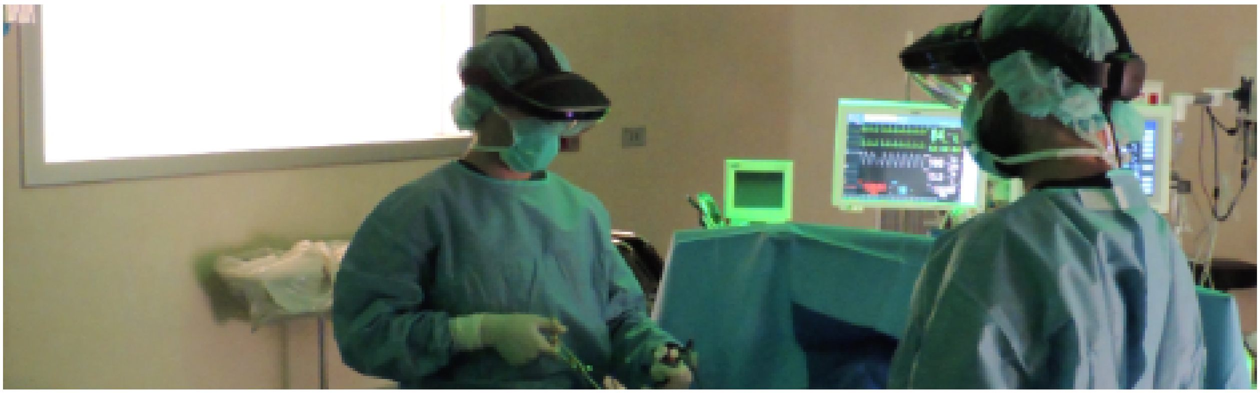
TÉCNICO LISBOA



LARA

Laparoscopy with Augmented Reality Adaptations

José Miguel Campos Gomes

Thesis to obtain the Master of Science Degree in

Telecommunications and Informatics Engineering

Supervisors: Prof. Joaquim Armando Pires Jorge
Prof. Daniel Simões Lopes

Examination Committee

Chairperson: Prof. Fernando Henrique Côrte-Real Mira da Silva
Supervisor: Prof. Joaquim Armando Pires Jorge
Member of the Committee: Prof. Manuel João Caneira Monteiro da Fonseca

November 2019

Acknowledgments

Em primeiro lugar, gostava de agradecer ao professor Joaquim Jorge pela oportunidade de poder fazer este projecto, e para os investigadores do VIMMI, nomeadamente o Maurício Sousa e o Daniel Simões Lopes. Sem a ajuda e conhecimento destas pessoas, não teria sido possível realizar esta tese. O meu agradecimento também vai para os colaboradores da fundação Champalimaud, em particular para o Dr. Nuno Figueiredo, a sua equipa médica, o Pedro Garcia da Silva e o João Frasão, pela ajuda prestada de forma a concretizar este projecto. Gostava de agradecer também aos meus amigos e colegas Ostap Kozak, Francisco Maria, Fernando Moreira, Luís Santos, André Soares, Pedro Maria e Diana Mendes pela ajuda prestada, pela companhia dada e pela sua amizade ao longo deste mestrado. Gostaria também de agradecer aos meus amigos do coração, Pedro Veloso e Alexandre Bernardo, pelo carinho, pelo apoio e pela amizade que me têm dado estes últimos anos. Pessoas como estas tornaram a faculdade menos difícil de enfrentar e deram-me alento para continuar. Finalmente, quero agradecer aos meus pais e à minha avó por todo o apoio dado durante o meu percurso escolar e académico. Sem vocês não teria conseguido chegar a este patamar da minha vida e por isso estou-vos grato. Dedico também esta conquista ao meu avô, alguém que tenho a certeza que ficaria muito orgulhoso de mim por ter chegado até aqui.

Obrigado a todos vós.

Abstract

One of the possible applications of Augmented Reality is surgery, especially laparoscopic surgery, which currently suffers from problems such as surgeon discomfort and fatigue caused by looking at a display positioned outside the surgeon's visual field, heightened by the length of the procedure. This fatigue is especially felt on the surgeon's neck, as it is strained from adopting an unnatural posture in order to visualise the laparoscopic video feed. Throughout this document we will present some works in the fields of Augmented Reality, as well as works in surgery and Augmented Reality applied to both surgery in general and laparoscopy in particular. Through a user and task analysis, we determined that users operate in a dimly lit environment, surrounded by monitors, and communicate through verbal commands and pointing gestures. We first performed a study on the effect of head-mounted displays on task performance, followed by the design and implementation of a multimodal interface that enhances the laparoscopic procedure, making it more comfortable for surgeons by allowing them to visualise the laparoscopic video regardless of neck posture, access patient imaging data without interrupting the operation and communicate with team members through the use of a pointing reticle.

Keywords

Laparoscopy; Augmented Reality; Head-mounted Display; Optical-see-through; Visual Field; Hand-eye Coordination

Resumo

Uma das possíveis aplicações de Realidade Aumentada é a cirurgia, em particular a laparoscopia, que actualmente sofre de problemas como desconforto por parte do cirurgião, e fadiga causada por estes terem de olhar para um monitor posicionado fora dos seus campos de visão, fadiga esta intensificada pela duração do procedimento. Esta fadiga é sentida especialmente no pescoço do cirugião, pois este é sobrecarregado devido à adopção de uma postura pouco natural para se poder visualisar o vídeo da laparoscopia. Ao longo deste documento iremos apresentar alguns trabalhos nos campos da Realidade Aumentada, bem como em cirurgia em geral e laparoscopia em particular. Através de uma análise de utilizadores e tarefas, determinámos que estes operam num ambiente pouco iluminado, rodeado por monitores, e que comunicam através de comandos verbais e gestos de apontar. Primeiro, conduzimos um estudo para determinar o efeito dos *Head-Mounted Displays* sobre a postura dos utilizadores, seguido pelo desenho e implementação de uma interface multimodal que melhora o procedimento laparoscópico, tornando a experiência mais confortável para os cirurgiões e permitindo que estes visualizem o vídeo laparoscópico independentemente da postura do pescoço, aceder a dados de imagens do paciente durante a operação e comunicar com os restantes membros da equipa através de um apontador.

Palavras Chave

Laparoscopia; Realidade Aumentada; Head-mounted Display; Optical-see-through; Campo Visual; Coordenação Óculo-manual

Contents

List of Figures

List of Tables

Acronyms

AR Augmented Reality

FLS Fundamentals of Laparoscopic Surgery

FOV Field of View

HMD Head-mounted display

MIS Minimally Invasive Surgery

MR Magnetic Resonance

MRI Magnetic Resonance Imaging

VR Virtual Reality

UI User Interface

1

Introduction

Contents

There is great potential in applying optical-see-through Augmented Reality (AR) to laparoscopic surgery, as doctors will be able to visualise both the task being performed and patient data simultaneously, while improving posture and comfort as a result of eliminating the need to stare at a monitor placed at a distance [1]. This type of visualisation is accomplished by enhancing a real-world setting with computer-generated information overlaid onto a screen, through which both real and virtual objects can be simultaneously observed and interacted with. It is possible that this can have a greater impact on surgeries than robot assistance ever did, as despite its potential, this type of procedure currently makes up to less than five percent of total laparoscopies.

AR has experienced a surge in popularity thanks to ever more powerful and ever less expensive hardware. However, applications of AR to many settings, including surgery, remain a challenging topic. We believe this stems from a lack of effort applied to the area of interfaces, as there is little work focused on it.

1.1 Motivation

The Food & Drug Administration has approved, on the 21st of September 2018, the use of Microsoft HoloLens in a preoperative surgical planning solution[1], proving the rise in popularity for this type of solution. But even though the technology may have matured in terms of implementation, with several commercial products available, such as the Microsoft HoloLens, the Samsung Gear VR and the Magic Leap One to name a few, there is little work done in the field of interfaces and interaction techniques.

1.2 Problem Description

As stated previously, surgery, and in particular laparoscopy, is one of the possible applications of AR. Laparoscopy is a type of minimally invasive procedure performed on the abdomen or pelvis, with the abdominal cavity being expanded with gas to permit the insertion and movement of laparoscopic instruments inside the body. Unlike open surgery, in laparoscopy there is a loss of direct visual contact with the organs, with the surgeons being required to use an endoscopic camera, which captures and feeds an image onto a display [1]. This results in a much more limited and restrictive experience compared to open surgery, as the surgeon's dexterity and ability to feel feedback from applying pressure on tissue is reduced by the laparoscopic instruments [2] [3]. The biggest problem with laparoscopic surgery, however, is hand-eye coordination, as surgeons have to look at screens placed outside the field of operation, which results in discomfort [3], affecting the surgeon's efficiency due to a disconnect between the visual and motor axis, because the surgeon cannot look at the instruments or hands and the field of surgery

[1] https://www.healthimaging.com/topics/advanced-visualization/fda-approves-augmented-reality-system-microsoft-hololens

simultaneously. To be successful, more training is required to adapt to this condition, as extra mental effort must be applied [2]. In addition, almost all these display screens are limited in sense that they do not support techniques to improve visual collaboration with the rest of the surgical team [4].

1.3 Objectives and Research Statement

In a first phase, we will perform a user and task analysis to gain a greater understanding of how users achieve their goals and how that can be improved. Next, we will perform a preliminary evaluation to assess whether using an Head-mounted display (HMD) impairs task performance in laparoscopic surgery. Finally, this work will aim to develop a multimodal interface prototype with surgical application for AR in order to mitigate some of laparoscopy's most serious problems.

We will develop a user interface for an Augmented Reality headset, presenting endoscopic video stream directly to the surgeons, as well as preoperative data for them to browse and analyse. This interface has to ensure a mainly hands-free experience in order to be usable even when a surgeon is holding tools in both hands. We can then highlight the research statement of this work as such:

> ***Using Augmented Reality in laparoscopy allows for improved performance in surgical procedures.***

1.4 Contributions

Considering the limitations of laparoscopy and the hurdles it places on surgeons, be it physically or mentally, our work offers the following contributions:

- **Solution for the neck tension problem** - our prototype offers a way to visualise the laparoscopic video while not forcing the surgeon to assume unnecessarily uncomfortable positions. Results have been extremely postive in this regard, with some users remarking this constitutes the most important aspect of the prototype.

- **A way to avoid interruptions during surgery** - From our field observations, we noticed that surgeons place down their tools to perform some secondary tasks, which interrupts the procedure. We designed our prototype to avoid this type of situations, with a completely hands-free approach, using both head gaze and foot movement as sources of input. Despite the novelty of the idea, users found it easy to get used to.

- **Improvements to surgeon-to-surgeon communication** - Communication between surgeons is ambiguous at best, with surgeons being barred from touching the screen due to sterilisation. Our prototype takes a very effective mechanism, the cursor, which is already found in robot-assisted laparoscopy, and introduces it to the more sustainable method that is traditional laparoscopy through the HMD headset. Leveraging this familiarity proved useful, with users commenting on how well it works, both in terms of activation as well as usage.

1.5 Research Context

The work developed in this dissertation is not an isolated effort: In fact, it is part of an ongoing partnership between Instituto Superior Técnico, INESC-ID and the Champalimaud Foundation to improve laparoscopic procedures. As such, the results stemming from our work will be used in future work done by INESC-ID researchers and IST Master's students alike.

1.6 Organisation of the Document

The organisation of this thesis is as follows: in Chapter 1 we have presented an overview for the topic of this thesis, explaining the current problem, as well as its goals and its research statement. In chapter 2, we will go over work done by other authors in the fields of AR interaction, then AR applied to surgery in general and focusing on AR applied in laparoscopy as well as different visualisation techniques concerning this type of surgical procedure. In chapter 3 we present an analysis of our users and the tasks they perform, conducted in partnership with surgeons from the Champalimaud Foundation. In chapter 4 we present the results of our preliminary evaluation on the effect of a HMD on user task performance, and in chapter 5 we present our prototype: what problems it solves, what was the design process that took place in order to solve them and the prototype's architecture: how every element fits in the bigger picture and their technical details. The prototype's evaluation is presented in chapter 6 and lastly, we conclude in chapter 7, while also discussing possible future work.

2

Related work

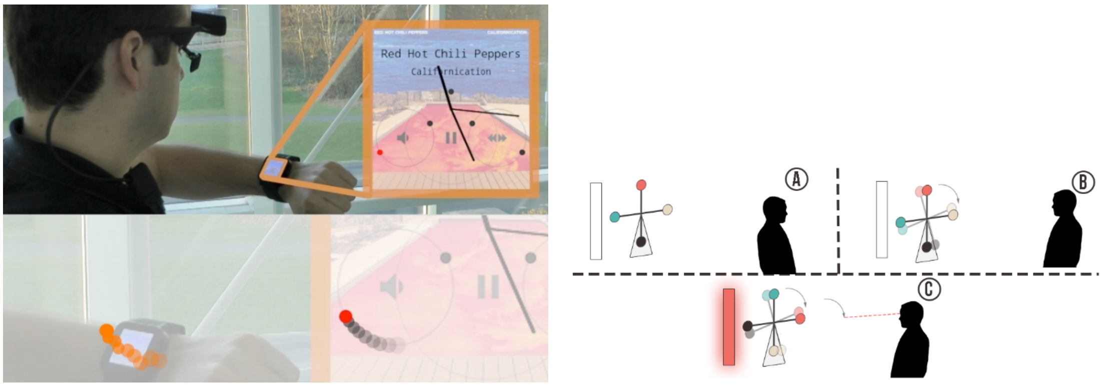
Contents

In the last decade, there have been works introducing new interaction techniques for Augmented Reality, attempts at incorporating the technology in the surgical field, including laparoscopy. With respect to laparoscopy, studies have been performed to better understand how the procedure is conducted, what are its limitations, and how they can be overcome using AR. These works explore different types of image visualisation, like different monitor positioning or usage of HMDs, as well as different sources of input and ways to communicate.

2.1 Augmented Reality Interaction and Usability

Knight and Baber [5] conducted a study to determine whether the use of an HMD causes users to alter their head position to a posture that demands greater effort. To do this, seven paramedics performed a simulated treatment of a patient with cardiac problems with dummies in two different scenarios: in the first, the dummy represented a fully conscious middle-aged man complaining of chest problems, while in the second the dummy represented a patient that had gone into shock, requiring the CPR technique to be performed. The participants performed both exercises, first in the baseline condition, with no HMD, and afterwards using a 0.12 kg Seattle Sight monocular transparent display. The authors concluded that wearing an HMD can indeed force wearers to modify their neck posture, placing their head and neck under increased levels of stress. The main cause for posture modification is the centre of gravity in the head being shifted to the front of the wearer's face, but Knight and Baber also hypothesise that users may alter their head positioning to see the image against a more uniform background, like a wall or the ground. Other reasons include poor fit, requiring the wearer to reduce slippage by balancing the HMD, and poor image Field of View (FOV), as the HMD's casing blocks vision outside the viewing window.

Esteves et al. [6] introduced a new interaction technique which enables hands-free input on smart watches. The idea is that the interface controls are moving and are selected by detection of smooth pursuit, the motion of the eyes following a moving stimulus. Since it requires moving targets to be detected, smooth pursuit is robust against false positives. The controls move in circular fashion, differ in phase offset, angular speed and direction. During testing, while users were first confused, after a brief period of experimentation, they easily learned to use the controls. Velloso et al. [7] also worked on this interaction technique, developing on the previous paper's idea and try to apply the same concept of animated targets and input detection through tracking of smooth pursuit to multiple objects in a shared space. These include a video-on-demand interface with animated controls on the screen, a music player with widgets projected onto a pair of speakers, a multi-coloured lamp controlled by a windmill with paddles corresponding to a different colour each, and a laser dot moving around a fan. Each device sends the coordinates of their targets to a central server, where the estimated gaze values from the tracker are also sent. Users considered this system to be easy to use as well, not tiring to their eyes and



(a) Esteves **(b)** Velloso

Figure 2.1: The works of Esteves and Velloso rely on tracking eye movement to activate different controls, in orbital motion.

the controls quick to activate. The ideas of these works are illustrated in fig. 2.1(a) and fig. 2.1(b).

Interactions on static real-world content was a topic worked on by Kim et al. [8], who introduced a prototype, shown in fig. 2.2, using Microsoft HoloLens, featuring four types of interactions: The first is details on demand, presenting details about whatever the user is pointing at, using head position tracking as a pointer. The second is highlighting, where the user emphasises relevant content by tapping on buttons or using voice commands. The third is filtering, hiding real-life content by overlaying patches with a background colour. The last is linked views, where upon highlighting a subset of data points from a given graph, a second graph with only that data is shown. Input is made available through voice commands with natural language and through gestures. Feedback is given visually, updating the view according to the interaction, and through audio, in the form of simple chimes.

Another interaction technique was introduced by Müller et al. [9], who explored the use of feet as a source of input, experimenting with foot-tapping. Their work consisted in having a HoloLens program instruct the users to tap a given target inside a semi-circular grid. This grid varied in number of rows, from one to three, as well as in number of columns, from two to six. Not only direct interaction was tested, indirect interaction was also a target of study. In direct interaction, the targets were presented on the ground, so users were required to look at the floor to interact with them. On the other hand, in indirect interaction, the semi-circular grid was displayed in front of them, as illustrated in fig. 2.3. The test measured both accuracy and efficiency, with direct interaction finding high accuracy results up to the highest condition, with three rows and six columns. Users commented that the HMD was easy to use and not tiring, compared to the default air-taps that HoloLens supports. However, they did complain about having to look down all the time. Indirect interaction was found to be not as accurate, even though it did not force users to look down. Accuracy suffered the most when using more than one row and tasks took more time to complete with each added row or column. However, this type of interaction found

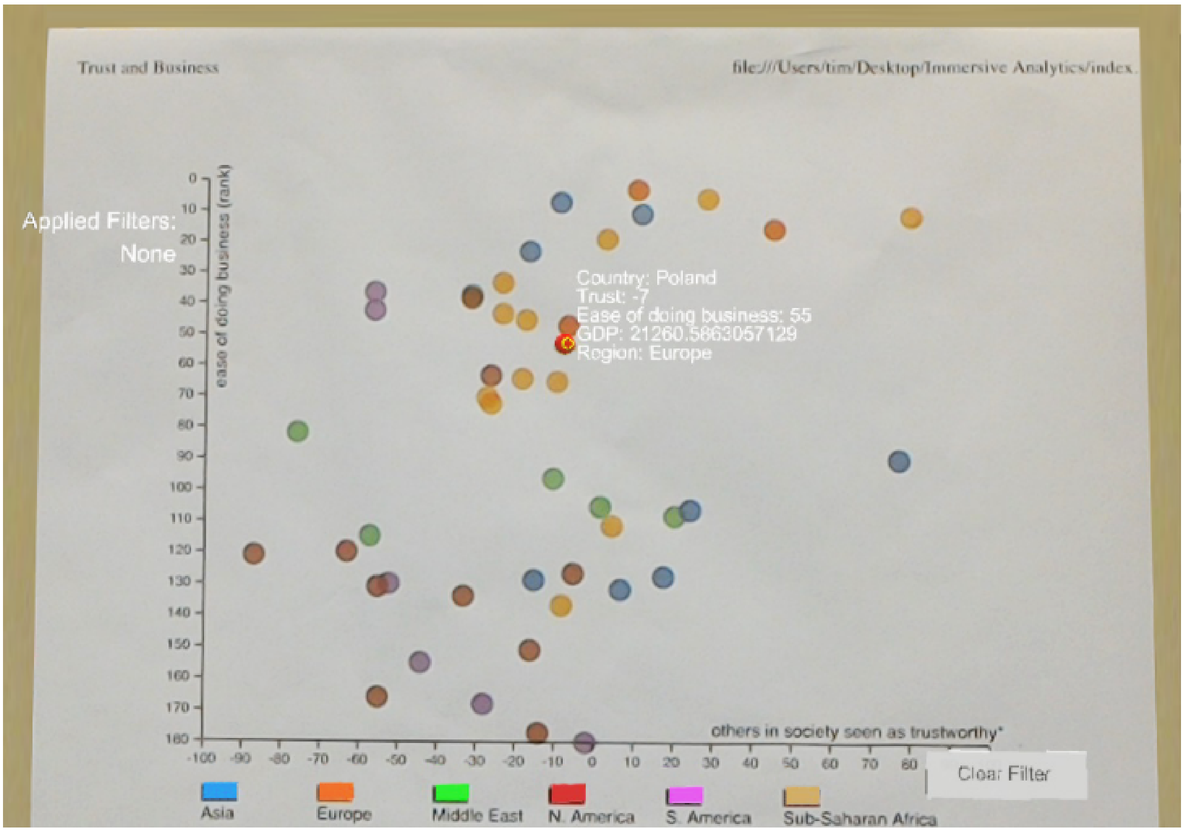


Figure 2.2: Kim used head position as a pointer to activate different controls or reveal additional information.

greater popularity among users, as they liked not having to look at the floor, and that the interaction felt easy to perform, especially if it were performed along with other tasks, due to not needing to use the hands. They also liked the radial placement of the targets. To best exploit this type of interaction the authors suggested favouring the target division into columns rather than rows as to prevent accuracy losses. They considered the use of direct interfaces to be best-suited to high accuracy interactions, with a large number of options, and proposed the use of indirect interfaces for longer-term interactions that need less accuracy, as well as in situations where a lower number of options is sufficient, or wherever there are constraints to the view, as in direct view the users have to look at the floor.

2.1.1 Discussion

In terms of usability, since Knight and Baber [5] performed their study in 2007, new devices have appeared, such as the Microsoft HoloLens 2[1], which claim to be immersive, as well as focus on comfort during extended periods of time and fitting well on the head. In theory, this should mitigate most problems listed by the authors, but we think this should be tested again, especially under the context of a

[1] https://www.microsoft.com/en-us/hololens/hardware

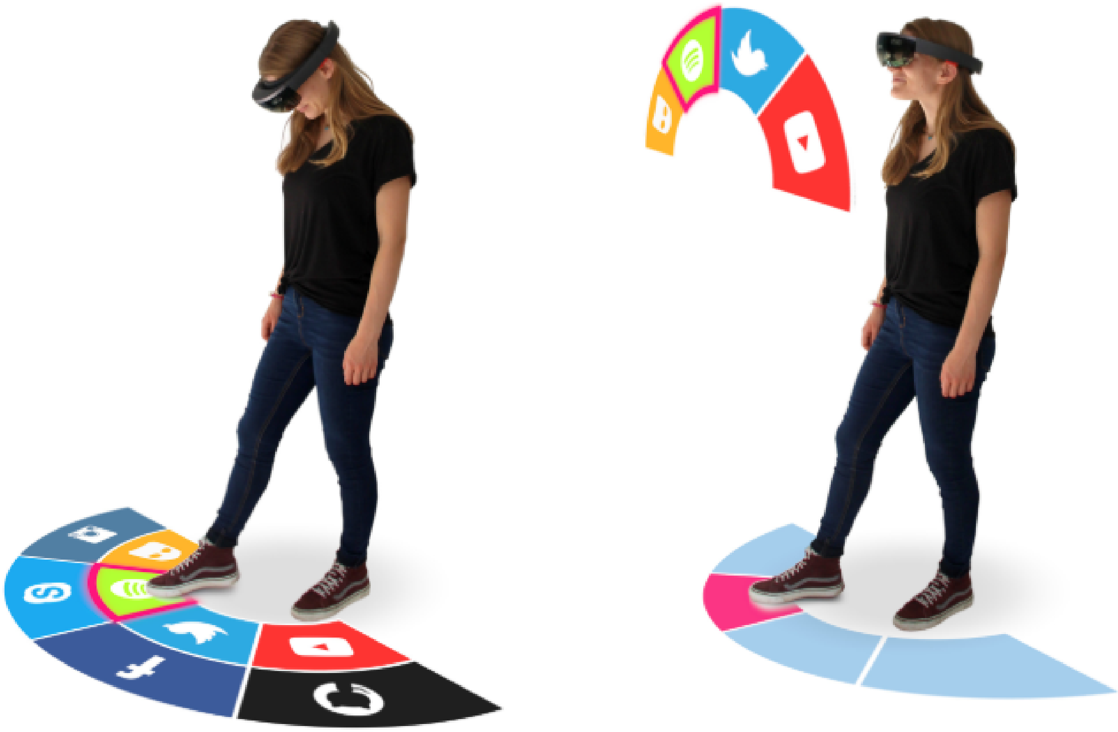


Figure 2.3: Müller et al. explored both direct and indirect interactions using foot taps.

laparoscopy.

The works of Esteves et al. [6] and Velloso et al. [7] aim at making controls more accessible, but unlike a smart watch screen, which is used for short periods of time, a HMD is used for much longer, and it may be tiring to have these controls always accessible and at the forefront instead of fading into the background like in the work of Velloso et al. If these controls were to be hidden on a side view, they would not differ much from simpler buttons which activate upon with a head gaze. Another consideration would be whether an eye tracker would fit inside a HMDs casing, which might be too compact.

With respect to interaction, the work of Kim et al. [8] introduced several features which seem useful for content manipulation, but because there were not performed any tests, it remains unclear how effective the concept really is. In terms of interaction techniques used, simple mechanics such as the head gaze stand out from the other two, voice commands and gestures, as using gesture controls in an operating environment could prove troublesome since it would require the users to take their hands off the tools. Voice interaction is considered to be state-of-the-art, but it is also considered to be limited when it comes to social environments. In a surgical environment, it can be disruptive to team communications. On the other hand, foot interaction as explored by Müller et al. cite [9] can prove useful, as it can prevent surgeons from needing to put down their tools. Given that surgeons have their view occupied with several things like tools, other operating staff, monitors and most importantly, the patient, perhaps it would be best to adopt an indirect type of interaction in this scenario. Additionally, surgeons already use pedals in the operating room, so it would not be advisable to make use of a foot tap, as it can result in their inadvertent activation, while still making use of foot movement.

2.2 Augmented Reality in Surgery

Watanabe et al. [10] developed a navigation system for brain surgery, where MRI or CT scans are superimposed onto the video image, captured by the back-facing camera of a Microsoft Surface Tablet PC, which is held by either the surgeon or an assistant. To help the placement of the 3D image on the screen, 6 trackers are spread throughout the room. Chen et al. [11] developed another surgical navigation system, which superimposes 3D preoperative data on the patient, but using a see-through AR headset instead. Pratt et al. [12] also explored overlaying 3D data on top of the patient using Microsoft HoloLens, but instead of trackers, the image is aligned manually. This is done by performing rotational and translational movements on the model until it matches the patient's anatomy in a satisfactory fashion. To switch between rotation and translation, it is required the use of a toolbar button or voice commands, while a 'air-tap and hold' gesture is used for 3D motion. Grinshpoon et al. [13] also approached the subject of 3D content manipulation, trying to address the problem of hands-free visualisation during operation without resorting to the user's feet. Using the voice recognition and head-tracking features of Microsoft HoloLens, it is possible to rotate and scale 3D content, as seen in fig. 2.4. 3D models are selected with the gaze and activated through voice commands, while manipulations are performed with head movements such as up, down, left or right.

On the other hand, Bautista et al. [14] compared the hand gesture recognition from the Meta glasses

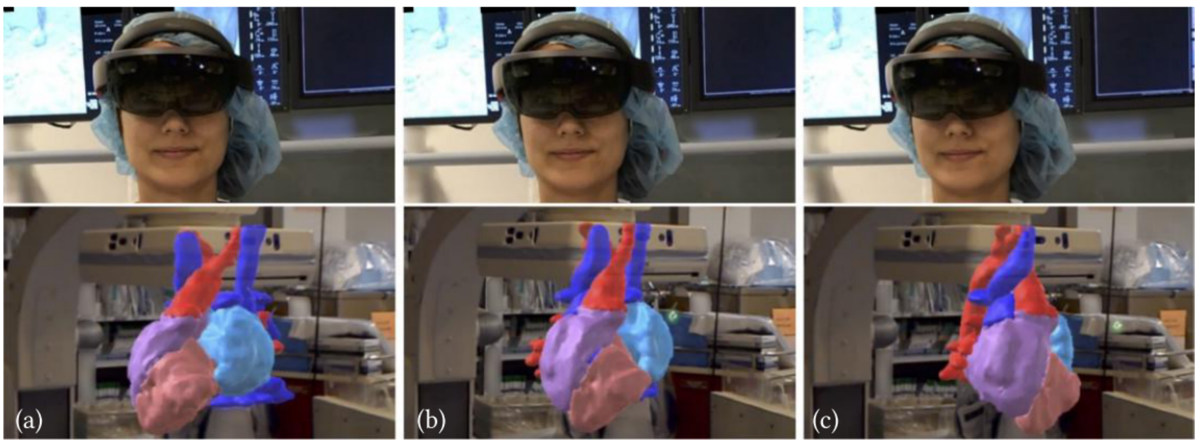


Figure 2.4: Grinshpoon et al. used head movements as a source of input for 3D image manipulation.

with the MYO arm-band in order to determine which was more suited to navigate an AR interface used for computer-assisted orthopedic surgery. This was done by analysing how well users learned device interaction gestures based on a training video, as well as navigating through the interface up to a given category, taking metrics such as time, number of errors and number of times the user required assistance. Results found the Meta to be more difficult to use, with users taking more time to perform the tasks and making more mistakes. Interestingly enough, with the Meta, discomfort was felt gradually rather than instantly.

2.2.1 Discussion

Watanabe et al. [10], Chen et al. [11] and Pratt et al. [12] have all developed ideas using superimposed 3D models onto the patient. However, applying these ideas to laparoscopy would not eliminate the need for an endoscopic camera, as it is still needed to observe possible bleeding, as well as changes in anatomy caused by incisions and the use of the bowel grasper, not to mention the expansion of the abdominal cavity with gas. In addition, in the work of Watanabe et al., with the tablet, there is no need to alternate the gaze between the surgical field and navigation screen, but it still physically obstructs the surgical field, as seen in fig. 2.5.

Grinshpoon et al. [13] attempted to circumvent HoloLens's limitations of a narrow FOV by always keeping the content in front of the user and treating the gaze like a joystick, instead of superimposing

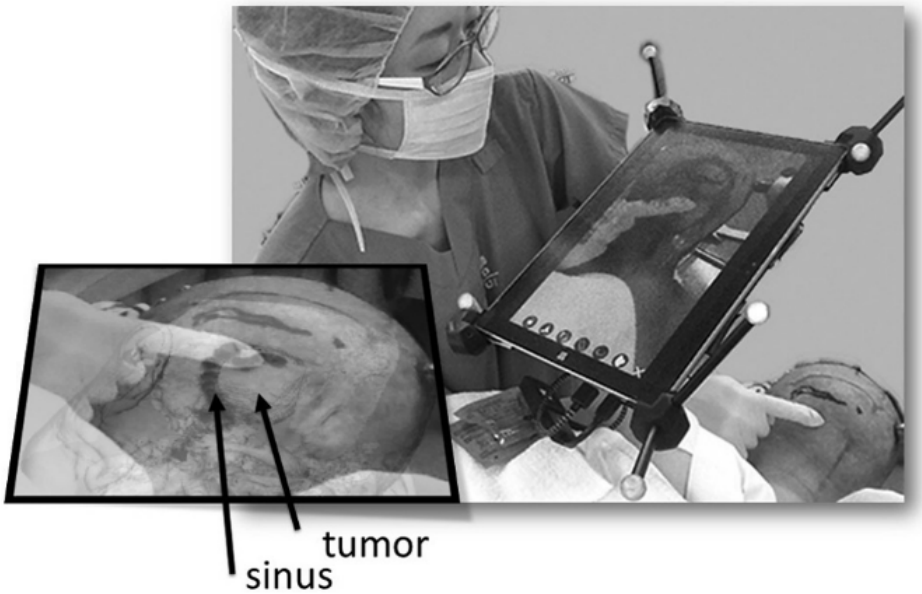


Figure 2.5: Watanabe et al. used a tablet to overlay medical imaging onto the patient in real-time.

the models like the previously mentioned works. However, spotty voice detection may be the Achilles' heel for the system, since it is essential to activate the manipulation modes. Given tests were still being conducted at the time, it is also uncertain how useful the system is.

2.3 Augmented Reality in Laparoscopy

Muratore et al. [15] made a listing of laparoscopy's limitations and propose the ideal image display in endoscopic surgery, mentioning the already existing Automated Endoscopic System for Optimal Positioning and the da Vinci Surgical System for their improvements in image quality and stability, but suggest superimposing preoperative and 3D imaging on the operating field to facilitate access to a target organ. For the future, the use of a 3D high definition image HMD is considered the ideal display system, citing the comfort of looking at the endoscopic image in any preferred head position, improving ergonomics and reducing neck strain. The use of an HMD is also seen as beneficial in the sense that it alleviates equipment clutter in the operating room. It is further noted the usefulness of individualised image manipulation features like zooming, which allows each surgeon to see the endoscopic video in the way they find most comfortable. Finally, the use of foot pedals is suggested in order to keep the hands free for using the surgical tools.

The work of Mentis et al. [16] also delves into the laparoscopic procedure itself, through analysis of video footage of Minimally Invasive Surgery (MIS) operations, in order to understand how surgeons instruct residents in seeing the body during laparoscopic surgeries and expose some of the effort surgeons put into them. The first way students are trained is in determining where and in which direction the camera and the tools are with respect to what they see. The second way is through guiding the hand, as the surgeon grabs the hand of the student while moving the camera in order for him to have better awareness of the instruments' movements. However, guidance also takes the form of pointing and giving verbal commands. Finally, the third way involves teaching students to envision what is not seen, inferring where hidden anatomy may be, based on what is seen on the video. Mentis et al. [16] suggest that training should allow the student to understand the images the same way the surgeon does, stating the need to augment the endoscopic video, either through an overlay, which would identify anatomy based on context, or a secondary screen that provided 3D models of the chest and abdominal cavity. This model could then be manipulated by both surgeons through gestures and voice control, while also identifying the camera's field of view on it, giving the student a greater understanding of positioning and helping him identify anatomy on screen.

Also on the topic of understanding and communication, Prescher et al. [17] conducted a study to find out whether using a navigation grid or a navigation pointer could help instructors in directing assistants to specific targets. The navigation grid consisted of a 3x5 coordinate grid, with each quadrant being assigned a number and a letter, while the cursor consisted of a fluorescent green dot integrated into a laparoscope and projected onto a display. Using a laparoscopic box trainer with 240 pins, each of the 24 subjects executed 15 tests where they had to locate 5 random targets in each one. The tests with no navigation tool had the instructor merely convey four directional commands: up, down, left and right. In the tests with the grid, the instructor specified the quadrant, then gave the same four commands for further orientation, while in the tests with the pointer, the instructor pointed at the target with the camera. Results demonstrate, through faster completion times, that the pointer is a superior tool for navigation and guidance, compared with both the grid and no tool. Feng et al. [18] also worked on screen augmentations as a means to improve communication, designing a telestration system for trainers to point or draw a sketch over a video for the trainee to see. Microsoft Kinect was used to implement this system. Gestures are used to control a small green circle acting as a pointer, and audio commands are used to switch between pointing and drawing, as well as to clear the screen. To evaluate the system's effectiveness, trainees performed a Fundamentals of Laparoscopic Surgery Box Trainer exercise, while the trainers gave them instructions and provided them with guidance using telestration.

Walczak et al. [19] evaluated whether the positioning of the monitor has an impact on laparoscopic performance. They had 52 participants execute an exercise in a custom-made simulator, where they had to pass a thread through 9 holes of different sizes. This exercise was performed four times in two different monitor positions, the first at eye level, 1.6m from the ground and 1m away from the subject, and the second 0.6m away from the participant, at an angle of 20º below eye level, shown in fig. 2.6. Time to execute the task was measured and participants were asked at the end of the test which position they preferred. Results show the time taken to perform the task was shorter when the screen was placed downwards, which corresponded to the position participants most preferred. This position allows users to flex the head at 15 to 45 degrees below eye level, which is the most comfortable position, as looking down improves eye lens accommodation and reduces eye weariness and headaches.

Maithel et al. [20] also evaluated the effect of the monitor, conducting a study to determine whether wearing a HMD improves task performance, or at least reduces muscle fatigue, comparing it in an operating scenario against the use of a traditional monitor placed at a lateral angle of 30º. 30 test subjects had to perform a triangle transfer task in four repetitions using a Computer Enhanced Laparoscopic Training System. This task was measured using depth perception, motion smoothness, response orientation, path length and execution time as parameters. It was proved that an HMD improved smoothness of motion, but performance in general was not found to be superior.

Batmaz et al. [3] compared four types of visualisation, direct vision, 2D fish-eye and undistorted view and 3D stereoscopic view, and studied their performance effects on a laparoscopic training exercise where the subjects were to place a small object in the centre of five targets, in a specific order. The 2D images were studied in two different positions, one with a monitor placed sideways, in a 45º offset, and

Figure 2.6: Walczak et al.'s second condition in their experimental setup had the screen positioned 20º below eye level, 60cm away from the participant.

another with the monitor placed straight ahead. The 3D stereoscopic view was implemented using a Oculus DK2 HMD, with the video following head movements. The results show 3D stereoscopic imaging does not have any performance edge over 2D, with objects being selectively coloured, facilitating depth perception. However, straight ahead monitor positioning did have performance benefits, as subjects took less time to perform the tasks as they felt less neck strain and more comfort.

Prescher et al. [21] also studied the effect of 3D viewing, and conducted a study to determine whether the use of a stereoscopic 3D display with glasses improved performance in trainees, with 32 subjects performing 10 repetitions of a peg transfer task, alternating between 3D and 2D displays. Time and number of dropped objects was measured and a questionnaire was completed by the subjects at the end of the test. The 3D display proved to reduce the time taken to complete the test as well as the number of dropped objects, while being generally preferred by the test volunteers.

Kihara et al. [22] developed a Virtual Reality (VR) system for use in real-world operation, combining a HMD with a 3D endoscope to provide the surgeon with high quality imaging right in front of him. The 3D HMD gives the feel of an open surgery and allows the visualisation of content regardless of head

Figure 2.7: The HMD in Kihara et al.'s work allows visualisation of the hands and tools by looking down with the eyes.

position, while direct vision is allowed by lowering the angle of sight. This prototype is shown in fig. 2.7.

Jayender et al. [23] worked on a mixed reality headset which integrates the image from the laparoscopic camera, a navigation system and diagnostic imaging, complemented by an audio feedback system. The system is implemented using a capture card to capture video from the laparoscope and the navigation system, a Unity application which renders the virtual environment into which the video is imported and a modified Oculus Rift DK2 headset to display the mixed reality environment. In it, each image is represented by a virtual monitor placed in front of the user, with the laparoscopic video being placed 15º below eye level, the navigation system placed lower, at an angle of 30º below eye level, and the diagnostic imaging placed to the side of the laparoscopic video feed. An audio navigation feedback system is also implemented, outputting sound based on the distance from the tool to the target. For interaction, pressing a foot pedal while placing a reticle on diagnostic images brings them closer, allowing them to be visualised in greater detail. To test this system, three different peg transfer exercises were employed, comparing it to a laparoscopic navigation with CT imaging approach. Time, accuracy, peg drops, incorrect peg selections and kinematic parameters such as velocity, acceleration and jerk were used to evaluate the tasks. While on the initial tests the results did not favour the system, the following tests demonstrated improved performance, reduced task time and reduced errors. A NASA Task Load Index questionnaire also demonstrated that the workload required from the user was significantly decreased.

2.3.1 Discussion

In the work of Muratore et al. [15], the authors emphasise the importance of the HMD, stating preoperative imaging could be individually manipulated through its use, as well as grant surgeons extra comfort by allowing them to see the laparoscopic image regardless of head positioning. Hands-free interaction is again considered, with the suggestion of using foot pedals instead. Also discussed is the issue of paradoxical imaging, which occurs when the camera faces the surgeon, causing movements with the tool to be appear inverted compared to the hand movements. However, in conversation with the surgeons of the Champalimaud Foundation, this does not appear to be an issue, since the surgeon can move around the patient in cases like this, which ensures the camera always faces the opposite direction. This freedom to move around also impacts the practicality of using foot pedals to ensure hands-free interaction, as the authors suggest, as the surgeon would have to either have the same pedals on multiple sides, or move the pedals around. In this case, exploring foot movement, as proposed by Müller et al. [9], could be more useful, as it would not incur in these situations.

In the work of Prescher et al. [17], the impact of the pointer in a real operating scenario may be lessened due to the fact that target selection is not random but rather contextual, meaning that the following targets may actually be located through description of what is being displayed on-screen. There is also the fact that the pointer is embedded in the laparoscope, which means that to move it, the camera must be moved. This causes the plane of view to change and forces the surgeon to readjust to the new perspective, losing any perceived depth beforehand. It would therefore be more useful if the cursor moved independently of the camera, possibly implemented on the HMD and moved with gestures, or with head tracking for a hands-free approach.

The telestration system of Feng et al. [18], shown in fig. 2.8, was generally well-received, with both trainer and trainee being positive about using it in training, but less so when considering real-world scenarios, as there are a couple of issues that need to be addressed: Firstly, the telestration is on a secondary display, disrupting the trainee's attention. Secondly, trainers complained about the lack of precision in Kinect's gesture tracking, requiring extra time to identify the inaccurate annotations on the monitor and discuss them with the trainee. Finally, the lack of a hands-free approach means the trainer has to put down his tools in order to produce the gestures, pausing the procedure.

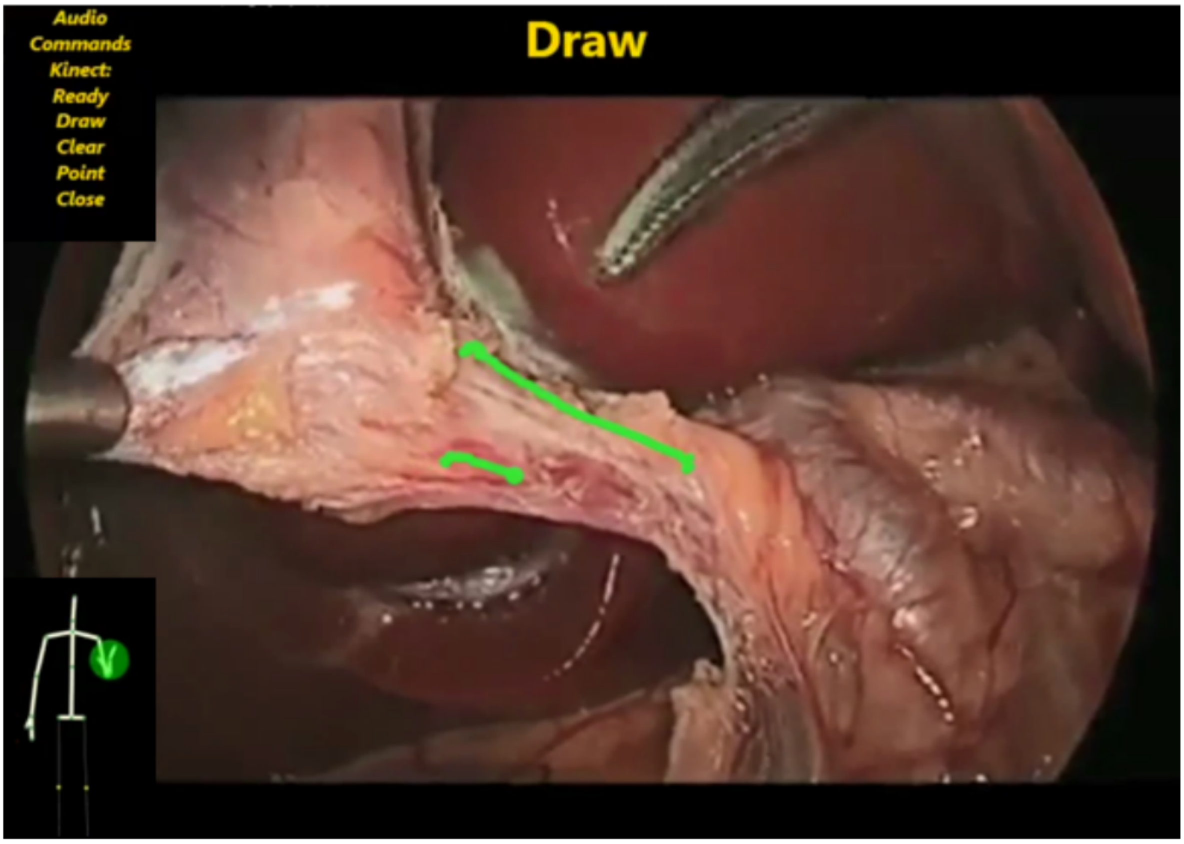


Figure 2.8: Feng et al. developed a telestration system to draw on the laparoscopic video, in an attempt to improve communications. The top left displays the available commands, invoked with voice activation via Kinect, while the bottom left displays the user's skeleton, detected through the Kinect's sensors.

The works of Walczak et al. [19], Maithel et al. [20] and Batmaz et al. [3] seem to support the usage of a HMD for laparoscopic surgery, with the video following the user's head movements. With this, users can assume their preferred head position instead of being forced to look sideways in order to see the video.

In the work of Kihara et al. [22], the surgeon can see his tools by looking down, but the non-see-through HMD blocks his sight, front and sideways, rendering him unable to perceive his surroundings, which may impair collaboration between team members. This can be mitigated by using an optical-see-through HMD instead. Operation time was also longer compared to the normal procedure, as the operation proceeded at a slower pace. However, at 420g, the device proved to be light enough to not cause discomfort even during the longest operation, which took 3 hours and 49 minutes, which proves the viability of HMDs in surgical procedures. The work of Jayender et al. [23] too provides an improvement concerning the visualisation aspect of the operation, with laparoscopic screen placement similar to the work of Walczak et al. [19], while also allowing interaction with patient data, bringing it closer to the user. These two works however, do not introduce anything in terms of collaboration, as they are more focused on a single person environment.

2.4 Overview

We revised work done in the fields of augmented reality interaction, augmented reality in laparoscopy, as well as other surgical procedures, in an attempt to better understand how to tackle the problem of improving the laparoscopic procedure by applying the usage of AR to it. As such, These works have been classified in terms of interaction sources and methods, visualisation methods and communication improvements. This classification is presented in table 2.1.

Of all presented works, in terms of interaction, the hand is the most commonly used body part, with the HoloLens's 'air-tap' gesture being used for control activation [8] [12], while hand movement is used for image manipulation [12] and drawing [18]. Voice commands is also a very popular option for control activation [8] [12] [13]. However, these interaction methods are difficult to use in an operating room. The first, the hand, requires the surgeon to pause the procedure before commencing the interaction, be it an air-tap, gesture or hand movement, which causes interruptions and disrupts the flow of the operation. The second, voice activation, may not be effective due to the already existing noise in the room. Furthermore, its lack of reliability may also cause disruptions in the flow of the procedure.

Other sources of input include the head, with a head gaze being used to select targets [8] [23]; it's a simple, yet effective method. Head movement is another approach, for when the content follows head movements [13]. A more elaborate approach takes the form of the eye gazing [6] [7], which was well-received by users, but may not transition well onto the surgical operating field: these controls would have to be displayed continuously and right in front of the user, unlike in the presented works, which could be distractive for users, but more importantly, it would take valuable space from the HMD's already limited field of view. In terms of feet, two different approaches emerge: using a foot pedal as a means to activate a selected control [23] and using foot movement to select and activate controls [9]. After comparing the two, we conclude using foot movement would be a more flexible choice, as it does not rely on extra hardware that is situated in a given position in space.

Regarding visualisation, most approaches use an optical see-through-HMD [8] [9] [11] [12] [13], mainly HoloLens, but these works are not applied specifically to laparoscopy, but AR and surgery in general. For laparoscopy, non-see-through and video-see-through HMDs are more prevalent [20] [3] [22] [23]. As previously stated, a non-see-through, and to some extent, a video-see-through HMD does not allow the user to perceive his surroundings, which can impair communication significantly, as there is a loss of perception for the rest of the team. Additionally, very few works regarding laparoscopy aim to improve communication [17] [18], with an emphasis being put on visualisation. With respect to using patient data/preoperative data interprocedurally, most works focus on displaying the patient data as 3D imaging rather overlaid onto the patient, instead of imaging such as MRI slices. In conclusion, we find that, as far as we know, there is no work incorporating Augmented Reality in laparoscopy that offers patient data consulting while supporting collaboration between team members.

2.5 Summary

In this chapter we presented some of the works done concerning interaction techniques in Augmented Reality, as well as the application of AR in surgical procedures, including works which delved deeper into the laparoscopic procedure itself, discussing how they fit into the scope of our own. In the next chapter, we will present the analysis of users and the tasks they perform.

Approach	Interaction	Visualisation	Communication
Esteves	Eye gaze		
Velloso	Eye gaze		
Kim	Air-taps Hand Gestures Voice Commands	Optical-see-through HMD	
Müller	Foot movement	Optical-see-through HMD	
Pratt	Air-taps Hand Gestures Voice Commands	Optical-see-through HMD	
Grinshpoon	Head Movement Voice Commands	Optical-see-through HMD	
Prescher		Traditional video monitor display	Pointer Navigation Grid
Feng	Hand movement	Additional video monitor display	Telestration
Maithel		Non-see-through HMD	
Batmaz		Non-see-through HMD	
Prescher		3D stereoscopic monitor display	
Kihara		Non-see-through HMD	
Jayender	Foot Pedal Head Gaze	Video-see-through HMD	

Table 2.1: Overview of related work.

3

User and Task Analysis

Contents

In this chapter, we will present the results of our user analysis, characterising them and the tasks they perform. This analysis was conducted through a user questionnaire, available in appendix A.1.1, filled out by seven members of the Champalimaud Foundation's surgical team. In addition to this questionnaire, we also performed presential observation of four laparoscopic surgeries on one of the Foundation's surgical room, and had conversations with no less than eight surgeons before and after the procedures, who explained us what was about to happen or what took place. During the surgery, nurses would give us more insight about the several stages of the surgery, or what was happening at that time.

3.1 Analysis Results

According to the questionnaire, whose results are available in appendix A.1.2, the surgical team is composed of mostly male members, with only one female surgeon in it, and their ages range from 34 to 42 years of age. Their cultural background is diverse, which elicits surgeries to usually take place in English. During surgery, there are at least six people involved in the procedure, shown in fig. 3.1: A head surgeon, who coordinates the entire procedure, one to four auxiliary surgeons, who mostly observe but also participate in parts of the surgery, a nurse solely responsible for passing the surgeons tools they may require throughout the operation, an anaesthetist keeping track of the patient's vital signs, a nurse supporting the anaesthetist and a circulating nurse. Additionally, a senior surgeon may come in and serve as advisor, providing insight and making remarks about what is being seen on camera.

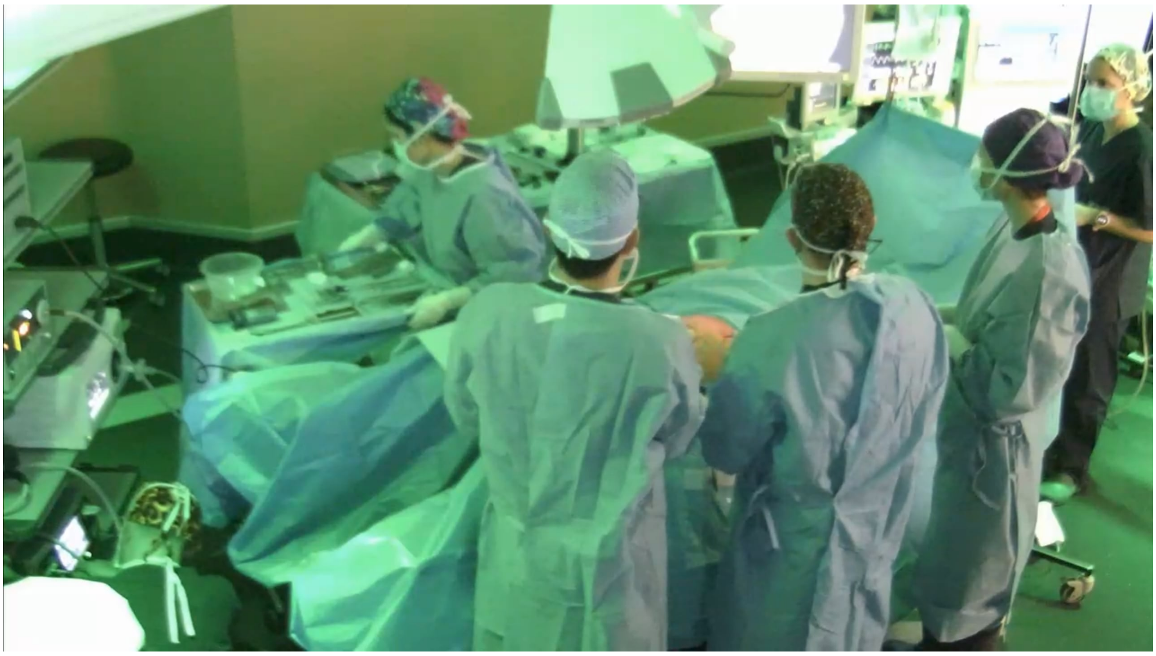


Figure 3.1: The operating room. From left to right: a roaming nurse (bottom left), a nurse responsible for the tools, the head surgeon, two assistant surgeons and the anaesthetist.

Figure 3.2: Top: Locking grasper. Bottom, from left to right: Trocar, laparoscope, locking grasper and needle driver. Notice the difference in the grasper and the needle driver's tips.

A laparoscopy begins long before the surgeons set foot in the operating room: The room must be setup by the nurses, who bring in the patient, who is already anaesthetised, as well as new tools. The patient's stretcher is loaded onto a structure placed on the floor and serves as an operating bed. Nurses cover up the patient in surgical drapes, leaving only the abdominal area exposed. A 1 to 2 centimetre incision is performed and the abdomen is inflated with carbon dioxide to allow surgical tools to be manoeuvred inside the patient. These tools, which enter the patient through trocars, include, but are not limited to, the laparoscope, which captures the patient's insides in a video that is displayed in the room's monitors; scissor, used to cut tissue; surgical mesh, used to support organs and help tissue repair; hook, which has an electrical current and, through triggering a pedal located underneath the operating table, is used to perform cauterisations; locking grasper, which is used to hold structures and move tissue, with the ability of holding its grasp without continuous user effort; needle driver, similar to the grasper but shorter, holds the needle in wound suturing and knot pusher, which is used to tighten knots in suturing. These tools are shown in the images above.

In parallel, surgeons have a preoperative meeting where they discuss their strategy for the surgery and consult patient data such as Magnetic Resonance Imaging (MRI)s (fig. 3.3) and computed tomographies. Consulting this before the operation is important because it is unfeasible to consult radiology data interoperatively. On one hand, the data is extensive, and on the other, it is required of the surgeon to abandon the operating table to, together with an assistant who is asked to come in, sit at a computer inside the operating room and browse the desired images. The surgeon then gives directions as to where to look and when to stop while the assistant handles the computer. Surgeons usually do not browse the images by themselves because, since they are sterilised, each interaction with non-sterile

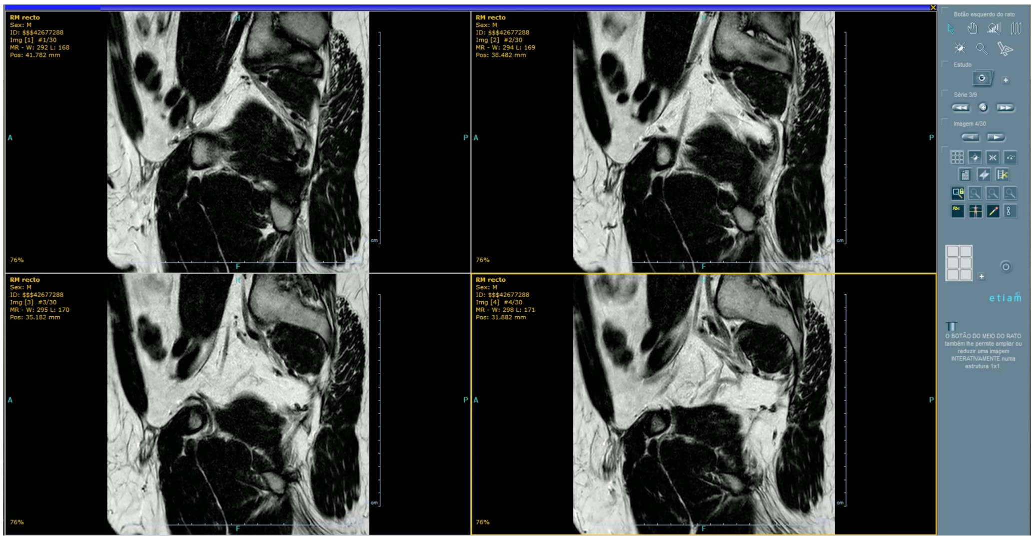


Figure 3.3: MR images consulted during the preoperative meeting.

equipment such as the computer would mean losing time putting on new garments and gloves, wasting additional resources.

The surgery itself is divided into three phases: In the first phase, the surgical team is tasked with locating the target structure, which we observed to usually be a tumour. To achieve this, surgeons must navigate inside the patient's abdomen and clear a path, using their tools, until the target is reached. This part usually lasts at least thirty minutes and can go up to three hours, as the tissue must be carefully cut and cauterised to prevent bleeding and damage. Next, the target structure is extracted through a larger incision, being removed in a more conventional manner, taking at least half an hour as well. In this part, the surgeons are looking down at the patient like in an open surgery instead of using the screens. Finally, in the last part, the remaining structures are placed back in the patient and incisions must be sutured and closed. This final part takes from thirty to sixty minutes. In the questionnaire, surgeons report that laparoscopies can take 2 to 3 hours and that longer operations last for 4 to 5 hours, which corresponds to our in-field observations. Surgeons also report that, in more extreme cases, they can take up to 11 hours, with shifts between personnel. Surgeries may also be aborted: in one of the surgeries we attended, for example, the procedure had to be cancelled because there was an unaccounted structure hindering access to the tumour.

The head surgeon usually starts the procedure and performs the most critical parts in the surgery, while explaining the student surgeons the steps he is taking. Eventually, he transfers control of the tools to a student and keeps guiding the surgery by handling the camera, adjusting its position as the procedure evolves. While holding it, the head surgeon continues his role of instructor, orally issuing commands, such as instructions to perform a cut, a suture, a cauterisation, or to grasp and hold a certain structure before doing something else. The instructor also points at the screen for the other surgeons to

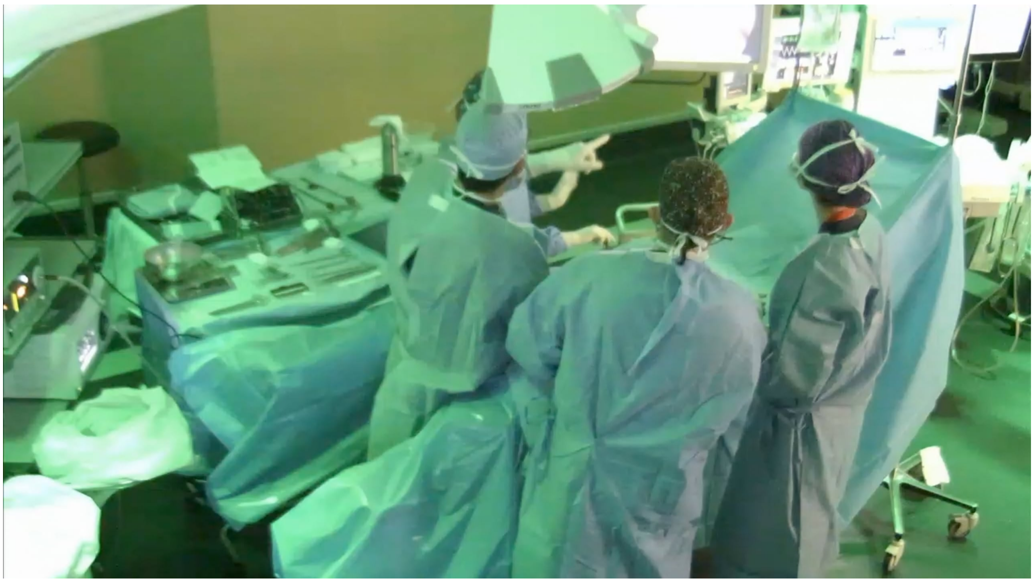


Figure 3.4: Doctors point at the screen to communicate.

understand what anatomical structure he is referring to and makes use of gestures for students to better understand the motion of the tools and envision cutting lines. Sometimes, pointing can also be done with the tools themselves, but even though it may be effective, it is not always correct, because if both hands are occupied, it implies letting go of a structure to point with the tool, or asking someone else to hold it. Additionally, pointing from a distance with the hand is ambiguous at best, as there is no clear way to tell where exactly a surgeon is pointing at, as can be seen in fig. 3.4.

Junior surgeons don't just learn from observing and participating in surgeries. In fact, before they do that, they first learn to do tasks like suturing, cutting and wielding a needle in Fundamentals of Laparoscopic Surgery (FLS) training exercises [24]. These exercises are done with a training box that features holes mimicking the incisions performed on a real patient, shown in fig. 3.5, through which the tools are passed, and include:

- **Peg transfer** - Using two locking graspers, participants must remove a peg from one side of the board with one grasper, transfer it in mid-air to the other grasper, and place the peg on the opposite side of the board, without dropping it in the process. If the peg is dropped, the last hand that held it must be the one to pick it up again. The transfer is then repeated for the remaining pegs on the board. Once they are all transferred, the participant must then transfer them all back to the original side of the board.

- **Precision cutting** - Participants must cut a circular pattern in a 2-ply piece of gauze, using one locking grasper and one pair of scissors, keeping their accuracy within 5 millimetres of the pattern. Participants are allowed to use tools on either side and to switch them.

- **Ligating loop** - Using a locking grasper, an endoloop and a pair of scissors, participants must place the endoloop, which is a pre-tied ligating loop, on a foam structure, keeping their accuracy within one millimetre of a black line drawn on the foam.
- **Suture with Extracorporeal/Intracorporeal Knot** - Participants must use two needle drivers, one pair of scissors and a long suture to close a slit in a penrose drain. A penrose drain is a type of soft rubber tube. The suture must be placed through two circular marks on the penrose, within an accuracy of a millimetre, with a needle and the slit must be closed with three knot throws. Finally, the participant must cut both ends of the suture. On the extracorporeal suture, the knots are thrown outside the box and pushed inside and tightened with a knot pusher. On the intracorporeal suture, however, no knot pusher is necessary, as knots are thrown inside the box and tightened with the graspers.

Unlike the real patient, however, the box can steadily hold a laparoscope, as the surgeon wielding the tools is usually assisted by another surgeon holding the camera. Another big difference is that there are no organs constraining tool movement, so both visualisation and movement are much easier. In real circumstances, while surgeons usually spend most of the operation in the same place, depending on the context of the situation they sometimes need to adjust their positioning in order to hold the tools in a more ergonomic fashion, but with the FLS training box, exercises can be completed while standing still.

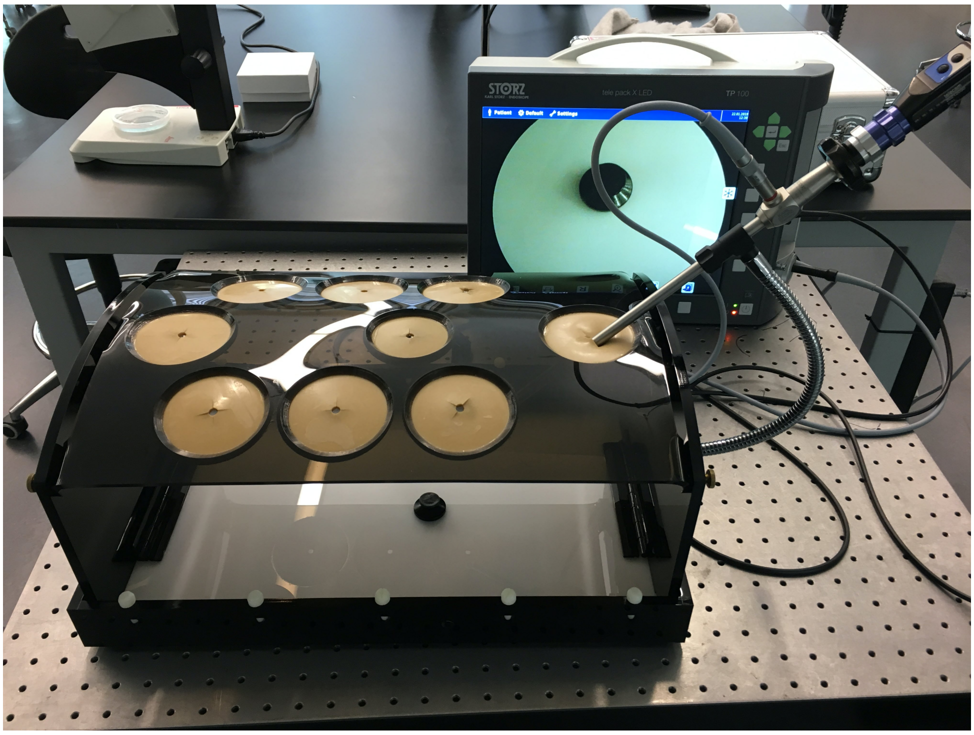


Figure 3.5: A training box with a laparoscope, whose video can be seen on the screen, is inserted through a centre hole and doesn't need to be held.

Figure 3.6: The inner area, with the computers tracking vitals signs to the left and a column with additional equipment to the right, where the pedals are also stored. The ceiling structure holds screens and auxiliary lights. A large screen embedded on the wall can be seen in the background.

Surgeons perform laparoscopy in an operating room with reduced lighting conditions, like the one in fig. 3.6. When the ceiling lights are turned off, green ambient lights help with visibility; this colour was chosen after multiple tests with different colours. When the surgery reaches the second stage and the laparoscope is not used, the room stays darkened, but auxiliary lights are used, as traditional ceiling lights are only used during preparation and turned on again when the surgery is finally over. In the operating room there is a perimeter delimited on the ground that separates two types of areas: the first is a restricted area, where the patient is located and only sterilised personnel, such as the people mentioned above, may enter. This area is also where the most critical medical equipment is located: A ceiling structure holds auxiliary lights, as well as six monitors which surround the patient, so surgeons can see the video feed from the laparoscope from all angles, although they usually all look at the same one in order to better communicate and understand each other. These monitors are adjustable, so they can be folded when the surgery is over to reduce clutter, or better positioned so the entire team can see it. Other equipment also includes the instrument tray, a column with auxiliary light controls and equipment where the laparoscope is plugged into, and monitoring equipment with four additional screens to keep track of the patient's vital signs. The heart rate monitor emits a periodic beep that is disturbed by the cutting tool, which causes interference by provoking a continuous sound akin to a lack of pulse. The outer area is where the nurses roam around and watch the operation; this is the area where non-sterilised personnel stay. The aforementioned senior surgeon stays here as well, making use of a larger screen that is embedded in the wall to give counsel to the operating team. This area also features some computers which are responsible for the room lighting and recording, as well as a whiteboard where data about the patient and the operation are included, such as age, gender, and tumour location. Lastly, there is a digital clock on the wall near the embedded screen, so surgeons can keep track of the procedure's duration.

Figure 3.7: On the left, the da Vinci Surgical System in the operating room. On the right, the surgical team calibrates the robot.

Aside from traditional laparoscopy, there is also robot-assisted laparoscopy, seen in fig. 3.7 and fig. 3.8. This type of laparoscopy is a bit different to the traditional method in the sense that two surgeons control a da Vinci Surgical System via consoles located in a separate room, and two people stay with the patient: an auxiliary surgeon who handles the laparoscopic instruments, loading them on and off the robot throughout the procedure, and a nurse responsible for passing the surgeon the instruments. The head surgeon communicates with the assistant via the da Vinci robot by placing a cursor on the video, in relevant locations pertaining to the task at hand, which might be, for example, an incision, or a suture. Communication also takes place orally, especially when the head surgeon wants to check in with the team in the operating room.

Robot-assisted laparoscopy is superior to traditional laparoscopy in terms of precision and comfort on the surgeon's part, as they can operate sitting down, and view the laparoscopic video in stereoscopy, but these benefits come with an enormous drawback: performing robot-assisted laparoscopy is currently very expensive, with each surgery costing around five to six thousand euros in maintenance, which makes it non-viable in terms of a cost-benefit ratio. Currently, robot-assisted laparoscopies account for no more than five percent of total laparoscopies performed, as they are mostly used in prostate surgery, where there is little room available to move instruments and little room for mistakes as well, as one could cost the patient his fertility. Robot-assisted surgery is also used in low rectal surgery, albeit as a means of practice for surgeons rather than out of necessity like the prostate one.

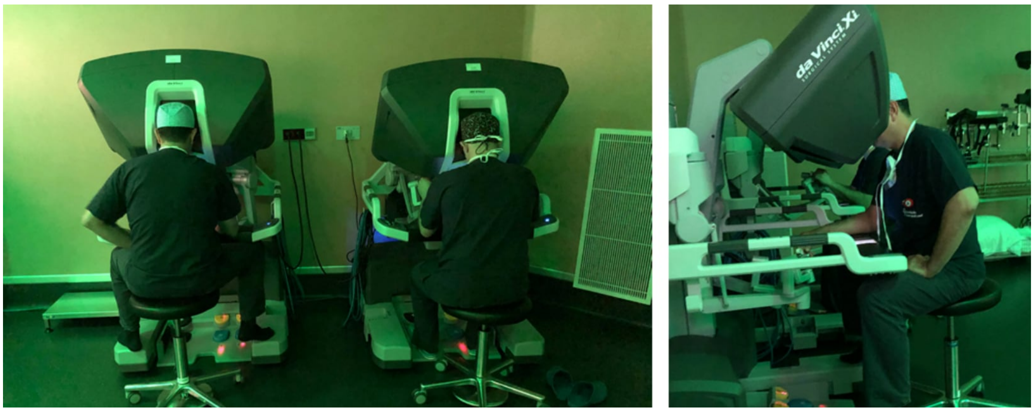


Figure 3.8: Surgeons operate with the da Vinci in a separate room, and as such they do not need to be sterilised.

As stated in section 1.2, laparoscopy is an intensive process, not just mentally but physically as well. The procedure is already very demanding in itself due to surgeons having to expend extra mental effort thanks to a lack of hand-eye coordination that is caused by indirect visualisation. That effort extends to the physical plane when we consider that they have look at the screen all the time, which places a continuous strain on their necks. Additionally, surgeons also currently face problems in communication. In fact, according to the inquired surgeons, just as they complain about difficulty in maintaining proper posture, so do they complain about not being able to let other surgeons know what part of the video they are pointing at, or to understand what others are pointing at as well.

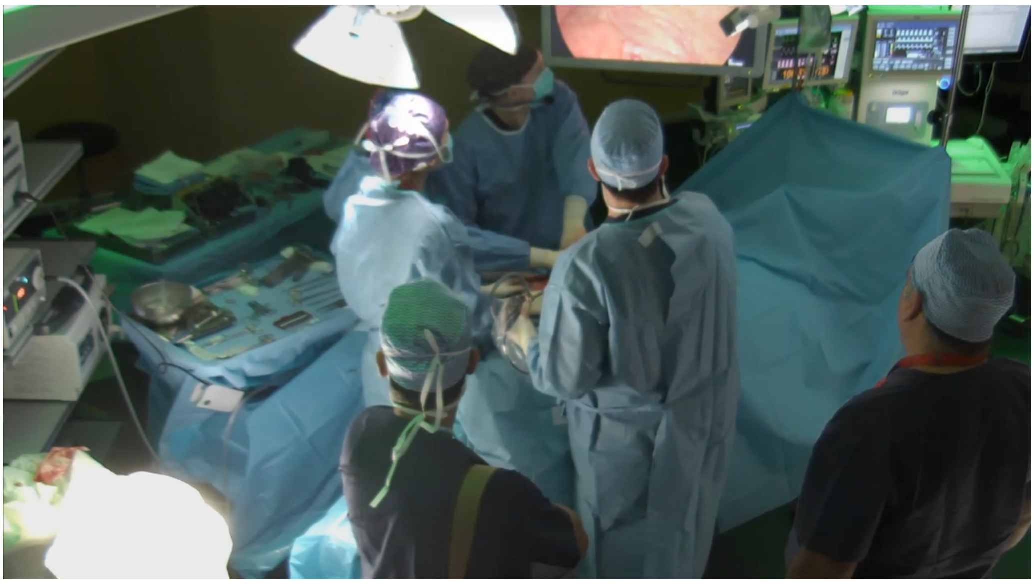


Figure 3.9: All doctors look to the same screen; sometimes, this means having to assume an uncomfortable position. On the right, we see the senior surgeon.

3.2 Design Requirements

Performing user and task analysis allows us to better understand the existing problems in the procedure of laparoscopy, while identifying several constraints and design requirements, which the solution will have to follow in order to address those problems. These problems, their requirements and the proposed solutions are presented in table 3.1.

3.3 Summary

In this chapter, we have presented our analysis of users and their tasks, performed through presential observation of laparoscopies, conversation with surgeons and medical personnel, and dissemination of questionnaires. From that, we have extracted problems, identified design requirements and proposed solutions for those requirements. In chapter 5 we present our approach to addressing these problems, while describing the design process that took place, up to its current iteration. Before that, in chapter 4 we will first perform a preliminary evaluation in order to assess whether using a HMD results in a difference in performance compared to using a monitor.

Problem Statements	Design Requirements	Design Solution
Visualising the laparoscopic video during extended periods of time is exhausting for the neck.	The solution should allow the user to adopt more comfortable neck postures instead of forcing the user to look to the side to see what the other surgeons are seeing.	Following display: The laparoscopic video follows user head movement, so users can look around and assume a neck posture that is more comfortable for them.
Current interactions surgeons have, such as pointing or consulting patient data, require them to let go of their tools, which interrupts the procedure.	Surgeons should have hands-free interactions in order to operate in an uninterrupted fashion.	Hands-free interaction: Every interaction is either done with the head or using the feet.
Browsing patient data interoperatively takes too long because it requires to call in an assistant, who browses the images for the surgeon.	Users should be able to look at patient data by themselves, without interrupting and adding extra time to the surgery.	Patient data image browser: users can look to the side to see and browse MR images from the patient.
Users may have to move around the patient in order to adopt better positions to hold their tools.	Interaction using the foot should not rely on pedals, as these would need to be moved around to cope with user movement.	Foot browsing: Users can use the foot to navigate the patient images, rotating it on its heel to change images faster or slower.
Pointing is unclear and ambiguous: different users have different interpretations of where a surgeon is pointing at.	Users should be able to point precisely and understand where other users are pointing at, regardless of position in the operating room.	Pointing reticle: users can place a reticle on both laparoscopic video and patient images, controlling it with head motion. This cursor is visible on other users' headsets.
Surgeons operate in a crowded area, as they are usually very close together.	Augmented space should present information close to the surgeon to prevent it from appearing intersected with a colleague.	Close quarters: Positioning of interface elements is no further than at an elbow's reach.

Table 3.1: Problem statements, design requirements and design solution for our prototype.

4

HMDs versus Monitor Displays: A Preliminary Evaluation

Contents

Our initial goal was to first assess how Augmented Reality displays could improve laparoscopic procedures in terms of lower duration, reduced physical stress and consequently less fatigue on the surgical team, without compromising task performance. To do this, we performed a preliminary evaluation with six users to evaluate the effect of an HMD on task time. These users are laparoscopic surgeons working for the Champalimaud Foundation.

4.1 Hypothesis

For our preliminary evaluation, we present our null and alternative hypotheses as follows:

Null hypothesis - There is no difference in terms of time between using a HMD and a monitor display.

Alternative hypothesis - There is a difference in terms of time between using a HMD and a monitor display.

4.2 Experimental Setup

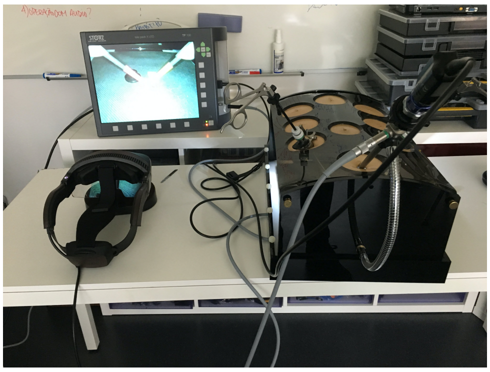
The preliminary evaluation sessions were conducted in an office inside the Champalimaud Foundation's building. To perform the laparoscopic exercise, standard laparoscopic tools were used in conjunction with a custom training box. To visualise the laparoscopic video, we used a Storz monitor display, illustrated in fig. 4.1(a), positioned at a distance, for the first condition. For the second condition, we used the Meta 2 HMD, shown in fig. 4.1(b), which merely displayed the laparoscope's video as well, without any kind of AR enhancement. Start time and stop time were announced orally.

4.3 Methodology

Before the evaluation began, users were explained the context of the session, as well as its goal, which was to measure user task performance under two conditions. Users were asked for permission to record the session for posterior data analysis.

For each condition, users performed five repetitions of a needle thread exercise. This exercised is used for training by surgeons of the Champalimaud Foundation in addition to the FLS training exercises mentioned in chapter 3, and consists in driving a needle through all holes in a board.

To help reduce possible bias associated with the order in which participants perform the task under which condition, the order users performed the tasks was dictated by a Latin square design: If the first user performed the tasks with the display monitor first, the second user would perform the tasks using first the Meta 2 headset.



(a) Experimental Setup

(b) Testing

Figure 4.1: The experimental setup for the preliminary evaluation.

4.4 Results and Discussion

We collected metrics on time taken (in milliseconds) and number of movements for each task. For each participant, we obtained the mean values for each of the two conditions, as we observed that task time decreased with each repetition. A paired-samples t-test was used to determine whether there was a statistically significant mean difference between the time taken to perform the exercise when participants used the Meta 2 HMD compared to the monitor display. Data are mean $\pm$ standard deviation, unless otherwise stated. There were no outliers in the data, as assessed by inspection of a boxplot. The assumption of normality was not violated, as assessed by Shapiro-Wilk's test (p=.741). Participants took less time to perform the exercise when using the Meta 2 (115.113 $\pm$ 31.972 s) as opposed to the monitor display (118.096 $\pm$ 64.909 s), a statistically insignificant decrease of 2.982 (95% CI, -47.178 to 41.212) s, $t(5) = -.173$, $p = .869$, $d = -.07$. The mean difference was not statistically significantly different from zero. Therefore, we reject the alternative hypothesis and fail to reject the null hypothesis.

Based on these results, we can conclude that using a HMD is neither better nor worse than the current procedure in laparoscopy.

4.5 Summary

In this chapter we have conducted a test to determine whether there is a significant difference between using a display monitor and using a HMD for laparoscopy. Our participants performed five repetitions of a needle thread exercise, for each condition, in a Latin square arrangement to prevent bias and, in the end, we concluded that there was no statistically significant difference between the two conditions.

The next chapter will introduce our prototype, which improves upon the current laparoscopic procedure by presenting important information to the surgeon in an augmented space.

5

The Prototype: A multimodal interface for minimally invasive surgery

Contents

Our developed prototype aims to solve the problems we observed in chapter 3 that surgeons currently experience during laparoscopic surgery. Furthermore, we wanted the prototype to be as unobtrusive as possible in order to make the surgery as uninterrupted as it could be. To achieve this, in developing the prototype, we took a fully hands-free approach so that surgeons do not have to put down their tools at all, allowing for a continuous surgical experience.

5.1 Laparoscopic Video

As stated in table 3.1, laparoscopy currently faces the glaring problem of monitor positioning. During surgery, screens are usually placed far away and at a uncomfortable angle, causing neck and eye strain over the course of a surgery, especially if it drags for longer periods of time. Given this, it was important to allow the surgeons some freedom in how they want to see the video, which led to a conclusion: The video, while visible, should follow user head movements so users do not have to reposition it in the augmented space, should they feel the need to assume another posture with the neck. We therefore implemented this in our prototype, illustrated in fig. 5.1, centring the video in the HMD's display and making it as large as possible without it extending beyond the borders.

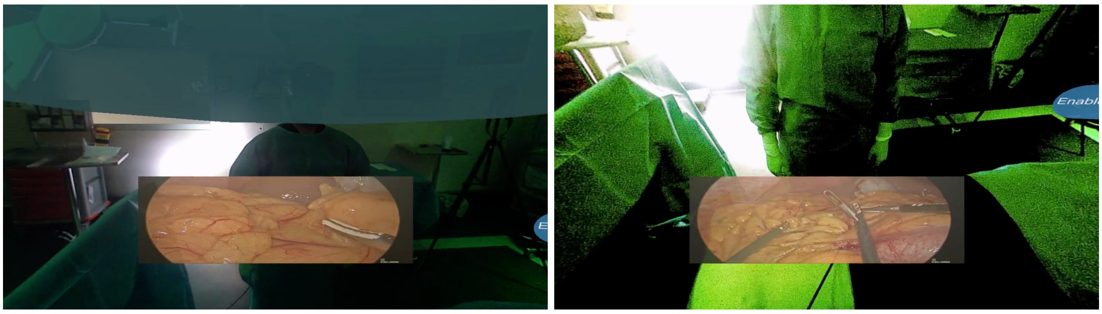


Figure 5.1: The user can change head positioning and still be able to see the video.

5.2 Patient Imaging

To continue the idea of content following the user, the initial approach consisted in using head gestures to activate different features. A downwards gesture would activate the patient data, an upwards gesture would activate the video and side gestures would trigger pointing functionalities. This would allow the surgeon to be positioned in any direction and still be able to look at the desired content. The idea would be that users could pick up on how to use the prototype by recalling the swipe gesture on smartphones.

To start implementing this, an example scene provided by Meta was adapted. In this example scene, cubes would change colour when gazed upon. In the adaptation, shown in fig. 5.2, the prototype would measure how many lines along a Y or X axis were gazed under a given amount of time in order to activate a given functionality. However, upon experimentation it was noted that these gestures either activated too easily by accident, or required too much effort and were thus uncomfortable to use. Additionally, while horizontal movements were not as uncomfortable as vertical ones, they displaced the HMD due to movement inertia, which meant the HMD had to be repositioned by hand after each head gesture. Finally, we also concluded these gestures would be awkward to make if the head was already assuming a certain position, for example, it would be difficult to make a *swipe left* gesture if the user was already looking to the left.

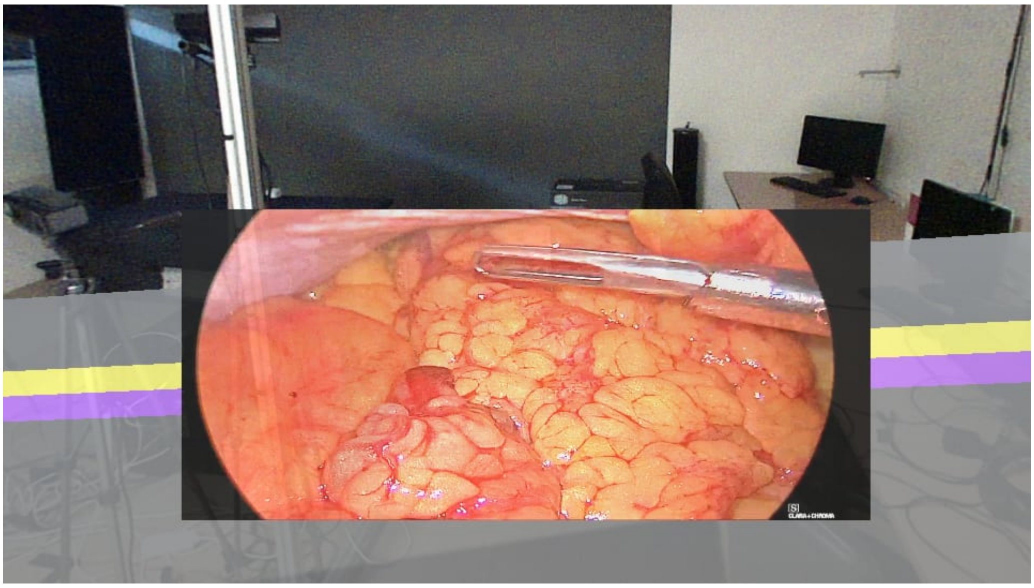


Figure 5.2: Adaptations to the example Meta scene. The idea was to activate a given number of lines under a given time threshold to switch between video and patient data.

It was then hypothesised that the patient data could be accessed by looking to either side, while the laparoscopic video could be visualised by looking forwards. We still intended to have the video follow the user's head movements, so they could look wherever they wanted or felt most comfortable and still be able to see the video and, in particular, take advantage of the headset's transparent screen and look at their hands as well. We therefore proposed the following: the user could look around and move their head within a given amplitude that the video would continue to be displayed, but after crossing that amplitude the video would be hidden to allow the patient data to be visualised. This idea is illustrated in fig. 5.3.

In an attempt to mimic the software surgeons already use in preoperative planning, as well as evoke a feeling of familiarity, two planes would be displayed, while the third would be able to be accessed with a control underneath the images. The patient imaging was initially meant to follow user head movements as well, but the idea was scrapped in favour of having them placed in the augmented world, as surgeons usually consult two images at once, and presenting both images at the same time meant these had to assume a smaller scale in order to fit the Meta 2's narrow field of view. By fixing them on the world, they can assume a larger scale, which means they can be viewed in greater detail, but still somewhat simultaneously by placing them close together and allowing the user to look around.

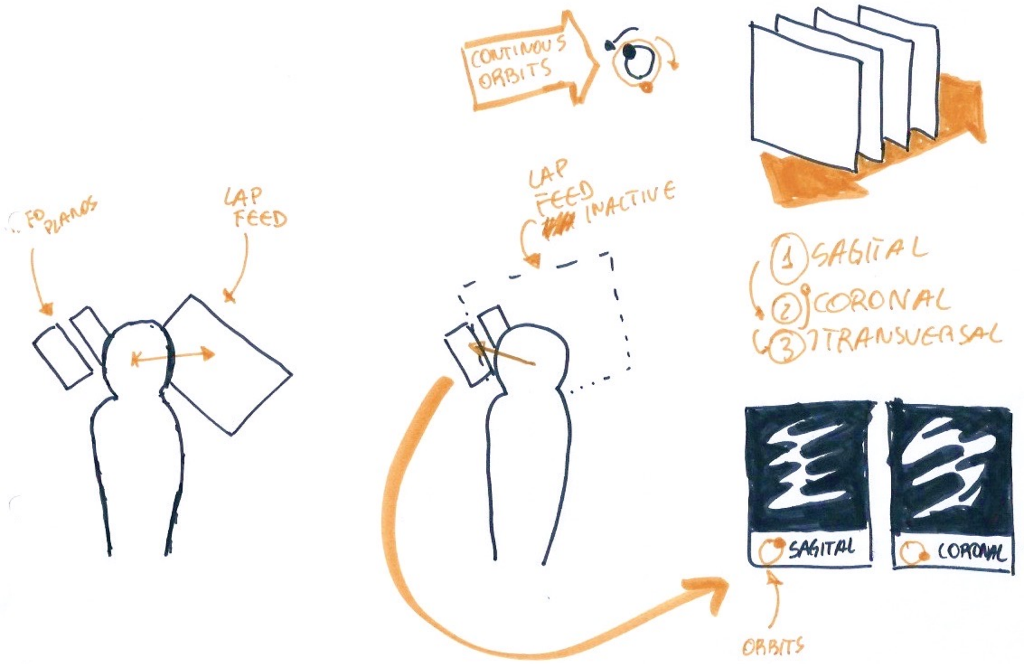


Figure 5.3: A concept of how the video and the data would be displayed. Because the video follows head movements, looking to the side to visualise patient data would hide it.

We thought about where to place the images, since we had the left and the right side available. Eventually, we settled on showing the images on both sides, letting users access medical imaging by looking to whichever side they found most comfortable. Following in the footsteps of Jayender et al. [23] and Walczak et al. [19], we positioned the images at an angle of 30 degrees below eye level, which also lets users look at the rest of the surgical team, as well as their surroundings.

With the placement decided, it was conceptualised that the image data sets could be navigated using arrows. The Meta 2 makes it possible to detect where the user is looking at with the head thanks to its accelerometer and gyroscope, which we used to activate these arrow controls. They were placed above and below each image, with the upwards arrow used to navigate to the next images and the downwards arrow used to navigate to previous images. These arrows reacted differently according to the area which was being gazed upon, as they were divided into four equally-sized sections, representing four different levels of speed. If the user gazed more to the bottom of the arrow, its corresponding section would be highlighted in a cyan contour and the image would change very slowly, at one image per second, but if they gazed to the last section at the pointy end, the images would change far quicker, at 16 images per second, which allowed a more swift traversal of the data set, while the second and third sections changed image at 4 and 8 images per second, respectively.

To aid in navigation, some User Interface (UI) elements were introduced: a progress bar informs users on how far they are in the data set and a label informs which plane is being viewed, as well as what is the current image's number. These elements were placed above the images for better organisation. A button below the images was also introduced in order to allow switching either of the two displayed planes to the third one. Finally, a small cross-shaped navigation reticle was implemented to facilitate target acquisition, helping the user understand where they are aiming with their head and to help them activate the interface controls.

This approach had an issue, however. Owing once again to Meta 2's narrow field of view, it was not possible to display an arrow along with its corresponding image, which in turn made it impossible to look at the data set while it was being browsed, as illustrated in fig. 5.4. In an attempt to mitigate this problem, a miniaturised version of the progress bar and the label were implemented next to the reticle. These elements would appear whenever the cursor pointed at an image or one of its arrows and this way, the user would always know which image and plane was being viewed. Unfortunately, not only did this not solve the issue, as users stop scrolling based on what they see on the images and not the image number itself, it also added more visual clutter, which called for a different approach, as making the arrows smaller would make them far more difficult to target, while not really reclaiming much space back for the user to actually see the images.

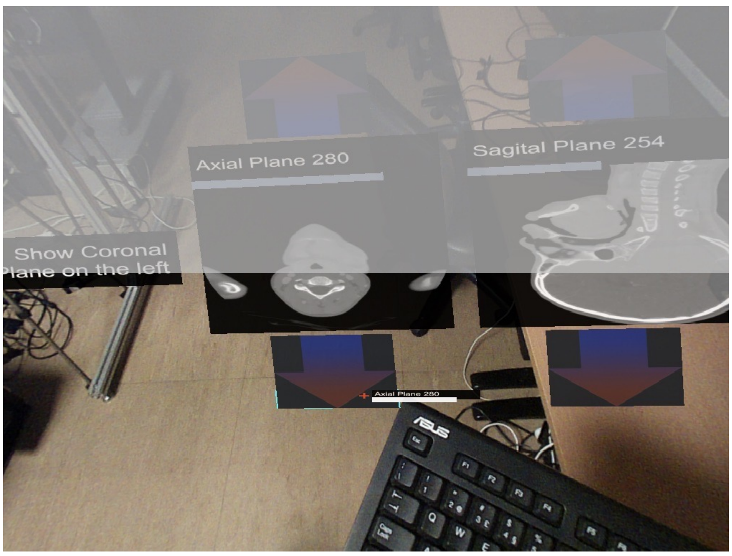


Figure 5.4: The problem with using arrows was that it became impossible to look at the image in its entirety. Note: The greyed out area represents an area that is not viewable when using the Meta 2, only on the application window.

We mentioned in table 3.1 that the interaction had to be hands-free, and keeping that intention we considered using the users' feet as a source of input. As stated in table 3.1, we could not rely on USB pedals, as we observed that not only surgeons could move around the patient, which meant the pedal had to be moved around as well, but they also had other pedals underneath the patient's bed, which could induce errors. Given this, we drew inspiration from the work of Müller et al. [9]: this work revolved around tapping with the foot to activate controls, which we deemed not viable due to the risk of activating a pedal inadvertently, but gave us the idea to implement image scrolling simply through the use of foot movements. Several ideas were initially explored, involving vertical and horizontal movements akin to swipe gestures, but these were dismissed as user balance was put into question because it required lifting the foot from the ground. We eventually settled on heel rotation, as it enables users to keep their balance and does not require them to lift their feet and balance themselves. Using heel rotation, turning their foot like a dial, users can rotate the foot to the left to access previous images and to the right to access the next ones, as shown in fig. 5.6(a) and fig. 5.6(b).

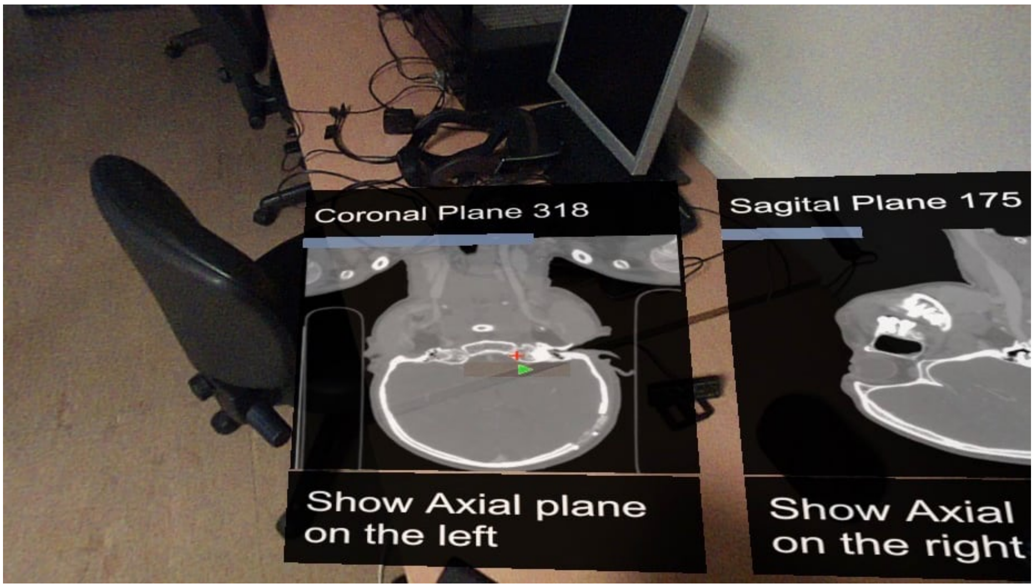


Figure 5.5: Abandoning the use of arrows in favour of foot detection resulted in a much less cluttered interface.

The more the foot was rotated from the starting point, the faster the images would change. However, this would only happen when the user was looking at the image, to prevent errors or accidental triggers, as well as let users move their feet freely when not using the mechanism and not feel constrained. Furthermore, to prevent the same feeling of constraint when looking at an image without the intent to change it, a dead zone was implemented, allowing for a small amount of movement to be done with the foot without resulting in image change. After that dead zone, images can be changed in three different speeds, represented by three differently-coloured icons in the interface: a green icon represents the image changing slowly at two per second; a yellow icon represents the image changing at a medium pace of four images per second; a red icon means the images are changing rapidly at ten per second. It should be noted that these changes happen continuously, meaning that the images change one by one instead of jumping after a full second, which helps giving the user a sense of movement and fluidity. This sense of movement is also enhanced by having the displayed icon fade in and out. These icons are positioned in a curved bar in accordance to the detected foot position, and the bar in turn is positioned around the reticle for accessibility, following the concept of indirect interaction presented by Müller et al [9].

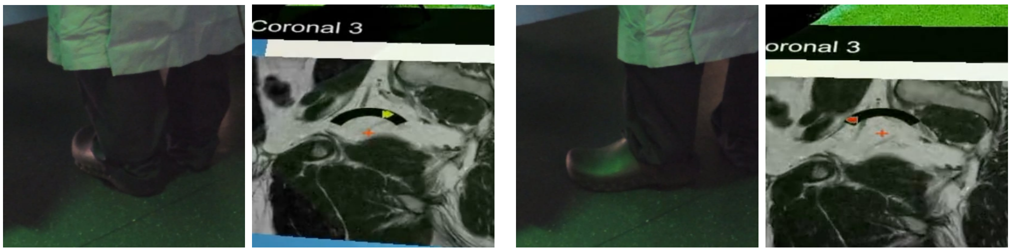


Figure 5.6: The curved bar helps users perceive the rotation their foot is assuming. Sensitivity was adjusted so users could reach both red icons without lifting the foot.

To implement foot movement detection, a mouse was introduced below the rubber clogs used by the surgeons. We first used a wired mouse, as they can have a significantly smaller form factor compared to wireless mice, to assess whether the rubber clog's sole could house it, before moving on to a wireless one, shown in fig. 5.8.

Figure 5.7: Before going straight to the wireless mouse, we wanted to test the feasibility of embedding a mouse on a clog with the smallest mouse we could find. We can observe the wireless mouse requests a larger portion of the clog, but as seen in the third figure, it does not interfere with its silhouette.

We decided to complement this interaction method and made the progress bar interactive, allowing it to change images when intersected by the user's gaze, akin to an application window's scroll bar. This would let the user go through the whole image data set quickly, roughly obtaining the desired position and then fine-tuning it with the foot. We also took into consideration that the user might accidentally hit the bar with the gaze, especially when looking up to activate the pointing mechanism, which we will refer in section 5.3. In order to prevent these undesired image changes, we developed a script, which we present in section 5.5, that lets the prototype understand whether the user has hit the bar by accident or intentionally. Finally, we considered that surgeons should all be on the same page when it comes to patient MR images, and so we implemented a synchronisation feature that, when a user changes images or planes, these changes also occur on other users' HMDs as well, preventing users from pointing at an image and each seeing the reticle placed on a different image.

We also attempted to replace the plane-switching buttons with foot gestures. For this, we thought of the works of Esteves [6] and Velloso [7] and tried using circular foot gestures to activate these controls, as illustrated in fig. 5.3. A clockwise rotation would change the plane on the right, and a counter-clockwise one would change the plane on the left. While the prototype recognised these movements well enough on their own, such was not the case when coupled with the heel rotation, as it would most times recognise a foot rotation in place of a heel turn and vice-versa. Thus, the gesture ended up not being implemented and the buttons were kept.

5.3 Pointing reticle

Because the physical screen is removed in lieu of the HMD, users lose the ability to point at it. Therefore, a new mechanism for pointing was necessary as well. Remembering the principle of hands-free interaction we elicited in table 3.1, we opted to use the head gaze as means of pointing, as a similar type of interaction had already been experimented with in during the development of the patient data browsing functionalities, with the reticle used for interface navigation being stable enough to be used as a pointing reticle as well. We chose not to use voice activation because we consider it to be an unreliable activation method, which is exactly what we want to avoid in a critical situation like a surgical context. Works like those of Kim et al. [8], Feng et al. [18], Pratt et al. [12] and Grinshpoon et al. [13] used voice commands, but these works did not evaluate how effective the use of voice commands were. Furthermore, we do not consider this type of interaction to be as fast as simply pointing at a control with the head.

When using the pointing reticle on the video, the video would be fixed in the augmented space so it could be properly pointed at, being placed at a centred position, instead of following the user's head movement, and while pointing at patient images, the pointing reticle would replace the red navigation reticle.

The pointing reticle is cyan-coloured and, in addition to the cross already present in the red reticle, features a coloured circumference. These differences let the user immediately know whether they are

Figure 5.8: While pointing, the video is fixed on the augmented space. Additionally, it is possible to interact with the images while pointing.

pointing or not and have greater contrast with the laparoscopic video, whose frames are comprised of mostly red and orange colours. On the other user's HMD, the reticle appears in a bright neon green, which also highly contrasts with the video, and helps distinguish which reticle belongs to which user, in the case both are pointing at the video.

For the reticle's activation, we decided to explore the only direction that was not yet in use: the space above both the video and the patient data. Looking up past a threshold and holding the gaze there for a brief moment would activate the pointing reticle, regardless of where the user was looking at, video or MR images, and looking up again would disable it. After experimentation, we considered the amplitude of 20 degrees above eye level to be an adequate value for the threshold, as it was reachable without activating by accident or requiring too much effort.

Despite the flexibility present in this interaction method, however, we considered the possibility that this movement could be too uncomfortable or just not practical for users to perform, and therefore we also implemented a small virtual button. This virtual button sits between the video and the patient data on both sides, so it would be easy to access, whether the user is looking at the video or the images.

5.4 Colour

5.4.1 Interface theme

When developing the prototype, the buttons for switching planes were found to be easily mistaken for simple labels, as it was not clear they were interactive. This can be observed in fig. 5.4 and fig. 5.5. Furthermore, there was also not clear for users that the navigation bar actually reacted when they gazed upon it. This was attributed to a lack of colour in the system, as everything was black and white. To solve this issue, and to let users easily understand what is interactive and what is not, interactive objects were coloured in blue, while informational text was left in black labels. Blue is commonly associated with calmness and trust, which should be critical to have in an operating environment.

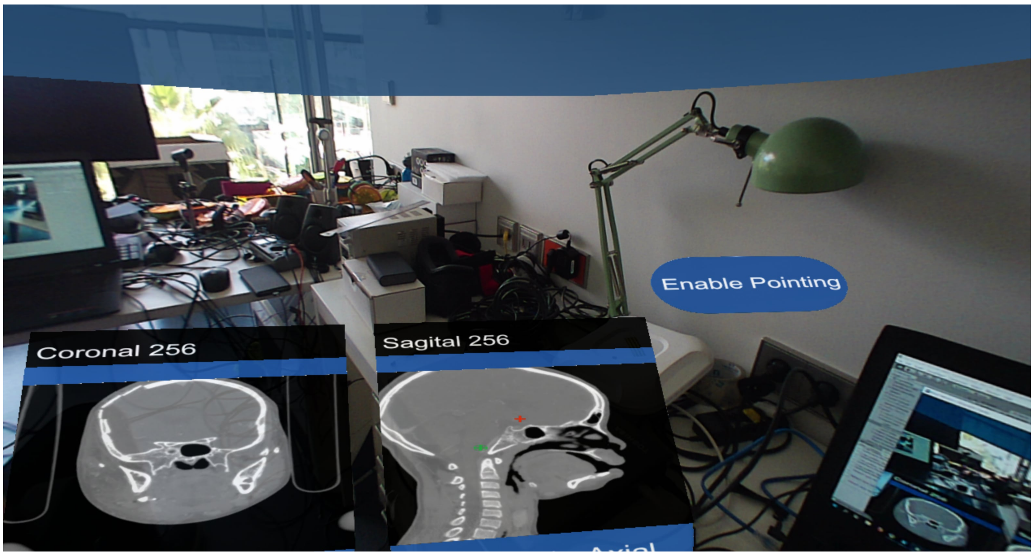


Figure 5.9: The button for pointer activation was positioned in a more upwards location to prevent being accidentally activated while the user transitions from the video to the images and vice-versa. The top of this picture contains the area which also activates pointing if the user gazes upon it. Additionally, the green pointer is present, indicating another user is pointing at the patient data.

5.4.2 Reticles

As stated in section 3.1, during surgery the room assumes a green tint, and therefore it was important to choose colours which would contrast with it. For the navigation reticle, which is always visible whenever the surgeon is not looking at the video, we chose to use red, as it directly contrasts with green. For the pointing reticle, however, the problem is more complex: not only it does it need to contrast with the room, it also needs to contrast with the video, which assumes tones of red and orange. We deferred to the video and chose two colours which would contrast with it: green and cyan. More specifically, green for the other user's reticle and cyan for the user's own reticle, which matches the blue chosen for the rest of the interface. We also added a black outline to these reticles to assure better contrast when viewing them against the greyscale patient images.

5.5 Prototype Architecture

In this section, we will present the prototype's architecture and go over its components in greater detail. The developed prototype is a Microsoft Windows 10 application developed in Unity 2017.4 using two hardware components: a Meta 2 HMD and a HP Z3700 wireless mouse embedded inside a rubber clog. The Meta 2 is a see-through Augmented Reality HMD which, due to its tethered nature, allows for usage during extended amounts of time, therefore lasting throughout an entire surgical procedure, regardless of its length.

The architecture is as follows, and is illustrated in fig. 5.10. Red is for hardware, blue is for image visualisation, green pertains to patient image viewing, yellow is for the pointing mechanism and brown is for the networking part of the prototype.

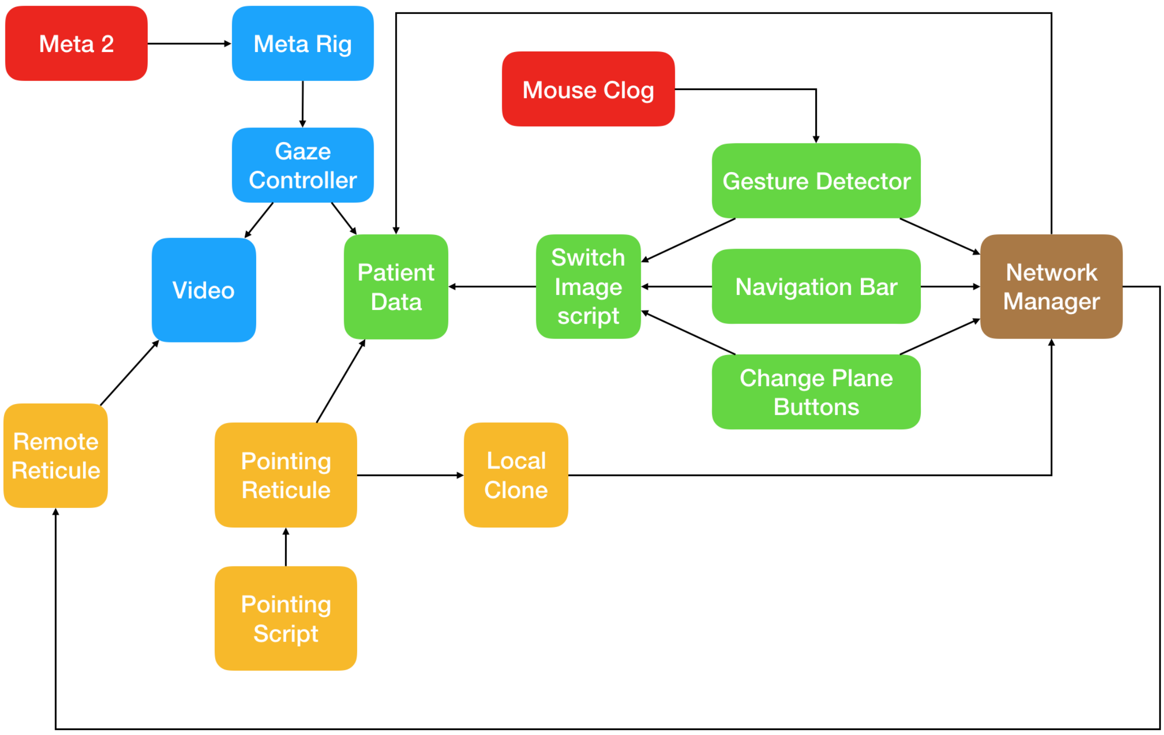


Figure 5.10: Conceptual overview of the prototype. Arrows indicate information flow.

Meta Camera Rig - The *MetaCameraRig* Game Object is a prefab made available by the Meta SDK and establishes the connection between the Unity application and the Meta 2 HMD. On one hand, it receives data such as position and rotation, which is then made available to the application through the prefab's transform, while on the other, it handles the scene cameras, being responsible for what is shown to the user via the display.

Gaze Controller script - This script is a simple, yet central part of the prototype, reading the rotation values of the Meta 2 in order to determine whether the laparoscopic video should be displayed or hidden to show the patient MR images. The threshold value after which the video is hidden is of 50 horizontal degrees for either side, granting the user the possibility to move their head with the video following the head motion in a range of 100 degrees. There is another threshold of 20 vertical degrees when looking up which, after crossing it for a short period of time, toggles the pointing mechanism.

Video Screen - This Game Object displays either video from a USB camera or from an existing file, if needed for demonstration purposes. The prototype displays the video from the laparoscope on the

HMD's transparent visor in an aspect ratio of 16:9, while making as much use of the Meta 2's screen real estate as possible.

Patient Data - There are two of these objects, one for each side, and they hold Game Objects and Canvas elements related to displaying information pertaining to patient imaging, such as labels with the current plane and image index, navigation bars, *change plane* buttons, as well as the patient MR images themselves.

Switch Image script - This script is responsible for loading the patient images from disk and displaying them on screen. It also holds all state pertaining to displaying the MR images: this includes current plane and index for both left and right image, number of total images per plane and path to the images. Finally, it also manages the navigation bar's appearance.

Navigation Bar - The navigation bar sits between the image and its label in the UI, and its functions are two-fold:

1. Show the user how far in the plane he or she has navigated, using the ProgressBar script. This script dynamically updates the Navigation Bar's size according to the current image being displayed. The need to perform this resizing programmatically stems from the fact that Unity is unable to simply scale objects on one side only. Thus, a rescale and reposition is required every time the next or previous image is loaded.

2. Allow for image navigation using the head gaze, with the BarNavigation script: the Navigation Bar is divided into ten equally-sized, but invisible segments. The choice to go with segments in lieu of casting a ray from the camera to the object, calculating the hit's local position and obtaining the corresponding percentage resulted from verifying in earlier experiments with the arrows that, although the HMD works reliably well with objects tightly stacked vertically, such was not the case horizontally, as the gaze could not be held on one object consistently, which would translate to images quickly jumping back and forth, especially in larger data sets. Thus, using ten horizontal but large segments was a preferable alternative. These segments correspond to a given percentage, ranging from five to ninety-five. Activating these segments loads the image on the index equivalent to the percentage attributed to said segment. Additionally, there is a safety feature concerning the activation of these targets: to prevent unintended activations, and to distinguish accidental gazes from intentional attempts at interacting with the bar, a small timer was implemented, waiting for the user to initially hold the gaze on the bar for 0.2 seconds before it becomes responsive. After this, they may drag the gaze left and right and the bar responds immediately, loading the image associated with the activated control. The user may also accidentally break gaze contact with the bar, and to prevent him or her from having to suffer that 0.2 second delay, which may lead to

feeling that the system is unresponsive, or responds unreliably, a second timer was implemented, distinguishing whether users stopped gazing at the bar accidentally or intentionally, allowing users a half-second to reestablish gaze contact so they can resume manipulating the bar without any delays.

Mouse Clog - The HP Z3700 wireless mouse fits underneath the sole of a rubber clog ceded by the Champalimaud Foundation. This clog is just like the ones used by surgeons during surgeries, and its sole was grinded with a Dremel rotary tool in order to fit the aforementioned mouse. Before its inlaying, the mouse was partially disassembled and its buttons were removed: the choice of removing the buttons was a result of initial testing, where buttons would accidentally be pressed when the clogs were simply being worn. Additionally, some plastic parts were also removed from the mouse to make it take less volume and fit more easily into the clog.

Gesture detector - The script for this Game Object reads horizontal input data from the mouse, calculating how far the mouse has moved from a starting position on each update cycle. This calculation starts when the user places the head gaze on either patient image and stops when the gaze is no longer upon the image. This value is then compared against a maximum value to obtain a percentage of movement. For the first ten percent, a dead zone is present in order to prevent accidental triggers, so nothing happens. Between ten and forty percent, images start progressing at a rate of two per second. Between forty and seventy percent, images progress at a rate of four per second. In the last thirty percent, the rate at which images change is of ten per second.

Colour Bar - Represents state pertaining to the gesture detector mentioned above. The bar is curved to better convey the user how much the foot has been rotated, with an icon being placed depending on position. The icon can be coloured green, yellow or red, depending on speed. Also depending on speed is the shape of the icon. While green, the icon is a play icon; on the yellow colour it is a fast-forward icon, representing greater speed. For the red colour, the fast-forward icon was adapted to feature a third play icon to represent an even faster speed. While visible, the icons fade in and out to confer a feeling of dynamism in the system.

Change plane buttons - These buttons sit below the image and change the plane from either side to the third, unselected one. For example, if the Axial and Coronal planes are shown, the button will switch the image from either to the Sagittal. When changing plane, the image index is kept to prevent the user from having to browse the entire plane again up to that point.

Navigation reticle - The navigation reticle is a red crosshair which appears whenever the laparoscopic video is not in view and pointing is not enabled. It is used to help the user target button controls using the head gaze.

Enable pointing buttons - There are three of them, one on each side and another one on top. The top one is disc-shaped, positioned in a manner that allows the user to look upwards for up to 20 degrees before it activates.

Pointing script - This script is responsible for handling system behaviour when the pointing feature is enabled, as well as the process of enabling or disabling it. When enabling pointing, the script fixes the laparoscopic video in the world in a centred position, hides the Navigation reticle and enables the Pointing reticle. When disabling pointing, the opposite is done: the Navigation reticle is again shown when appropriate, the Pointing reticle disappears and the video resumes following head movements. While pointing is active, the script acquires the Meta Camera Rig Game Object's rotation values and places the reticle 0.7 units away from it, so it is always facing the user at the same distance, while not colliding with the remaining objects in the world, which are 0.8 units away from the camera. After positioning the reticle, it calls a Network Manager function to broadcast its transform to other users.

Pointing reticle - Unlike the Navigation reticle, the Pointing reticle is always visible regardless of whether the user is looking at the video or not. For better visibility, it features an additional circumference around the crosshair, it is coloured in cyan so it better contrasts with the laparoscopic video, and the black outline help visibility when pointing to patient images.

Remote reticle - The Remote reticle represents what the other user is pointing to. It is similar in form to the Pointing reticle, but is coloured in green instead of cyan.

Local Clone - When the user points at the screen, the Local Clone Game Object is used to copy the reticle's transform in order to send the local position in lieu of the global position. Because the Local Clone is a child object of the Video Screen, a user that is not pointing can look around, have the video follow his or her head movements and still see the other user's reticle pointing in the correct position. If the user points at patient data, the Local Clone's local position is not used and the Pointing reticle's global position is sent instead.

Network Manager - Not the Network Manager Game Object that is included in the Unity Editor, but rather an empty object with the same name. Responsible for all communications between clients. Uses Unity Networking classes NetworkServer and NetworkClient to streamline communications. Due to the architecture of these classes, a user is required to act as the Server for the session, while the other joins as a Client. However, the messages that each exchange are identical and are processed identically as well. These messages include loading images and switching planes, as well as positioning the Remote reticle, hiding the Remote reticle when the other user disables turns off Pointing. To connect the users, all that is necessary is the IP address of the computer acting as Server and that they are all connected to the same network.

5.6 Summary

In this chapter we presented our proposed solution to the issues identified in chapter 3. We have described the approaches taken in addressing the issues of presenting the laparoscopic video to the user, presenting and allowing the user to browse patient data, namely MR images, and aid the user in pointing effectively at the laparoscopic video. We presented the three main components to the prototype which address these challenges: the display that enables the user free head movement while visualising the laparoscopic video; the patient data image browser, which allows users to analyse MR images in the Axial, Coronal and Sagittal planes using two complementary interaction methods; and the pointing reticle, which allows for more effective communication between the surgical team. Finally, we presented the architecture to the solution, delving deeper into how each of the three main components are tied together. In chapter 6, we will describe the methodology used to evaluate our solution.

6

Evaluation

Contents

In order to evaluate how the users perceived the implemented features in the prototype, in terms of usefulness and usability, we used a think-aloud protocol to conduct qualitative evaluation sessions with users, which will be described in detail in this chapter.

6.1 Experimental Setup

The qualitative evaluation sessions were conducted in an office inside the Chamaplimaud Foundation's building, under slightly reduced lighting conditions. The prototype, executed on the Unity Editor version 2017.4, ran on a ASUS Strix laptop, with an Intel Core i7-6700 HQ CPU, 16 GB of RAM, a Nvidia GeForce GTX 1070 graphics card with 8 GB of VRAM and Windows 10 installed. To better emulate the prototype's usage in the context of a real operation, the prototype played a 720p video footage of a laparoscope, recorded in a previous surgery, while the displayed MR images are PNG files, converted from a set of anonymised DICOM images, pertaining to a rectum magnetic resonance, using IrfranView. While experimenting the prototype, users were asked to hold two laparoscopic instruments, which were partially inserted into a custom-made laparoscopic exercise training box, as seen in fig. 6.1. Having the users hold the tools not only helped mimic the context of an operation, but it also prevented them from accessing the patient's MR images by rotating their torso, which would make the task much easier.

6.2 Methodology

In this subsection we describe the methodology used to perform a qualitative evaluation of the prototype. Each user session was expected to last thirty minutes, with the following phases:

1. **Introduction** - In this first phase, users were explained the context of this session, as well as its goal, which was to evaluate the prototype itself and not the user. Users were then asked permission to record the session, with the goal of posteriorly analysing the footage for better understanding of the given feedback.

2. **User profile questionnaire** - To begin, users were asked to fill out a user profile questionnaire, which is available in appendix A.2.1.

3. **Prototype explanation and demonstration** - The prototype's features were explained and demonstrated. In the end, users were asked if there was something they would like to be explained again.

4. **Prototype Exploration** - Users put on the prototype and were given ten minutes to explore it freely and informally, with the session moderator giving suggestions to try certain features users had not tried before. Users were also advised to talk out loud what they were trying to do in case they were having trouble doing something.

Figure 6.1: User in the exploration phase of the evaluation. Each hand holds a tool to confer a greater sense of authenticity and to prevent the user from rotating the torso as easily.

5. **Questionnaire and Interview** - After the ten minute mark, users were told to halt the experimentation and take off the headset, before being instructed to fill out a user preferences questionnaire. This questionnaire was then followed by a semi-structured interview, in order to better appraise the user's experience and opinion regarding the prototype. Both the user preferences questionnaire and the semi-structured interview are available in appendix A.3.1 and appendix A.3.3, respectively.

6.3 Participants

The participants are laparoscopic surgeons working for the Champalimaud Foundation who came in either randomly or at an agreed time. The set of users comprises 8 participants, one of them female, with ages ranging from 33 to 52. All of them have at least seven years of experience performing traditional laparoscopy, averaging at eleven. In terms of robot-assisted laparoscopy, none of them surpass five years of experience. However, on the topic of HMDs, almost all of them report having very limited experience, with only one having used them more than once. The full results gathered from the questionnaire are available in table A.2.

6.4 Results and Discussion

In this subsection we present the results of the qualitative evaluation, focusing first on each component of the prototype and presenting general opinions and impressions afterwards. We take into consideration not only the results of the user preferences questionnaire, but also responses given during the semi-structured interview and impressions shared during the prototype exploration phase. The full results for the user preferences questionnaire are available in table A.3.

6.4.1 Laparoscopic Video

Initial impressions regarding the laparoscopic video were positive: surgeons found the video following the user's head movements useful and easy to visualise, as seen in table 6.1, with some participants remarking how ergonomic it is, compared to the current way of looking at the video. Other participants found it to be the core benefit of the entire prototype.

Statement	Median (IQR)
It was easy to visualise the video	6 (0)
I liked having the video follow my head movements	5 (1)

Table 6.1: Responses to the questionnaire regarding the laparoscopic video.

However, the main argument against the video is its display size: Participant 1 complained that the video looked unfocused. P6 also wished the video would take a larger amount of the user's FOV, and P3 felt that the video looked small compared to the screens used in the surgical room. P4 also felt that the video was better on a larger screen, but also considered that the number of people passing by in the background could work against it, which is something that may not happen with the HMD if the user is looking down.

6.4.2 Patient Data

With respect to the patient data images and their navigation through the navigation bar, the foot and the button, users liked the interaction mechanisms, as demonstrated in table 6.2. Users were all able to perform all types of image interaction autonomously, although one of the users did not understand how to stop the images.

In terms of image navigation, users found the navigation bar and the foot to be complementary mechanisms, as they could make an approximation of what they wanted with the bar and make a fine adjustment with the foot. The foot itself was found to be comfortable, as surgeons were already used to using the pedals in the surgical room. In addition, participant 8 felt that the prototype could be improved if, when navigating on one plane, the other plane kept up. This participant also criticised the navigation

Statement	Median (IQR)
It was easy to examine the patient images	6 (1)
I liked where the images were positioned	5 (2)
It was easy to browse the images along the axis using the bar	6 (2)
It was easy to browse the images along the axis by rotating the foot	6 (0)
It was easy to switch between axial, coronal and sagittal planes with the button	6 (0)
It was easy to understand how fast the images were changing	6 (1)
I liked having different speeds to navigate the images	6 (1)

Table 6.2: Responses to the questionnaire regarding the patient data images and the interaction with them.

bar, saying that when controlling it, the focus is on the bar and not the image, thus preferring the foot and complimenting how the foot rotation and scroll speed were being represented.

Displaying two images as a means to mimic the imaging software used in the preoperative meeting was a generally accepted idea, although one recurring piece of feedback was the option to hide one of the images and enlarge the other one.

Due to good image definition, users found that it would be viable to use the prototype in real surgical environments. They remarked that it is not always necessary to consult images, but when it is, the prototype would prove useful.

In the end, users liked how easy it was to consult the images and how it is not necessary to call an assistant, which gives the surgeon a feeling of control, as an assistant might not know exactly what to look for. In particular, participant 4 said "It's better to see it by yourself instead of asking an assistant to scroll, or to focus", calling the foot mechanism "quite nice".

6.4.3 Pointing

Feelings regarding the pointing mechanism were positive as well. Users had no problem telling the red reticle, used to aid the user in activating interface elements, from the blue reticle, which is also used in pointing. The data obtained in the questionnaire regarding the pointing mechanism is summarised in table 6.3.

Statement	Median (IQR)
It was easy to activate the pointing reticle by looking up	5 (1)
It was easy to activate the pointing reticle by using the button	6 (0)
I could easily understand whether I was pointing or not	6 (0)
I could easily understand where/what I was pointing at	6 (0)

Table 6.3: Responses to the questionnaire regarding the pointing mechanism.

Regarding the two activation methods, using the button is majorly preferred, with only two users leaning towards looking up instead. The most common gripe regarding looking up entails having to make a more extreme movement, where the button is easier to reach. As such, participant 5 felt that the button is easier to activate because it is closer, but remarked that by looking up it is not necessary to think about where to look. Participant 4 felt looking up requires too much effort, preferring the button as well, saying “you look at the button and you immediately know when it works, it’s blue and then you can point”, while complimenting the prototype for allowing one user to point and still allowing the other to look around.

With respect to the reticle itself, users felt that it was easy to notice and to control, and that unlike the current procedure of pointing at the screen, it was unambiguous. In particular, participant 1 felt that the pointing reticle is simple and easy to learn and use, and useful to communicate with colleagues and to explain certain things to them. Participant 7 noted that there is not much difference compared to pointing with the robot, and participant 1 felt the pointing reticle was clear and praised the choice in using the colour cyan, not only because it is similar to the reticle present in the Da Vinci robot, but also because there are no human structures in this colour and therefore it clearly contrasts with them, while also saying it contrasts well with the MR images. Participant 6 felt the pointing mechanism is more beneficial to users who aren’t holding surgical tools, such as surgeons wielding the laparoscope, while on the other hand participant 8 commented that the reticle is useful even when holding tools, as it sometimes is necessary to release a noble structure, or ask someone else to hold it, in order to use the tool to point, calling the reticle “extremely useful”.

6.4.4 General Impressions

Impressions regarding the prototype in general were positive, as summarised in table 6.4. Overall, users were receptive to the idea of using the prototype in a surgical environment. One user noted that it was perfectly possible to observe the video and look at the tools, while another said “everything was quite easy to get, intuitive and natural”, praising the control activation times. Participant 5 also felt that the interaction was almost natural, that it could be learned in two or three minutes and that nothing activates by accident. Participant 1 stated that, unlike some other devices the Champalimaud Foundation has, the prototype is easily usable by everyone.

Statement	Median (IQR)
Overall, I liked using the prototype	5 (0)
I think the prototype would help making laparoscopic surgery less fatiguing	5 (1)

Table 6.4: Responses to the questionnaire regarding the prototype in general.

6.5 Limitations

Some drawbacks about the prototype were noticed: There were many grievances about the weight of the Meta, with participants 1, 4, 6 and 8 complaining about it. Participant 8 further specified that the problem lied not so much in the headset weighing too much, but rather how the weight was distributed: very unevenly and completely at the front, which corroborates Knight et al.'s findings [5]. In addition, participant 4 likened the prototype's weight effect to using a surgeon's lamp. However, participant 2 commented that the HMD is actually comfortable and that over time the user forgets that they're wearing a HMD, comparing it to a PlayStation VR. Participant 3 also remarked that in terms of comfort, it is not more uncomfortable than the current procedure.

Some users also complained that the laparoscopic video itself looked unfocused, while others thought it looked too small. In terms of image quality, we can attribute this to the footage being presented in 720p, rather than the 1920x1080 resolution that the surgeons claim the operating room screens use, as well as the video itself not being in the highest quality, with the visibility of block artefacts throughout the footage. The quality of the MR images however, was widely praised as being very clear and sharp.

Lastly, there were also problems in terms of calibration: The Meta 2 places the objects in the game world according to its initial calibration values, meaning a poor positioning of the headset while launching the application would result in the user observing the objects as if they were tilted. This also sometimes resulted in the user seeing the patient data presented in one side much closer than the other and posteriorly complaining that the data was very hard to reach. When facing this situation, users were then asked to visualise the data on the other side, which they in turn reported was very easy to access.

6.6 Summary

In this chapter, we have presented the evaluation of the prototype, describing the methodology employed to perform a qualitative evaluation of the prototype, as well as its results. We verified that the prototype was widely accepted, being considered easy to use and having clear benefits, while also noticing some technical shortcomings. In the next chapter, we conclude our work and discuss possible future work as a follow-up to this one.

7

Conclusion

Contents

Laparoscopy is a diagnostic procedure used to examine organs inside the abdomen. It's minimally invasive in the sense that it requires small incisions. These small incisions are required to insert the laparoscope, a long and thin tube with high intensity light and high resolution camera. This camera sends video to a computer screen, allowing surgeons to see inside the patient's body in real time, without needing open surgery. Laparoscopy is usually employed when non-invasive methods such as ultrasounds, CT scans and MRI scans don't provide enough information in detecting abdominal problems. Compared to open surgery, laparoscopy creates smaller scars that result in smaller and fewer incisions, lesser tissue damage, lesser pain following the operation procedure, which in turn results in a lesser requirement for analgesics and shorter hospital stay due to a faster recovery time.

Despite these benefits, laparoscopy suffers from problems like discomfort and mental fatigue, caused by the monitor being placed outside the surgeon's field of vision, forcing him or her to uncomfortably position the neck for an extended period of time. It also suffers from a lack of hand-eye coordination, as surgeons have to take the eyes off the tools they handle in order to visualise the laparoscopic video. There is also the problem of communication: surgeons point at the screen to get their point across, be an instruction or an explanation, but because the screens are placed far away the exact location they are pointing gets lost in ambiguity. Lastly, there is the issue of visualising preoperative data: currently, the procedure for visualising data such as an MRI is to call in an assistant, who will sit down with the surgeon to control the images, while the surgeon commands him to go back and forth until the desired image is found.

We have made a study on the conditions in which the laparoscopic surgeons perform their tasks, analysed areas of improvement and designed our prototype in an attempt to improve upon those aspects. Firstly, the developed prototype allows the user to visualise the laparoscopic video while looking at the tools by having it follow user head movements. Secondly, it allows patient data to be consulted during the operation, without needing to interrupt the procedure and call in an assistant, by merely requiring the surgeon to look to either side. Lastly, it allows surgeons to precisely point at the screen in a non-ambiguous and more viable fashion than using the finger or the laparoscopic tools.

In our evaluation, we found that users were receptive to the innovations brought forth by our work, showing excitement about the fact that their issues are being mitigated. The capability to observe the laparoscopic video has the potential to reduce the physical effort required by surgeons, the displaying of MR images cuts time losses whenever the need to consult an image arises and the ability to accurately target anatomic structures on the screen improves understanding between team members.

Finally, we believe that, by streamlining the visualisation of important data, as well as team communication, our work has the potential to change the laparoscopic procedure to one that does not require as much mental and physical effort as it requires now.

7.1 Future work

In this dissertation, we presented a new way to perform laparoscopic surgery that mitigates some shortcomings that currently plague the procedure. We believe our work provides a solid foundation upon which laparoscopy can eventually benefit from the application of AR. As such, there are a number of ways in which the current work can be expanded into. There are also ways in which the current prototype can be improved upon. We list some possibilities below:

1. **Immersive visualisation** - While the scope of this work focused on the visualisation of laparoscopic video with the liberty to look around and assume different neck postures, it would be interesting if users could look around to observe the inside of the patient in an immersive environment.

2. **Stereoscopic visualisation** - The 2D visualisation removes the perception of depth from the user, which increases their mental effort in understanding how the tool is moving. Stereoscopic 3D mitigates this problem and improves task performance [21], therefore it would be useful if the technical constraints could be solved in order to bring this type of visualisation to the prototype.

3. **Preoperative calibration** - Akin to the preparation procedures employed by nurses, the prototype would benefit from storing user preferences specific to each surgeon, which would in turn be loaded during said procedure. These preferences could specify parameters such as screen and image distance and size.

4. **Integration with surgeons who are not using a headset** - In our observations we found that sometimes senior surgeons would come in and give counsel on the operation at hand. This type of help consisted in the surgeon pointing at the monitor embedded in the wall. Currently, there is no way to convey what that surgeon is pointing to. We can suggest adapting the work of Sousa et al. [25] to implement this type of functionality.

5. **Migration to a different HMD** - As stated in section 6.5, there are some issues and limitations with the Meta 2 glasses, and since the Meta company has shut down and left the product unfinished in its development stage, we question whether these issues will eventually be sorted out or not. Because of this, we believe the work could continue to evolve in a more stable manner if a change in hardware occurred. We thus propose the migration to Microsoft HoloLens 2[1], as it is untethered, which should help reducing clutter in the operating room, but it also iterates upon the previous product by investing heavily in ergonomics, benefiting from a weight reduction and adjustment of the centre of gravity more to the centre of the head.

[1] https://www.microsoft.com/en-us/hololens/hardware

Bibliography

[1] S. Bernhardt, S. A. Nicolau, V. Agnus, L. Soler, C. Doignon, and J. Marescaux, "Automatic localization of endoscope in intraoperative CT image: A simple approach to augmented reality guidance in laparoscopic surgery," *Medical Image Analysis*, vol. 30, pp. 130–143, 2016.

[2] M. Leite, A. F. Carvalho, P. Costa, R. Pereira, A. Moreira, N. Rodrigues, S. Laureano, J. Correia-Pinto, J. L. Vilaça, and P. Leão, "Assessment of laparoscopic skills performance: 2D versus 3D vision and classic instrument versus new hand-held robotic device for laparoscopy," *Surgical Innovation*, vol. 23, no. 1, pp. 52–61, 2016.

[3] A. U. Batmaz, M. De Mathelin, and B. Dresp-Langley, "Seeing virtual while acting real: Visual display and strategy effects on the time and precision of eye-hand coordination," *PLoS ONE*, vol. 12, no. 8, pp. 1–18, 2017.

[4] R. R. D. C. K. Henry Fuchs1, Mark A. Livingston1, J. R. C. P. R. S. H. D. Keller1, Andrei State1, and M. Anthony A. Meyer, "Augmented Reality Visualization for Laparoscopic Surgery," *Federation bulletin / Federation of State Medical Boards of the United States*, vol. 50, pp. 316–322, 2006. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/14051687

[5] J. F. Knight and C. Baber, "Effect of Head-Mounted Displays on Posture," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 49, no. 5, pp. 797–807, 2007. [Online]. Available: http://journals.sagepub.com/doi/10.1518/001872007X230172

[6] A. Esteves, E. Velloso, A. Bulling, and H. Gellersen, "Orbits: Gaze Interaction for Smart Watches using Smooth Pursuit Eye Movements," *Proc. UIST '15*, no. 1, pp. 457–466, 2015. [Online]. Available: http://dl.acm.org/citation.cfm?doid=2807442.2807499

[7] E. Velloso, M. Wirth, C. Weichel, A. Esteves, and H. Gellersen, "AmbiGaze: Direct Control of Ambient Devices by Gaze," pp. 812–817, 2016. [Online]. Available: http://dl.acm.org/citation.cfm?doid=2901790.2901867

[8] T. Kim, B. Saket, A. Endert, and B. MacIntyre, "VisAR: Bringing Interactivity to Static Data Visualizations through Augmented Reality," 2017. [Online]. Available: http://arxiv.org/abs/1708.01377

[9] F. Müller, J. McManus, S. Günther, M. Schmitz, M. Mühlhäuser, and M. Funk, "Mind the Tap," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*, 2019, pp. 1–13. [Online]. Available: http://dl.acm.org/citation.cfm?doid=3290605.3300707

[10] E. Watanabe, M. Satoh, T. Konno, M. Hirai, and T. Yamaguchi, "The Trans-Visible Navigator: A See-Through Neuronavigation System Using Augmented Reality," *World Neurosurgery*, vol. 87, pp. 399–405, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.wneu.2015.11.084

[11] X. Chen, L. Xu, Y. Wang, H. Wang, F. Wang, X. Zeng, Q. Wang, and J. Egger, "Development of a surgical navigation system based on augmented reality using an optical see-through head-mounted display," *Journal of Biomedical Informatics*, vol. 55, pp. 124–131, 2015.

[12] P. Pratt, M. Ives, G. Lawton, J. Simmons, N. Radev, L. Spyropoulou, and D. Amiras, "Through the HoloLens™ looking glass: augmented reality for extremity reconstruction surgery using 3D vascular models with perforating vessels," *European Radiology Experimental*, vol. 2, no. 1, p. 2, 2018. [Online]. Available: https://eurradiolexp.springeropen.com/articles/10.1186/s41747-017-0033-2

[13] A. Grinshpoon, G. J. Loeb, and S. K. Feiner, "Hands-Free Augmented Reality for Vascular Interventions," no. Figure 1, 2018.

[14] L. E. B. R, G. P. F, and M. F. M. G, "Augmented Reality User interaction to Computer Aided Orthopedic Surgery system," 2018.

[15] C. S. Muratore, B. A. Ryder, and F. I. Luks, "Image display in endoscopic surgery History of surgical endoscopy," pp. 349–356, 2007.

[16] H. M. Mentis, A. Chellali, and S. Schwaitzberg, "Learning to See the Body: Supporting Instructional Practices in Laparoscopic Surgical Procedures," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2113–2122, 2014.

[17] H. Prescher, D. E. Biffar, C. A. Galvani, J. W. Rozenblit, and A. J. Hamilton, "Surgical navigation pointer facilitates identification of targets in a simulated environment," in *Simulation Series*, vol. 46, no. 10, 2014, pp. 246–252. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84908300463&partnerID=40&md5=ad45e1a201ed67044a103f89bf855270

[18] Y. Feng, J. Chhikara, J. Ramsey, and H. M. Mentis, "Perceived Usefulness and Acceptance of Telestration in Laparoscopic Training."

[19] D. A. Walczak, D. Pawełczak, P. Piotrowski, P. W. Trzeciak, A. Jędrzejczyk, and Z. Pasieka, "Video display during laparoscopy – where should it be placed?" *Videosurgery and Other Miniinvasive Techniques*, vol. 1, pp. 87–91, 2015. [Online]. Available: http://www.termedia.pl/doi/10.5114/wiitm.2014.47434

[20] S. K. Maithel, L. Villegas, N. Stylopoulos, S. Dawson, and D. B. Jones, "Simulated laparoscopy using a head-mounted display vs traditional video monitor: An assessment of performance and muscle fatigue," *Surgical Endoscopy and Other Interventional Techniques*, 2005.

[21] H. Prescher, D. E. Biffar, J. Rozenblit, and A. J. Hamilton, "The comparison of high definition versus stereoscopic display on standardized fundamental laparoscopic skill procedures," *Summer Computer Simulation Conference, SCSC 2014, Part of the 2014 Summer Simulation Multiconference, SummerSim 2014, July 6, 2014 - July 10*, vol. 46, no. 10, pp. 346–351, 2014.

[22] K. Kihara, Y. Fujii, H. Masuda, K. Saito, F. Koga, Y. Matsuoka, N. Numao, and K. Kojima, "New three-dimensional head-mounted display system, TMDU-S-3D system, for minimally invasive surgery application: Procedures for gasless single-port radical nephrectomy," *International Journal of Urology*, vol. 19, no. 9, pp. 886–889, 2012.

[23] J. Jayender, B. Xavier, F. King, A. Hosny, D. Black, S. Pieper, and A. Tavakkoli, *A Novel Mixed Reality Navigation System for Laparoscopy Surgery*. Springer International Publishing, 2018, vol. 2878. [Online]. Available: http://link.springer.com/10.1007/b93810

[24] G. Fried, "FLS Manual Skills Written Instructions and Performance Guidelines," *FUNDAMENTALS of LAPAROSCOPIC SURGERY*, no. February, p. 4, 2014. [Online]. Available: http://www.flsprogram.org/wp-content/uploads/2014/03/Revised-Manual-Skills-Guidelines-February-2014.pdf

[25] M. Sousa, R. K. Dos Anjos, D. Mendes, M. Billinghurst, and J. Jorge, "Warping deixis: Distorting gestures to enhance collaboration," *Conference on Human Factors in Computing Systems - Proceedings*, no. Chi, pp. 1–12, 2019.

A

Appendices

A.1 User and Task Analysis

A.1.1 User and Task Analysis questionnaire

User Questionnaire

This questionnaire's purpose is to better understand the Champalimaud Foundation's laparoscopic surgeons, as well as the tasks they perform, and shouldn't take more than 5 minutes. We thank you for your collaboration and inform that this questionnaire is completely anonymous and used only for investigation purposes.

* Required

1. **Age** *

___

2. **Gender** *

Mark only one oval.

- ◯ Female
- ◯ Male
- ◯ Other

3. **What experience do you have with Head-mounted Displays?** *

Mark only one oval.

- ◯ A lot, I use them regularly
- ◯ Some, I've used them but not too many times
- ◯ Little, I've used them once or twice
- ◯ I have never used this type of device before

4. **How many surgeons are usually around the patient during the surgery?** *

Mark only one oval.

- ◯ <3
- ◯ 3-5
- ◯ >5

5. **How much time does a laparoscopy usually take?** *

___

6. **Of this time, how long does it take until there is tissue extraction and the endoscope isn't being used?**

___

7. **Finally, how long does it take from when the laparoscope is used again until the end of the procedure?**

Untitled Section

8. **What type of data is consulted during pre-operative planning?** *

9. **How many images are usually consulted during planning?** *

10. **Would it be useful to have this data available while the surgery is underway?** *
Mark only one oval.
    - Yes, all of it
    - Yes, but only a few that are selected
    - No
    - Other: ____________

11. **What type of difficulties do you most feel while carrying out a laparoscopy?** *
Check all that apply.
    - Looking at the monitor while controlling the tools
    - Letting others know what part of the video I'm pointing at/talking about
    - Understanding what part of the video others are talking about
    - Maintaining a comfortable posture throughout the surgery
    - Other: ____________

12. **How useful would it be to perform surgery with a 3D camera (while not using the da Vinci robot)?** *
Mark only one oval.

	1	2	3	4	5	
Completely useless	◯	◯	◯	◯	◯	Extremely useful

13. **If a pointing system were to be implemented, would it be more useful if the pointer were controlled with hand gestures or through hands-free interaction (like head movement tracking)?** *

Mark only one oval.

- Hand gesture interaction
- Hands-free interaction

14. **How useful would it be to have a telestration system (drawing on video) to complement a pointing system?** *

Mark only one oval.

	1	2	3	4	5	
Completely useless	◯	◯	◯	◯	◯	Extremely useful

15. **Is there anything else you feel that could be improved in laparoscopy with the use of Augmented Reality? (optional)**

Powered by Google Forms

A.1.2 User and Task Analysis questionnaire results

	Question						
User	1	2	3	4	5	6	7
1	42	Male	I have never used this type of device before	<3	5	3h30	45m
2	34	Female	I have never used this type of device before	3-5	4h		
3	37	Male	I have never used this type of device before	<3	1h	25m	5m
4	36	Male	I have never used this type of device before	3-5	2h	30m	30m
5	33	Male	Little, I've used them once or twice	3-5	5h	3h	30m
6	40	Male	I have never used this type of device before	3-5	5h	1h	1h
7	39	Male	Little, I've used them once or twice	<3	3h	2h30	15m

	Question		
User	8	9	10
1	MRI Pelvic or liver	120	Yes, all of it
2	CT scan, MRI..	15	Yes, but only a few
3	Imaging	10	Yes, but only a few
4	CT, MRI, History	20	Yes, but only a few
5	MRI/CT scan and colonoscopy	20	Yes, but only a few
6	MRI, blood tests	100	Yes, but only a few
7	Radiology images, clinical reports, data from consultation	2	Yes, but only a few

	Question				
User	11	12	13	14	15
1	b) c) d)	3	Hand gesture interaction	5	Weight issues with head-mounted displays
2	b) d)	3	Hands-free interaction	3	
3	b)	4	Hands-free interaction	3	
4	b)	4	Hand gesture interaction	4	Magnification and zoom in and out
5	b) d)	3	Hands-free interaction	5	Voice indications transformed to text
6	d)	4	Hand gesture interaction	4	Less foggy image
7	d)	5	Hands-free interaction	5	Merging the radiology images with what you see during surgery

Table A.1: Responses to the questionnaire. Values in question 11 correspond to the order of the available answers.

A.2 User Profile

A.2.1 User Profile questionnaire

User Profile

* Required

1. **Age** *

   ____

2. **Gender** *

   Mark only one oval.

   - ( ) Female
   - ( ) Male
   - ( ) Other

3. **What experience do you have with Head-Mounted Displays?** *

   Mark only one oval.

   - ( ) A lot, I've used them at least five times in the last month
   - ( ) Some, I've used them more than twice
   - ( ) Little, I've tried them on
   - ( ) I have never used this type of device/What is a head-mounted display?

4. **What experience do you have in performing conventional laparoscopic surgery (in years)?** *

   ____

5. **What experience do you have in performing robot-assisted laparoscopic surgery (in years)?** *

   If less than a year, introduce a fractioned year (example: 0.5 years).

   ____

Powered by Google Forms

A.2.2 User Profile questionnaire results

	Question				
User	1	2	3	4	5
1	33	Male	Some, I've used them more than twice	7	1
2	41	Male	Little, I've tried them on	11	0.5
3	35	Female	Little, I've tried them on	8	1
4	39	Male	Little, I've tried them on	12	0.5
5	52	Male	Little, I've tried them on	10	3
6	39	Male	Little, I've tried them on	10	2
7	38	Male	Little, I've tried them on	10	4
8	42	Male	Little, I've tried them on	17	5

Table A.2: Responses to the User Profile questionnaire

A.3 User Preferences

A.3.1 User Preferences questionnaire

User Preferences Questionnaire

Please rate the following affirmations using a 1-6 scale, with 1 representing STRONGLY DISAGREE and 6 representing STRONGLY AGREE

* Required

Laparoscopic video

1. **It was easy to visualise the video** *

   Mark only one oval.

	1	2	3	4	5	6	
Strongly disagree	◯	◯	◯	◯	◯	◯	Strongly agree

2. **I liked having the video follow my head movements** *

   Mark only one oval.

	1	2	3	4	5	6	
Strongly disagree	◯	◯	◯	◯	◯	◯	Strongly agree

Patient Images

3. **It was easy to examine the patient images** *

   Mark only one oval.

	1	2	3	4	5	6	
Strongly disagree	◯	◯	◯	◯	◯	◯	Strongly agree

4. **I liked where the images were positioned** *

   Mark only one oval.

	1	2	3	4	5	6	
Strongly disagree	◯	◯	◯	◯	◯	◯	Strongly agree

5. **It was easy to browse the images along the axis using the bar** *
Mark only one oval.

	1	2	3	4	5	6	
Strongly disagree	◯	◯	◯	◯	◯	◯	Strongly agree

6. **It was easy to browse the images along the axis by rotating the foot** *
Mark only one oval.

	1	2	3	4	5	6	
Strongly disagree	◯	◯	◯	◯	◯	◯	Strongly agree

7. **It was easy to switch between axial, coronal and sagittal planes with the button** *
Mark only one oval.

	1	2	3	4	5	6	
Strongly disagree	◯	◯	◯	◯	◯	◯	Strongly agree

8. **It was easy to understand how fast the images were changing** *
Mark only one oval.

	1	2	3	4	5	6	
Strongly disagree	◯	◯	◯	◯	◯	◯	Strongly agree

9. **I liked having different speeds to navigate the images** *
Mark only one oval.

	1	2	3	4	5	6	
Strongly disagree	◯	◯	◯	◯	◯	◯	Strongly agree

Pointing

10. **It was easy to activate the pointing reticule by looking up** *
Mark only one oval.

	1	2	3	4	5	6	
Strongly disagree	◯	◯	◯	◯	◯	◯	Strongly agree

11. **It was easy to activate the pointing reticule by using the button** *
Mark only one oval.

	1	2	3	4	5	6	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	⬭	Strongly agree

12. **I could easily understand whether I was pointing or not** *
Mark only one oval.

	1	2	3	4	5	6	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	⬭	Strongly agree

13. **I could easily understand where/what I was pointing at** *
Mark only one oval.

	1	2	3	4	5	6	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	⬭	Strongly agree

Overall experience

14. **Overall, I liked using the prototype** *
Mark only one oval.

	1	2	3	4	5	6	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	⬭	Strongly agree

15. **I think the prototype would help making laparoscopic surgery less fatiguing** *
Mark only one oval.

	1	2	3	4	5	6	
Strongly disagree	⬭	⬭	⬭	⬭	⬭	⬭	Strongly agree

16. **If you replied "Disagree" or "Strongly disagree" to any of the questions above, please succinctly explain why, below.**

17. **Feedback about the prototype**

A.3.2 User Preferences questionnaire results

	Question														
User	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	5	3	6	6	6	6	6	6	5	6	6	6	6	5	4
2	6	6	6	4	6	5	6	6	5	4	6	5	6	5	5
3	6	4	6	6	6	6	6	6	6	5	6	6	6	5	4
4	6	5	5	4	2	6	5	4	5	4	6	6	6	4	4
5	6	5	5	4	6	6	5	5	6	6	5	6	6	6	6
6	6	5	6	6	4	6	6	5	5	5	5	5	5	5	5
7	6	6	5	5	6	6	6	6	6	6	6	6	6	6	5
8	5	6	5	5	3	4	6	1	6	5	6	6	6	5	3

Table A.3: Responses for the User Preferences questionnaire

A.3.3 Semi-structured interview

1. We were told that the software used by the surgeons in preoperative planning uses only two out of three planes, which the prototype tries to emulate. Would it be preferable if the prototype showed the three all at once? Or perhaps just one instead of two?
2. Would you be able to use the prototype to see the patient's MRIs or X-ray scans effectively during the surgery?
3. To change the image, you can use the bar above it or control the mouse with the foot. Do you feel the bar is enough to find what you need to locate or did you find the mouse more useful?
4. The prototype allows you to begin pointing either by looking up or activating the button located on either side. Which is more useful, in your opinion?
5. How do you compare the prototype's pointing mechanism with the current practice of pointing at the screen?
6. What were your greatest difficulties in using the prototype, if any?
    (a) If you had difficulties using the foot, do you feel if you had a differently-sized shoe you would be able to perform the gesture in a more adequate fashion?
7. What did you like the most about the prototype?
8. What could be improved upon in the prototype?
9. Do you feel this prototype is more comfortable than how laparoscopy is currently conducted? (would a lighter HMD improve the procedure that much)