 5.11 Center position and circle size errors per head-gaze filter. 70
Figure 5.12 Dot Positioning Error Filters. 70
Figure 5.13 Diameter error TL1 comparison 71
Figure 5.14 Task Levels Performance Comparison 72
Figure 5.15 Center positioning error TL3 comparison 73
Figure 5.16 Dot positioning error TL3 and TL2 comparison 73
Figure 5.17 Left: Peg Transfer Trials. Right: Thread Passing trials. 75

LIST OF TABLES

Table 2.1 AR Interaction methods for surgery in related works. 34
Table 4.1 Reticle stabilization modes 52
Table 5.1 Task Level and Trials Sequence Example. 58

CONTENTS

1 Introduction **19**

1.1 Approach and Goals ... 21

1.2 Organization Overview ... 22

2 Related Work **25**

2.1 Remote Collaboration with XR technologies ... 25

2.2 VR and AR in Healthcare ... 27

2.3 Surgery XR Applications ... 29

2.3.1 Laparoscopy ... 31

2.3.1.1 Visualization and Interaction ... 31

3 Challenges in the Design of an Asymmetric XR Interface for Remote Collaboration in Laparoscopy **35**

3.1 Overview ... 35

3.2 Planned Modes of Utilization ... 37

3.3 Initial Planned Application ... 37

3.3.1 AR Interface Layout ... 38

3.3.2 AR Interaction Design ... 40

3.3.3 Planned VR Remote Interface ... 41

3.4 Challenges ... 42

4 Final AR Interface for Laparoscopy Surgeries **45**

4.1 Hardware Design ... 45

4.2 Interface Layout ... 46

4.3 Interaction design ... 49

4.3.1 Head-Gaze Pointing ... 51

4.3.2 Selection and manipulation ... 52

4.3.3 Research Questions ... 54

5 USER EXPERIMENT: AR Affordances for Laparoscopy **55**

5.1 Overview ... 55

5.1.1 Design ... 56

5.1.2 Protocol ... 57

5.1.3 Data collection ... 59
5.1.3.1 Accuracy from images ... 59
5.1.3.2 Duration times ... 60
5.1.3.3 Subjective self-evaluation ... 62
5.2 Results ... 62
5.2.1 Outliers ... 63
5.2.2 Task Load ... 63
5.2.3 Learning effect ... 64
5.2.4 Time-performance per head-gaze condition ... 65
5.2.5 General annotation precision ... 67
5.2.6 Annotation precision per head-gaze condition ... 69
5.2.7 Annotation accuracy per task Level ... 71
5.2.8 Qualitative Data ... 73
5.3 Discussion ... 74
5.3.1 Time and learning ... 74
5.3.2 Interaction performance ... 75
5.3.3 Head-gaze performance ... 76
5.3.4 User impressions ... 77
5.3.5 Limitations ... 77
5.4 Future Work ... 78

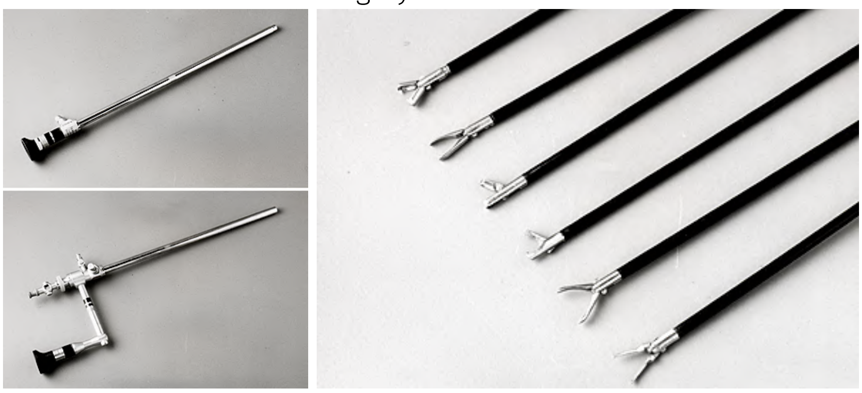
6 General Conclusions **79**
6.1 Remarks ... 81
6.2 Limitations and Future Work ... 81

APPENDIX A — Resumo expandido **91**
A.1 Desafios no planejamento de uma aplicação em realidade mista ... 91
A.2 Interação em Interfaces AR para Laparoscopia ... 91
A.3 Experimento com usuários ... 92
A.4 Resultados ... 92
A.5 Conclusão ... 93

1 INTRODUCTION

Minimally Invasive Abdominal Surgery (MIAS) can provide advantages over traditional open surgery, such as shorter recovery time, less blood loss, and less postoperative pain and complications (Squirrell et al., 1998). This is due to the tiny incisions where the surgeons can access the abdomen and pelvis with the aid of a camera (laparoscope, fig. 1.1-left) and perform the surgery using special sterile instruments, such as graspers, scissor, dissector (fig. 1.1-right). Starting with cholecystectomy surgeries, laparoscopic approaches were introduced in the 1980s and quickly showed advantages and opened for expansion to other types of surgery, such as colorectal surgery, appendectomy, liver surgery, and a variety of different procedures (Bittner, 2006). Due to these advantages, MIAS has still seen exponential growth and worldwide adoption.

Figure 1.1 – Laparoscopic instruments: Right: Laparoscope camera. Left: Different types of surgery instruments.



Source: Gould et al. (2019)

However, laparoscopic surgeries are more complex procedures to perform and to learn (Mentis, Chellali and Schwaitzberg, 2014a; Rodrigues et al., 2012). They impose constraints for surgeons, such as the manipulation space, the angle of view, and a decreased touch sensation. Even experienced surgeons face difficulties, such as reduced tactile feedback, deviations, and rotations away from the target horizon of laparoscopic cameras, that are often held by less-experienced surgeons, creating delays and loss of concentration.

During training and procedures, surgeons must comprehend where and what to look at, how to manipulate tissue, where to cauterize, and the directions of the instruments with the strength to be applied, despite operations field limitations (Mentis, Chellali and Schwaitzberg, 2014b). The learning curve is more prolonged and requires different types of training using bench models to simulate specific actions and develop technical skills and precision in using tools (Scott et al., 2000). The novice surgeon's laparoscopy skills can be improved through virtual reality (VR) simulators, which provide visual feedback on the deformation of tissue and organs (Basdogan, Ho and Srinivasan, 2001).

Additionally, in the operating room (OR), the surgical team is not looking at the intervention target as with open surgery, but instead, at monitors located 2 to 3 meters away, positioned at different angles/distances for each team member, requires an extra conscious effort to correct body posture (Supe, Kulkarni and Supe, 2010). Furthermore, this setup limits access to preoperative images and planning data during surgery. When these data need to be consulted, the surgeon cannot quickly leave the clean operation area and thus requests help from assistants who hold and manipulate these materials following verbal instructions, which is far from ideal. Finally, and more importantly, communication between surgical team members can be difficult and awkward. Since team members hold instruments and execute complex tasks, pointing to "areas of interest" (Sevdalis et al., 2012) and disambiguating pointing references without losing time and concentration is difficult. It also can introduce pauses and hinder team communication and synchronization, increasing the potential for lost time and mistakes.

An operating room is also a complex environment, and limited space around the table makes it problematic to enlist specialists' help or instruct students, who could be located several feet away on the non-sterile field or farther away in different hospitals/countries (Sebajang et al., 2006). Generally, novice surgeons receive instructions through verbal explanations and deictic words from experienced surgeons, such as the tasks to be performed using the instruments and where to accomplish them. The pointing locations, gestures, and movements can be challenging to understand (Mentis, Chellali and Schwaitzberg, 2014a), considering there is no absolute referential and direct view of the procedure but a visualization through the camera and the 2D monitor. Moreover, looking at a video monitor and other external information, e.g., vitals, imposes postural strains and focus dispersion.

In parallel, AR and VR devices have become accessible, lightweight, and powerful while improving in quality and resolution, making them increasingly suitable for portable, high-definition, hands-free, and non-intrusive usage in surgical settings. AR applications have been developed and evaluated, demonstrating effectiveness in laparoscopic skills training (Barsom, Graafland and Schijven, 2016). Some applications use AR head-mounted displays (HMD) to place objects in the real world and the surgical video, helping explain the training and execution of procedures. These systems can guide and identify the points of interest during the operation through shared appointments (Heinrich et al., 2021), complementing verbal explanations and helping to reduce the complexity of the learning. In other cases, the application combines the views of surgery, vitals, and preoperative exams in the field of view to minimize the interruptions and help with focus during the procedure (Janabi et al., 2020).

VR applications typically focus on the sense of presence in the operation room and the role-playing scenario using 3D immersive environments with virtual representations. Most training applications are aimed at technical skills in surgery, where the user can manipulate and interact with 3D models of anatomical parts and tissues. In the simulation of laparoscopy surgeries, the realism of both the tissue simulation and the physical interactions with instruments is crucial. Existing systems use special devices to provide haptic feedback and focus on the visual fidelity of the simulated behaviors. These technologies have potential in the health professional's education, improving postintervention knowledge and skills (Kyaw et al., 2019). Virtual reality HMDs can also be helpful to guide and train in remote collaboration applications for surgeries (Gasques et al., 2021), which can connect surgeons and improve communication. In this dissertation, we explore techniques to improve communication using AR, and the challenges of connecting and developing an MR application for remote collaboration for laparoscopy surgeries.

1.1 Approach and Goals

This work presents the design space for interaction in intraoperative AR for laparoscopy. We do so throughout the development and evaluation of a demonstrator that can improve the ergonomic conditions (e.g., using AR goggles as screens) in the operating room. Besides displaying the laparoscopic video on a virtual screen (approximately 24 inches, considering a positioning distance of

50 centimeters), wherever the surgeon prefers to place it, the system provides access to preoperative images and planning data. The demonstrator also allows for real-time deictic communication through annotations, icons, drawings, and other visual indications that may help with collaborative tasks (Sarmiento et al., 2013; Sarmiento et al., 2014). The surgeons can make such annotations and direct trainees and assistants to look at the critical spots. We describe the research, design, and test of the AR interface, but also present the challenges of a future VR interface for remote participants that will connect with the local AR interface user, integrating both interfaces in an XR application. The main goal of this work, based on a review and analysis of available literature on AR and VR for laparoscopy, is to develop and evaluate an AR Interface for laparoscopic surgeries and its interaction methods. We aim to improve the communication between the surgical team and the expert to novice surgeons using the AR Interface intraoperatively. The presented study focuses on the head-gaze developed methods for stabilization combined with a hand device for navigation, interaction, and drawing in panels of an AR Interface. In long-term planning, the AR Interface should be able to be linked with a VR Interface for remote collaboration in laparoscopic surgeries.

1.2 Organization Overview

We start with a search for related work, understanding the principles of AR, VR, and XR, and passing through definitions and uses of remote collaboration using these technologies. We searched for simulators that use virtual and augmented reality in medicine, their techniques, and applications. We also present the current state-of-art related to the use of AR in laparoscopy surgeries and training procedures, besides the state-of-art of remote collaboration in general surgery.

After the review of related work, we present in Chapter 3 an overview of the planned interactions and utilization scenarios of the complete XR application. We describe a developed prototype of the AR interface, which we used to collect information in a preliminary demonstration and conversation with stakeholders, a laparoscopic professor and surgeon, and other physicians. The chapter also introduces our planned VR remote interface with the challenges of developing the XR application's connection, visualization, and interaction. Based on this first phase of the work, we finish the chapter presenting the motivation and our focus on the constraints of the AR Interface for intraoperative utilization.

In Chapter 4, we present the final design of our AR interface with the developed and applied interaction techniques. We describe the hardware to simulate the training environment and the devices used for interaction, including a novel proposed hand device: A ring with two buttons, which the user wears under the sterile gloves and uses the buttons to interact with the interface. Closing the chapter, we present the research questions we want to address with the proposed interface and its interaction methods.

Chapter 5 presents a study that we conducted with voluntary participants (n=32) in simulated annotation tasks to understand the impact of head-gaze and hand-based annotation techniques in augmented reality. The chapter describes the user experiments, protocol, results, and discussion based on several analyses performed with the collected data. Finally, in Chapter 6, we present our general conclusions, prominent findings, and the possibilities for future work.

2 RELATED WORK

In this chapter, we will describe the consolidated concepts in the literature and present the state of the art in XR technologies for the medical field and remote collaboration. We first give context to the technologies and HCI concepts adopted in the work, such as mixed reality and remote collaboration. Then, we explain the utilization of simulators and MR applications in the medical and surgery context. Finally, we concentrate on laparoscopic surgeries, exploring the variety of simulators and previous work, focusing on the recent results using AR, their interface, interactions, and limitations.

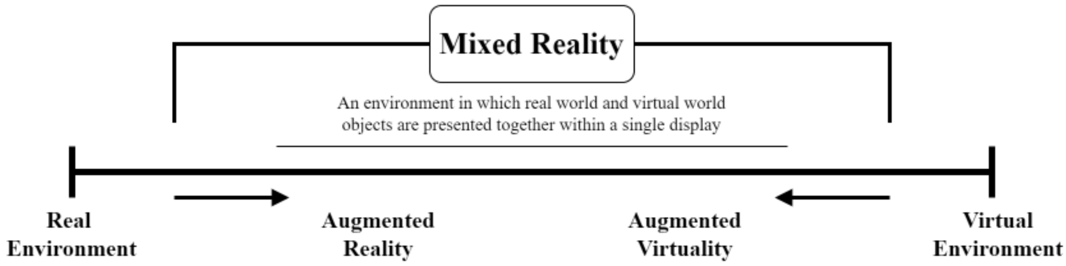
All AR, VR, and MR approaches have become constructive grounds for cutting-edge research across various domains (Saad, Bennis and Chen, 2019). The entertainment industry constantly pushes the boundaries of storytelling and user engagement through VR and AR experiences. Engineers leverage these technologies for design and prototyping. These immersive technologies are harnessed in healthcare for surgical simulations, telemedicine consultations, and even mental health therapy. These technologies also provide a revolutionary way of learning in education (e.g., using gamification (Lampropoulos et al., 2022)), where students can delve into complex concepts through interactive augmented and virtual environments.

Although there is no unanimity about the definition of mixed reality in the literature, we assume the Reality-Virtuality (RV) Continuum model (Milgram and Kishino, 1994; Milgram et al., 1995), and will refer to mixed reality (MR) as the combination obtained from overlaying digital objects on a view of the real world, independently of their occlusion relations (fig 2.1). We also assumed Extended Reality (XR) to be an umbrella term that includes all VR, AR, and MR technologies.

2.1 Remote Collaboration with XR technologies

In the present landscape, remote collaboration using augmented reality (AR), virtual reality (VR), and mixed reality (MR) has emerged as assertive tools for handling the evolving needs of the modern interconnected world. In conditions where physical presence and co-located work are not always possible or suitable, these immersive technologies bridge geographical gaps, offering users a realistic

Figure 2.1 – Representation of the Reality-Virtuality Continuum.



Source: Adapted from Milgram and Kishino (1994)

sense of sight separated by extensive distances. Previous works using AR and VR to improve remote collaboration have been implemented for several contexts, differing in the paradigm (AR or VR), setup settings, interaction methods, which information is shared, and other particularities of each domain.

The applications also can differ in the MR space symmetry, which refers to the interaction paradigm within the Reality-Virtuality Continuum (Milgram and Kishino, 1994; Milgram et al., 1995). The space symmetry definition is related to the paradigm where the remote and local users are located, i.e., in a symmetric MR space, the users are in the same paradigm, such as VR. While in an asymmetric MR space, the users are in different paradigms, such as VR-AR or AR-VR.

Fidalgo et al. (2023) reviewed several works on remote assistance and training using mixed reality, approaching its common topics and differences. The identified applications domain includes multiple fields, such as solutions for industrial and maintenance tasks, assembling, repair, and production line planning. Other specific areas are also included, such as sports, education, and navigation. Furthermore, other works compile specific interaction methods and visualization in remote collaboration using mixed reality, such as annotations (Borhani, Sharma and Ortega, 2023), view interfaces for AR-VR communication (Chang, Lee and Yoo, 2023), gaze-supported modalities (Jing et al., 2022), virtual replicas (Tian et al., 2023), and asymmetric AR-VR approaches (Grandi, Debarba and Maciel, 2019).

Spatial-temporal information is one of the most essential settings explored in related works. In most of them, the temporal information is defined as synchronous, where distant users are connected at the same time and executing the task or training together. However, some solutions explore recording 3D information and guiding and replicating the task performed with asynchronous recorded data (Chidambaram et al., 2021). On the other hand, different approaches are using MR to exhibit spatial information during training and remote collaboration. Some techniques use 360 cameras to stream the passive spatial local information through a video to a remote user, using a VR head-mounted display (Piumsomboon et al., 2019; Lee et al., 2019). Other solutions combine and evaluate 360 videos with a 3D model reconstruction of the physical local setup to allow the remote user to manipulate virtual objects, insert elements in specific positions in the actual local environment, and share visual cues, such as pointing and annotations to guide and solve the collaborative tasks between AR and VR environments (Teo et al., 2020; Kim et al., 2019).

Kumaravel et al. (2019) implement a bi-directional MR environment to coach and train using depth cameras and AR HMDs. Users have the same roles, equal setup, and possible configuration, sharing their environment and tasks and switching quickly between the four modes available. Jing et al. (2022) explore the impact of sharing gaze behaviors in collaborative mixed reality, conducting two user studies to evaluate the interfaces and the effect of this approach, finding that the gaze behavior enables the communication to be less verbally complex, therefore lowering collaborators' cognitive load while improving mutual understanding. Similarly, other works explore various interaction strategies in remote collaboration, such as speech to visualize the shared gaze (Jing, Lee and Billinghurst, 2022) and virtual replicas sharing gestures and avatars during tasks (Wang et al., 2023).

2.2 VR and AR in Healthcare

Over the past few decades, we have seen a growing use of virtual reality (VR) and augmented reality (AR) technologies in various human mental and physical health applications. For instance, some applications aim to treat phobias and help individuals recover from severe mental conditions. One example is presented by Raya et al. (2023), who developed an application for pediatric intensive care units (ICUs) to treat and rehabilitate children suffering from delirium. This application provides external stimuli and can potentially serve as a nonpharmacological therapy.

Non-surgical medical procedures have also benefited from the wide range of advantages provided by virtual reality simulators (Reznek, Harter and Krummel, 2002; Bianchi, Zanatta and Rieder, 2020; Gerup, Soerensen and Dieckmann, 2020). These non-surgical procedures are usually translated as patient-healthcare professional encounters, where empathy and communication skills play an essential role in providing a satisfactory experience to the patient (Caldwell, 2019). They are also important for medical doctors and nurses, so one finds the use of virtual patients for training both professionals (Jimenez, 2022; Patel et al., 2023; Kleinheksel, 2014). Clinical skills for health care have also been explored, as we worked on a correlated project, described in Negrão et al. (2023), a conference paper where we developed and conducted a think-aloud study in a pediatric case, where the users have to perform essential physical examinations and the anamnesis, compile all the information and determine a diagnosis and a treatment (see Fig. 2.2).

Figure 2.2 – Simulation of a pediatrician appointment. The doctor is conducting the information collection and anamnesis (left). The physician performs cardiac auscultation on the patient (center) as part of the physical exam. As one of the rules of biosafety, she is washing her hands in the sink (right).

Source: Negrão et al. (2023).

Car et al. (2022) mapped and reviewed a series of works that use virtual, augmented, and mixed reality in undergraduate medical education, searching for effectiveness in the use of measurement instruments with valid evidence in randomized controlled trials (RCTs). These works also show that less common outcomes included participants' attitudes, satisfaction, cognitive load, learning efficacy, engagement or self-efficacy beliefs, emotional state, competency developed, and patient outcomes. Kyaw et al. (2019) also presented a review that shows a slight improvement in the effectiveness of VR in learning compared to traditional methods. Finally, Aliwi et al. (2023) also reviewed the role of immersive VR and AR in health, locating that most studies with these technologies improve medical communication, although user tolerability limitations were identified.

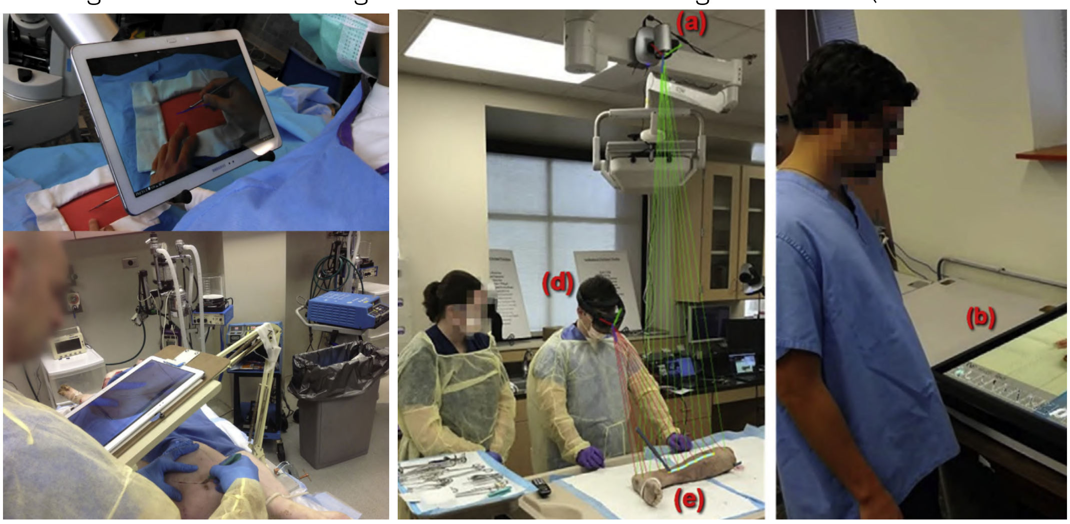
2.3 Surgery XR Applications

AR and MR have been applied to different tasks in the surgical environment, mainly using AR with head-mounted displays, such as Microsoft Hololens, for learning anatomy (Souza et al., 2020), training procedures (Izard et al., 2018), planning surgeries, and displaying information during surgery. Surgery planning applications and prototypes explore the visualization of 3D reconstructions of organs in holograms, which the surgeons can have, based on preoperative exams, better visualization of the patient anatomy, structures, depth, and planning the interventions (Mojica, 2018; Sánchez-Margallo et al., 2021).

VR applications commonly focus on training surgical technical skills, which involves visual and haptic feedback from the manipulation the subjects perform on the virtual organs (Lungu et al., 2021). Such applications aim at improving psychomotor skills, such as hand-eye coordination, spatial orientation, and manipulations, and have targeted mostly minimally-invasive surgeries like laparoscopy (Grantcharov et al., 2004; Huber et al., 2018) and colonoscopy (Wen et al., 2018).

Another advantage of AR interfaces is the possibility of designing different 2D and 3D objects and videos and positioning them at other environmental locations. Using an AR headset, the surgeon can view real-time video of the minimally invasive surgery with complementary information, such as CTs and other preoperative exams, that the surgeon cannot manipulate without leaving the sterile field (Janabi et al., 2020).

Figure 2.3 – Left: 1st STAR Version: remote collaboration using tablet for both remote and local surgeons. Center and Right: 2nd STAR Version using an AR HMD (HoloLens Gen-1).



Source: Andersen et al. (2016) and Rojas-Muñoz et al. (2020)

AR applications use 2D interfaces for remote collaboration and teaching through annotation in a tablet during the procedures. Loescher, Lee and Wachs (2014) have investigated an AR approach that uses a tablet suspended between the trainee surgeon and the operation field to display annotations from the remote surgeon, similar to the first version of STAR (System for Telementoring with augmented reality, Fig. 2.3-left), where these annotations are suspended directly into the field of view of trainee surgeons (Andersen et al., 2016). Rojas-Muñoz et al. (2020) improve the STAR system (Fig. 2.3-center, 2.3-right), using an AR head-mounted display to receive remote guidance directly in the field-of-view of the local surgeon through an AR head-mounted display, improving the usability and extends the area of work.

Cofano et al. (2021) evaluate the gains from AR in spine surgery and the practical application as remote assistance, using teleconferencing software and some types of cues in AR for local users. In ARTEMIS (Gasques et al., 2021), it is possible to aggregate multiple points of view of the surgery, like the camera from an AR head-mounted display and other cameras put in the operation room. The expert remote surgeon is in a VR environment, visualizing a reconstructed 3D model of the patient's body, with a set of functionalities to indicate and communicate cues to the local novice surgeon.

Mixed reality in healthcare for remote collaboration has been studied, including remote surgical consultations and training like ARTEMIS (Gasques et al., 2021) and Tadlock et al. (2022). Other applications, such as telemedicine and telementoring, use AR in real-time as an alternative to traditional 2D impersonal telecommunication platforms (Dinh et al., 2023). These technologies have also been explored in specific medical fields, such as training for medical staff in low-resource environments using AR (Hale et al., 2022); Spine medicine (Morimoto et al., 2022) and remote medical assistance in emergencies (Garcia et al., 2023). Finally, some works have conducted systematic reviews, examining developments in healthcare industries, covering domains such as telemedicine, clinical care, education, and others (Bansal et al., 2022).

2.3.1 Laparoscopy

In laparoscopic surgeries, all the interventions are guided and visualized through the camera inside the patient, turning this video feed into one of the most crucial points in an application for training, in which the visualization of the operation field must be clear. The physical training using the instruments is also the focus of work, which is challenging due to the limited space, motion angle, handling, and the 2D visualization of 3D procedure.

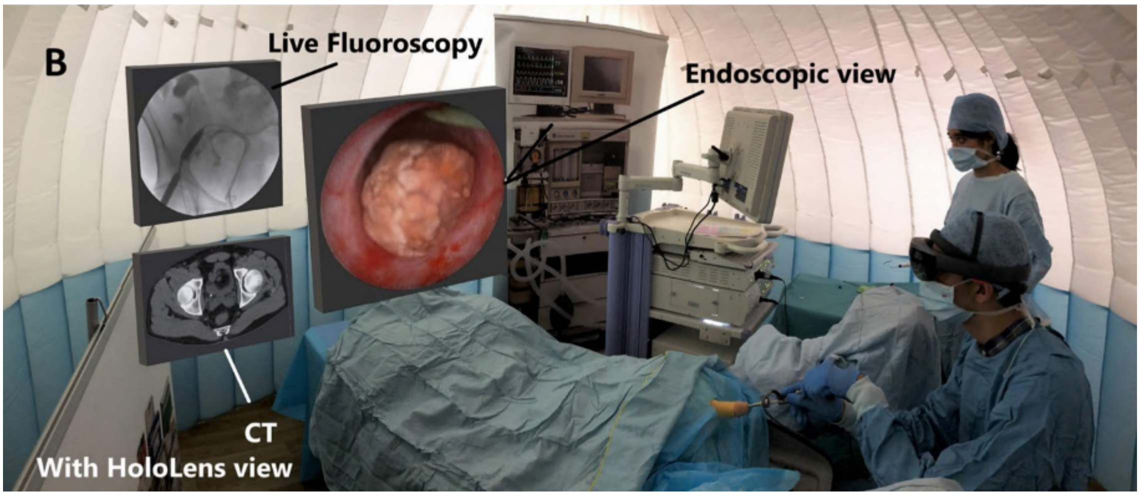
The traditional method for initial training and acquiring skills for a video laparoscopy is using trainer boxes that recreate the scenario in a closed box with inanimate synthetic objects inside, and the trainee interacts with surgical instruments (Derossis et al., 1998). Simulators using virtual reality (VR) for laparoscopy training have also been developed over the years and can be a more realistic solution in the surgical environment (Munz et al., 2004). Virtual simulators for laparoscopy have also been developed and demonstrated to be helpful in psychomotor skills training (Aggarwal, Moorthy and Darzi, 2004). Recently, with advances in (VR), the simulation software seeks to improve immersion and a sense of realism in visual feedback by inside body 3D realistic rendering, organs deformation (Qian et al., 2015), and sense of touch through haptics (Lin et al., 2016).

2.3.1.1 Visualization and Interaction

Augmented reality applications have been developed and evaluated, indicating the effectiveness of using head-mounted displays (HMDs) to place 3D objects in the real world to help explain during the procedures (Barsom, Graafland and Schijven, 2016). These systems can identify important information for guiding and learning, such as the points of interest during the operation through shared appointments (Heinrich et al., 2021), complementing verbal explanations and helping to reduce the complexity of the understanding. In other cases, the application combines the views of surgery, vitals, and preoperative exams in the field of view (Fig. 2.4) to minimize the interruption and support with focus (Janabi et al., 2020).

The HoloPointer (Heinrich et al., 2021) uses the head-gaze direction from Hololens to provide a local shared reference pointing in the 2D operation video monitor triggered by voice commands. All the staff in the operation room can

Figure 2.4 – An example of a simulated view of what the operator wearing the HoloLens sees.



Source: Janabi et al. (2020)

view the exact point that the surgeon wants to show on the monitor, reducing errors and resulting in an economy of movement. In another work, the eye-gaze is tracked and used for free-hand drawing and pointing in laparoscopy video to guide novice surgeons through a standard 2D video monitor (Feng et al., 2020). This system also is triggered and controlled by voice commands.

AR systems for laparoscopy also use hand gestures and voice commands to control the actions to select and manipulate objects. In some works, the voice is combined with gaze direction (head-gaze and eye-gaze) as a trigger to control annotations and pointing functions. While voice commands can be confusing in the operation room, hand gestures demand that the surgeon pause the procedure to interact with the interface. Foot pedals are familiar to surgeons, as they are already used during laparoscopic procedures for cautery activation. Thus, the feet can be considered for interaction in augmented reality, such as using pedals to trigger and confirm visualization and navigation actions through the interface (Jayender et al., 2018). Some adaptations allow the surgeon to perform many functions with the foot, such as slide through previous image exams in the interface panels, allied to the head-gaze for selecting and pointing (Zorzal et al., 2020).

There are no guidelines, however, on how to apply head-gaze, foot actions, and hand actions to control AR interfaces during surgery. In table 2.1, we present a comparison of the interfaces, devices, and interaction methods of some works that use AR interactions for surgery in general. Designers struggle with trial and error to make usable interactions in a challenging environment where a narrow field of view, low accuracy tracking and motion control, and high attention demand undermine the accuracy and efficiency of the operation. In this dissertation work, we approach some of these issues, exploring the design space and characterizing the problems to be solved in an AR system for intraoperative assistance and the challenges of planning a VR system to collaborate for laparoscopy.

Table 2.1 – AR Interaction methods for surgery in related works.

Work	Interface	AR Equipment	Input Actions
Mojica (2018)	AR surgery planning using 3d reconstructed structures.	HoloLens 1st Gen	3D structure manipulation using voice commands and navigation through the slices of MRI using hand gestures.
Heinrich et al. (2021)	Co-located collaborative AR interface for pointing in a standard 2D laparoscopic video feed	HoloLens 1st Gen	Voice commands to select, move, hide the annotations and the shared head-gaze point.
Zorzal et al. (2020)	AR environment with panels to visualize pre-operative images and the laparoscopic feed	Meta2 HMD	Head-gaze to point and select, heel foot rotation to scroll images.
Kumar et al. (2020)	Co-located collaborative AR interface for heart planning intervention with 3D reconstructed structures.	HoloLens 1st Gen	The interaction are performed by airtap with a toolbox (scale, rotate, move).
Zhang et al. (2019)	AR visualization for neurosurgery, mapping of brain tissue, intracranial vessels, nerves, tumors, and their relative positions using MRI.	HoloLens 1st Gen, Huawei M5 tablet	All manipulation of holographic models were manipulated by mobile tablet and visualized in the AR HMD.

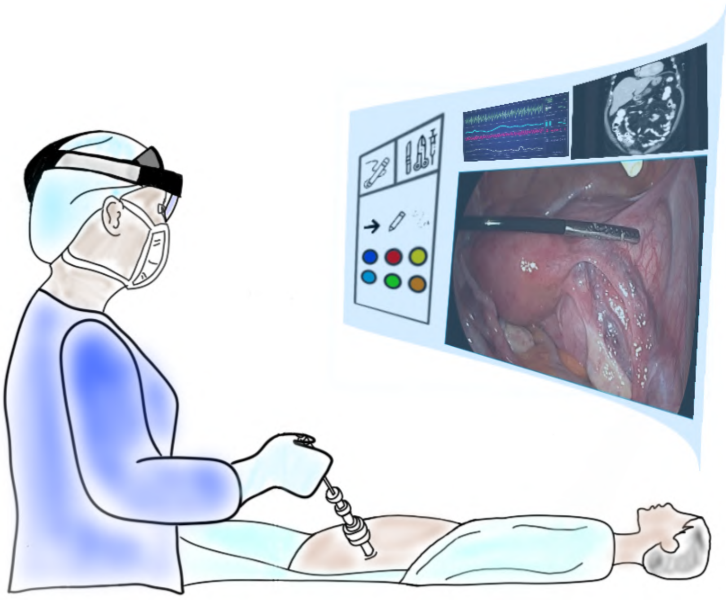
Source: The Author

3 CHALLENGES IN THE DESIGN OF AN ASYMMETRIC XR INTERFACE FOR REMOTE COLLABORATION IN LAPAROSCOPY

3.1 Overview

In our initially planned scope, the MR application is divided into two main interfaces: (1) an immersive VR interface for remote expert surgeons with the visualization of the local operating room and the view of laparoscopy; (2) an AR interface for the local surgeon performing the procedure, visualizing the video from surgery, preoperative exams, and the shared guiding information from the remote surgeon.

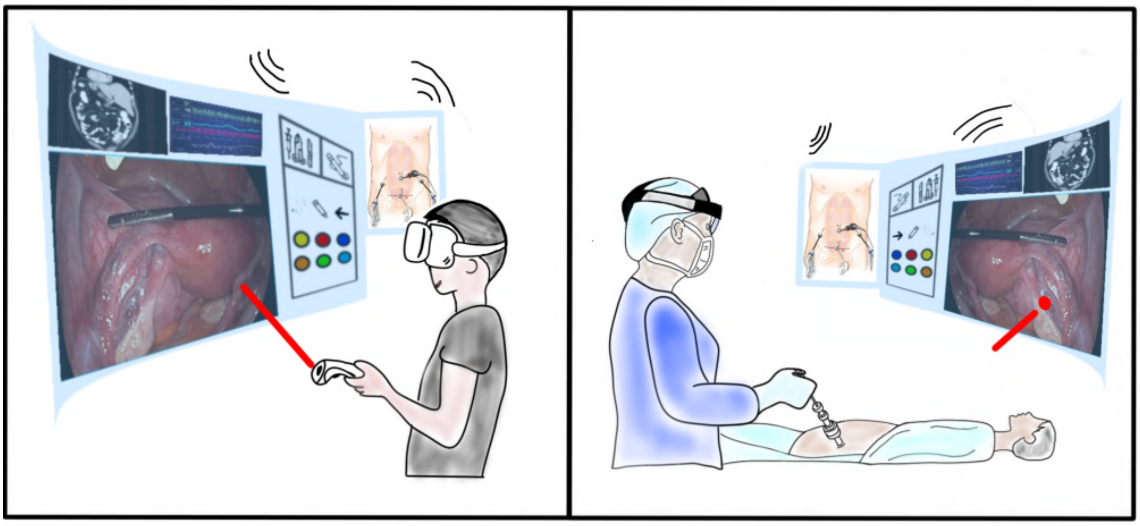
Figure 3.1 – A sketch for initial augmented reality interface.



Source: The Author

Considering where the most interesting interface possibilities are, we assumed that the AR interface was the primary focus of development and evaluation, which the surgeon can use in training, and procedures, and share the OR, laparoscopy view, and all the current information with the stakeholders remotely or even co-located in the operation room. Therefore, the first phase is the development of the interface for the augmented reality HMD, with the main functions and tools that trainee surgeons can use during the procedure without the VR interface for remote collaboration, as the sketch in the figure 3.1. We also focus on the evaluation of the use of this interface in standalone mode. The AR interface is thus the focus of the remainder of this dissertation, even if the broad context of the collaborative setup is also mentioned.

Figure 3.2 – A sketch representing the complete system. Left: VR interface for the remote surgeon. Right: AR interface for a local surgeon.



Source: The Author

The second phase, which is not covered in this dissertation, is the development of the interface for the virtual reality HMD. This will follow the guidelines used in the first phase, focusing on improving the communication between the surgeons through the interface, which was planned always to be connected with the AR interface. In that future setup, experienced surgeons at a distance guide a local surgeon, or in an alternative mode, several remote students can visualize a surgeon using the AR Interface and performing the procedure, as the sketch in the figure 3.2. The third and last phase, also not covered here, consists of the connection of the interfaces and an evaluation experiment focusing on improving communication during a training session guided by the remote expert surgeon and analyzing the impact of internet latency and other factors in this communication. Due to the challenges faced with AR technology planning and development, the second and third phases are out of this dissertation's scope and are mentioned here only to provide context. They will be explored in future projects.

In the following sections of this chapter, we describe the initially planned application, its modes of utilization, our prototype, its layout, and the development challenges. We also describe the first contact of the prototype of the AR interface with stakeholders and the valuable information we achieved during this demonstration session. However, due to a set of challenges described in the last section of this chapter, we decided to switch our research focus to the AR interface and its interactions in the INTRA scenario.

3.2 Planned Modes of Utilization

In the concept of the interfaces, we planned three modes of utilization, combining or not the VR and AR Interfaces. The main objective is not to limit the application to be used only for collaboration mode or only for local mode, opening the possibilities that better fit the surgeon's needs. The three modes are described as:

INTRA Scenario: The surgeon visualizes and interacts in augmented reality with the live video of laparoscopy and can access preoperative exams, planning data, and vitals positioned in the real environment.

INSIDE-OUT Scenario: The students watch and visualize remotely in VR the procedure performed by the preceptor using and interacting with the AR interface, the same interface described in the INTRA Scenario.

OUTSIDE-IN Scneario: An expert surgeon using VR guides a local novice surgeon using AR to perform complex procedures using interactions and deictic communication between the interfaces. In this case, we planned to have a 360º camera filming and streaming the video from the local operation room, in which the remote surgeon can view the surgical team and the actions.

Although we have planned to start all the modes of utilization, in the following sections, we will focus on planning and evaluating the interactions and the augmented reality UI for the INTRA and INSIDE-OUT Scenarios. In parallel, other research groups focus their work on the virtual reality UI (Taweel et al., 2023).

3.3 Initial Planned Application

Prior work shows that AR/VR can provide several types of support for laparoscopic surgery. We are exploring interactions on AR interfaces to enhance the surgeon's communication with local and remote stakeholders in VR when needed and to remove the need for communication where it can benefit the operation outcome. To learn about these interactions, we propose a minimalist AR interface design. To guide our strategy, we assumed a small set of requisites and constraints that are common to laparoscopic procedures:

- the surgeon works standing;
- there is an inside-the-body video feed;
- preoperative images must be consulted intraoperatively;
- the surgeon wears sterile gloves;
- there are stakeholders in the OR and potentially remotely;
- the surgeon needs to indicate locations and objects in the surgical field to the stakeholders;

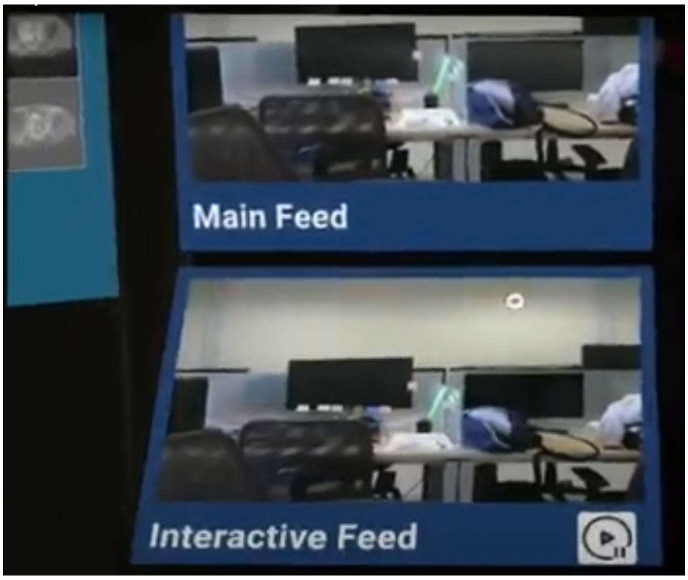
Although the complete work aims to implement the interface in actual laparoscopy surgeries, we understand that the application needs to pass several tests and adjust before a trial in the real environment. Thus, besides the requisites and constraints set before, we planned our tests and the first experimentation of our interfaces using a white training box for laparoscopy, which the students use to learn how to interact with the instruments, the angle, and camera control.

3.3.1 AR Interface Layout

In the development of the first version of the interface, we tried to include the maximum amount of information around the surgeon, having two large panels only for laparoscopy, in which one can never be paused and is a real-time procedure, where the surgeon can quickly react to unexpected behavior in the procedure without any interruptions or annotations placed above the video. Another panel is where the surgeon can freeze, annotate, and save a screenshot without any concerns, besides the panel for MRI, CT, and all other information that the surgeon wants to access during the procedure. We also thought that the more significant number of panels and information would help the users, both in the standalone use, to not interfere in the clean video feed of surgery and have all the information available, and in the remote collaboration use, where the remote expert user could manipulate and rearrange the panels and the information in the best way to guide the local surgeon.

We use the game engine Unity[1] 2019.4.20 f1 LTS version for development, the latest version with support for Windows XR Plugin with *Holographic Remoting*. This plugin provides an integration to connect and stream the application in the editor mode for the Microsoft HoloLens 1st Generation[2] (HL1) available in our lab. The HL1 has some limitations, and a major one is a 30-degree field of view (FOV), which is smaller than the 54 degrees in the Hololens Gen-2 and the average VR HMDs of over 90 degrees. This limitation causes the user to be unable to see virtual objects anywhere in their natural field of view. The holograms and interface elements appear only on a centered rectangle, requiring head turning to explore other areas. We also utilize Mixed Reality ToolKit 2 (MRTK)[3], a Microsoft-driven project, which includes scripts, assets, and features to improve the AR/VR software development, especially for Hololens.

Figure 3.3 – First Version With Two Video Panels



Source: The Author

Using demonstrative videos and data for preoperative exams, we demonstrate the first version of the interface (fig 3.3) to surgeons and physicians in the HCPA (Clinical Hospital of Porto Alegre), which is a teaching hospital linked with the Federal University of Rio Grande do Sul. In this opportunity, we have planned the following set of questions to ask the surgeons:

[1]https://unity.com/
[2]https://learn.microsoft.com/en-us/hololens/hololens1-hardware
[3]https://learn.microsoft.com/en-us/windows/mixed-reality/mrtk-unity/mrtk2/?view=mrtkunity-2022-05

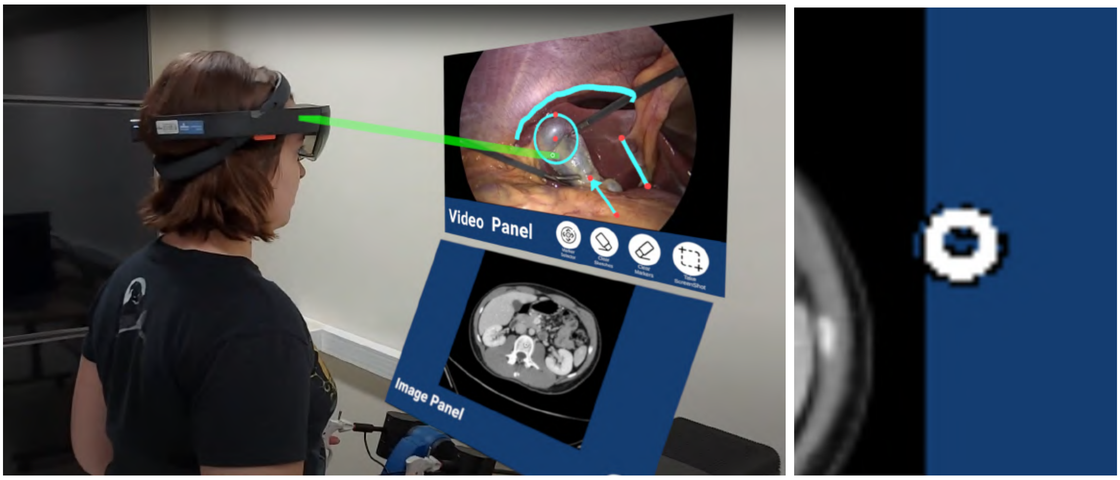
- What is the best position of the monitor with the surgery laparoscopy video?
- Anchor the virtual monitor in the surgeon's FOV or anchor in a specific point in the real world?
- The user would like to choose when anchoring in the environment or in their FOV?
- Two equal virtual monitors from surgery can be helpful? Pausing one of them can be useful?
- Which information is more relevant during the surgery?
- Drawing and placing annotation in the surgery video can be helpful?
- How is the connection of the laparoscopic equipment in the operation room?

In unstructured interviews, we noted that a lot of panels and information can be confusing for the surgeons, and it is not necessary to use a second panel only for pausing and annotating; they can do this in the main video feed without freezing in any situation. The location of the second video was replaced by the image and data panel, which reduces the movement of the head and optimizes the utilization of the interface during the surgery, besides providing a concise and efficient interface. Therefore, we conclude that they need to access the laparoscopic live feed from inside the body constantly and to access preoperative data sporadically. Other information, such as vitals, is monitored by other team members.

3.3.2 AR Interaction Design

We have studied several methods for interactions with the AR interface without interfering in the sterile field and demand that the surgeon take the hands off the instruments during the surgery. In the literature review, we found several works that present voice command solutions. However, the local team communication and the future focus on remote collaboration, where the surgeons must communicate with each other remotely, must be considered in the design of interaction. Hand gesture(Kumar et al., 2020) solutions are also debatable due to the necessity for taking off the hands of the instruments and the inaccurate capture of movements, which can be awkward. Also, from the literature, we found some adaptations using foot interactions, where the surgeon can access different information and perform interactions according to the foot movements and with foot adaptation for confirmation. Furthermore, there are approaches using head-gaze and eye-gaze to point out areas of interest and show deictic commands for remote collaborative users, which fits well with our purposes.

Figure 3.4 – Head-Gaze representation and the mapped reticle for interaction.



Source: The Author

Therefore, we decided to move on with a head-gaze approach, which is more ecological to the operating room and still an effective way to point areas, draw, and annotate in the video-lap panel (fig. 3.4), although in some cases might be discomfort and inaccurate. Annotations are helpful in learning and teaching the procedures. We dismissed the eye-gaze due to the unavailability of equipment, leaving it for future works. Together with head-gaze to point and select, we understand that foot pedals are familiar to the surgeons and can be used to confirm actions pointed by the head-gaze, besides having a shorter learning curve. Nonetheless, as they already have to use pedals during the surgery to cauterize, more pedals can be confusing and prone to errors. Consequently, we planned and included an evaluation of a new hand device, described in the section 4.1 as a ring device with the same function as the pedal, to confirm actions pointed by the head gaze.

3.3.3 Planned VR Remote Interface

In our planned scope, the VR Interface is similar to the AR interface, with different possibilities of interaction and control of information, as the user is not with the hands busy and can interact simply using the VR controllers with techniques such as ray casting. We also planned to include a video and panel control, where the VR remote user can manipulate the position of the panels to help the local surgeon. In addition, the remote experienced surgeon can also view the annotation and the pointing of the head-gaze from the local surgeon, which can help to understand how to help the novice surgeon.

3.4 Challenges

During the exploration process of the AR Interface, we faced a couple of challenges that we had to understand and explore before starting the development of the VR Interface and the connection, as the limitations of interaction methods and hardware, the implementation and integration of the hardware with the HL1. Thus, due to the previous necessary evaluation and the limitation imposed by the available hardware, we have focused on evaluating the interaction methods of the AR Interface for use standalone in the INTRA mode.

Although AR systems can help in the training of laparoscopic surgeries and have been demonstrated in see-inside-the-body applications, intraoperative AR systems require interaction. They could not succeed and interfere in the course of the procedure, due to the learning curve of how to interact effectively with visual data and the interface during surgery.

The Microsoft HoloLens (HL) is commonly used for developing and evaluating augmented reality applications in related works. Some of the AR applications for laparoscopy surgeries use voice commands and hand gestures as their primary method of interaction, both supported by HoloLens. However, using voice commands may be less accurate, and it may become confusing during procedures, as well as hand gestures, which make the surgeon pause the process and take their hands from the surgical instruments to interact with the interface and objects. Through a visual ray-cast cursor, the head-gaze method comes up as a solution to pointing specific positions and performing actions in the AR interface. We understand that these interaction methods have contributed to enhancing the learning of laparoscopy surgeries using augmented reality.

However, during the planning and first tests of the interface prototype, we noted that the trembling of the direct head-gaze and the narrow field of view of

HL1 appeared as a challenge, which we focused on finding and evaluating possible solutions, as described in the following chapters and the user experimentation of this dissertation.

4 FINAL AR INTERFACE FOR LAPAROSCOPY SURGERIES

In this chapter, we describe our final setup after several design iterations. A user evaluation is presented in Chapter 5.

4.1 Hardware Design

To simulate the live video from a laparoscopic source, we use a Logitech c525 webcam to receive live video data and stream it to the AR interface. However, the software permits selecting and configuring the system setup to receive live video information from any source connected to the desktop, such as a training laparoscopic attached to a USB port or a capture card. Moreover, it is possible to use recorded video and static images as a source for evaluation and demonstration purposes. We also built a similar white box using expanded polyethylene to understand how the surgeons are positioned and how we can place the interface in the environment.

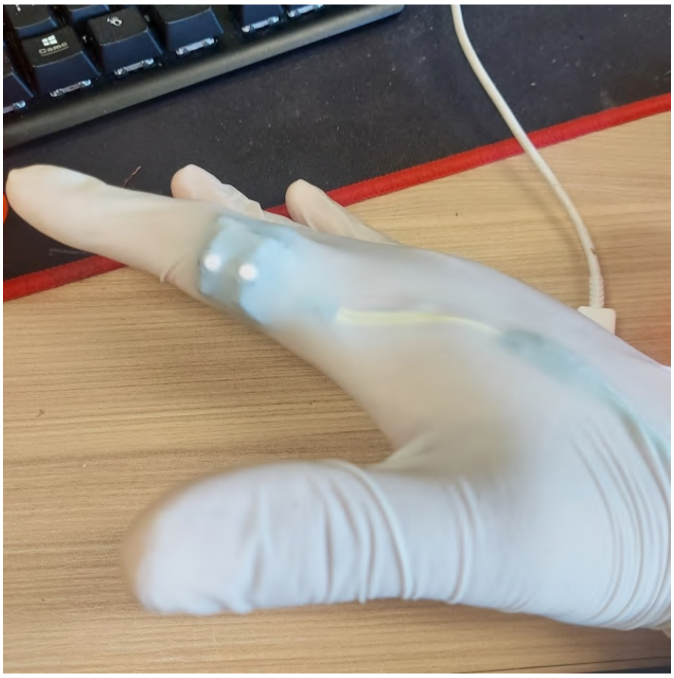
The interface interaction in the laparoscopy AR system is a challenging planning and development step, as the surgeons focus on the procedure and use their hands to operate the surgical instruments. Therefore, to interact with our interface, the user uses the head-gaze direction provided by Hololens, with a graphic reticle as a cursor (see Sec. 4.2 for details). Input is also needed for confirming options. Thus, we had to adapt physical props to provide a fast and precise clicker for confirmations.

We designed two options. One is a pair of foot pedals similar to those already used in operating rooms. Surgeons also are familiar with pedals, additional pedals may be cumbersome. So, we provide a second option to allow hand interaction without releasing the surgical instruments. The device is an index finger ring with two 5mm push buttons on top of it, which the surgeon can wear under the sterilized gloves (see fig. 4.1. The buttons are reachable with the thumb without taking the hand off the instruments.

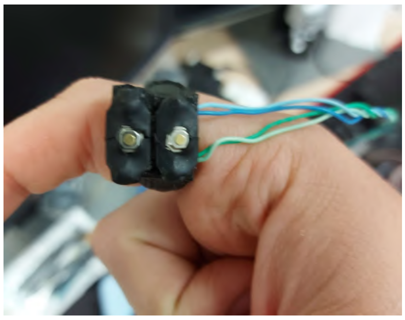
We built the pedals in the prototype using a two-part structure 3D printable model published by *Adafruit* on Thingiverse [1]. We built a 3D model adapter in Blender 3.0 to use a 12mm push button in each pedal. A metal spring binds

[1] https://www.thingiverse.com/thing:2829107

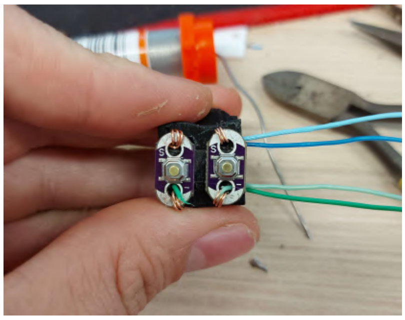
Figure 4.1 – The ring device with two buttons.



(a) Ring under glove.



(b) Without glove



(c) Scrapped ring

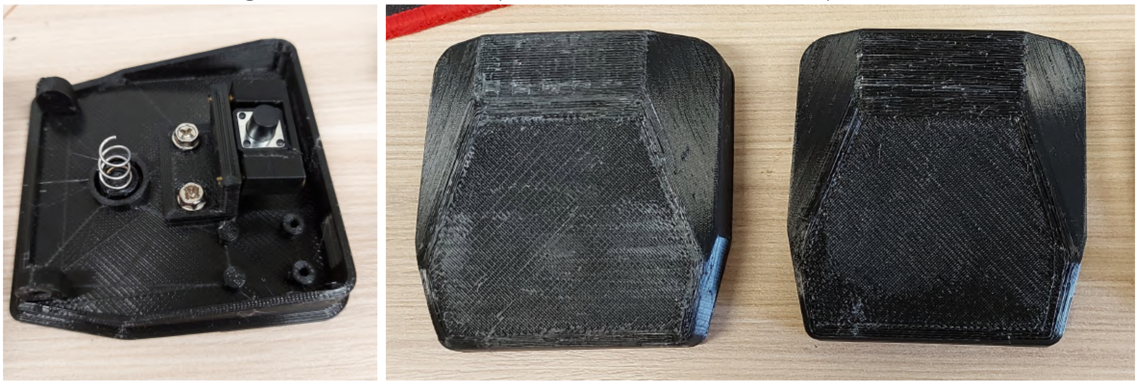
Source: The Author

the bottom part with the top part of the pedal, besides non-slip pads placed under the pedals (fig. 4.2). The ring is also a 3D model built in Blender, inspired by simple signet rings to place the buttons and orifices to pass the cables. Both devices are printed in 2.85mm PLA using a BCN3D Sigma 3D printer. Their connection with the system is through an Arduino Due connected to a desktop, which interprets the actions and executes them on the AR Interface equally for both devices. However, while planning the user experiment, we acquired commercially manufactured wireless three-foot pedals (see fig 4.3).

4.2 Interface Layout

As virtual augmentations can be placed anywhere around the user, we prioritize displaying the surgery laparoscopic video and preoperative images in two adjacent panels viewed as holograms that the user can place wherever they wish around them in an egocentric perspective. Furthermore, the surgeon must be attentive to the real environment, the operation room, and the assistants during the procedure. Thus, our interface design avoids cluttering the view with unnec-

Figure 4.2 – Initial 3D printed handcrafted foot pedals.



Source: The Author

Figure 4.3 – Commercially manufactured wireless foot pedals (we use only two of three inputs).

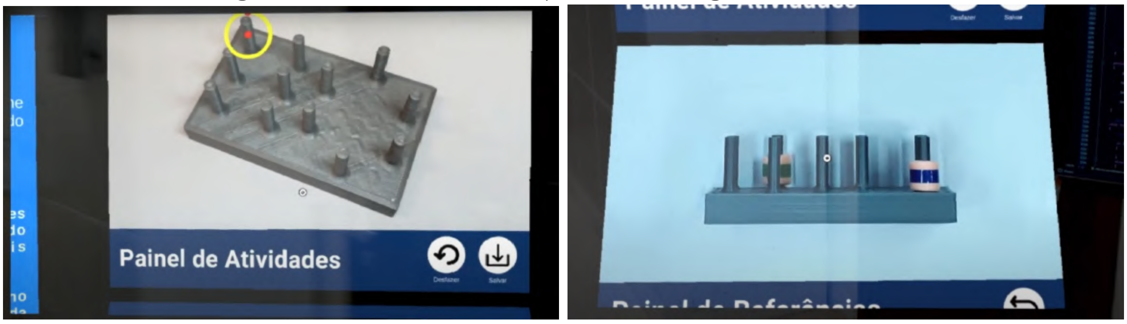
Source: The Author

essary information and objects that could disturb concentration. Each panel has specific functions, interactions, and virtual controllers with different actions. An overview of the panels can be seen in Fig. 4.5. Detailed views are presented in Figs. 4.8 and 4.9. A view of panels using the HL1 is presented in figure 4.4.

The most significant panel of the interface contains the real-time video from the laparoscopy procedure and its functionalities, initially positioned straight in front of the user's field of view. Right below the video panel is the image panel that compiles folders with preoperative exams, such as MRI, CT, and any planning sketches that could be helpful. Finally, in our experimental version, we included two lateral help panels containing usage indications for the interaction controls. More panels can be included for specific minimally invasive surgeries, such as a vitals panel for cardiac procedures.

The laparoscopic video could be easily enlarged in the user's field of view to show more detail, which is challenging in conventional monitor-based la-

Figure 4.4 – The interface panels through the HL1 view.



Source: The Author

Figure 4.5 – A macro view of the panels in the interface.

paroscopy. In our current setting, for evaluation purposes and to cope with the limited Hololens FoV, we maximized the video panel to fill all the user's augmented view, avoiding distractions and unnecessary information in the interface and improving the visualization of the procedure with the annotations. The user can also have a better ergonomics arrangement of the panels, changing the position of the interface using the head-gaze direction. The user can also move the head forth and back to zoom in and out naturally. The FoV issue tends to disappear as the technology evolves and with the recent releases on MR head-mounted displays that can perform in both fully immersive environments and in mixed real-world plus virtual elements (video-see-through).

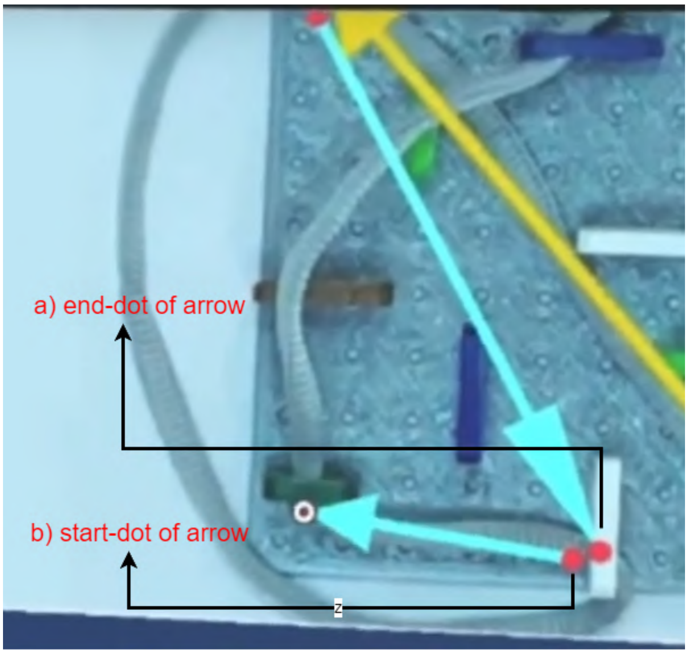
4.3 Interaction design

The higher-level user actions can be summarized as:

1. manipulate medical instruments;
2. observe video feed;
3. check for a specific preoperative image;
4. signal a location in the video;
5. circle an object in the video;
6. save a screenshot;
7. adjust the location of the hologram.

Items 1 and 2 are inherent to the surgical task and are not supported by the interface. The other five tasks can be accomplished by a pointing plus confirmation metaphor with direct manipulation using the panels presented in Sec. 4.2 above. We placed the interactive actions on the panels, where the video feed, preoperative images, and the interface's virtual controllers are placed. The Help Panels are for experimental purposes and are not interactive, just for content visualization.

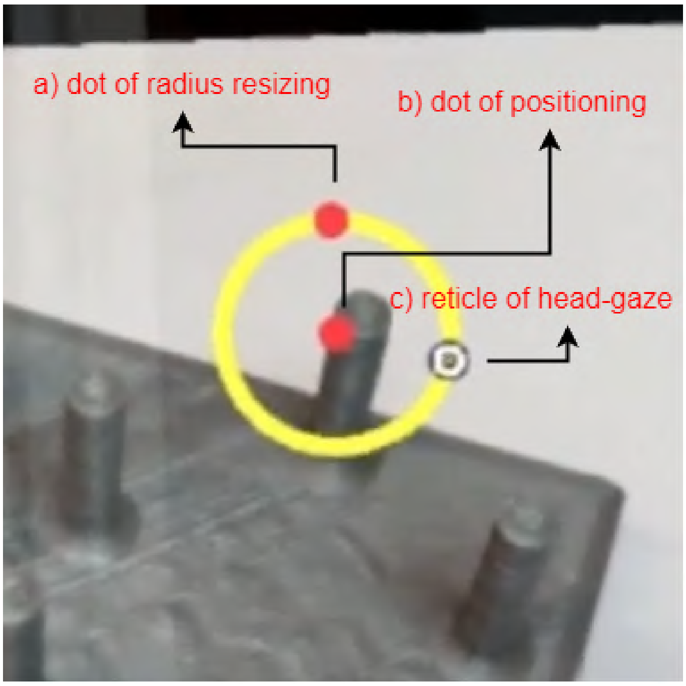
Figure 4.6 – The arrow annotation and its manipulation points. The line manipulation is equal to the arrow.



Source: The Author

While confirmation relies on the ring or pedal buttons already described, we argue that a head-gaze-controlled reticle is a suitable solution for pointing. Notice that the hands are busy and that using the surgical instrument tip for pointing will interfere with the actions significantly and have limited scope. Few other options remain, such as foot pointing, voice, and facial expressions. None of them has reported great accuracy in previous works. Head-gaze pointing, in turn, has been extensively explored (Grinshpoon et al., 2018; Trejos et al., 2015; Asao et al., 2021).

Figure 4.7 – The circle annotation and the manipulation dots for positioning and resizing.



Source: The Author

The laparoscopy live video panel implements the annotation metaphor. It allows free drawing in the live video, which is paused while drawing, as well as three standard markers: line, arrow, and circle. All markers work similarly, using the left pedal or the left button of the ring device to confirm the actions and the head-gaze reticle to provide the direction (selection locus). Two points are inserted for the line and arrow by pressing, dragging, and releasing the button, as is shown in figure 4.6. For the circle, the first point indicates the center (fig.4.7a), and the second provides the radius(fig.4.7b). Two red dots in the markers help in the delimitation and can be used to manipulate existing markers by grabbing and moving. The first and second dots change the length and direction of lines, the arrowhead position, and the length of arrows. Selecting and holding the center dot of a circle changes its position, and the radius dot changes its radius.

Finally, the free draw is the most straightforward annotation and does not permit repositioning or editing.

Some virtual controllers (fig. 4.8) that help set up the annotations tool are placed below the video in the panel, with a virtual button to take screenshots of the live video, including the annotations created. To use these controllers, the user has to point the head-gaze reticle in the controller and press the left pedal or the left button in the ring device. Further, next to the screenshot button are two clean buttons, one to clear the free draw and another to clear the standard markers from the video. Moreover, the last control opens a floating circle menu to select the marker for the annotation tool (see fig 4.8-right).

4.3.1 Head-Gaze Pointing

The head-gaze is the primary method to select locations in our interface. It is captured from the HL1 orientation. The typical circular reticle pointer is shown where the gaze vector intercepts objects within the view, allowing for a mouse-pointer-like pointing. However, the reticle pointer provided by the head-gaze direction is sensitive to tracking inaccuracies and user-generated jittering. It can be challenging for inexperienced users in AR systems to accurately move and position the reticle on the interface controllers and use the annotation tool. Thus, we apply different stabilization strategies to make the reticle control smoother. We explore the effects of these strategies in a user test (Sec. 5.1.1). We implemented three modes (free, average, scale) and two levels (low, high) for this stabilization, resulting in five conditions: free, average-low, average-high, scale-low, and scale-high, as summarized in Table 4.1.

The average filtering approach stores the n previous head-gaze directions and calculates a simple average. It renders the reticle at the respective averaged hit position. The scaling approach works by scaling down the head-reticle movement ratio, which is 1 when not scaled. We applied different scales for the vertical and horizontal directions to fit the display aspect ratio. On the disadvantages side, the average will add up latency while the scaling will decouple the center of the view from the target. If exaggerated, scaling will cause the user to turn so much the head that the target cannot be seen. As the display area follows the head, the target might even be cut out of the display area.

Table 4.1 – Reticle stabilization modes

Method	Level	Parameters
Free	-	Direct head-gaze mapping
Scaled	Low	V-Rot * 0.825, H-Rot * 0.75
	High	V-Rot * 0.75, H-Rot * 0.5
Average	Low	mean(last 20 frames position)
	High	mean(last 40 frames position)

Source: The Author

4.3.2 Selection and manipulation

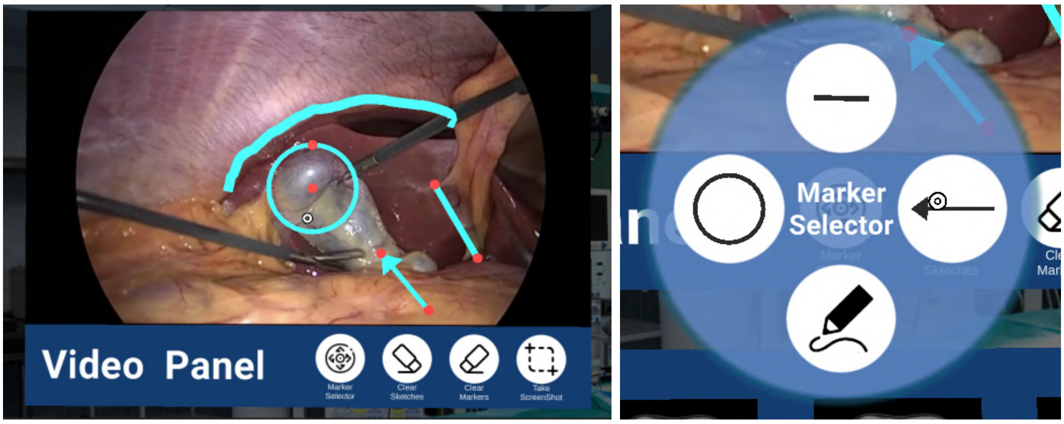
Each user task can be performed in the interface by pointing and clicking on elements that appear in the holographic panels described in Sec. 4.2. In such a way, we designed each task as a combination of direct selection and manipulation actions, privileging efficiency.

In the live video panel, four actions can be initially triggered by selecting the respective button: place an annotation, clear markers, clear sketch, and screenshot (see fig. 4.8). When $annotation$ is selected, a circular menu with four options allows choosing among line, arrow, circle, or free sketch. Lines, arrows, and circles are considered markers. Markers can be deleted separately from the free sketch because surgeons often need to save markings, while the free sketch is used for incidental communication only.

After choosing the desired marker, each has a specific set of placement/size actions. For $arrow$, the user first selects the tail position and then moves to select the arrowhead location. For $line$, the process is similar. For $circle$, the user selects the center location and then moves away to choose the radius. These markings can also be edited after creation. Red circular picking points appear at the line and arrow extremities and the center and periphery of circles (see fig. 4.8). Free sketching is applied by holding the button/pedal while moving. All selections described here are made with the left button/pedal.

Finally, the last possible action with the video panel involves all the panels in the interface. It supports the need of the user to modify the placement of the set of holograms to the most convenient location for each user and each moment of the operation. When the head-gaze reticle is in the video panel and the right button or pedal is pressed, the whole interface will follow the user's head-gaze to anywhere they wish until the right confirmation is activated, fixing the interface

Figure 4.8 – Left: The video panel with annotations.
Right: Floating Marker Selector Panel.

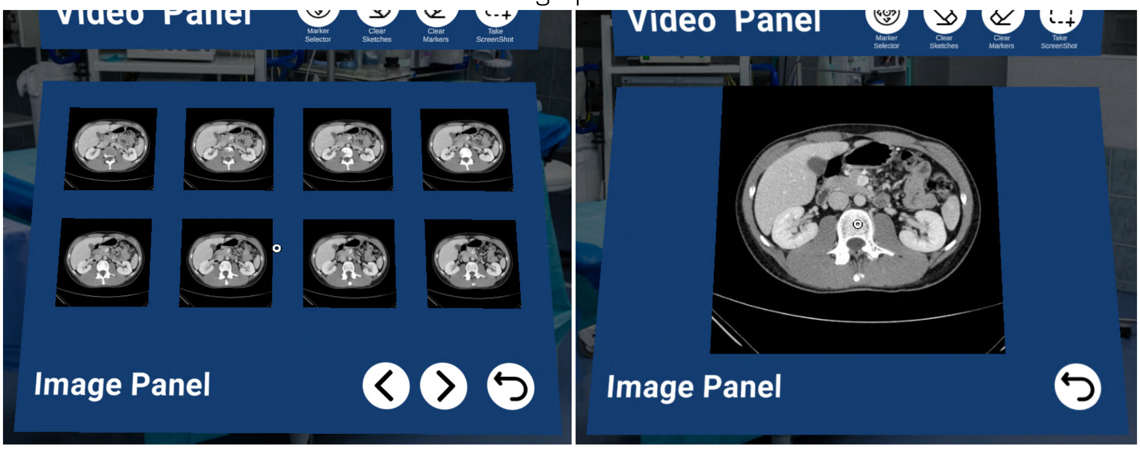


Source: The Author

in the space. Besides setting up the most ergonomic location, this function also allows the surgeon to maintain the panels locked to the head like a heads-up display, which can be helpful to keep all the attention on the surgery, even when the head moves.

The actions in the image panel (fig. 4.9) are more straightforward; the user can navigate through folders and the image grid by pointing the reticle to the virtual controllers and selecting the folders and images by clicking on them. When an image is selected and presented in the panel size, the user can scroll to the next or previous images using the right and left buttons while the reticle is on the image. There is also a virtual back button to return to the image grid.

Figure 4.9 – Left: the grid-view in the image panel. Right: Full view of a picture in the image panel.



Source: The Author

4.3.3 Research Questions

While we designed this interface accounting for the user's needs and general good practices of human-computer interaction, several design aspects must be evaluated before they are final. Focusing on interaction and communication improvement provided by the AR interface, we will evaluate the following aspects in this work:

- The effects of averaging or scaling in head-gaze interaction;
- How much scaling or averaging time is best;
- The effects of interacting with the foot or the hand;
- How much overhead the change of focus among panels imposes;
- How much communication improves with the use of the interface.

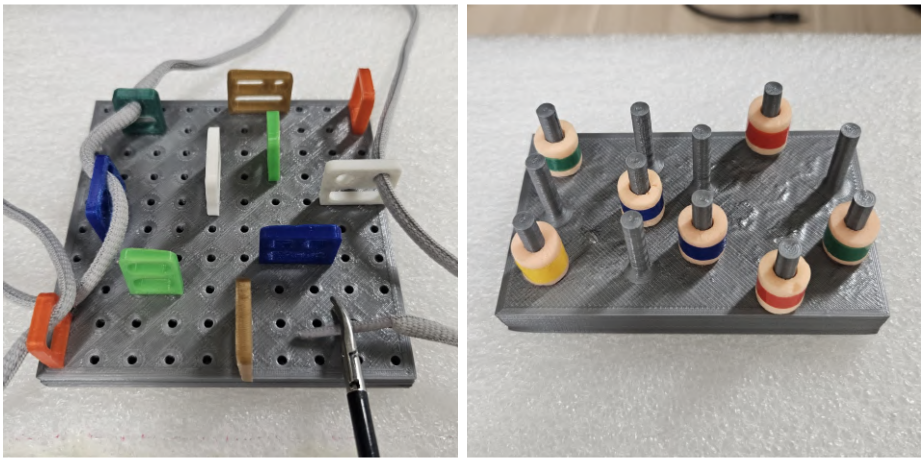
Among these aspects, we focus on the first two in the following chapter of our user-study.

5 USER EXPERIMENT: AR AFFORDANCES FOR LAPAROSCOPY

5.1 Overview

Following the final developed AR interface and challenges presented in the chapter 4, instead of using a real laparoscopy setting with doctors and patients, which would be too premature at this research stage, we developed two physical tasks similar to laparoscopy actions using 3D-printed problems. These tasks are adaptations based on the Fundamentals of Laparoscopic Surgery (FLS) (Ritter and Scott, 2007; Hafford et al., 2013) and the McGill Inanimate System for Training and Evaluation of Laparoscopic Skills (MISTELS) (Vassiliou et al., 2006), which are considered the gold standard for training and assessment of laparoscopic skills.

Figure 5.1 – Laparoscopic Physical Tasks. Left: Thread Passing. Right: Peg Transfer



Source: The Author

The *peg transfer task* (fig. 5.1-right) is where the user transfers small rings using a couple of laparoscopic tools from one peg to another in a peg board. The pegboard model that we used has been published in Thingverse[1] by *SpaceChild*, and the rings we manufactured using simple EVA (Ethylene Vinyl Acetate) with colored adhesive tapes to differentiate the pegs. The other task is thread passing (fig. 5.1-left). The surgeon has to pass a thread through a sequence of portals on a board. We create the 3D models of the portals with different patterns using Blender 3.5 and adapted the 3D model of the board where the portals stand from a model published by *adamtal* in Thinverse[2]. To simulate the thread, we used a simple shoelace.

[1] https://www.thingiverse.com/thing:4974050
[2] https://www.thingiverse.com/thing:2145069

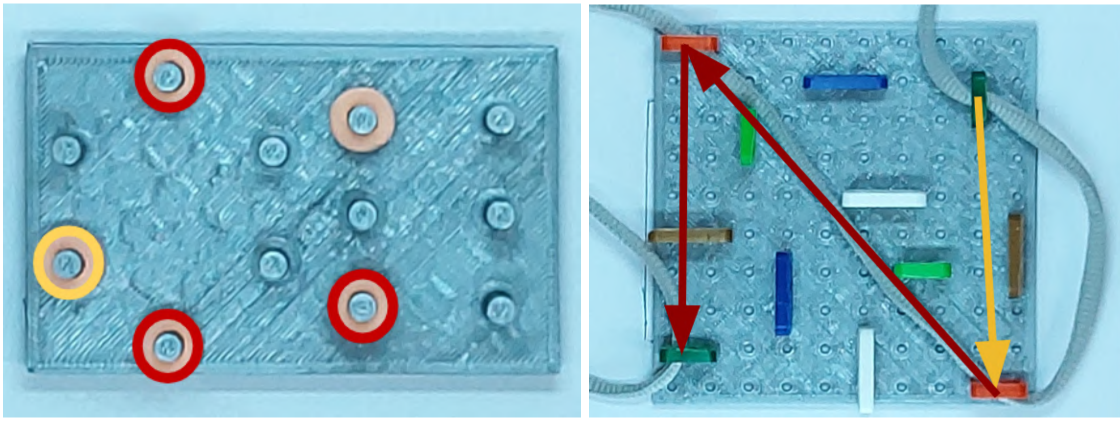
5.1.1 Design

To find an outcome to our research questions, we need to assess the effect of using the scaling or the averaging approaches on user performance. We also need to evaluate the most suitable level of scaling and averaging. We obtain four conditions using two levels for each (low and high). Adding the baseline with direct mapping, we obtain five interface conditions. Besides the five $interface$ conditions, we have another independent variable: $task$. This can be peg or $thread$, with three different difficulty level problems, which gives us three conditions for each task. A $problem$ is a specific configuration of rings in pegs (more rings means more difficult) or a sequence of portals traversed by the thread (the more portals, the harder). We have conducted a pilot test with two VR/AR experts and a within-subjects design to check for feasibility/duration, and the thirty ($5 * 2 * 3 = 30$) unique combinations per user were demonstrated to be exhausting and too prolonged, which could affect the measures. Therefore, we decided to conduct the experiments with a between-subjects design for $task$, which required half of the unique trials per user: 15.

For this test, we removed the actual surgical tools manipulation so that the users will only make annotations on the video. Thus, the population for this test does not need to be composed of surgeons. However, we asked participants to stand with their hands at the laparoscopy instruments, even though they were not performing the physical manipulation. As the hand is dominant for manipulation with the general population, and the foot pedals are also already common to surgeons, this test will be conducted with the ring device only to reduce the number of trials and focus on the measurement of annotation precision and the interface interaction in AR. Therefore, the main hypothesis of the user study are:

- The head-gaze stabilization methods are more accurate for interaction and drawing;
- Using the head-gaze stabilization methods, the users can perform the tasks with less time;
- The communication of surgeons is improved by the interface and its functionalities.

Figure 5.2 – Problem examples for the two tasks. Left: Peg transfer task, where the red circles are positioned above the rings, except the green one, and the yellow is a reference (In this picture angle, no one ring color is visible). Right: Passing thread task, the yellow arrow is a reference, and the arrows connect the points that the thread passes through.



Source: The Author

5.1.2 Protocol

First, we collected demographic information in a questionnaire to understand participants' previous experience with augmented reality systems and non-conventional interactions using hand confirmation and head movements. Based on our pilot tests, we divided the tasks in a between-subjects design due to the long duration and effort spent performing two tasks in a single test. Thus, the participants executed the trials of only one of the two tasks; the first half completed the trials of *Peg Transfer task*, and the other half completed the trials of the *Thread Passing task*.

To be correctly carried out, either task needs information stored in photographs of the solution. These photographs are quasi-orthogonal views of the problem solution. They are available from the preoperative images panel, simulating tomography data in a surgical setting.

In the *Peg Transfer task*, the users see an empty peg board in a semi-isometric perspective. They have to annotate a circle on each pin that contains a ring in the image panel, except for the green ring (The colors are visible in the fig. 5.1 right). One ring is pre-annotated with a yellow circle to provide an example during the trial. In the *Thread Passing task*, participants see an empty portal board in a semi-isometric perspective and were instructed to trace an arrow connecting the points where the thread passed through in the provided images.

Participants were explicitly instructed to disregard any curves or variations in the thread and focus on identifying the starting and ending points of the thread path through the portals. The first connection path was marked with a yellow arrow to assist participants, serving as a guide to indicate the initial thread position and the direction to follow in each trial.

Before starting the trials, all participants passed an adaptation and training stage, in which they visualized all the panels in the AR interface and learned the functions of the panels and how to interact, point, and annotate in the video panel using the head-gaze and the ring buttons. The participants also performed a training trial with a less difficult level of tasks, in which they had to understand the problem and the steps to finish the trial. All training trails are performed using the standard approach for the head-gaze reticle control, the direct mapping mode, without stabilization and compensation.

Table 5.1 – Task Level and Trials Sequence Example.

Task Level	**Easy**	**Moderate**	**Hard**
Circles to Draw	1	2	3
Arrows to Draw	2	4	8

Trial	**Reticle Methods Randomized**		
1 - 3	AvgLo	AvgHi	Free
4 - 6	AvgHi	Free	SLo
7 - 9	Free	SLo	SHi
10 - 12	SLo	SHi	AvgLo
13 - 15	SHi	AvgLo	AvgHi

Source: The Author

Within each task, other two elements determine each trial: the task level (1 to 3), determining the execution order from the easiest to the most challenging; the head-gaze reticle control approach (average-low, average-high, scaling-low, scaling-high, or direct mapping) randomized by a Latin square. The trial commences when participants have read the instructions and accessed the designated folder within the image panel containing the photographs corresponding to the current task level. The user must view and interpret the images, draw the annotations on the video feed panel accordingly, and click the save screenshot button to complete the trial. Table 5.1 summarizes the trial parameterization in an example. Only six unique problem boards are used, three for each task and one for each level. These problems are repeated five times in the same order but with a different reticle control approach, resulting in 15 trials that allow all subjects to perform all possible combinations. A user performs a combination of one head-gaze condition and task difficulty. Notice that the table shows an example and that each reticle method column is an upward rotation of the previous one. For the next participant, the last column is rotated in such a way that trial 1 will use the SLo condition.

After performing the 15 unique trials, the participants answer a post-test questionnaire to collect subjective factors, such as perceptions about the reticle control approaches and the interaction methods. Besides, we administered the NASA Task Load Index form (NASA TLX) (Hart and Staveland, 1988)) to collect the effort and workload per task.

5.1.3 Data collection

The application automatically stores the time in seconds to complete each trial, starting when the user opens the folder in the image panel and ending when the user saves by clicking on the button on the task panel. Each trial also yields a saved image with the circles/arrows on the board.

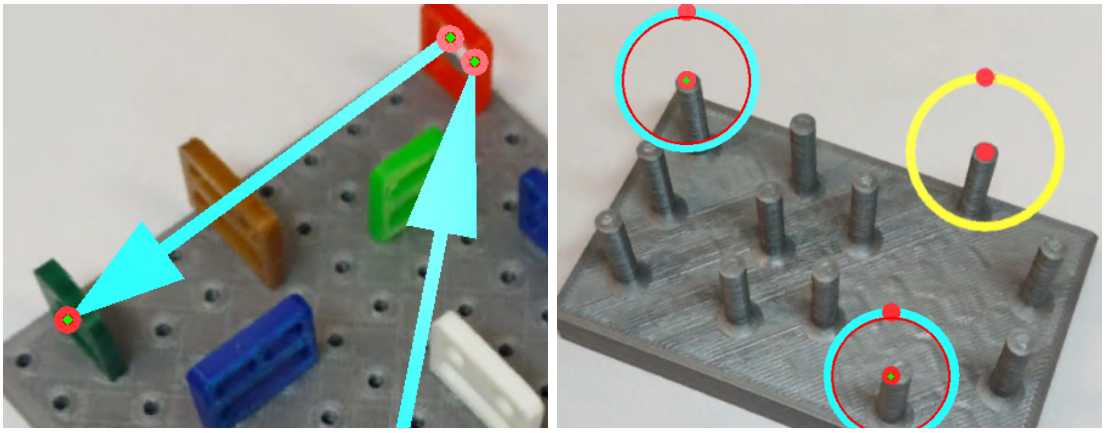
5.1.3.1 Accuracy from images

Accuracy is based on the center dot of the circles, the radius, and the starting and ending dots of arrows. The application does not keep the positions during the execution. Thus we post-processed the images to detect and estimate the variables. All annotations in the panel contain two red dots. So, we used a Hough Circles Transform algorithm from OpenCV (Bradski, 2000) to calculate pixel positions of circle centers in the images. We created Ground Truth images for each task and level, pushed them on the detector script, and used the output to calculate the positioning error in all user trials.

From the Peg Transfer Task pictures, we evaluate the error of positioning the circle's center at the top of the indicated pin and the difference between the measured radius with the one that comprehends the full length of the peg on which the ring is placed. The measurement was automatized by image processing and compared with a ground truth image, as mentioned above. We decomposed the distance measured into vertical and horizontal components for individual analysis. We also checked the conformity of the solution with the instructions passed.

In the Thread Passing Task, we extracted the position of each arrow's starting and ending dots from the images. These are compared with the polygons corresponding to the portals to detect if the dots are inside the portal where the thread passes. Being inside the correct portal means the solution conforms with the instructions. However, in this task, there was no exact point where the user should set the starting and ending dots of the arrows, in such a way that the accuracy computation is more tricky. Moreover, it was impossible to place a dot on top of another as the interface would think that the user's intention is to edit that point instead. Thus, we compute the distance between the incoming and outgoing dots and subtract the minimum distance possible, which is equal to two radii.

Figure 5.3 – Dots and Circles Detection



Source: The Author

Figure 5.3 shows the detection of dots and circles. In both tasks, dots are detected, and the algorithm highlights the center as green in the picture. Besides, the circumference of circles is highlighted as red, which is used to calculate the radius and diameter. In addition to the time measurement, the extracted data from images allow the improve the comparisons and analysis with other extracted data.

5.1.3.2 Duration times

The system logs the total time of the trial, including the annotation in the video panel, navigation, and image referencing times. For a more detailed analysis, we dissect time data into finer parts, separating the annotation duration from the duration of the other actions.

This detachment will permit us to compare the performance per trial and analyze the performance evolution along the sequence of trials. The challenge is how to differentiate when the user annotates or checks the images and other information. To do this, we adapted principles of the Fitts law (MacKenzie, 1992; Accot and Zhai, 1997). The Fitts law defines the time to point a location as a function of the target distance and size. It is commonly represented by the equation 5.1, where a is the reaction time constant, b is the performance index, D is the distance, and S is the size of the target. Our experiment tasks involve pointing locations to place annotations, which is modeled by Fitts law, but it also includes other actions that are not target-dependent. Thus, we used the equation 5.2 to understand and analyze our time data.

$$t = a + b * \log_2 \frac{D}{S+1} \tag{5.1}$$

$$t(H) = a + (H * x) \tag{5.2}$$

The total time t depends on the hardness H (1, 2, or 3) and includes a, the time spent browsing images, navigating the interface, and making decisions, and x, which is the time to place an annotation marker in the task panel. To determine x, we have to estimate a, which is constant and represents the start/stop time of the task, i.e., the time that does not depend on the pointing of targets and is applicable for all trials of all users.

We estimate a by checking the mean differences in total duration between task levels $H = [1, 2, 3]$. As H is the number of annotations in the task, we have:

$$x \approx \overline{t(3)} - \overline{t(2)} \approx \overline{t(2)} - \overline{t(1)} \tag{5.3}$$

$$a \approx \overline{t(1)} - x \approx \overline{t(2)} - (2 * x) \approx \overline{t(3)} - (3 * x) \tag{5.4}$$

where $\overline{t(H)}$ is the mean time for H task hardness; x represents the mean of time spent performing the annotations; a is the constant that represents the time spent in other tasks than performing annotations.

5.1.3.3 *Subjective self-evaluation*

To measure the user experience, we surveyed the participants with a questionnaire about the system's use and perception, including the raw NASA TLX (Hart, 2006) to measure the workload and the effort. We asked about the perception of the reticle accuracy and responsiveness, the general feeling of the operation of the devices, and the interface limitations. Furthermore, we asked about the perceived level of accomplishment of the tasks and the confidence in using the head-gaze to draw circles or arrows in the task panel. The participants were also free to leave comments and suggestions about the system and interface interaction.

5.2 Results

The experiments were conducted in a controlled interaction laboratory. The system in this phase runs in the Unity Editor, with the Hololens Gen-1, through a wireless connection. The computer that runs the system is a Dell XPS 8930 with i7-8700, 16GB RAM, and a Geforce GTX 1070. The hand interactions were performed with the ring in the user's right-hand index finger, and the participants were guided to hold the laparoscopy instruments in a white box during the test.

A total of thirty-two (32) volunteers participated in the experiment, 16 for Thread Passing Task and 16 for Peg Transfer Task. Most are students or professors of our institute's computing courses (93.8%), and ages ranged from 21 to 57 years ($M \approx 25.6$, $SD \approx 7.4$). The 75% of the users reported never having experienced augmented reality through a head-mounted display, and 21.9% used it less than five times prior to the experiment. Only 3% of the participants reported having experienced AR with a headset more than five times. When asked about prior experience with virtual reality head-mounted displays, 87.6% of the users utilized it less than five times (43.8%) or never experienced it (43.8%). All the following statistical analyses and graphs were made using the R[3] language and the Tidyverse[4] packages: ggplot2, dplyr, and tidyr.

[3]https://www.r-project.org/
[4]https://www.tidyverse.org/packages/

5.2.1 Outliers

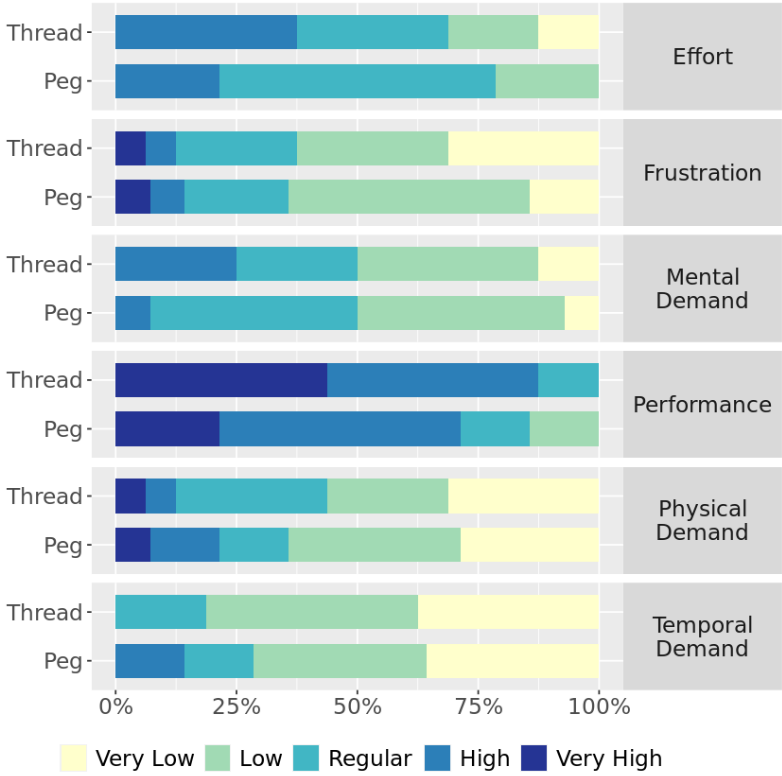
Starting with the raw data, we first identified as outliers the users whose performance skewed significantly from the norm and consequently are outside our scope of target users. A correlation between time measurement and the precision data of annotation was evidenced by the contrast between these users compared with others.

We removed $User1$ data due to the distinct behavior from all other users, both in terms of time (All Users: $MD = 39.48$, User 1 $MD = 67.61$) and the trials' accuracy in the center positioning (All Users: $MD = 4.66$, User 1 $MD = 23.87$) and diameter size of the circles (All Users: $MD = 12.97$, User 1 $MD = 25.86$). Thus, we consider that was a misunderstanding in the test protocol. Looking at the trials solution, we identify that $User3$ executed 7 of 15 trials incorrectly and demonstrated inconsistent performance in the time data (All Users: $MD = 39.48$, User 3 $MD = 70.8$) beyond the accuracy in the correct trials annotation data, differing from all other users in the circle center positioning (All Users: $MD = 4.66$, User 3 $MD = 9.51$) and diameter size(All Users: $MD = 12.97$, User 3 $MD = 30.63$).

5.2.2 Task Load

We applied the NASA TLX questionnaire to measure workload perception of the tasks. Fig. 5.4 illustrates the TLX scores by subscale. The similarity between the two tasks is notable in almost all evaluations. No significant difference between tasks was found when applying Mann-Whitney. Besides, low charges of mental, physical, and temporal demand reflect less frustration and a high perception rating of performance. However, the users consider that the effort to perform the tasks is regular to high, which can be associated with the number of trials and the repetitive tasks.

Figure 5.4 – Nasa Task Load Index for each Task



Source: The Author

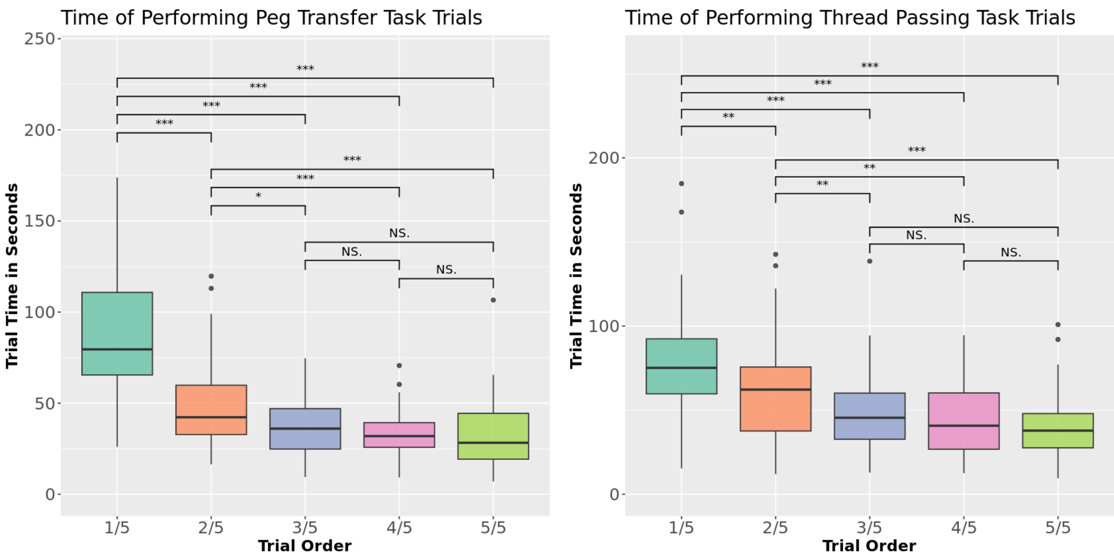
5.2.3 Learning effect

Due to the low previous experience of the participants with the techniques used in our interface, a steep learning curve is expected, which would affect the analysis. We intend to focus the analysis on the performance after initial learning. To determine when the learning phase ends, we divided the 15 sequential trials into five equivalent groups and compared these groups with each other.

Shapiro-Wilk did not allow us to assume the normality of the distribution for either of the tasks. Thus, we performed a Mann-Whitney U test comparing the ten pairs of groups, as illustrated in Fig. 5.5. Notice that the first two groups are significantly different between them and the other three groups (***: $p < 0.001$, **: $p < 0.01$, *: $p < 0.05$), while the three remaining groups do not show a significant difference among them (NS: $p > 0.05$). We conclude that most of the learning occurs in the initial six trials, after which the users are familiarized with the interface and interaction controls.

This allows us to discard the first six trials and keep the remaining nine for the remainder of the analysis. However, we noticed signs of fatigue among the users in the last trials due to the effort and demand required from participants. Thus, we have also discarded the last trial from further analyses. Consequently, after learning and adapting, we kept the stable data of the eight middle trials (trials 7 to 14, out of 15) in the following analyses.

Figure 5.5 – Time Groups. Left: Peg Transfer. Right: Thread Passing.



Source: The Author

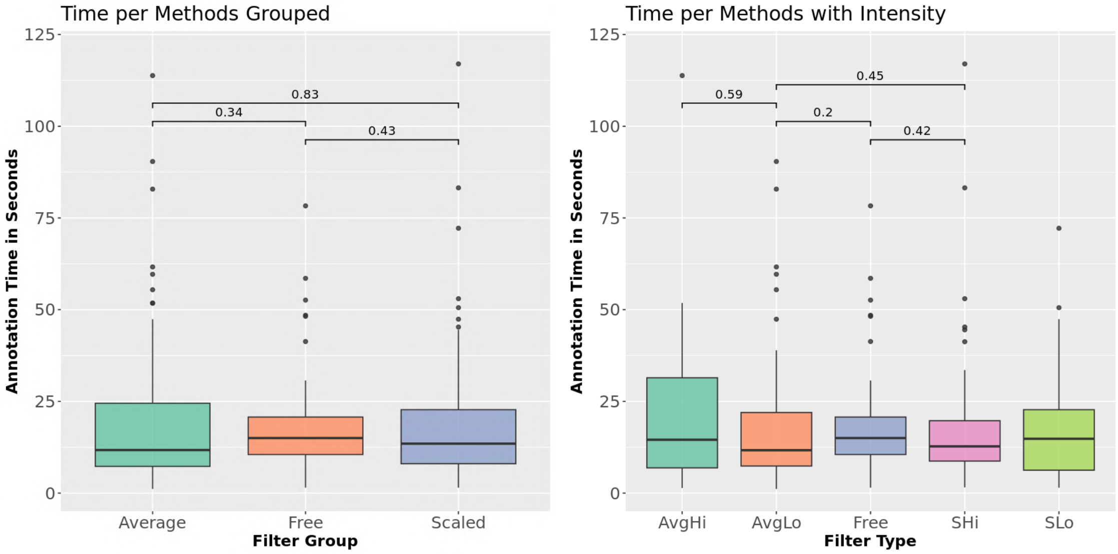
5.2.4 Time-performance per head-gaze condition

Having isolated the time spent per annotation, as described in Sec. 5.1.3.2, we can conduct comparisons of the impact between different head-gaze conditions on isolated annotations, regardless of the task level. The result of a Shapiro-Wilk normality test on these data did not allow us to assume a normal distribution. Thus, we applied a Mann-Whitney U test. For the Peg Transfer task, the difference between filter groups is not significant ($p > 0.05$), and the median times are very close, as shown in the Fig. 5.6 left, even though the Free method showed a higher median time than the Average filter (Average: $MD = 11.43, SD = 21.55$. Free: $MD = 13.78, SD = 17.45$). The same occurs when we compare the filter types, with a slight variation among conditions. The tendency on the medians mainly occurs between Free ($MD = 13.78, SD = 17.45$) and Average-Low ($MD = 10.4, SD = 22.55.$), as is shown in Fig. 5.6 right.

In the Thread Passing, Mann-Whitney did not find significant differences as well. Fig. 5.7 illustrates the weak effect of head-gaze conditions on time, where

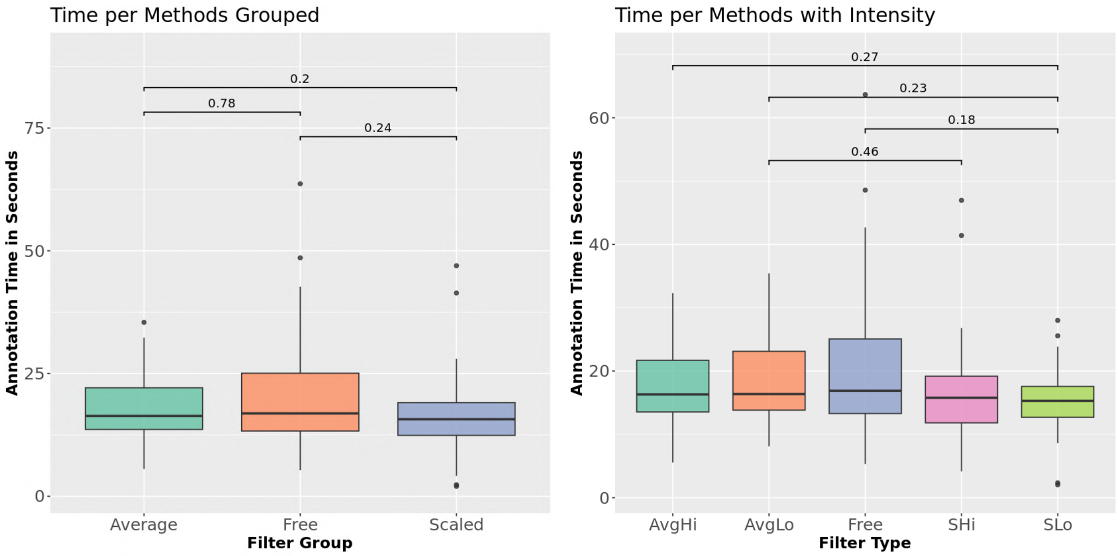
Scaled vs. Free and Scaled vs. Average show slight non-significant differences on the medians favoring Scaled. In Fig. 5.7 right, we see that Scaled-low tends to be slightly faster than the other filters, but no significant difference could be demonstrated.

Figure 5.6 – Head-Gaze Methods Peg Transfer Time Variation.



Source: The Author

Figure 5.7 – Head-Gaze Methods Thread Passing Time Variation.



Source: The Author

5.2.5 General annotation precision

To evaluate the precision and accuracy of the annotation tasks, we analyzed the annotation positions and radii extracted from the saved images. Some of the Peg Transfer images were not correctly saved due to cached data and could not be used. In total, 33 ($\approx 7.33\%$ of all 450 considered images and $\approx 15.71\%$ of peg considered images) saved pictures were corrupted and not included in any analyzes.

Considering all the valid data extracted from images, we analyze three measures for Peg Transfer task:

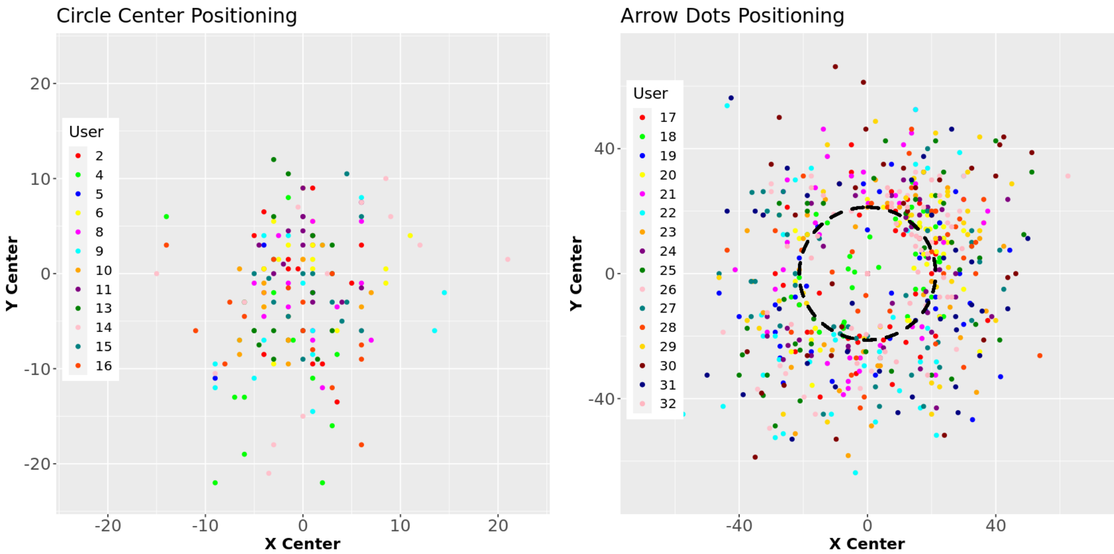
Correctness/accuracy: the correct peg is annotated

Center precision: distance of the circle center to the center of the peg's top face

Radius precision: difference between the expected radius to the measured radius, which should comprehend the height of the peg.

For the Thread task, we analyze the precision of positioning of the starting and ending points of each arrow in the board, as well as correctness/accuracy according to the passed instructions and guiding images.

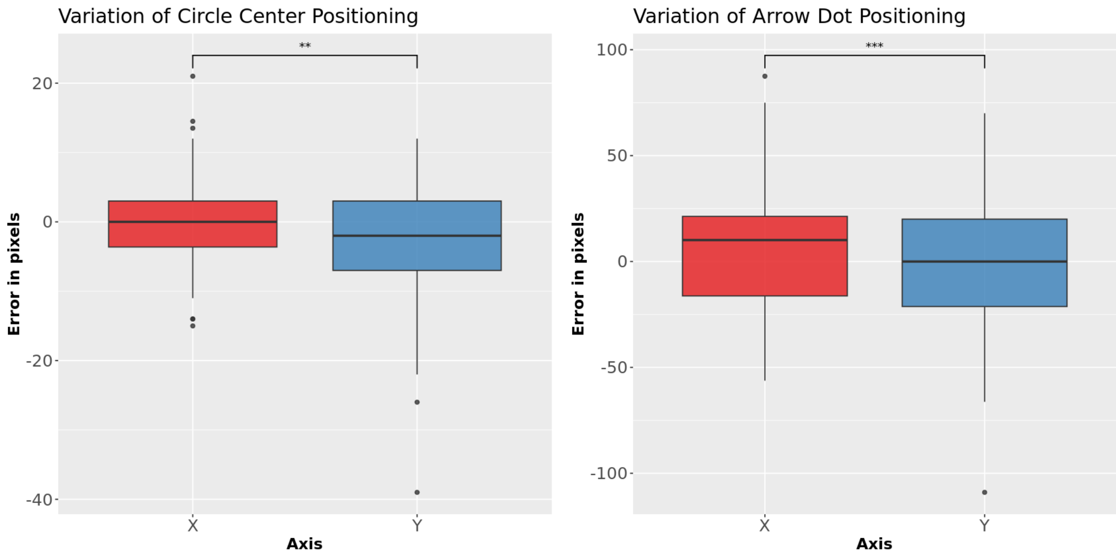
Figure 5.8 – Circle center and arrow end positioning errors per axis. Left: circle positions are more scattered vertically. Right: arrow end positions are distributed around 2 radii from the center with a larger vertical spread than horizontal.



Source: The Author

Firstly, we identified the trials representing incorrect solutions. In all the experiments, 15 of 417 ($\approx 3.59\%$) valid trials were incorrect, of which seven are from the 177 trials of the Peg Transfer task ($\approx 3.95\%$), and eight are from the 240 trials of the Thread Passing task($\approx 3.33\%$). The low error rate shows that the interface is effective for placing annotations and that the participants understood the task.

Figure 5.9 – Circle center and arrow end positioning errors.



Source: The Author

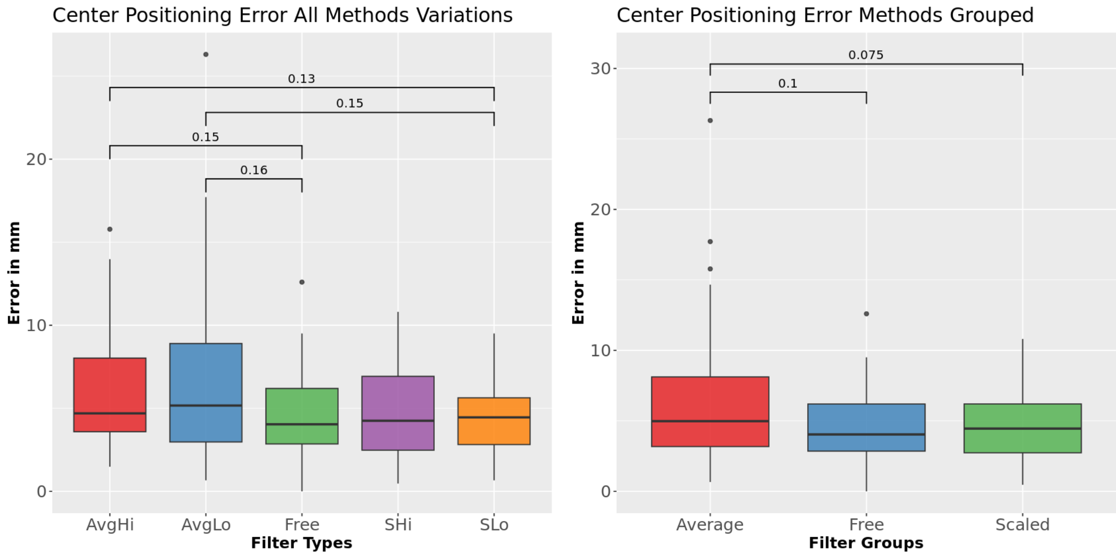
Surprisingly, we observe that the errors in the circle center position in the Peg Transfer task are more notable in the vertical axis than the horizontal axis, with a tendency to be lower. This is visible in the plot of Fig. 5.8-left. To further explore this observation, after Shapiro-Wilk, we use a Mann-Whitney test that detected a statistically significant difference ($p = 0.0024$) of the error between the two axes, as depicted in Fig. 5.9-left.

Doing the same analysis for the Thread Passing task, the result is similar to the Peg task. Looking at Fig. 5.8-right, the dot plot is not very clear due to the range of dots being positioned around a circumference. This is an effect of the limitation of the interface in allowing starting and ending arrow tips to be superimposed. However, when isolating the vertical error and the horizontal error, a Mann-Whitney test, detected a significant difference ($p = 0.0004$) between the errors in each axis for this task.

5.2.6 Annotation precision per head-gaze condition

Here we analyze how the circle center and radius, and the arrow ends, are affected by different head-gaze mappings. An offset of 21 pixels is used when comparing incoming and outgoing arrow ends not to penalize the error imposed by the interface limitation already mentioned. Comparisons of the head-gaze filters, their combinations, and the task levels are proposed. Shapiro-Wilk demonstrated that these data do not follow a normal distribution.

Figure 5.10 – Center Positioning Error Filters.



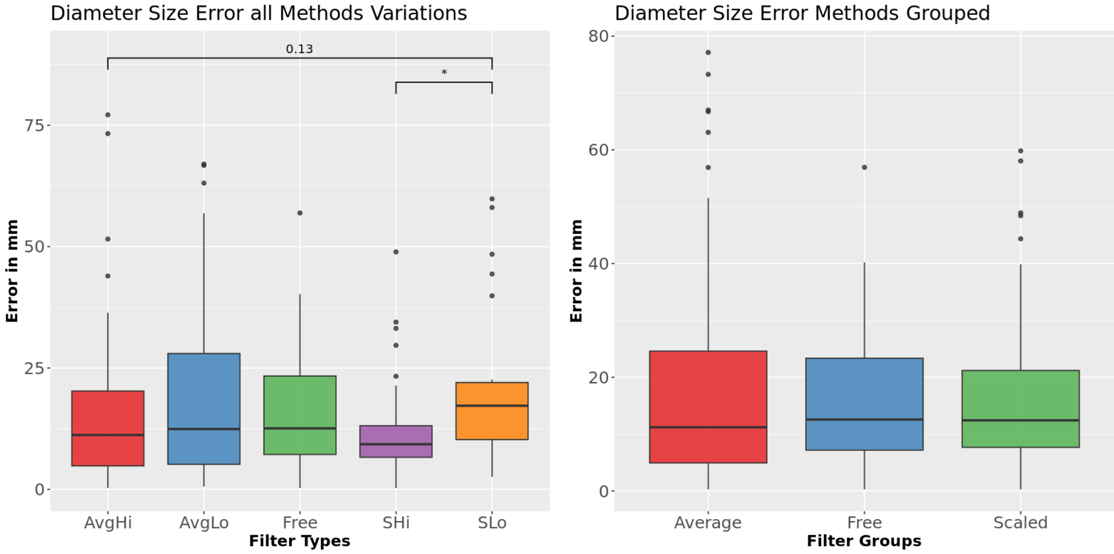
Source: The Author

The circle center errors were grouped by filter type (Avg, S) and filter intensity (Hi, Lo). A Mann-Whitney U test showed that there was no statistical significance between any filter type or group, although we observed a trend favoring Scaled-Low (See Fig. 5.11-left).

Although it is possible to visualize in the figure 5.10 left, the groups' variation in the millimeters errors in center positioning. When we compare the groups only by the method, disregarding the intensity, also there's no statistical significance, despite the highest variation of Average methods and the near 0.05 p-value of comparison with the Scaled methods, as shown in figure 5.10 right.

For the comparison of circle radii, the general picture is the same, except for the comparison between the two intensities of the Scaled methods, where the Low mode demonstrated a significant lower error in comparison with the High mode ($p - value < 0.05$)) after Mann-Whitney. The figure 5.11-right shows

Figure 5.11 – Center position and circle size errors per head-gaze filter.



Source: The Author

the variations and this contrast. However, when the groups are defined only by the method, the Scaled data is aggregated, and there's no significant difference between any method.

In the Thread Passing task, we did not find any statistical difference between the filter types and intensities. Nevertheless, some trends have been observed, as significance almost reached the threshold in some cases, especially between AvgLo ($MD = 5.84$) and Free ($MD = 6.47$) with $p = 0.059$ and AvgLo ($MD = 5.84$) and AvgHi ($MD = 7.49$) with $p = 0.06$.

Figure 5.12 – Dot Positioning Error Filters.

Source: The Author

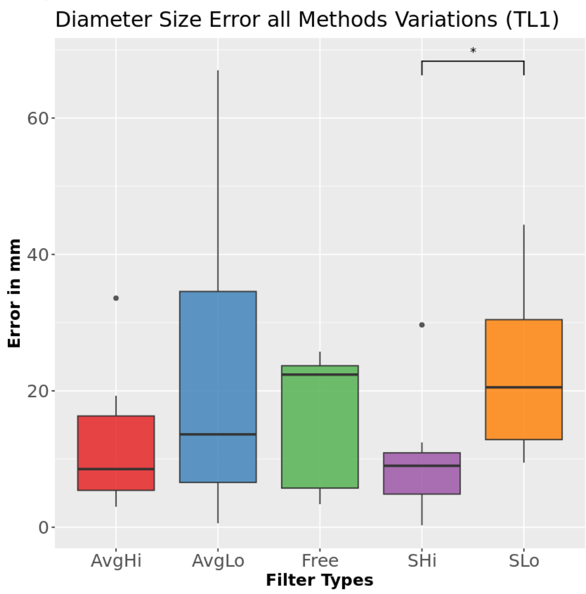
The direct mapping of head-gaze to interact and annotate the dots of arrows presents a higher and more variable error than the average low and high, as shown in the figure 5.12 left, which can indicate that the average method helps the users in annotation precision in the task panel. The free method errors also variables more than the Scaled with low intensity. In the figure 5.12 right, when we consider only the methods grouping the intensities, it is possible to see that the Free method still presents more errors than the other two. However, it's not demonstrated any statistical difference between the groups.

5.2.7 Annotation accuracy per task Level

Arguably, the annotation precision may be affected by the number of consecutive annotations (task level) in a single trial. Thus, we grouped the individual annotations by the task level of the trial each annotation belongs to. Three groups are formed with various numbers of elements as there are three times more annotations that belong to level three tasks than level one. However, Mann-Whitney did not find significant differences between groups for circle center positioning (Fig. 5.14-left) and circle radius size, indicating the effect of task level on errors is negligible.

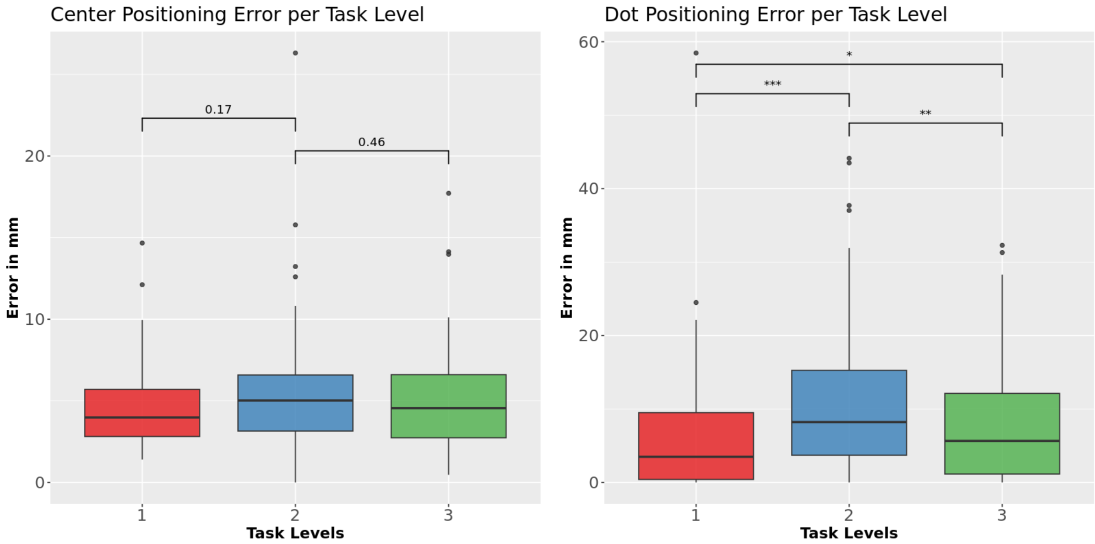
Differently, for the Thread task, Mann-Whitney found the levels significantly affect the error, as shown in Fig.5.14-right.

Figure 5.13 – Diameter error TL1 comparison



Source: The Author

Figure 5.14 – Task Levels Performance Comparison



Source: The Author

Furthermore, we also analyze the precision of each head-gaze filter at each task level individually. Isolating the task levels, we found significant differences. Starting with the trials with $task-level = 1$ for the Peg task, we obtain a significant difference in radius error between the two intensities of the Scaled method. The error is smaller in the High mode than in the Low mode, as is shown in Fig. 5.13. This confirms what is shown in Fig. 5.11-left for all data. For the circle center positioning error, no significance was found for $task - level = 1$. The same occurs for $task - level = 2$, where we could not find any significance both for circle positioning data and radius errors.

Nevertheless, for $task - level = 3$, we ran Mann-Whitney again, and a significant difference was found between the Average and the Scaled methods in the circle center position error, more specifically between the Average-High and the Scaled-Low conditions. The Average method, especially in the high mode, presented a higher error than the Scaled method, especially in the low mode, which is visible in Fig. 5.15.

For the Thread Passing task, we also ran the Mann-Whitney test with the arrow end positioning error, and we found significance in just one comparison. With $task - level = 3$, the Average Low method showed lower error than the Free head-gaze method, as depicted in Fig. 5.16-left. Fig. 5.16-right shows that some results almost showed significance, but this was not enough to be conclusive about filter conditions in $task - level = 2$.

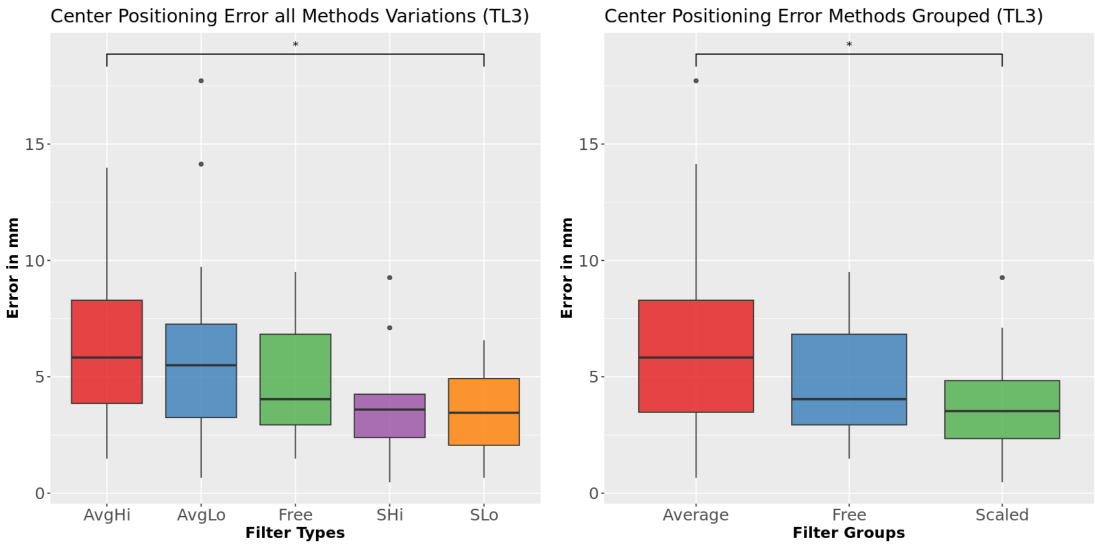


Figure 5.15 – Center positioning error TL3 comparison

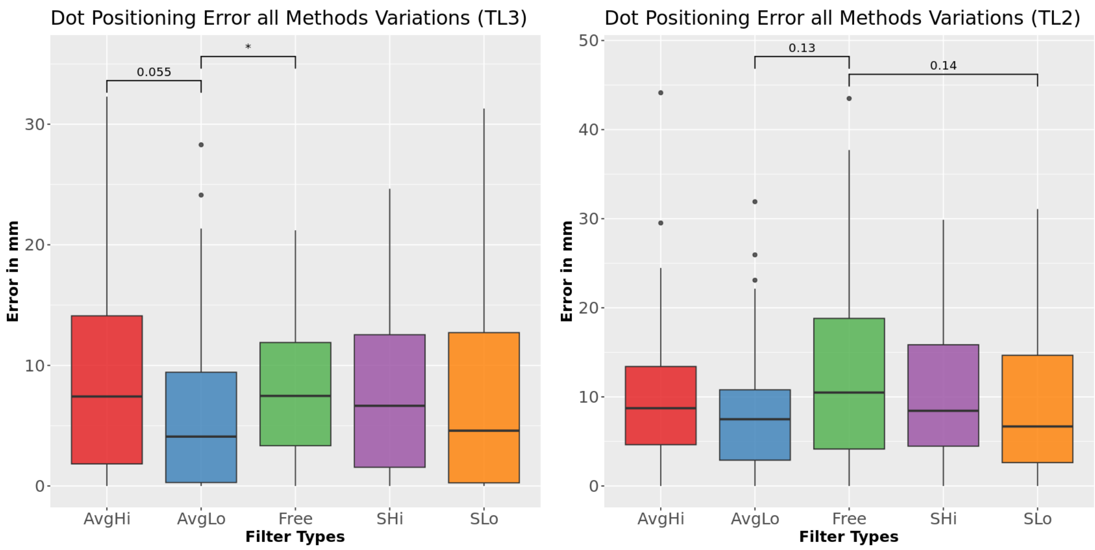


Figure 5.16 – Dot positioning error TL3 and TL2 comparison

5.2.8 Qualitative Data

The users answered a post-test questionnaire, and we evaluated some topics to understand ergonomics and interaction methods. We asked first about the perception of the head-gaze applied methods and if they feel any impact in the reticle delay, trepidation, and difficulty in solving the task trials, besides the inconvenience of the button hardware utilization and the fine interaction with the head-gaze. The users, in general, demonstrated that they perceived the impact of the methods, in which ten related that they felt a delay in the reticle at least some times in the trials, and 17 felt it in the majority of trials. When we asked about the trepidation, 19 users felt that the reticle was trembling in the majority of trials and seven at least some time in trials. In general, the users felt confident in solving the problems and did not sense difficulty in performing annotations in the task panel.

In our questionnaire, we also asked about the hardware used to interact with the interface, the Hololens Gen-1, and the Ring device with buttons in hand. About the ring, most users revealed that it was not cumbersome to use it to confirm interactions, and only two reported that it was moderately uncomfortable. However, the answers about the interactions using the head-gaze direction with Hololens Gen-1 are very distributed. Only two users reported as very uncomfortable interacting with head-gaze using Hololens; eight reported as regularly uncomfortable, and the other 20 were divided into not discomforting and a little uncomfortable.

5.3 Discussion

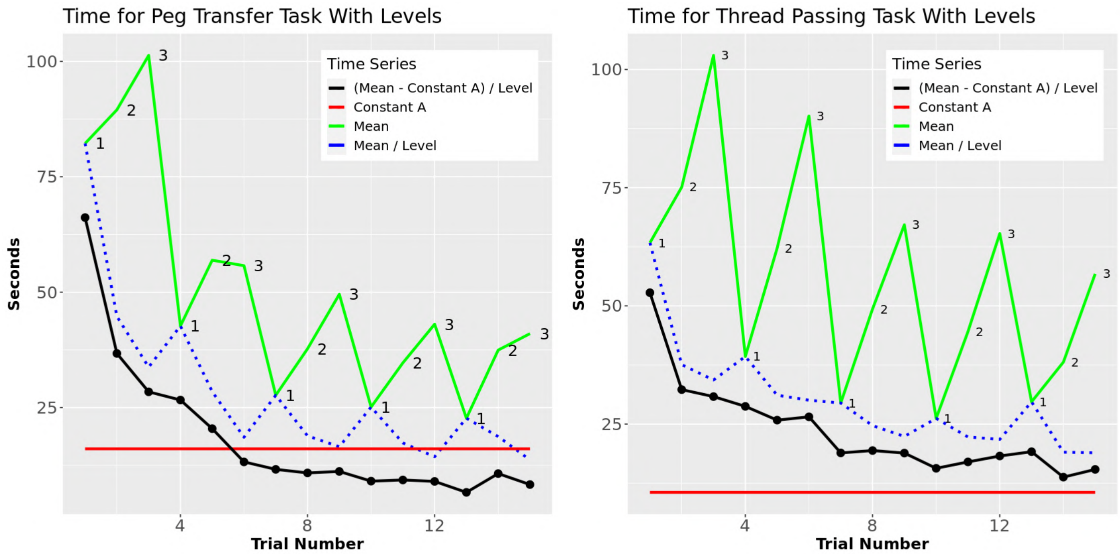
In this section, we discuss the potential reasons for the results found and general participant feedback, besides the discussion about the main findings and proposed methods. The main areas to focus on are the users' performance and experience, the limitations of the application, and the impressions of the utilization of this type of application in the laparoscopy surgery field.

5.3.1 Time and learning

In Fig. 5.17, we display time data and how it tends to decrease and stabilize through the execution of the trials. The green line is the raw mean for all users for each trial in the sequence. Ups and downs reflect the task levels in the respective trial, which repeats the pattern 1, 2, 3 five times, indicating how duration is impacted by the number of annotations in the trial. Notice that level two is not twice the time as level one, and level three is not thrice that time, especially in the Peg Transfer, indicating the presence of a constant time not dependent on the number of annotations. Indeed, while dividing each duration by the level (number of annotations) could be expected to smoothen the curve, the blue dotted line shows that this is not the case and that some residual time is not level-dependent. These observations triggered the need to estimate the constant time a (see Sec.5.1.3.2), here represented by the red line. When discounting a and level from the trial's time, we obtain the black line that represents the time per annotation used in our analysis.

Although the start-stop time of the task, represented by a, differs between the tasks, it is visible in the plot of Fig. 5.17 that the participants spent more time in the first six trials, where they are becoming accustomed to the interface and the interaction methods. This difference between tasks indicates that the time spent in image reading and navigation is task-dependent. As of the seventh trial, the mean trials' time stabilizes, which permits more analysis of the user's performance and utilization. The rule applies similarly for the time when we analyze it by task hardness, and the black dotted line in the figure 5.17 graphics shows that the times just for annotating in the trials are equivalent between the three levels in both peg transfer and thread passing tasks.

Figure 5.17 – Left: Peg Transfer Trials. Right: Thread Passing trials.



Source: The Author

5.3.2 Interaction performance

The impact of the change of focus between panels on usability was one of our research questions. It can be understood by means of the correctness obtained. The majority of users performed all the trials correctly, following the specifications and particularities of each tasks levels and its images, which demanded attention and concentration between panels.

The high number of correctly completed trials also demonstrates that the use of the ring device with buttons on the hand, in addition to the head-gaze, was effective in accessing the folders and images necessary, and effective in making annotations on the task panel. This occurred despite the hardware limitations and the arguably tiresome standing position with the hands on the laparoscopic instruments during the whole experiment.

5.3.3 Head-gaze performance

When we look at the head-gaze filtering methods, we see some differences among conditions in terms of annotation precision. We expected that the Average and Scaled methods, in both modes, would improve the head-gaze interaction with the user, making it less complicated to use the reticle and avoiding trembling and errors by fast movements in the direct mapping. This expectation was not confirmed as there is no clear advantage of filtering over Free in most cases. Only specific comparisons show significant differences as demonstrated. In the Thread task, the Average-low method performed better than the Scaled and Free. It was shown that, in this task, the average methods grouped also performed better.

However, when we look for the Peg task results, even if the difference was not significant, it is shown that the error trend is inverse to the Thread task. The Scaled method performed better than the Average and Free in both circle center positioning and radius. The nature of the annotation tasks can explain this difference between the Peg and Thread tasks. When the Peg task requires positioning and size of specific positions and sizes, the Thread task requires continuity in the thread path and in the point that the thread must pass, which could explain why a method is better than one for a specific task.

The error difference between the X-axis and the Y-axis is a major finding in the analysis of the head-gaze data. In both tasks, the data spread more along the Y-axis than along X. One possible cause is the weight of the Hololens gen-1, which can make it more difficult to hold a head-gaze position vertically. Another possibility is the placement of the panels. The task panel is always in front, and the image panel is always below, which obliges the user to vary more on the vertical axis to access the images during the solution of the task trial.

When grouping the data by task level, we observe that level does not impact the positioning error for the Peg task but does impact the Thread task. The number of annotations made by the user in each level is different in both tasks, as indicated in Table 4.1. This could be the explanation if it were not for the most significant error presented being in level 2 of the tasks and not in level 3. Therefore, we assume that it is a combination of the number of annotations with the arrangement of the images, and the path that must be followed, since in the Peg task the user does not need to care about direction and continuity, as in the Thread task.

5.3.4 User impressions

From the performance data, we notice that the users successfully interacted, understood, and solved the task trials. Accessing guide images, instruction texts, and interface controls was not problematic when using them. However, when we look at the interaction using head-gaze, we notice that the applied methods may not be balanced in the best way since people noticed the delay, the trembling, and the difficulty of reaching points in more attempts than expected.

The ring device positioned in the hand, with the buttons for interaction, was well applied and accepted by the users since most of them did not feel uncomfortable using it, even when standing up and with their hands on the instruments. Therefore, the combination of head-gaze with the use of hands for confirmation interactions can be recommended, as it is accurate in the selection of points of interest necessary for quick interactions, such as an undo button (suggestion given by several users during tests).

5.3.5 Limitations

Although the reticle head-gaze methods did not improve performance as we expected, the methods showed that they can help. The calibration and application of these methods can be improved to enhance this potential. Therefore, adjustments and different levels for the modes of methods, or even a combination of the methods, can lead to better results. The range of parameters was also limited by the Hololens specification, such as the FoV, and the impossibility of testing methods using eye-tracking to interact and perform annotation in the laparoscopy feed.

Another limitation is associated with the task measurements. We had to develop an image detection algorithm to solve a feature that could be implemented directly in the application, avoiding a possible corruption of saved images, as there was in the Peg task data.

5.4 Future Work

The current user experiment is limited to the interactions in augmented reality, focusing on stabilizing head-gaze pointing for annotations and utilizing a new input hand device under the sterilized gloves, the ring device. However, for the next steps, we have planned another user experiment, with the surgical residents as target users, who are performing laparoscopic tasks and interacting with the interface at the same time. The main objective is to consolidate the improvement of communication and learning using the AR interface and the head-gaze as a pointing interaction method, besides comparing the foot device in a simulated real environment with the ring device. Among the aspects we wish to evaluate, there are:

- The workload of interacting while performing the task;
- The usability of consulting planned data to solve the task;
- The effects of interacting with the foot or the hand;
- The improvement of communication using the interface during the task.

For this next experiment, we plan to use the developed laparoscopic physical training tasks with a white box, simulating all the learning environments to evaluate the interaction techniques and the aspects listed above.

6 GENERAL CONCLUSIONS

In this work, we presented an overview of the literature and related work that explores the principles of remote collaboration, AR, VR, and MR for health environments, and specifically, both topics applied to laparoscopic surgeries. We described in Chapter 3 the primary perspective for planning and development of the MR application, its modes of utilization for laparoscopic surgeries, and the challenges that we faced and turned into research questions.

We started developing the AR interface focus in the INTRA Scenario, where the surgeon has to wear the AR HMD and can interact with the interface during the procedure, drawing annotations, saving essential frames of the surgery, and accessing preoperative exams and planning information. Although the common interaction method in related works was voice and hand gestures, in our scenario, we aimed to reduce interruptions and communication confusion during the procedure. Thus, through a visual ray-cast cursor, the head-gaze method came up as a solution to pointing and selecting specific positions and actions in our AR interface. However, the trembling and inaccurate head-gaze direction appeared as the main challenge of this work, besides the narrow FOV from Hololens Gen-1.

In Chapter 4, we presented all the details of our AR application, the hardware design where we describe the hand-device and foot pedals planned and developed. The interface comprises panels with the necessary information, and the interaction design follows high-level user actions, such as observing the laparoscopic video feed, annotating the video, and saving a frame from the procedure. We present our head-gaze pointing techniques to reduce the trembling and optimize the FOV: The Scale Method, scaling down the head-reticle movement ratio, which is 1 when not scaled. The Average Filtering Method stores the n previous positions of the raycasted position and calculates a simple average to a filtered new position. After the proposal and development, we planned a user experiment to evaluate the techniques and the interface.

The user experiment in Chapter 5 focused on evaluating hand and head-gaze affordances using augmented reality for annotation during laparoscopy surgeries and similar settings. The main study presented aimed at understanding the impact and improvements that head-gaze average filtering and the scale method can bring to perform annotations in the laparoscopy video feed in a virtual monitor positioned straight in front of the user. The user experiment was performed in a between-subject protocol with 32 volunteers recruited in the Institute of Informatics. We have achieved a series of statistical analyses, evaluating different aspects of the utilization, such as time, task completeness accuracy, and the focus of our development: the impact of utilization of the head-gaze direction for interaction and the head-gaze filtering methods proposed in this work.

During the analysis, we started processing the time data and selecting the centered trials from the users, disregarding the initial six trials and the last one, where we considered that the users were in the adaptation phase or exhausted because of the time spent during the test, we also have presented visualizations of the learning curve and the statistical difference between the trials. Looking for the annotation precision data, we have found that the task misunderstanding, represented by the incorrect solutions of the trials, was only 3.59% in general, disregarding the corrupted and outliers data. Our main finding is related to the precision axis of positioning annotations, where the Y-axis was more sensible to the head-gaze and showed more position error than the X-axis.

The collected data on annotation error did not allow us to find statistically significant differences among the head-gaze methods tested, even when we grouped by technique and disregarded the condition level. However, when we isolate them by task level, some variables have shown statistically significant lower error rates using the filter methods than the direct input from head-gaze. Examples are the dot positioning error from the Thread Passing Task and the diameter size error with the easy task levels.

We also have presented in results and discussion a general view of qualitative data and the user experience during the tasks, including a task load measurement, where we found that the users are confident about their performance and reported a low level of physical demand, besides a medium to high effort to perform the tasks. The most important question about using the ring device was that the users revealed that it was not cumbersome to use to confirm interactions. And the head-gaze experience with HL1 is variable among users.

Concluding the user experiment chapter, we presented the limitations of the user study and the interface, and a brief explanation about the next experiment to evaluate the utilization of the interface by the surgeons while they are performing physical tasks.

6.1 Remarks

The approach presented demonstrated that the interface can be helpful and applicable, and the ring device proposed can also be included as an interaction tool by the surgeons. It also addressed the application of such an interface to streamline surgeons' access to instructions, vitals, planning data, and preoperative exams in an augmented virtual monitor, which can help during training and actual procedures with remote assistance. While the interaction head-gaze filtering methods demonstrated significant differences only in some instances, the overall precision obtained is suitable for their utilization in accordance with the latency and framerate of the device.

6.2 Limitations and Future Work

In our research, although we had planned the VR Interface and the possible communication between the interfaces, we still needed to develop the initially proposed remote VR interface, limiting the user study and evaluating variables only to the local AR Interface and its interactions. We faced a couple of challenges in the development using the HoloLens Gen-1, such as the software and hardware compatibility.

We only partially achieved some of our user experiment goals. However, considering the valuable utilization case of the interface, we plan a second experiment with improvements in the interface and interaction, switching the AR HMD from HoloLens Gen-1 to a video see-through XR HMD, which improves the development and compatibility with the devices, although might present new challenges of implementation, such as integration with new devices and communication between users. We briefly describe the possible future user experiment in Chapter 5, where we aim to test with surgery students and compare our developed ring hand device with the foot pedals. We also plan to build and connect the VR interface to evaluate improving communication and remote collaboration in laparoscopic procedures.

REFERENCES

ACCOT, J.; ZHAI, S. Beyond fitts' law: models for trajectory-based hci tasks. In: **Proceedings of the ACM SIGCHI Conference on Human factors in computing systems**. [S.l.: s.n.], 1997. p. 295–302.

AGGARWAL, R.; MOORTHY, K.; DARZI, A. Laparoscopic skills training and assessment. **Journal of British Surgery**, Oxford University Press, v. 91, n. 12, p. 1549–1558, 2004.

ALIWI, I. et al. The role of immersive virtual reality and augmented reality in medical communication: A scoping review. **Journal of Patient Experience**, SAGE Publications Inc., v. 10, 1 2023. ISSN 23743743.

ANDERSEN, D. et al. Medical telementoring using an augmented reality transparent display. **Surgery (United States)**, Mosby Inc., v. 159, n. 6, p. 1646–1653, 6 2016. ISSN 15327361.

ASAO, T. et al. Development of a drawing application for communication support during endoscopic surgery and its evaluation using steering law. **Applied Sciences**, Multidisciplinary Digital Publishing Institute, v. 11, n. 10, p. 4505, 2021.

BANSAL, G. et al. Healthcare in metaverse: A survey on current metaverse applications in healthcare. **IEEE Access**, Institute of Electrical and Electronics Engineers Inc., v. 10, p. 119914–119946, 2022. ISSN 21693536.

BARSOM, E. Z.; GRAAFLAND, M.; SCHIJVEN, M. P. Systematic review on the effectiveness of augmented reality applications in medical training. **Surgical endoscopy**, Springer, v. 30, n. 10, p. 4174–4183, 2016.

BASDOGAN, C.; HO, C.-H.; SRINIVASAN, M. Virtual environments for medical training: graphical and haptic simulation of laparoscopic common bile duct exploration. **IEEE/ASME Transactions on Mechatronics**, v. 6, n. 3, p. 269–285, 2001.

BIANCHI, I.; ZANATTA, A. L.; RIEDER, R. Augmented reality in medical teaching-learning process content: A systematic review. In: **2020 22nd Symposium on Virtual and Augmented Reality (SVR)**. [S.l.: s.n.], 2020. p. 129–133.

BITTNER, R. Laparoscopic surgery—15 years after clinical introduction. **World journal of surgery**, Springer, v. 30, p. 1190–1203, 2006.

BORHANI, Z.; SHARMA, P.; ORTEGA, F. R. Survey of annotations in extended reality systems. **IEEE Transactions on Visualization and Computer Graphics**, IEEE Computer Society, 2023. ISSN 19410506.

BRADSKI, G. The OpenCV Library. **Dr. Dobb's Journal of Software Tools**, 2000.

CALDWELL, G. The process of clinical consultation is crucial to patient outcomes and safety: 10 quality indicators. **Clinical Medicine**, Royal College of Physicians, v. 19, n. 6, p. 503–506, 2019. ISSN 1470-2118. Available from Internet: <https://www.rcpjournals.org/content/19/6/503>.

CAR, L. T. et al. **Outcomes, Measurement Instruments, and Their Validity Evidence in Randomized Controlled Trials on Virtual, Augmented, and Mixed Reality in Undergraduate Medical Education: Systematic Mapping Review**. [S.l.]: JMIR Publications Inc., 2022.

CHANG, E.; LEE, Y.; YOO, B. A user study on the comparison of view interfaces for vr-ar communication in xr remote collaboration. **International Journal of Human-Computer Interaction**, Taylor and Francis Ltd., 2023. ISSN 15327590.

CHIDAMBARAM, S. et al. ProcessAR: An augmented reality-based tool to create in-situ procedural 2D/3D AR Instructions. In: **DIS 2021 - Proceedings of the 2021 ACM Designing Interactive Systems Conference: Nowhere and Everywhere**. [S.l.]: Association for Computing Machinery, Inc, 2021. p. 234–249. ISBN 9781450384766.

COFANO, F. et al. Augmented Reality in Medical Practice: From Spine Surgery to Remote Assistance. **Frontiers in Surgery**, Frontiers Media S.A., v. 8, 3 2021. ISSN 2296875X.

DEROSSIS, A. M. et al. The effect of practice on performance in a laparoscopic simulator. **Surgical Endoscopy**, v. 12, p. 1117–1120, 1998. ISSN 1117-1120.

DINH, A. et al. **Augmented Reality in Real-time Telemedicine and Telementoring: Scoping Review**. [S.l.]: JMIR Publications Inc., 2023.

FENG, Y. et al. Virtual pointer for gaze guidance in laparoscopic surgery. **Surgical endoscopy**, Springer, v. 34, n. 8, p. 3533–3539, 2020.

FIDALGO, C. G. et al. A survey on remote assistance and training in mixed reality environments. **IEEE Transactions on Visualization and Computer Graphics**, IEEE Computer Society, v. 29, p. 2291–2303, 5 2023. ISSN 19410506.

GARCIA, F. M. et al. Health-5g: A mixed reality-based system for remote medical assistance in emergency situations. **IEEE Access**, Institute of Electrical and Electronics Engineers Inc., v. 11, p. 59016–59032, 2023. ISSN 21693536.

GASQUES, D. et al. Artemis: A collaborative mixed-reality system for immersive surgical telementoring. In: **Conference on Human Factors in Computing Systems - Proceedings**. [S.l.]: Association for Computing Machinery, 2021. ISBN 9781450380966.

GERUP, J.; SOERENSEN, C.; DIECKMANN, P. Augmented reality and mixed reality for healthcare education beyond surgery: an integrative review. **Int J Med Educ**, v. 11, p. 1–18, 2020. Available from Internet: <http://www.ijme.net/archive/11/augmented-reality-for-healthcare-education/>.

GOULD, J. et al. Instruments and devices used in laparoscopic surgery. **Literature review current through: Jan**, 2019.

GRANDI, J. G.; DEBARBA, H. G.; MACIEL, A. Characterizing asymmetric collaborative interactions in virtual and augmented realities. In: IEEE. **2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)**. [S.l.], 2019. p. 127–135.

GRANTCHAROV, T. P. et al. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. **British Journal of Surgery**, v. 91, n. 2, p. 146–150, 2004. ISSN 00071323.

GRINSHPOON, A. et al. Hands-free interaction for augmented reality in vascular interventions. In: **2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)**. [S.l.: s.n.], 2018. p. 751–752.

HAFFORD, M. L. et al. Ensuring competency: are fundamentals of laparoscopic surgery training and certification necessary for practicing surgeons and operating room personnel? **Surgical Endoscopy**, Springer, v. 27, p. 118–126, 2013.

HALE, A. et al. Remote training for medical staff in low-resource environments using augmented reality. **Journal of Imaging**, MDPI, v. 8, 12 2022. ISSN 2313433X.

HART, S. G. Nasa-task load index (nasa-tlx); 20 years later. In: SAGE PUBLICATIONS SAGE CA: LOS ANGELES, CA. **Proceedings of the human factors and ergonomics society annual meeting**. [S.l.], 2006. v. 50, n. 9, p. 904–908.

HART, S. G.; STAVELAND, L. E. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In: **Advances in psychology**. [S.l.]: Elsevier, 1988. v. 52, p. 139–183.

HEINRICH, F. et al. Holopointer: a virtual augmented reality pointer for laparoscopic surgery training. **International journal of computer assisted radiology and surgery**, Springer, v. 16, n. 1, p. 161–168, 2021.

HUBER, T. et al. Highly immersive virtual reality laparoscopy simulation: development and future aspects. **International journal of computer assisted radiology and surgery**, Springer, v. 13, p. 281–290, 2018.

IZARD, S. G. et al. Virtual reality as an educational and training tool for medicine. **Journal of medical systems**, Springer, v. 42, p. 1–5, 2018.

JANABI, H. F. A. et al. Effectiveness of the hololens mixed-reality headset in minimally invasive surgery: a simulation-based feasibility study. **Surgical Endoscopy**, Springer, v. 34, n. 3, p. 1143–1149, 2020.

JAYENDER, J. et al. A novel mixed reality navigation system for laparoscopy surgery. In: SPRINGER. **International Conference on Medical Image Computing and Computer-Assisted Intervention**. [S.l.], 2018. p. 72–80.

JIMENEZ, F. A. Can Virtual Patient Simulation Be Used in Substitution of Traditional Clinical Hours in Undergraduate Nursing Education? A Review of the Evidence. **White Paper**, p. 15, 2022.

JING, A. et al. Comparing gaze-supported modalities with empathic mixed reality interfaces in remote collaboration. In: . [S.l.]: Institute of Electrical and Electronics Engineers Inc., 2022. p. 837–846. ISBN 9781665453257.

JING, A.; LEE, G.; BILLINGHURST, M. Using speech to visualise shared gaze cues in mr remote collaboration. In: . [S.l.]: Institute of Electrical and Electronics Engineers Inc., 2022. p. 250–259. ISBN 9781665496179.

JING, A. et al. The impact of sharing gaze behaviours in collaborative mixed reality. **Proceedings of the ACM on Human-Computer Interaction**, Association for Computing Machinery, v. 6, 11 2022. ISSN 25730142.

KIM, S. et al. Evaluating the Combination of Visual Communication Cues for HMD-based Mixed Reality Remote Collaboration. In: **Conference on Human Factors in Computing Systems - Proceedings**. [S.l.]: Association for Computing Machinery, 2019. ISBN 9781450359702.

KLEINHEKSEL, A. J. Transformative learning through virtual patient simulations: Predicting critical student reflections. **Clinical Simulation in Nursing**, Elsevier Inc, v. 10, n. 6, p. e301–e308, 2014. ISSN 18761399. Available from Internet: <http://dx.doi.org/10.1016/j.ecns.2014.02.001>.

KUMAR, R. P. et al. Use of mixed reality for surgery planning: Assessment and development workflow. **Journal of Biomedical Informatics**, Elsevier, v. 112, p. 100077, 2020.

KUMARAVEL, B. T. et al. Loki: Facilitating remote instruction of physical tasks using Bi-directional mixed-reality telepresence. In: **UIST 2019 - Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology**. [S.l.]: Association for Computing Machinery, Inc, 2019. p. 161–174. ISBN 9781450368162.

KYAW, B. M. et al. **Virtual reality for health professions education: Systematic review and meta-analysis by the digital health education collaboration**. [S.l.]: JMIR Publications Inc., 2019.

LAMPROPOULOS, G. et al. Augmented reality and gamification in education: A systematic literature review of research, applications, and empirical studies. **applied sciences**, MDPI, v. 12, n. 13, p. 6809, 2022.

LEE, G. A. et al. A User Study on MR Remote Collaboration Using Live 360 Video. In: **Proceedings of the 2018 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2018**. [S.l.]: Institute of Electrical and Electronics Engineers Inc., 2019. p. 153–164. ISBN 9781538674598.

LIN, B. et al. **Video-based 3D reconstruction, laparoscope localization and deformation recovery for abdominal minimally invasive surgery: a survey**. [S.l.]: John Wiley and Sons Ltd, 2016. 158–178 p.

LOESCHER, T.; LEE, S. Y.; WACHS, J. P. **An Augmented Reality Approach to Surgical Telementoring**. [S.l.], 2014.

LUNGU, A. J. et al. A review on the applications of virtual reality, augmented reality and mixed reality in surgical simulation: an extension to different kinds of surgery. **Expert Review of Medical Devices**, Taylor and Francis Ltd., v. 18, p. 47–62, 2021. ISSN 17452422.

MACKENZIE, I. S. Fitts' law as a research and design tool in human-computer interaction. **Human-computer interaction**, Taylor & Francis, v. 7, n. 1, p. 91–139, 1992.

MENTIS, H. M.; CHELLALI, A.; SCHWAITZBERG, S. Learning to see the body: supporting instructional practices in laparoscopic surgical procedures. In: **Proceedings of the SIGCHI conference on human factors in computing systems**. [S.l.: s.n.], 2014. p. 2113–2122.

MENTIS, H. M.; CHELLALI, A.; SCHWAITZBERG, S. Learning to see the body: Supporting instructional practices in laparoscopic surgical procedures. In: **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems**. New York, NY, USA: Association for Computing Machinery, 2014. (CHI '14), p. 2113–2122. ISBN 9781450324731. Available from Internet: <https://doi.org/10.1145/2556288.2557387>.

MILGRAM, P.; KISHINO, F. A taxonomy of mixed reality visual displays. **IEICE TRANSACTIONS on Information and Systems**, The Institute of Electronics, Information and Communication Engineers, v. 77, n. 12, p. 1321–1329, 1994.

MILGRAM, P. et al. Augmented reality: a class of displays on the reality-virtuality continuum. In: DAS, H. (Ed.). **Telemanipulator and Telepresence Technologies**. SPIE, 1995. v. 2351, p. 282 – 292. Available from Internet: <https://doi.org/10.1117/12.197321>.

MOJICA, C. M. M. A prototype holographic augmented reality interface for image-guided prostate cancer interventions. In: **Eurographics Workshop on Visual Computing for Biomedicine**. [S.l.: s.n.], 2018.

MORIMOTO, T. et al. **XR (Extended Reality: Virtual Reality, Augmented Reality, Mixed Reality) Technology in Spine Medicine: Status Quo and Quo Vadis**. [S.l.]: MDPI, 2022.

MUNZ, Y. et al. Laparoscopic virtual reality and box trainers: Is one superior to the other? **Surgical Endoscopy and Other Interventional Techniques**, v. 18, n. 3, p. 485–494, 3 2004. ISSN 09302794.

NEGRãO, M. D. et al. Design and think-aloud study of an immersive interface for training healthcare professionals in clinical skills. In: **Symposium on Virtual and Augmented Reality (SVR'23), November 6–9, 2023, Rio Grande, Brazil**. [S.l.: s.n.], 2023.

PATEL, D. B. et al. Computer-Supported Experiential Learning-Based Tool for Healthcare Skills. **IEEE Computer Graphics and Applications**, IEEE, v. 43, n. 2, p. 57–68, 2023. ISSN 15581756.

PIUMSOMBOON, T. et al. On the shoulder of the giant: A multi-scale mixed reality collaboration with 360 video sharing and tangible interaction. In: **Conference on Human Factors in Computing Systems - Proceedings**. [S.l.]: Association for Computing Machinery, 2019. ISBN 9781450359702.

QIAN, K. et al. Virtual reality based laparoscopic surgery simulation. In: **Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST**. [S.l.]: Association for Computing Machinery, 2015. v. 13-15-November-2015, p. 69–78. ISBN 9781450339902.

RAYA, L. et al. Development of a Virtual Reality Tool for the Treatment of Pediatric Patients in the ICU. **IEEE Computer Graphics and Applications**, v. 43, n. 2, p. 69–77, 2023. ISSN 15581756.

REZNEK, M.; HARTER, P.; KRUMMEL, T. Virtual reality and simulation: Training the future emergency physician. **Academic Emergency Medicine**, v. 9, n. 1, p. 78–87, 2002. ISSN 10696563.

RITTER, E. M.; SCOTT, D. J. Design of a proficiency-based skills training curriculum for the fundamentals of laparoscopic surgery. **Surgical innovation**, Sage Publications Sage CA: Los Angeles, CA, v. 14, n. 2, p. 107–112, 2007.

RODRIGUES, S. P. et al. Suturing intraabdominal organs: when do we cause tissue damage? **Surgical endoscopy**, Springer, v. 26, p. 1005–1009, 2012.

ROJAS-MUÑOZ, E. et al. The System for Telementoring with Augmented Reality (STAR): A head-mounted display to improve surgical coaching and confidence in remote areas. **Surgery (United States)**, Mosby Inc., v. 167, n. 4, p. 724–731, 4 2020. ISSN 15327361.

SAAD, W.; BENNIS, M.; CHEN, M. A vision of 6g wireless systems: Applications, trends, technologies, and open research problems. **IEEE network**, IEEE, v. 34, n. 3, p. 134–142, 2019.

SÁNCHEZ-MARGALLO, J. A. et al. Application of mixed reality in medical training and surgical planning focused on minimally invasive surgery. **Frontiers in Virtual Reality**, Frontiers, p. 144, 2021.

SARMIENTO, W. J. et al. Awareness of other: Evaluating the impact of proximity cues in collaborative tasks. In: **2013 IEEE Virtual Reality (VR)**. [S.l.: s.n.], 2013. p. 63–64.

SARMIENTO, W. J. et al. Measuring the collaboration degree in immersive 3d collaborative virtual environments. In: **2014 International Workshop on Collaborative Virtual Environments (3DCVE)**. [S.l.: s.n.], 2014. p. 1–6.

SCOTT, D. J. et al. Laparoscopic training on bench models: better and more cost effective than operating room experience? **Journal of the American College of Surgeons**, Elsevier, v. 191, n. 3, p. 272–283, 2000.

SEBAJANG, H. et al. The role of telementoring and telerobotic assistance in the provision of laparoscopic colorectal surgery in rural areas. **Surgical Endoscopy and Other Interventional Techniques**, Springer, v. 20, p. 1389–1393, 2006.

SEVDALIS, N. et al. Quantitative analysis of intraoperative communication in open and laparoscopic surgery. **Surgical endoscopy**, Springer, v. 26, p. 2931–2938, 2012.

SOUZA, V. et al. The effect of virtual reality on knowledge transfer and retention in collaborative group-based learning for neuroanatomy students. In: **2020 22nd Symposium on Virtual and Augmented Reality (SVR)**. [S.l.: s.n.], 2020. p. 92–101.

SQUIRRELL, D. M. et al. A randomized, prospective, blinded comparison of post-operative pain, metabolic response, and perceived health after laparoscopic and small incision cholecystectomy. **Surgery**, v. 123, n. 5, p. 485–495, 1998. ISSN 0039-6060. Available from Internet: <https://www.sciencedirect.com/science/article/pii/S0039606098001433>.

SUPE, A. N.; KULKARNI, G. V.; SUPE, P. A. Ergonomics in laparoscopic surgery. **Journal of minimal access surgery**, Medknow, v. 6, n. 2, p. 31–36, 2010.

TADLOCK, M. D. et al. Mixed reality surgical mentoring of combat casualty care related procedures in a perfused cadaver model: Initial results of a randomized feasibility study. **Surgery (United States)**, Elsevier Inc., v. 172, p. 1337–1345, 11 2022. ISSN 15327361.

TAWEEL, A. et al. Survivrs: Surround video-based virtual reality for surgery guidance. In: **2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)**. [S.l.: s.n.], 2023. p. 191–195.

TEO, T. et al. Exploring interaction techniques for 360 panoramas inside a 3D reconstructed scene for mixed reality remote collaboration. **Journal on Multimodal User Interfaces**, Springer Science and Business Media Deutschland GmbH, v. 14, n. 4, p. 373–385, 12 2020. ISSN 17838738.

TIAN, H. et al. Using virtual replicas to improve mixed reality remote collaboration. **IEEE Transactions on Visualization and Computer Graphics**, IEEE Computer Society, v. 29, p. 2785–2795, 5 2023. ISSN 19410506.

TREJOS, A. L. et al. Randomized control trial for evaluation of a hands-free pointer for surgical instruction during laparoscopic cholecystectomy. **Surgical endoscopy**, Springer, v. 29, n. 12, p. 3655–3665, 2015.

VASSILIOU, M. et al. The mistels program to measure technical skill in laparoscopic surgery: evidence for reliability. **Surgical Endoscopy And Other Interventional Techniques**, Springer, v. 20, p. 744–747, 2006.

WANG, P. et al. Behere: a vr/sar remote collaboration system based on virtual replicas sharing gesture and avatar in a procedural task. **Virtual Reality**, Springer Science and Business Media Deutschland GmbH, v. 27, p. 1409–1430, 6 2023. ISSN 14349957.

WEN, T. et al. Colonoscopy procedure simulation: Virtual reality training based on a real time computational approach. **BioMedical Engineering OnLine**, v. 17, p. 9, 2018. Available from Internet: <https://doi.org/10.1186/s12938-018-0433-4>.

ZHANG, Z.-y. et al. Preliminary application of mxed reality in neurosurgery: Development and evaluation of a new intraoperative procedure. **Journal of Clinical Neuroscience**, Elsevier, v. 67, p. 234–238, 2019.

ZORZAL, E. R. et al. Laparoscopy with augmented reality adaptations. **Journal of biomedical informatics**, Elsevier, v. 107, p. 103463, 2020.

APPENDIX A — RESUMO EXPANDIDO

Neste trabalho, apresentamos uma visão geral da literatura e trabalhos relacionados que exploram os princípios de colaboração remota, Realidade Virtual, Aumentada e Mista para ambientes médicos e, especificamente, ambos tópicos aplicados a cirurgias laparoscópicas. Descrevemos no Capítulo 3 o panorama inicial do planejamento e desenvolvimento da aplicação de RA, seus modos de utilização para cirurgias laparoscópicas e os desafios enfrentados que acabaram se tornado questões de pesquisa.

A.1 Desafios no planejamento de uma aplicação em realidade mista

Começamos a desenvolver a interface AR com foco no cenário INTRA, onde o cirurgião utiliza o óculos de realidade aumentada e tem a possibilidade de interagir com a interface durante o procedimento, desenhando anotações necessárias, salvando imagens importantes da cirurgia, além de acessar exames pré-operatórios e informações do planejamento cirurgico. Embora o método de interação mais utilizado em trabalhos relacionados tenha sido a voz e os gestos com as mãos, em nosso caso, procuramos reduzir possivéis interrupções e erros comunicação durante o procedimento. Assim, por meio de um cursor visual gerado por um raio de apontamento a partir da direção da cabeça, o método *head-gaze* surge como uma solução para apontar e selecionar elementos e posições específicas na interface em RA. No entanto, a sensibilidade aos movimentos da cabeça, que geram uma direção imprecisa, aliado do campo de visão limitado dos óculos de RA Hololens Gen-1 apareceram como o principal desafio deste trabalho.

A.2 Interação em Interfaces AR para Laparoscopia

Apresentamos no capítulo 4 todos os detalhes da aplicação AR desenvolvida, o design do hardware onde descrevemos o desenvlvimento do anel adaptado com botões e os pedais planejados para as confirmações necessárias das interações. A interface visual é composta por painéis com as informações necessárias baseadas nas interações do usuário, como observar o vídeo laparoscópico, realizar anotações no vídeo e salvar determinados pontos em imagem ou vídeo do procedimento. Apresentamos também as técnicas de *head-gaze* técnicas que utilizamos para compensar o movimento e a sensibilidade do apontamento com a cabeça, afim de reduzir os erros durante a utilização e otimizar o campo de visão limitado do HL1. Os métodos são: O Método de Escala, reduzindo a proporção de movimento realizado pela cabeça em relação ao cursor visual, no qual a proporção padrão é 1 quando não dimensionada. O Método de Filtragem pela Média armazena as n posições anteriores da posição do cursor visual e calcula uma média simples para uma nova posição filtrada. Após a proposta e desenvolvimento, planejamos um experimento com usuários com o objetivo de avaliar as técnicas aplicadas e a interface desenvolvida.

A.3 Experimento com usuários

O experimento com usuários apresentado no capítulo 5 concentra-se na avaliação da usabilidade das técnicas de *head-gaze* e interação utilizando o anel com botões em realidade aumentada para anotação durante procedimentos de laparoscopia e configurações semelhantes. O principal estudo apresentado teve como objetivo compreender o impacto e as melhorias que os métodos de Escala e Filtragem pela Média podem trazer ao realizar anotações no feed de vídeo de laparoscopia em um monitor virtual posicionado diretamente na frente do usuário. O experimento do usuário foi realizado em um protocolo *between-subjects* com 32 voluntários do Instituto de Informática da UFRGS. Realizamos uma série de análises estatísticas com os dados coletados, avaliando diferentes aspectos da utilização, como tempo, precisão e completude da tarefa e o impacto da utilização dos métodos de *head-gaze* propostos neste trabalho.

A.4 Resultados

Durante a análise, iniciamos o processamento dos dados de tempo e selecionamos os dados das tentativas centrais dos usuários, desconsiderando as 6 tentativas iniciais e a última, onde entendemos e analisamos que os tempos dos usuários refletem uma fase de adaptação e possivel fadiga (no caso do último tempo), apresentamos no capítulo os gráficos que representam essa curva de aprendizado e a diferença estatística entre os grupos de tentativas. Observando os dados de precisão das anotações, constatamos que as tarefas foram claramente descritas e eram possíveis de serem solucionadas, o que é representado pelas soluções incorretas dos ensaios, onde houve apenas 3,59% de erro no geral, desconsiderando os dados corrompidos e os dados de usuários removidos por estarem fora do escopo do trabalho. Nossa principal descoberta está relacionada a diferença das posições dos eixos das anotações realizadas, onde o eixo Y foi mais sensível ao mapeamento da direção da cabeça e apresentou uma taxa de erro de posicionamento maior do que o eixo X.

Os dados de erros do posicionamento das anotações utilizando métodos de *head-gaze* desenvolvidos neste trabalho não apresentam significância estatística entre si nos dados gerais, mesmo quando agrupados por método e desconsiderando o nível deles. No entanto, quando isolamos pela dificuldade da tarefa, alguns dados se mostraram diferença estatística com menor taxa de erro usando os métodos propostos do que o mapeamento direto do movimento da cabeça, como o erro de posicionamento dos pontos das setas na tarefa de passagem de linha e o erro de dimensionamento dos círculos na tarefa de transferência de argolas na dificuldade fácil.

Também apresentamos nos resultados e discussão uma visão geral dos dados qualitativos e da experiência do usuário durante as tarefas, incluindo uma medição de carga de trabalho, onde constatamos que os usuários estavam confiantes em seu desempenho e relataram baixo nível de exigência física e temporal. Porém, também relataram um esforço médio a alto para realizar as tarefas. Os usuários também revelaram que não era complicado usar o anel para confirmar interações. Concluindo o capítulo do experimento com usuário e resultados, apresentamos as limitações do estudo e da interface, além de uma breve explicação sobre o próximo experimento para avaliar a utilização da interface pelos cirurgiões durante a execução das tarefas físicas.

A.5 Conclusão

A abordagem apresentada demonstrou que a interface pode ser útil e aplicável, e o hardware proposto também pode ser incluído como ferramenta de interação pelos cirurgiões. Também foi demonstrado que a aplicação de tal interface para agilizar o acesso dos cirurgiões a instruções, sinais vitais, dados de planejamento e exames pré-operatórios em um monitor virtual aumentado, que pode auxiliar durante treinamentos e procedimentos reais com assistência remota. Embora os métodos de *head-gaze* desenvolvidos tenham demonstrado diferenças significativas apenas em certos casos, a precisão geral obtida é adequada para sua utilização de acordo com a latência e a taxa de quadros do dispositivo de realidade aumentada.