

Augmented Reality Display with Hands-free Control for Surgical Procedure Enhancement

An Electromyographic-based approach

Margarida Veloso Branco Lopes Ramalho

Thesis to obtain the Master of Science Degree in

Engenharia Biomédica

Supervisors: Prof. Ana Luísa Nobre Fred
Prof. Hugo Humberto Plácido da Silva

Examination Committee

Chairperson: Prof. João Miguel Raposo Sanches
Supervisor: Prof. Hugo Humberto Plácido da Silva
Member of the Committee: Prof. Daniel Jorge Viegas Gonçalves

May 2023

I declare that this document is an original work of my own authorship and that it fulfills all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Acknowledgments

I would like to express my gratitude and deepest appreciation to Professor Hugo Plácido da Silva and Professor Ana Nobre Fred for their continuous support, enthusiastic feedback, suggestions and conjoined brainstorming along this process. Professor Hugo is a truly remarkable and inspiring person, and I am incredibly lucky to have had him as my supervisor. His everlasting optimism, encouraging words and patience with me were fundamental for the successful completion of this chapter, especially after I initiated my professional journey. I would also like to thank Dr. Carolina Rodrigues for her co-supervision and availability along the experimental process, as well as Dr. Manuel Azancot de Menezes for receiving me in his office and giving me valuable feedback. Special thanks to Ana Sofia Carmo for her support and assistance with the inclusion of her EpiBOX project, and to all the participants who willingly contributed their time and effort to participate in the data collection process. A word of appreciation also goes to Instituto de Telecomunicações (IT) for their support throughout this work, and for including me in their circle whenever I went to work there.

This thesis marks the end of five unforgettable years, in which I feel like I learned a lot, not only academically but also in terms of who I am today. I had the chance to meet some of my dearest friends, who made this adventure the most special. To Laura, Carolina, Beatriz and João, thanks for colouring my days and being there for me every step of the way, even when I came to you with the most random thoughts and doubts. Also, to Vicente and the rest of our little band, our messy jam sessions are times I am going to cherish and take with me.

I would also like to thank my friends from home, who have been there for me not only during this thesis adventure, but throughout all these years, you make me happy. To Mariana, my partner in crime, thanks for showing me the brighter side of life, and for going along with our crazy (and at times, admittedly, not the most thought-through) plans. Lastly, to my amazing family, especially to my lovely parents, I feel truly lucky to have you both by my side, thank you for always supporting me and being my rocks. You keep teaching me what unconditional love means, and although you see life from different perspectives, you really are my role models, and I honestly hope I can someday be a little bit like you.

With love,
Margarida Ramalho

Abstract

The constant evolution of surgical interventions in medicine has saved countless lives and improved overall patient health. However, the increasing complexity of variables and information in the operating room has created challenges for surgical teams, potentially impacting intervention outcomes. Particularly in high-risk procedures like cardiac surgeries, preventable errors often result from teamwork or system constraints. Improving the access, visualization, and integration of information in the operating room is a crucial challenge to enhance the flow of surgery and minimize disruptions that can impact patient safety. In this dissertation, an EMG-based hands-free Augmented Reality (AR) system is proposed to assist the surgeon and auxiliary team during surgery, in visualizing patient information on demand, in real-time. The system is composed by a head-mounted Augmented Reality see-through headset, which displays relevant clinical information about the patient being intervened (e.g., vital signals, previous medical imaging). The displayed information is controlled by the user, through specific forehead movements, captured by facial electromyography. Studies were carried out to assess optimal EMG acquisition and processing, as well as different electrode models were evaluated. A mobile application, *ARSurgery*, was developed as the interface for the AR system, including the algorithm responsible for detecting and classifying facial muscle inputs. The developed system was tested with two groups of subjects, including surgeon doctors, obtaining a very satisfactory performance, with mean precision and recall rates of 0,951 and 0,988, respectively.

Keywords

Augmented Reality, Surface Electromyography, Onset Detection, Operating Room, Signal Processing, Hands-free Control

Resumo

A contínua evolução da medicina no âmbito das intervenções cirúrgicas tem salvo inúmeras vidas e melhorado a saúde dos pacientes. No entanto, a crescente complexidade de variáveis e informação no bloco operatório apresenta novos desafios para as equipas cirúrgicas, podendo inclusivamente afetar a intervenção. Particularmente em procedimentos de elevado risco, parte das falhas consideradas evitáveis resultam de constrangimentos na equipa ou no sistema. Melhorar o acesso, a visualização e a integração da informação no bloco operatório é um desafio fundamental para melhorar o fluxo da cirurgia e minimizar interrupções que possam afetar a segurança do paciente. Nesta dissertação, é proposto um sistema de Realidade Aumentada (RA) *hands-free*, baseado em eletromiografia (EMG) facial, para auxiliar o cirurgião e respetiva equipa, durante a cirurgia, na visualização de informação sobre o paciente, em tempo real. O sistema é composto por um visor de RA transparente, ajustado à cabeça do cirurgião, que exhibe informações clínicas relevantes sobre o paciente intervencionado (e.g., sinais vitais, imagiologias prévias) quando necessário. As informações exibidas são controladas pelo utilizador através de movimentos faciais específicos na zona da testa, capturados por eletromiografia. Foram realizados estudos para otimizar a aquisição e processamento do sinal EMG, bem como testados e avaliados diferentes modelos de elétrodos. Foi desenvolvida uma aplicação móvel, *ARSurgery*, como interface para o sistema de RA, incluindo o algoritmo responsável pela deteção e classificação dos movimentos faciais captados. O sistema desenvolvido foi testado em dois grupos de indivíduos, incluindo cirurgiões, tendo sido obtido um desempenho bastante satisfatório, com valores médios de precisão e sensibilidade de 0,951 e 0,988, respetivamente.

Palavras Chave

Realidade Aumentada, Eletromiografia de Superfície, Segmentador, Bloco Operatório, Interação Homem-Máquina, Controlo Mãos-livres

Contents

1	Introduction	1
1.1	Problem Description	3
1.2	Motivation	4
1.3	Research Goals	6
1.4	Achievements	7
1.5	Organization of the Document	8
2	Related Work	9
2.1	Augmented Reality	11
2.2	Hands-free Control Modes	12
2.2.1	Speech-based Control	13
2.2.2	Eye-based Control	13
2.2.3	Brain-Computer Interfaces	15
2.2.4	EMG-based Control	16
2.3	Electromyography	17
2.3.1	Muscle Physiology	17
2.3.2	EMG Signal	18
3	Methodology	21
3.1	Workflow	23
3.2	Requirements and Hardware Restrictions	23
3.2.1	AR Headset Requirements	24
3.2.2	Choosing the Appropriate Type of sEMG Electrodes	27
3.2.3	Facial Expressions Chosen as Input for the HCI	28
3.2.4	Electrode Placement on the Face	29
3.3	Application Structure and Displayed Pages	31
4	Implementation	33
4.1	BITalino Integration	35
4.2	Final Prototype	36

4.2.1	Application description	36
4.2.2	Smartphone - AR Headset connection	39
4.2.3	Signal processing and onset detector	40
4.2.4	Methodology validation	40
5	Results	43
5.1	Characterization and optimization of the EMG signal and hardware	45
5.1.1	Electromyographic signal pre-processing	46
5.1.2	Electromyographic signal quality	46
5.1.3	Detector's hyperparameters	57
5.1.4	Onset detector implementation	57
5.1.5	AR system prototype validation	58
5.2	Study with the interest group (surgeon doctors)	62
6	Conclusion	65
6.1	Conclusions	67
6.2	Future Work	67
	Bibliography	69
A	Informed Consent and Request for Data Collection	75

List of Figures

1.1	Illustration of a standard operating room (OR) layout during cardiac surgery, emphasising the monitors scattered around the OR displaying different data sources.	6
2.1	Illustration of the different types of Mixed Reality Technologies.	11
2.2	Illustration of different eye-based controlled devices: wearable eye-tracking headset; EOG-based HMI application.	14
2.3	Illustration of an EEG-based Brain-Computer Interface.	16
2.4	Illustration of a Motor Unit (MU) with emphasis on its main components. Schematic representation of a motor unit with n muscle fibers.	18
2.5	Time domain and frequency domain representation of an EMG signal.	20
3.1	Schematic diagram of the Augmented Reality System main components	23
3.2	Illustration of the AR headset superimposed with the surgeon's commonly required equipment.	24
3.3	Standard surgical apparatus used by surgeon doctors during surgery.	26
3.4	ViP-display (version 1) AR headset, manufactured by ©ROTACIONAL.	27
3.5	Illustration of the facial commands	29
3.6	Augmented Reality head-mounted display with attached EMG control system. Illustration of equipped surgeon doctor using the developed system.	30
3.7	ARSurgery displayed pages.	32
4.1	BITalino (r)evolution Board, featuring its assembled version.	35
4.2	Architecture of the (optional) connection between ARSurgery and the BITalino measuring patient's biosignals, through the EpiBOX interface.	36
4.3	ARSurgery displayed pages, with sequential guidance for the setup phase.	38
4.4	Illustration of ARSurgery portrait and landscape modes.	39
4.5	Examples of ARSurgery displayed notifications and warnings.	39

4.6	EMG raw signal - without applying noise attenuating filters. EMG signal filtered with a Weighted Moving Average FIR filter, with a window size of 100 samples.	40
4.7	Montages used in experimental tests 1 and 2.	41
5.1	Illustration of the experimental protocol for the acquisition of the electromyographic signals.	45
5.2	Illustration of the EMG signal before and after the pre-processing: conversion to mV units, subtraction of the baseline mean and full wave rectification.	47
5.3	Illustration of the EMG signal acquisition using two types of electrodes simultaneously - Ag/AgCl pre-gelled electrodes (Eg), and Ag/AgCl raised profile dry electrodes (Er).	48
5.4	Statistical comparison of EMG test trials using Ag/AgCl pre-gelled electrodes (Eg), and Ag/AgCl raised profile dry electrodes (Er).	49
5.5	Statistical comparison of EMG test trials using raised profile (Er) and flat profile (Ef) Ag/AgCl dry electrodes. Statistics include all subjects without signal filtering.	51
5.6	Statistical comparison of EMG test trials using raised profile (Er) and flat profile (Ef) Ag/AgCl dry electrodes. Statistics are based on Subject 1, with signal filtering using an LMS adaptive filter.	52
5.7	Precision and Recall curves for Hodges and Bui EMG onset detector, considering different filters applied to input EMG signal.	56
5.8	Illustration of an EMG signal segment and its resulting test function, as well as the corresponding computed threshold.	58
5.9	Illustration of the movements' sequence followed by participants in Test 1 and Test 2.	59
5.10	Performance metrics results of experimental tests including True Positive Rate, False Positive Rate, False Negative Rate, and EER / BER Movement mismatch Rates.	60
5.11	System Usability Scale scores, respective percentiles and grades.	64

List of Tables

3.1	Summary of the ViP-display AR headset main specifications	27
3.2	Summary of PLUX Biosignals' EMG sensor main characteristics	28
3.3	Illustration of the electrodes used in the signals' acquisitions and summary of their main characteristics	28
4.1	Summary of BITalino (r)evolution Board BT/BLE main specifications	36
5.1	Summary of participants characteristics	45
5.2	Metrics computed from the signal, depending on the signal's region	47
5.3	Number of trials performed by each subject, with each of the three types of electrodes - Er, Ef and Eg.	50
5.4	Relative difference (%) between metrics' mean values obtained for Ef and Er electrode models, while considering all subjects, and without applying any filter to the signals. . . .	53
5.5	Relative difference (%) between metrics' mean values obtained for Ef and Er electrode models, while considering one subject, and applying an adaptive filter on the signals. . . .	53
5.6	Summary of participants data - number of participants, sex and age - discriminated for both study groups - control group and test group	59
5.7	Total number of experimental trials performed by the participants, discriminated by the test type - Test 1 and Test 2	59
5.8	Two-part survey applied to the test group	63

Acronyms

ADC	Analog to Digital Converter
AR	Augmented Reality
AP	Action Potential
AV	Augmented Virtuality
BCI	Brain-Computer Interface
BER	Brief Eyebrow Raise
EB	Eye Blink
ECG	Electrocardiography
EDA	Electrodermal Activity
EEG	Electroencephalography
EER	Extended Eyebrow Raise
EMA	Exponential Moving Average
EMG	Electromyography
EOG	Electrooculography
ERP	Event-related Potential
FD	Flow Disruption
FIR	Finite Impulse Response
FN	False Negative

FP	False Positive
IIR	Infinite Impulse Response
GMA	Gaussian Moving Average
HMI	Human-Machine Interface
LMS	Least Mean Square
MU	Motor Unit
MUAP	Motor Unit Action Potentia
MMSC	Mathematical Model for a Static Contraction
OR	Operating Room
sEMG	surface Electromyography
SMA	Simple Moving Average
SNR	Signal-to-Noise Ratio
SUS	System Usability Scale
TP	True Positive
W	Single-eye Wink
VR	Virtual Reality

1

Introduction

Contents

1.1 Problem Description	3
1.2 Motivation	4
1.3 Research Goals	6
1.4 Achievements	7
1.5 Organization of the Document	8

1.1 Problem Description

The continuous advancements in medicine, specifically in the realm of surgical procedures, have played a pivotal role in improving the overall well-being of patients and, in numerous instances, even saving lives. However, this evolution, promoted by technological advances in equipment and procedures, has led to an exponential growth in the complexity of variables and information to take into account while performing surgery. The Operating Room (OR) has become an extremely complex environment, currently involving the collaboration of many different areas of expertise and personnel (e.g. anesthesiologists, surgeons, nurses, perfusionists, among others), each focused on particular tasks and responsibilities. In addition, throughout the surgical procedure, there is an ongoing necessity to constantly verify and monitor various types of information, both pre-operative (such as medical records and previous imaging modalities) as well as real-time physiological data / vital signs. This complex information flow and coordination of tasks by the different specialists in the OR can potentially cause disruptions in the flow of the surgery, which may lead to technical errors and less favorable outcomes for the patient being intervened [1].

This is especially noticeable in the case of higher-risk procedures, such as cardiac surgeries. The incidence of surgical adverse events among the latter is close to 12%, compared with 3% in other surgical procedures, being that 54% of these incidences were posteriorly considered preventable [2,3]. Thus, it is extremely important to find ways to minimize these preventable errors, especially since they are often not related with technical skills or training, but instead are associated with teamwork or system failures.

Particularly, when it comes to the information flow in the OR during surgery, currently the latter is still mostly static, with several monitors and displays being scattered along the room, and with the surgeon and auxiliary team having to actively seek the information. This shifts their concentration between the patient and the information source, and inclusively at times forces them to move around the OR. These interruptions, designated as Flow Disruptions (FDs), create a break in the flow of the surgery, and can contribute to hand-eye coordination challenges as well as dispersion of attention [1], possibly leading to technical errors. Therefore, one of the current challenges, especially in this rapidly advancing technological landscape, lies in finding ways of facilitating the access and visualization of information, and creating solutions to integrate it, so that the medical team can more easily acquire the knowledge they need, without disrupting the surgery flow.

1.2 Motivation

In the Operating Room (OR), the environment comprises the physical space, the equipment, and the individuals involved, such as staff members and patients. During a surgical procedure, the physical layout of the OR has a direct impact on the process, and must be optimized to ensure maximum ergonomic efficiency within the environment, so the medical team can interact efficiently and safely. However, ergonomics is still considered suboptimal with respect to patient safety in the OR [4–6], with layout-related disturbances accounting for more than half of the surgical Flow Disruptions (FDs) [7].

Alongside unfavorable patient outcomes, flow disruptions can also contribute to increase the duration of the surgery (up to 32%), which subsequently incurs greater expenses [8, 9]. Since the study of surgical adverse events is so important to prevent future incidents, and these are often correlated with FDs that happen along the surgery, there has been a great effort to understand and evaluate the main causes of the latter. To better characterize FDs, several authors have divided them in different categories [4, 6, 8, 10–12]. Because most categories overlap across studies, this work adopts the division made by Palmer et al. (2013) in which FDs can arise from any combination of six classes: (1) Communication; (2) Usability; (3) Physical layout; (4) Environmental hazards; (5) General interruptions; and (6) Equipment failures [4]. Each category and subdivision is described in Table 1 of Palmer et al.'s study (2013, p. 1070).

Specifically regarding the access to information in the OR, it currently remains responsible for a portion of these FDs. For example, the need to check for pre-operative exams of the patient in an external computer causes a temporary interruption in the surgery (general interruption FD), and can also create difficulties in communication and lead to the increase of the environmental noise level (communication-related FD). On the other hand, if the OR is a smaller space and the monitors are scattered around it, this may obstruct or prevent movements by the staff (layout related FD). Furthermore, when the surgeon is checking the patients' vital signs in real-time on a monitor, or even during an image-guided (e.g., angiography-guided) procedure, there are temporary losses of focus as well as of the visual field of the patient being intervened. These can result in unwanted hands movements, that may further lead to technical errors.

This work focuses on tackling information-related FDs, finding suitable alternatives to access information in the OR. Improving this aspect may have direct consequences in the reduction of FDs during surgery, contributing to the general improvement of the surgery process and outcome, and to the decrease in surgery time. However, when thinking about alternative ways for the surgeon and/or auxiliary medical team to view the necessary information, other than the standard monitors distributed around the OR, the initial challenge that emerges is determining how surgeons will interact with the information source (i.e., display) and effectively control the contents being presented. Moreover, since surgeons

usually have both hands occupied during procedures while manipulating surgical tools, this adds an additional layer of complexity.

One possible solution for the latter restriction is to make the interaction between the user and the display hands-free. There are different ways to control technology without using hands, such as voice control, using limbs / hands movements, among others. These however, although being common Human-Machine Interfaces (HMIs), are not ideal for the surgical environment. The voice control is not adequate due to the noise level already present in the OR, which makes it difficult to use. The limbs movements, especially hands movements, can be counterproductive in relation to what the surgeon is performing surgically. Hence, for this project, the input selected to control the display of information were the user's facial expressions, because it offers less constraints than the others methods, particularly in this context of application. One method to achieve this type of control is resorting to surface Electromyography (sEMG) sensors, placed on the skin above the muscle, that detect electrochemical changes associated with muscle activation (in this case, facial muscles), indicating the performance of facial gestures.

Regarding the anatomical placement of the display, in order to build an alternative that effectively mitigates potential attention-related FDs caused by fixed monitors, it is preferable to present the information within the surgeon's field of vision. The most efficient way to achieve this would be having a display that accompanies the surgeons head movements, ensuring that it remains within their line of sight whenever necessary. For this purpose, this work explores the utilization of an Augmented Reality headset, since this type of system can be fixed on the user's forehead, directly in front of his eyes.

Figure 1.1 shows a typical OR layout for cardiac surgery. Although the size and design of the OR can vary, it generally constitutes a complex environment, with a high number of different visual displays distributed along the room.



Figure 1.1: Illustration of a standard operating room (OR) layout during cardiac surgery, emphasising the monitors scattered around the OR displaying different data sources ¹.

1.3 Research Goals

Building upon the main findings outlined in Section 1.2, the research goal in this work is to devise a novel system based on Augmented Reality (AR) head mounted displays and hands-free control using EMG, as a way of minimising attention-related FDs and improving surgical procedures. The main research questions addressed, can be summarised as follows:

1. How can access to information in the Operating Room be facilitated, namely the information that is currently displayed on static monitors scattered throughout the room?
2. Having a system that allows easy access to this information, how can the surgeon interact with it to control the contents being presented to him, knowing that both his hands are typically occupied and that he has specific surgical equipment placed on himself (limiting the available space for additional devices)?
3. How can this interaction be performed in a safe way, without compromising the sterility level required in the surgical environment?

The proposed solution addresses directly these questions. It enables the relevant clinical information about the patient (both pre-operative as well as real-time physiological data / vital signs) to be presented in line of sight. Furthermore, the surgeon and auxiliary medical team can navigate through the content during surgery with minimal FDs.

¹ Adapted from commons.wikimedia.org, with author: Pfree2014; minor changes applied to the image.

The AR system can be controlled through forehead gestures, resorting to electromyographic signals, collected by sEMG sensors included in the AR headset. Both the gestures and the control mechanism should be user-friendly and easy to operate, to guarantee that the system actually assists the doctors in viewing the information they need, rather than introducing additional entropy into the already complex process that is the surgery.

For this purpose, another goal of the project is to evaluate which type of forehead movements results in a more efficient control of the system, and to build the appropriate algorithm to identify them in the sEMG signal. Finally, one last component of the project consists in designing an experimental protocol to: (1) Validate the developed prototype; (2) Ensure the system is functioning properly; and (3) Evaluate its overall performance, in a simulated environment.

1.4 Achievements

This project included several components, both hardware and software related. The different tasks were successfully accomplished throughout this project, namely:

- Development of a mobile application for the visualization of clinical relevant information. This app is the interface between the user and the Augmented Reality (AR) display, and its contents are mirrored in real-time on the latter.
- Development of an interactive algorithm based on facial surface electromyography (sEMG), for hands-free control of the mobile app, and therefore, the AR system display.
- Design of an experimental protocol for the acquisition and analysis of facial sEMG data. This study was approved by the Ethics Committee of Instituto Superior Técnico (Statement nº 19/2021). Implementation of this protocol on a group of subjects, resulting in the creation of a sEMG signals' dataset.
- Interface of the mobile app with a general-purpose biomedical data acquisition platform for vital signs monitoring.
- Building of a hands-free EMG-controlled AR system full prototype, including the integration of sEMG sensors, compatible with standard surgical apparatus.
- Design of an experimental protocol to validate the developed prototype and evaluate its performance. Implementation of this protocol in two distinct groups of subjects: (1) a control group composed by regular individuals, and (2) an expert group composed by surgeon doctors.
- Interviews with different surgeon doctors, to gather requirements for the AR system, as well as possible constraints (related to the surgical environment), to consider upon building the solution.

1.5 Organization of the Document

This chapter outlined the context and motivation behind the present work, as well as the the main objectives and contributions accomplished. In Chapter 2 the relevant background concepts are introduced, with the main principles behind muscle physiology and key concepts on electromyography described. Free-hands modes of human-machine interaction are presented. Chapter 3 introduces the methodology for the development of a novel hands-free Augmented Reality system to be integrated in the Operating Room. Requirements and constraints are identified for the proposed solution. Chapter 4 details the technical implementation of the system and choices made for the final prototype, as well as the methodology validation protocol. Chapter 5 showcases the findings of the evaluation conducted on the different elements of the system, including the system's overall performance. Chapter 6 then outlines the main conclusions and proposes the future work.

2

Related Work

Contents

2.1 Augmented Reality	11
2.2 Hands-free Control Modes	12
2.3 Electromyography	17

2.1 Augmented Reality



Figure 2.1: Illustration of the different types of Mixed Reality Technologies: AR, AV and VR ¹.

Augmented Reality (AR) was created to simplify users' life, and refers to the integration of digital information - in the form of text, graphics, audio and other virtual enhancements - with the user's environment, in real time. AR technologies therefore include all systems that enhance the real world by superimposing computer-generated information on top of it.

With the recent advances in display and optical technologies, together with the continuously evolving digital processors, AR technologies have been emerging and increasingly applied to different industries, such as health care, education, engineering design, manufacturing, retail and entertainment [13].

Particularly in the medical field, AR technologies have offered a new approach for treating patients, explaining complex medical situations to patients and their relatives, educating and training medical professionals, and also for planning surgeries [14].

The concept of Augmented Reality can be viewed as part of the Reality-Virtuality (RV) Continuum framework. This framework, originally proposed by Paul Milgram and Fumio Kishino in 1994 [15], depicts the spectrum between real-world environments (reality) and virtual environments (virtuality). This taxonomy was recently revised to adapt to the context of modern technology and mixed reality modalities that have emerged over time [16], and currently comprises four main categories:

- **Physical Reality:** This category represents the real-world environment in which we naturally exist.

¹ Image retrieved from <https://covrisolutions.wordpress.com/2020/08/08/the-reality-virtuality-continuum/>.

It encompasses our physical surroundings, objects, and interactions without any augmentation or virtual elements.

- **Augmented Reality:** This category represents environments where virtual elements are overlaid onto the real world, allowing users to perceive and interact with both real and virtual objects simultaneously. Examples of AR include head-up displays, where digital information is projected onto a transparent screen in the user's field of view.
- **Augmented Virtuality:** In AV where the primary focus is on virtual elements with minimal real-world presence. In AV, virtual elements are dominant, but real-world objects or data are integrated into the virtual environment to enhance the overall experience or provide additional context. An example of AV is a virtual training simulation that includes real-world data or objects to provide context or improve realism.
- **Virtual Reality:** VR environments provide fully immersive experiences by completely replacing the user's real-world perception with a virtual environment. Users typically wear a head-mounted display and interact with the virtual world through input devices like controllers or hand tracking.

In essence, while VR fully transports the user into a virtual world, disconnecting them from the physical reality, AR enhances the perception of reality by overlaying virtual objects onto the user's live view [17]. This distinctive characteristic makes AR the most suitable technology to use upon this project, as it is crucial to ensure that the surgeon's senses, particularly vision, remain unobstructed and uncompromised during surgery.

2.2 Hands-free Control Modes

For the proper functioning and control of a technological device, its user must be able to communicate and interact with it, giving his input to the machine. This interaction, designated by Human-Machine Interaction (HMI), requires a user interface, that can include different input devices and modes. Common examples of HMI are the computer, in which typically the user uses a keyboard and a mouse as input devices to control it, or even smartphones, where the control is mainly done via touchscreen.

However, there are instances in which using the hands to control the different technologies is not the best option nor the most efficient, especially when the operator's hands must remain engaged in their primary task. Furthermore, individuals with physical impairments may face challenges or limitations in using conventional hand-dependent control mechanisms. Hence, in recent years there has been an emphasis on the development of alternatives that are hands-free, i.e., input modalities that don't depend on using the hands [18].

In this section the main hands-free input modalities will be briefly discussed, as well as some of their advantages and disadvantages.

2.2.1 Speech-based Control

Speech-based control typically uses machine learning methods to map the user's input speech waveform into corresponding text or a discrete output. These outputs are then used to identify a specific command that will trigger corresponding actions.

This type of control mechanism is very efficient for applications in which the speech commands are used to navigate through menu items, or perform direct control tasks (e.g. stop / resume), since isolated or connected words are sufficient as inputs. In these cases, changes in the computer display (e.g. icon highlighting) can serve as feedback to show the user that the voice command was well recognized.

On the other hand, continuous speech recognition systems that do not require pauses between words, are more adequate for applications in which the operator uses the system to fill out information fields, report forms (e.g. medical transcriptions) and related assets [19, 20]. For these cases, the recognized speech should appear in the display the user is interacting with, in order to give him feedback about the intended action.

However, there are still several challenges with voice control systems. One of them being background noise interference that can limit the system performance in high and dynamic noise environments, which is the case of the OR during surgical procedures [21]. To overcome this, the microphone should be put near the mouth of the user, which entails other hardware complications. Another challenge, particularly for single word commands, concerns the design of the input vocabulary. The commands must be simple and straightforward, so they don't add up complexity and do not distract the operator, but, on the other hand, they should also be specific enough so that parallel dialogues do not interfere with the system, and vice-versa.

For these reasons, this type of user interface, although hands-free, has constraints that limit its application in the control of the Augmented Reality system developed in this project.

2.2.2 Eye-based Control

Eye-tracking control technology is based on devices that measure the movement and position of the eye in real time. This type of user interface is better suited for control activities in which the user is looking at a display, since the direction of the gaze acts as a real-time pointer instead of traditional input sources such as the computer mouse. Eye dependent interfaces are especially useful for people with motor disability (e.g. persons with Amyotrophic Lateral Sclerosis - ALS), since they often present severe motor impairment, making it difficult to use standard input devices as controllers to communicate with the technology. There are different methods to track eye movements:

- Video-based tracking, using image processing to optically detect one (or more) features of the eye, and tracking them to map the eye' position. However, these type of methods impose some constraints in terms of computing power and conditions of use, such as lighting, the presence of spectacles, shape of the eyes, etc. Furthermore, many require the camera to be head-mounted - as illustrated in fig. 2.2a - which adds hardware and usability constraints [22].
- Using special contact lenses equipped with technology that can facilitate eye-tracking (e.g. encapsulated photodetectors). This method offers less constraints, but is more intrusive since it requires the placement of contact lenses directly on the eyes [23].
- Measuring the electrical potential of the muscles that position the eye, to infer about their movement, i.e., Electrooculography (EOG). EOG is based on the existence of an electrostatic field that rotates with the eye. By recording small differences in the skin potential around the eye resorting to electrodes, its position can be detected [18]. EOG-based eye-tracking is quite straightforward to implement, making it a good option when the aim is to execute simple commands, since the detection of eye movement - up/down, left/right, and eye-blinking - is discrete [24]. However, for optimal measurement of the EOG signals, ideally the electrodes should be placed near the upper and lower lids of the eye (as illustrated in Figure 2.2b) - to detect vertical movements - or on the external canthi, to detect horizontal movements. This aspect prevents the EOG-based eye tracking from being a good candidate to implement in this project, since having electrodes placed near the eye during extended periods of time can be uncomfortable for the surgeon, and can also disturb the correct functioning of other surgical apparatus, such as the magnifying glasses.

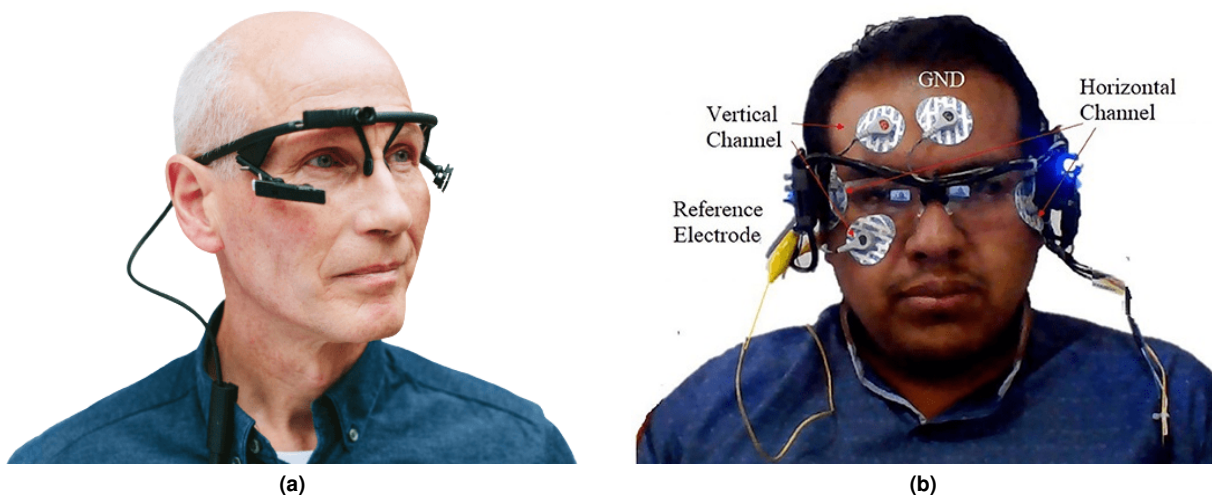


Figure 2.2: Illustration of eye-based controlled devices: On the left, a wearable eye-tracking headset, manufactured by Pupil Labs². On the right, an EOG-based HMI application [25].

² Image retrieved from <https://pupil-labs.com/products/core/>

2.2.3 Brain-Computer Interfaces

Brain-Computer Interfaces (BCIs) are systems that establish a direct communication between the brain and an external device (e.g., computer, prosthetic limb), without relying on traditional physiological pathways like muscles or nerves. This type of interfaces recognize the user's intention of command by detecting and interpreting brain activity. The activity of the brain can be measured with sensors implanted inside the body (invasive BCIs) or with external sensors (non-invasive BCIs). For non-invasive BCIs, this is achieved by placing a set of surface sensors on the scalp of the user, and measuring either electric activity (with electroencephalography, EEG), magnetic activity (with magnetoencephalography, MEG), or metabolic activity (e.g., with functional near-infrared spectroscopy, fNIRS) [26]. fNIRS measures changes in blood oxygenation to infer brain activity, but its low temporal resolution limits its ability to capture rapid changes accurately, and therefore, is not frequently used for BCIs. Conversely, EEG and MEG allow higher temporal resolution. However, although MEG provides higher spatiotemporal resolution than EEG, it also requires a large and expensive magnetometer, hence, generally, for non-invasive BCIs, EEG is the preferred technology, as illustrated in Figure 2.3.

Electroencephalographic (EEG) activity is assessed by positioning electrodes at various points on the scalp. The electrical signals are recorded as voltage differences between two electrodes, typically with one serving as a reference for the others. Measuring the activation of a specific brain region requires positioning electrodes as close as possible to that area, and therefore electrodes placed at different cortical positions allow measuring different neural processes. EEG allows recording different patterns of brain activation: Spontaneous brain activity, which is measured in the absence of any explicit task or stimulus, and induced activity, that emerges as a response to an event, such as a sensory stimulus or a specific action. Thus, EEG-based BCI systems rely on detecting changes in the brain patterns produced as a response to some voluntary or involuntary mental command. One of these types of neural processes that BCIs can use as inputs are Event-related Potentials (ERPs), which appear as a response to external sensory stimuli. These include P300, Steady-state evoked potentials, and Error-related potentials. On the other hand, BCIs can also use neural processes that do not require explicit stimuli, being instead associated with internal brain events, for e.g., Event-related desynchronization/synchronization (ERD/ERS), and Slow cortical potentials (SCP) [27]. Although EEG-based BCIs are a promising technology with benefits in various fields, they have several limitations, such as lower signal resolution compared to other methods, susceptibility to artifacts and noise contamination, and the need for proper electrode placement and signal processing techniques. Furthermore, EEG signals may be affected by individual variations, making it necessary to calibrate the system for each user, and require a larger training period [28]. These added constraints, both hardware and processing-wise, particularly the wiring setup, which involves placing the electrodes along the scalp, may jeopardize the necessary sterility and interfere with the medical apparatus.



Figure 2.3: Illustration of an EEG-based Brain-Computer Interface. ³.

2.2.4 EMG-based Control

Electromyography (EMG) based control involves using muscle electrical activity to interface with and control different devices and technologies. The EMG signals are captured by placing surface electrodes on the skin above the interest muscle(s), and analysing the electrical signals produced during muscle activation. This type of signal is particularly useful for HMI, assuming the subject does not have any impairments that hinder the necessary muscle activity.

Different studies have proven the efficiency of EMG-based HMI across diverse fields, including in Prosthetics and Assistive Devices [29–31], Wearables and Applications [32] and Proactive Healthcare and Wellness [33, 34].

EMG-based control allow the monitoring and interpretation of muscle activity in various body parts such as limbs, hands and face, and, unlike EEG-based control, which typically requires a larger number of electrodes, can be achieved with a relatively smaller number of electrodes. Furthermore, EMG-based control provides the ability to distinguish commands based on the duration of the stimulus, offering an additional level of flexibility and control options. By analyzing the temporal characteristics of EMG signals, such as the duration of muscle activation, users can perform different actions and commands, resulting in a more adaptable and versatile control interface. For this reasons, EMG is an efficient and practical choice for implementing control interfaces.

³ Image retrieved from https://www.resna.org/sites/default/files/conference/2017/emerging_technology/Obiedat.html.

2.3 Electromyography

EMG consists in measuring the electrical activity of skeletal muscles, resorting to electrodes placed directly on the muscle, or on the skin above it - surface electromyography (sEMG). In the context of pathology, this technique can be used to assess the health of muscles and the nerve cells that control them (motor neurons), helping to diagnose a variety of neuromuscular diseases, motor problems, nerve injuries, or degenerative conditions, such as ALS and muscular dystrophy. However, EMG can also be used in biomedical related applications, such as prosthetics control, grasp recognition, exoskeletons, and human computer interaction [35].

2.3.1 Muscle Physiology

The electrical activity capable of being measured by the electrodes, is produced from two states of a skeletal muscle - at rest, and during contraction. When the skeletal muscle is at rest, each of the muscle cells has an electric potential of approximately -80 mV (resting potential). However, when the muscle is contracted, it generates an electric potential in its Motor Unit (MU). MUs represent the anatomical and functional element of the neuromuscular system, and designate each group of muscle fibers and corresponding motor neuron. These electric potential differences are produced when a motor neuron activates a neuromuscular junction, by sending two intracellular action potentials in opposite directions. These action potentials are then propagated by depolarizing and re-polarizing each one of the muscle fibers [36]. Hence, the sum of the intracellular action potentials of all muscle fibers of a MU is called a Motor Unit Action Potential (MUAP), and the EMG signal captured by the electrode is a linear summation of several trains of MUAPs. The MUAP waveform can be characterized by a number of parameters related to the structure and physiology of the MU. The most commonly analyzed parameters are the MUAP's amplitude, duration, phase, turn and baseline. Figure 2.4a illustrates a (1) motor neuron (and its axons), a (2) muscle (composed by a set of muscle fibers), and (3) the connection between the axon terminals, and each muscle fiber, through the neuromuscular junction. Figure 2.4b shows a MUAP, being recorded by an EMG electrode, as the summation of the action potentials of all the single muscle fibers present in the recording uptake area of the electrode. It also shows the MUAP typical waveform, and the main parameters that can be calculated.

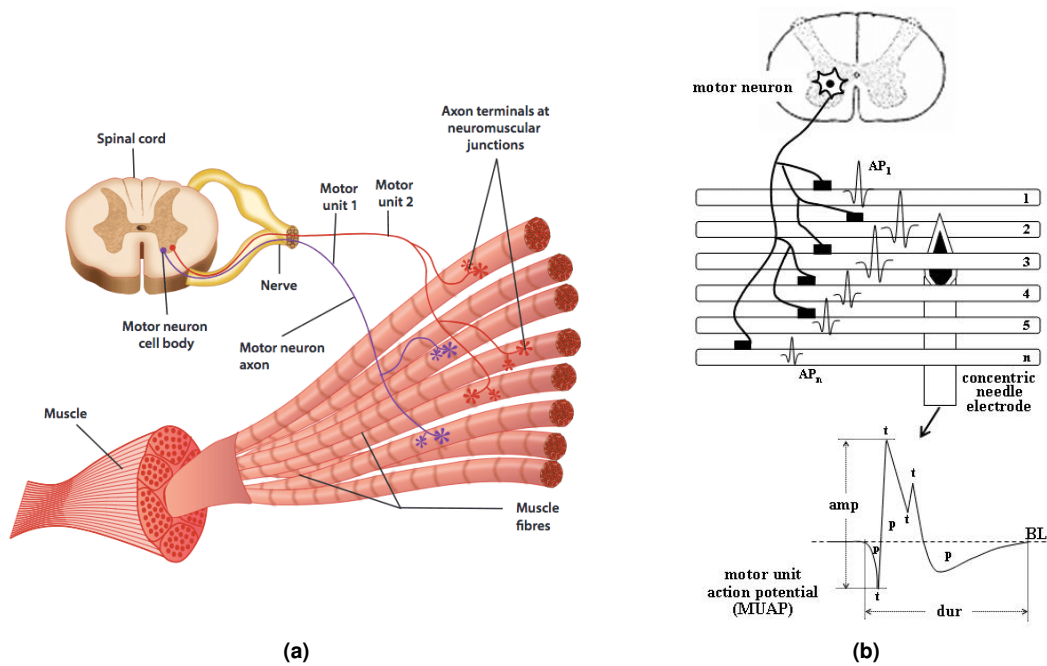


Figure 2.4: (a) Illustration of a Motor Unit (MU) with emphasis on its main components - one motor neuron, connected with group of muscle fibers ⁴. (b) Schematic representation of a motor unit with n muscle fibers [36]. The algebraic summation of the AP of all the single fibers present in the recording uptake area of the electrode ($AP_1+AP_2+\dots+AP_n$) generates the motor unit action potential (MUAP). The main parameters of the MUAP waveform are indicated: amp = amplitude; dur = duration; p = phase; t = turn; BL = baseline.

2.3.2 EMG Signal

As described in Section 2.3.1, EMG is a biosignal that results from the electrical variations generated by the MUAPs during muscle contraction. It is worth mentioning that there are two types of muscle contraction - static and dynamic. The first one happens when the joints do not move, and so the lengths of the adjacent muscle fibers do not change (e.g. maintaining the hand contracted to hold a cup of coffee). The second occurs when the joints are in motion, which causes the muscle fibers to also change their lengths (e.g. waving the hand to do a salutation gesture). The EMG signal can be modeled as a stochastic process that depends on both the static and the dynamic types of muscle contraction. The Mathematical Model for a Static Contraction (MMSC) can be described as a stationary process, because the mean and covariance remain approximately the same over time, and the EMG signal depends exclusively on muscle force. Equation (2.1) represents the MMSC, where N is the number of active MUs, $s_i(t)$ is the train of impulses that indicate the active moments of each MU, $m_i(t)$ are the MUAPs of each MU, and $*$ represents the convolution function.

⁴ Image retrieved from <https://awesomeopensource.com/project/iandanforth/pymuscle>.

$$EMG(t) = \sum_{i=1}^N s_i(t) * m_i(t) \quad (2.1)$$

On the other hand, the dynamic contraction is a non-stationary process, given by Equation (2.2), where $a(t)$ represents the intensity of the EMG signal (information signal), $w(t)$ corresponds to the unit-variance Gaussian process representing the stochastic aspect of the EMG (carrier signal), and $n(t)$ is the noise from the sensors and physiological artifacts [37].

$$EMG(t) = a(t)w(t) + n(t) \quad (2.2)$$

However, the signal measured by the electrodes will greatly depend on their placement in relation to the muscle fibers being analysed - the closer to the muscle fiber the electrode is placed, the better the EMG signal acquired. In the clinical context, the EMG signal is often even measured with the electrode inserted directly on the target muscle, through a needle, which is an invasive procedure, as illustrated in Figure 2.4. In this case, the area of the recording surface has typically around 1 mm of radius, and comprises about 100 muscle fibers. Although each MU has hundreds of muscle fibers, these are widely spread throughout the cross section of the muscle, which makes the needle electrode record only around 4-6 fibers from each MU [38].

The waveform and firing rates of the MUs captured in the EMG recording can give relevant diagnostic information. When it comes to using EMG for human computer interaction, the signal is not recorded directly on the target muscle, but instead, on the skin surface overlying said muscle, resorting to surface Electromyography (sEMG), to make the process less invasive and overall more feasible in out of the lab contexts. However, this surface recording entails some limitations, particularly, the intensity and overall quality of the acquired signal. With sEMG, excitation level is acquired from a large area including several MUs, being mostly used to measure superficial, large, and easily accessible muscles, hence, selective recording of deep muscles is not possible [39]. Also, due to the relatively large pick-up area of the EMG skin electrodes, unwanted signals from neighboring muscles may be recorded i.e., crosstalk. To minimize this error source, aspects such as electrode placement in relation to the target muscle, electrode size, and inter-electrode distance must be closely pondered. Another factor that influences the sEMG signal quality is the presence of other tissues between the target muscle and the electrode, namely, adipose connective tissue, and skin. The passage of the EMG through the tissues attenuates the signal's amplitude and frequency, and depends on the muscle to skin distance [40].

Furthermore, especially during dynamic muscle contractions, the movement of the muscle in relation to the overlying skin can introduce errors and unintentional variations in the EMG signal.

Lastly, two other factors that can affect the EMG signal relate to the properties of the sEMG electrodes and amplifiers used - namely, the type of adhesion the electrodes have with the underlying skin. If the

amplifiers designed to amplify the captured signal, and the process of converting it from analog to digital (A/D conversion), are deemed acceptable, the Signal-to-Noise Ratio (SNR) - measure that reflects the signal's quality - depends mainly on the electrode to skin contact. The electrodes' contact surface can be either the metal they are constituted by (e.g. in the case of dry electrodes), or they can be covered by an electrolytic, conductive gel layer (e.g. in the case of pre-gelled electrodes). This layer is added with the aim of reducing the impedance on the electrode-skin interface, which allows to measure the signal with a higher intensity (and therefore higher SNR), less distortions, and a lower power line interference [41,42].

Although the described limitations can impact the quality of the EMG signal recording, nowadays it is possible to use sEMG in a wide range of biomedical applications, provided the requirements for its optimal usage are accomplished. The average EMG signal measured with surface electrodes has amplitudes ranging between 0 and 10 mV (peak-to-peak) prior to amplification, and frequencies between 10 and 250 Hz [43, 44].

Figure 2.5: Time domain (top) and frequency domain (bottom) representation of an EMG signal.

3

Methodology

Contents

3.1 Workflow	23
3.2 Requirements and Hardware Restrictions	23
3.3 Application Structure and Displayed Pages	31

3.1 Workflow

Building upon the description of the existing solutions, and motivated by the lack of suitable hands-free control methods for AR headsets in the context of OR, as characterized in Chapter 2, the present chapter aims to detail how the proposed solution was developed and implemented. Figure 3.1 outlines the main components of the proposed approach.

Figure 3.1: Schematic diagram of the Augmented Reality System main components: Augmented Reality head-mounted display; BITalino for EMG control of the headset, with connected sEMG electrodes; ARSurgery mobile application. Optional components for direct acquisition of patient's biosignals: EpiBOX interface; BITalino connected to the patient

The developed AR system is composed by different components that interact with each other:

- An optical see-through AR Headset with hands-free control, which should be placed and adjusted on the user's head;
- A mobile application - ARSurgery - whose screen is mirrored on the AR headset display in real-time, i.e., the content being displayed for the user is the application itself;
- A BITalino with an EMG sensor using three electrodes (bipolar montage and reference electrode), to collect data from which the commands used to interact with the ARSurgery app will be detected [45];
- The mobile application also allows to connect a second BITalino measuring biosignals directly from the patient, if desirable. This connection is made through a Raspberry Pi device, configured for that purpose (EpiBOX). EpiBOX connects to the patient's BITalino and ARSurgery via Bluetooth and WiFi technologies, respectively.

The system and the choices made its development will be further explained in the next sections.

3.2 Requirements and Hardware Restrictions

The system was designed having in mind two main aspects: on the one hand, the characteristics and features that were found to be key for its appropriate functioning, and, on the other hand, the design requirements and restrictions existing within the OR and surgery context. Different aspects and features were considered due to the specific context of application, namely: (1) The Augmented Reality headset requirements, to be able to integrate it in a surgical environment without disrupting the surgery's flow; (2) The type of sEMG electrodes being used to capture the control signals from the user; and (3) The facial triggering signals chosen as input for the HCI.

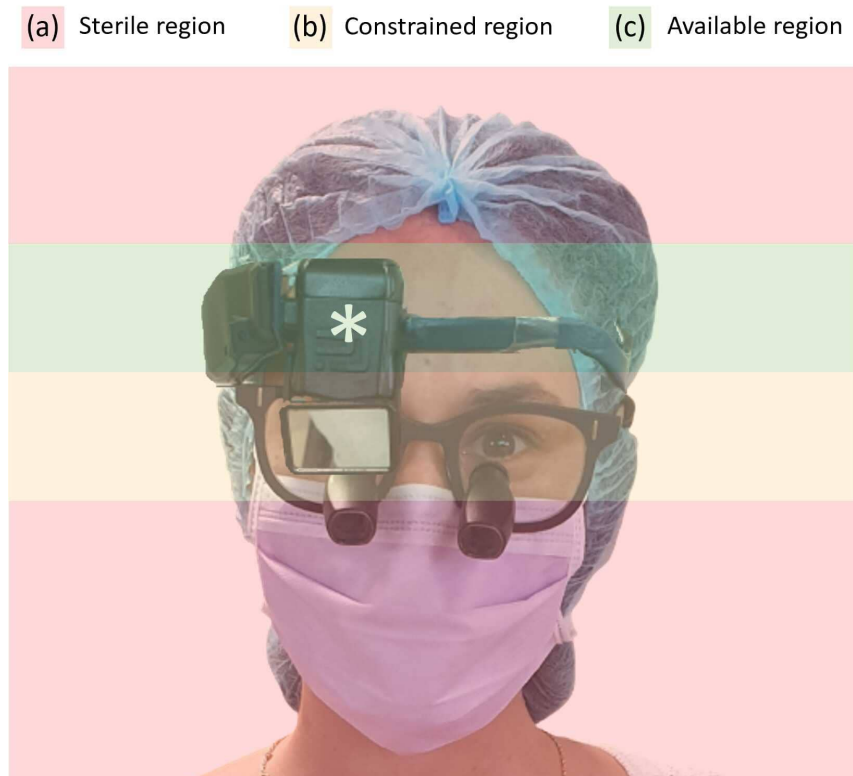


Figure 3.2: Illustration of the AR headset superimposed with the surgeon's commonly required equipment. The figure divides the surgeon's head / face in three different regions according to usability constraints and available space: (a) Sterile region, (b) Constrained region and (c) Available region.

Figure 3.2 illustrates the AR headset superimposed with the surgeon's commonly required equipment. In the figure, the surgeon's head / face is divided in three different regions according to usability constraints and available space: (a) Sterile region - areas that must remain protected with surgical cap and mask in order to maintain a sterile environment; (b) Constrained region - area where the magnifying glasses and/or surgical headlamp occupy the surgeon's face, hence being an already obstructed area; (c) Available region - mostly unobstructed area, where the AR headset and EMG electrodes may be placed. It is important to note that the headset component comprised by the projector and the display, marked with an (*) in the figure, when placed on the user's head, occupies part of the forehead (either above the left or right eye).

3.2.1 AR Headset Requirements

Regarding the main device of this system - the Augmented Reality headset - to be used by the surgeon doctor and / or auxiliary medical team: during surgery, the head-mounted device must be completely **see-through**, to allow a clear view without any obstructions in the field of vision. Additionally, this device must be **easy to both put on and remove**. Firstly, to avoid unnecessary time delays in pre- and post-

surgery. Also, in case of emergency or malfunctioning of some part the system, it is imperative that the surgeon is able to easily remove the headset so it does not further disrupt the procedure. Ultimately, the technology being user-friendly promotes its acceptance by the users, so the simpler the device is to put on, the better. It is also important for the device to be **light and adjustable**, so that it lays comfortably on the surgeon's head and does not increment his fatigue, especially since it will be maintained for the duration of the procedure. Finally, a key factor to consider was the surgical apparatus and headwear the surgeon already has to use when performing surgery. This includes surgical clothing (scrubs, gown and cap) and surgical mask. Also, depending on the type of surgery, the doctor may use magnifying glasses and / or surgical headlamp. In Figure 3.3 examples of these surgical items are presented.



(a) Surgical mask



(b) Surgical scrub cap



(c) Magnifying glasses



(d) Surgical apparatus

Figure 3.3: Standard surgical apparatus used by surgeon doctors during surgery.

These equipment, particularly the magnifying glasses and headlamp, occupy a significant part of the surgeon's forehead, so the chosen AR headset must be compatible with those devices, and adjust on the head in such manner that does not interfere with them.

Lastly, a feature that was considered as a nice-to-have for the headset, was having an integrated camera. Having a camera merged in the headset allows for further developments in terms of the features the system can offer to the user. Examples of these features will be detailed in Section 6.2.

Taking all these requirements into account, different AR headset models were analysed, and for the prototype it was decided to use the ViP-display (version A1) AR Headset manufactured by ROTACIONAL¹, a portuguese company dedicated to the development of electronic devices, particularly, this type of AR Headsets. This model was selected since it is completely see-through, fully compatible with the use of eyeglasses, adjustable to different users' head anatomy, and also has adjustable brightness / contrast, which allows to adapt to various light environments. ViP-display - the chosen AR headset model chosen for the development of this prototype - is shown in Figure 3.4 and its main specifications are presented in Table 3.1.

¹Available at <https://rotacional.com/vipdisplay/>.



Figure 3.4: ViP-display (version 1) AR headset, manufactured by ©ROTACIONAL.

Table 3.1: Summary of the ViP-display AR headset main specifications

ViP-display AR headset main specifications		
Display	Resolution	1920×1080p (FullHD)
	Color Depth	RGB 16 Million colours
	Contrast Ratio	>10 000:1 (color)
	Maximum Luminance	≥ 100 to 500 Cd/m ²
	FOV (diagonal)	≥ 30°
	Eye Relief	Adjustable (worn Eyeglasses Simultaneously)
Luminance Control	Brightness adjustment	Indoor and Outdoor
	Contrast adjustment	Indoor and Outdoor
Connection	Communication	Wi-Fi
	Power Supply	Built-in (autonomy for 1 work shift)
Measurements	Dimensions	185×44×228 mm (w × h × l)
	Weight	111 g

3.2.2 Choosing the Appropriate Type of sEMG Electrodes

Regarding the type of sEMG electrodes to be used in this system, there was the necessity to choose dry electrodes, instead of the more standardly used gelled self-adhesive electrodes. Although the latter type typically performs better in terms of the EMG signal quality that it records, it is also less practical to apply on the skin and requires a more careful maintenance resulting also in higher costs. The choice of using dry electrodes in this project was made considering different studies supporting that signal quality obtained by dry Ag/AgCl electrodes, although slightly inferior to pre-gelled Ag/AgCl electrodes, is satisfactory, and has been progressively used in HCI applications [46]. The electrodes were connected via 3-lead UC-E6 cable to a BITalino EMG sensor ². The main characteristics of the EMG sensor are presented in Table 3.2.




²Available at <https://www.pluxbiosignals.com>

In terms of the dry Ag/AgCl electrode model used in the developed prototype, two different models were tested and compared. The two models differed in terms of shape and distance to the skin – one being more elevated (Er) and the other one, more proximal (Ef) to the skin. In addition, for comparison purposes, tests were also performed with standard gelled self-adhesive electrodes (Eg), considered as the reference in this study, since they are typically more conductive than the non-gelled type. The different electrodes are illustrated in the Table 3.3. Results for these comparison tests will be presented and discussed throughout Section 5.1.2.

Table 3.2: Summary of PLUX Biosignals' EMG sensor main characteristics

Parameter	Value
Gain	1009
Range	± 1.64 mV (with VCC = 3.3 V)
Bandwidth	25-480 Hz
Consumption	~ 0.17 mA
Input Voltage Range	2.0-3.5 V
Input Impedance	7.5 GOhm
CMRR	86 dB

Table 3.3: Illustration of the electrodes used in the signals' acquisitions and summary of their main characteristics

Electrode denomination	Illustration	Sensor material	Conductive gel	Durability	Diameter
Raised profile - Er		Ag/AgCl	No	Reusable	11 mm
Flat profile - Ef		Ag/AgCl coated polymer	No	Reusable	10 mm
Gelled (reference) - Eg		Ag/AgCl coated polymer	Yes	Disposable	24 mm

3.2.3 Facial Expressions Chosen as Input for the HCI

To choose the type of movements that act as inputs for navigating through the content displayed in the AR headset screen, different aspects were taken into account. Firstly, the control mechanism has to be hands-free, and interfere the least possible with the primary tasks the user is performing. Limb movements are not the most appropriate since usually the surgeon and auxiliary medical team already have them occupied. Thus, it was decided to use facial commands, i.e., facial EMG signals, as inputs for the AR system. There are a range of facial movements commonly used in HCI applications, such as

lower lip raising, left and right side smirks, closing lips, clenching molar teeth, etc. [47, 48], but these involve placing electrodes on the lower half of the face in order to get useful signal. On the other hand, choosing movements that move muscles of the face's superior half, allows to also place the electrodes on that region (Figure 3.2), which is more convenient since the electrodes may then be integrated in the AR headset headband. Hence, three different commands were selected for analysis as potential commands: single-eye winking (W), brief eyebrows' raising (BER) and extended eyebrows' raising (EER). These three command movements are illustrated in Figure 3.5. Additionally, since periodic blinking is a physiological human behaviour, this movement was also included in the analysis, to study its impact on the acquired electromyographic signal and consequent interaction with the AR system.

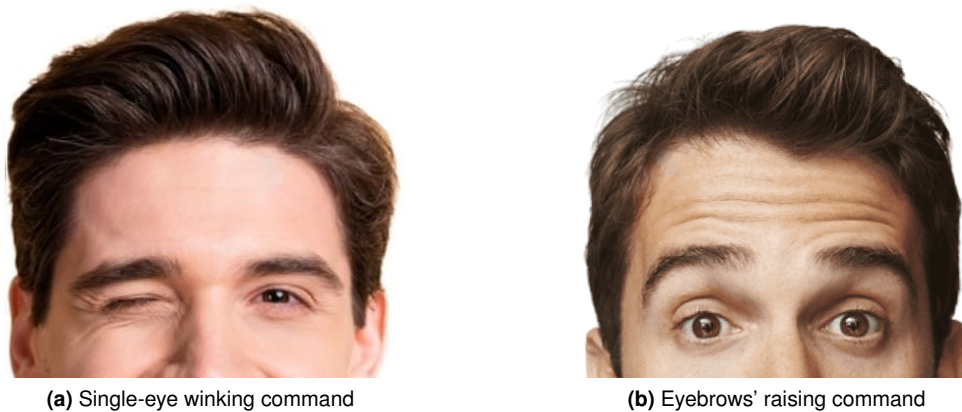


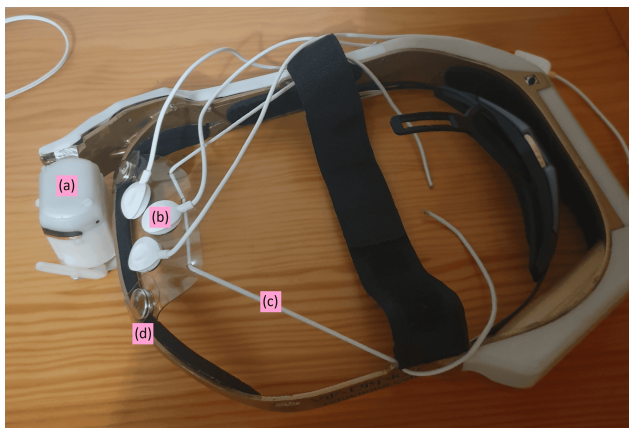
Figure 3.5: Illustration of the facial commands chosen as input for the AR system: single-eye winking, brief eyebrows' raising and extended eyebrows' raising.

3.2.4 Electrode Placement on the Face

Building upon the logic followed to choose the most adequate input commands, the electrode placement was also chosen considering the usability requirements for this system. In order to minimize the impact of the AR system in the surgeon's primary tasks, and due to the limited space available in the surgeon's head (as described in Section 3.2.1) it was determined to integrate the electrodes in a position prone for integration in the AR headset itself, within the headband. Furthermore, as shown in Figure 3.2, the headset itself already occupies a portion of the user's forehead, hence, it is desirable to place the electrodes in the medial region, between the two *frontalis* muscles, and above the *procerus* muscle.

It is worth mentioning that bipolar montages of EMG electrodes benefit from placing the electrodes adjacent to each other, along the direction of the muscle being measured, while placing a reference electrode further away, in an electrically neutral tissue. In this type of montage, the two electrodes capture two distinct potentials in the muscle of interest, with respect to the reference electrode, being then differentially amplified, thus eliminating common components in the two signals.

However, for HCI applications, and because the purpose of measuring the EMG signal is to distinguish between muscle rest and muscle activation in the temporal domain, and not necessarily to gather information about the EMG signal shape, the best approach in terms of usability is to place the electrodes horizontally, with the reference electrode slightly above. Figure 3.6a shows the AR headset with an attached support substrate built for the purpose, that fastens to the headband with snap buttons, where the electrodes can be adjusted according to the user's forehead. Particularly, it shows (a) the headset's projector, (b) the EMG electrodes, (c) an elastic band which, in this prototype, adjusts the support substrate to the user's forehead, and (d) the snap buttons that adjust the case to the headband. Figure 3.6b illustrates the surgeon using the modified headset together with their remaining surgical equipment. The adequacy and quality of EMG signals being acquired with this specific montage were assessed through comparison tests, that will be presented and discussed in Section 5.1.2.



(a) ViP-display AR headset with attached electrodes



(b) Surgeon using the modified AR Headset

Figure 3.6: AR head-mounted display with attached EMG control system. The left image depicts the main components of the system: (a) AR headset projector; (b) bipolar montage of EMG electrodes to control the system; (c) adjustable strap to maintain the electrodes correctly placed on the user's forehead; (d) support substrate with snap buttons. to fix the EMG electrodes. The right image illustrates a surgeon using the developed system, together with the surgical apparatus used during surgery.

3.3 Application Structure and Displayed Pages

The mobile application, named **ARSurgery**, developed as the interface for the AR display, was built assuming two phases: a (1) setup phase, and (2) the display phase itself, which is hands-free. The setup phase, leads the user in connecting the different devices to the app (via Bluetooth and WiFi), and allows the user to calibrate the app in order to optimize the EMG commands recognition. It is worth mentioning that the auxiliary devices (BITalino and EpiBOX) should be switched on beforehand, and be ready to allow connections. Additionally, the smartphone where the app is running must have Bluetooth and WiFi enabled. After the devices are connected and the app is conveniently calibrated, the user can proceed to control the displayed information on the app, and consequently, on their AR headset display, in real-time.

The hands-free control mode allows the user to switch between the different pages of the app, each one containing relevant clinical information about the patient being intervened, e.g., their biosignals, or previous imaging exams. The built prototype, illustrated in Figure 3.7, includes two pages among which the user can navigate: (1) Biosignals Viewer and (2) Monitor Viewer. The screen designated Biosignals Viewer continuously displays the patient's biosignals being acquired (through a dedicated BITalino connected to the patient). The second screen, designated Monitor Viewer, continuously presents the view from the phone's camera. This feature can be used to emulate one of the monitors present in the OR, chosen according to the needs of the surgeon, by positioning the phone's camera in front of said monitor. The user can switch between these two pages, and a third one that is completely dark, acting as a hands-free switch-off for the head-mounted AR display.

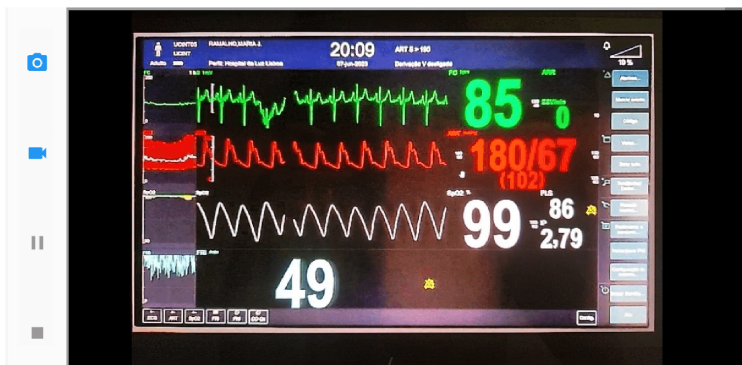
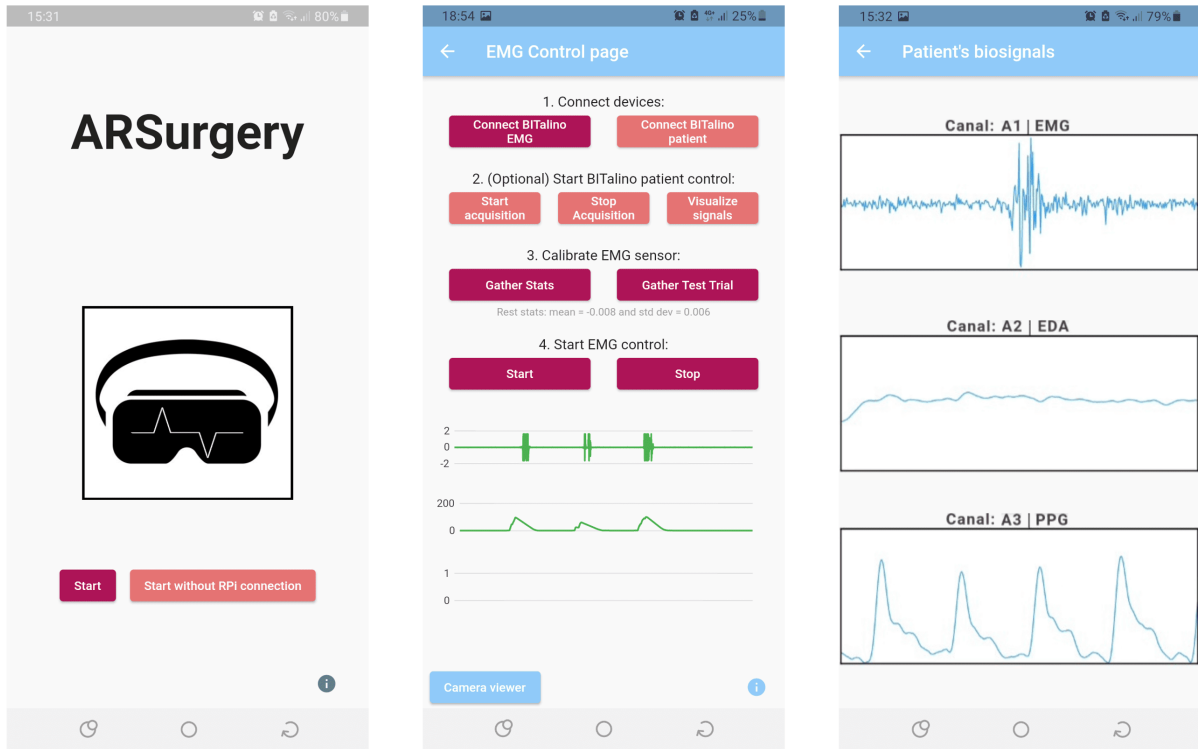


Figure 3.7: ARSurgery displayed pages. At the top, from left to right: Home page; Calibration Page; Patient's Biosignals Viewer page. At the bottom, Camera Viewer page.

4

Implementation

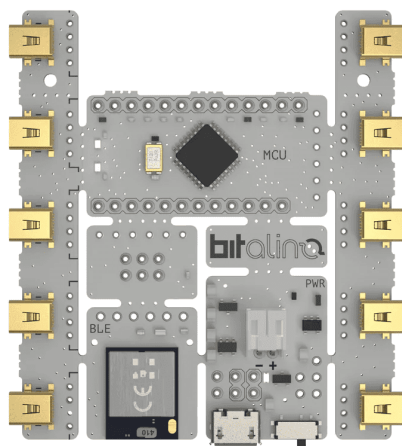
Contents

4.1 BITalino Integration	35
4.2 Final Prototype	36

This Chapter outlines the implementation of the prototype, providing a comprehensive overview of the choices made throughout its development and detailing the technical aspects of the implementation process. Furthermore, it describes how the different components of the methodology were validated and explains the evaluation of the system's performance.

4.1 BITalino Integration

To acquire the EMG signal responsible for the app control, there needs to be an interface which collects the data, and sends it to the app in real-time. There are different options available that are able to serve this purpose, but considering the usability factor and the context where the system is to be used, wireless technologies are more advantageous. Hence, it was decided to use a *BITalino (r)evolution Board BLE*¹, to collect the signal arising from the EMG electrodes. Its main characteristics are presented in Table 4.1. This all-in-one board has all the needed electronic blocks for biosignal acquisition pre-connected, and can connect via Bluetooth Low Energy (BLE) to the device handling the captured data, in this case, the smartphone with ARSurgery mobile application. BITalino connects to the EMG sensor via micro-USB, and is fully portable as it already comes with a rechargeable LiPo battery.



(a) BITalino (r)evolution Board BT/BLE



(b) BITalino (r)evolution Board BT/BLE protected with a 3D printed PLA enclosure

Figure 4.1: BITalino (r)evolution Board, featuring its assembled version.

Furthermore, BITalino can record a variety of biosignals additionally to the EMG, including Electrocardiography (ECG), Electroencephalography (EEG) and Electrodermal Activity (EDA). With this feature, a second BITalino device can be used to record biosignals of the patient being intervened, and present them directly on the mobile app (and therefore on the surgeon's AR headset), without the need to emulate the OR monitor in which they are being displayed.

¹ Available at <https://www.pluxbiosignals.com/products/bitolino-revolution-board-kit-ble-bt?variant=41622003941567>.

For this, it is necessary to connect the second BITalino to the ARSurgery app. Since most smartphones do not allow pairing two Bluetooth devices simultaneously, a workaround solution was to connect the optional BITalino measuring patient's biosignals to the mobile app via Wi-Fi. As the BITalino board used in this project does not have built-in Wi-Fi connection, the alternative was to use an adaptation of *EpiBOX* [49] as the interface between this second BITalino and ARSurgery. EpiBOX was developed as an automated platform that enables the long-term acquisition of biosignals. This platform uses a *Raspberry Pi* as the recording unit (EpiBOX Core), and is coupled with a Python software (PyEpiBOX) responsible for the data communication, acquisition configuration, and storage, and with an Android mobile application (EpiBOX App) providing the interface for the interaction with the system. The EpiBOX platform was adapted for this project, to incorporate the visualization of the data coming from the patient's BITalino, directly in ARSurgery.



Figure 4.2: Architecture of the (optional) connection between ARSurgery and the BITalino measuring patient's biosignals, through the EpiBOX interface.

Table 4.1: Summary of BITalino (r)evolution Board BT/BLE main specifications

BITalino main specifications	
Sampling Rate	1, 10, 100 or 1000Hz
Analog Ports	4 in (10-bit) + 2 in (6-bit) + 1 auxiliary in (battery) + 1 out (8-bit)
Digital Ports	2 in (1-bit) + 2 out (1-bit)
Communication	Bluetooth or BLE
Range	up to 10m (in line of sight)
Sensors	EMG; ECG; EDA; EEG; ACC; LUX; BTN
Size	100x65x6mm
Battery	500mA 3.7V LiPo (rech.)

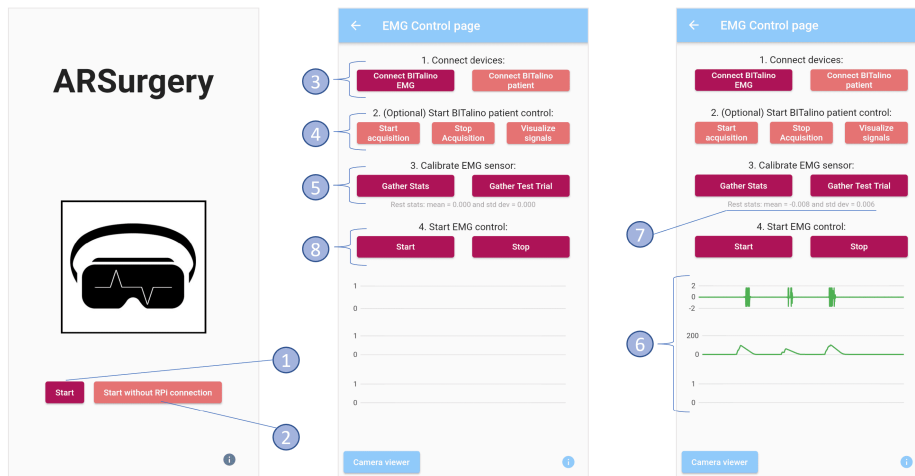
4.2 Final Prototype

4.2.1 Application description

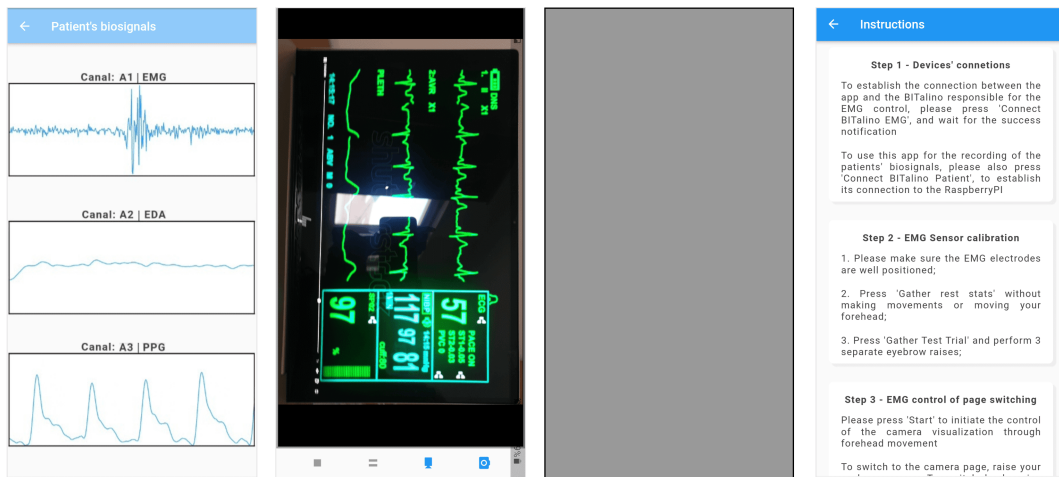
In the setup phase, illustrated in Figure 4.3a, the user should start by connecting the devices of the system to the app. By pressing the top left and right buttons (Figure 4.3a - ③), the app automatically

connects, respectively, to the BITalino measuring the control EMG and to the BITalino acquiring biosignals from the patient, in case it exists. It is important to note that the homepage has two start modes the user can choose from - with or without measuring patient's biosignals (Figure 4.3a - ① or ② respectively). After the BITalino devices are connected, the EMG sensor from the user's forehead should be calibrated. For this step, the user must press the calibration buttons (Figure 4.3a - ⑤) and follow the indications that appear on the screen. For the "Gather Stats" button, this consists in remaining still for 7 seconds while the app retrieves some characteristics of the EMG signal during rest (indicated in Figure 4.3a - ⑦). It is worth mentioning this step is only completed if the standard deviation value obtained is inferior to 0.2, otherwise, the user must readjust the electrodes and do another trial. This value was decided based on the average standard deviation of the test trails performed beforehand. As for the "Gather Test Trial", the steps consist in performing three separate eyebrows' raising movements, one of the triggering movements chosen as a command for the application. On the lower half of the calibration page, the EMG signals being acquired for the calibration are presented in real-time (Figure 4.3a - ⑥). This can give the user a visual feedback of their actions, and inform them about the intensity and similarity of his commands. After this, the user may press "Start EMG control" (Figure 4.3a - ⑧) and the application is ready to be controlled by the user in a hands-free manner.

Resorting to the facial commands, the user can navigate between three different screens in the ARSurgery displays - Patient's biosignals viewer, Monitor viewer, and black screen - previously described in Section 3.3. The Monitor viewer, in addition to display the OR chosen screen in real-time, also allows to photograph and/or record said screen. This feature can become useful for keeping track of the patient's health state during surgery, and for subsequent teaching purposes, especially if an outlier procedure takes place. Additionally, future advances in this system that use an AR headset with integrated camera, may take advantage of this recording option to actually film the surgery from the user's point of view. Furthermore, ARSurgery includes an instructions page (Figure 4.3b - right), where the user can find a summarized explanation for each step.



(a) ARSurgery displayed pages for the setup phase, with sequential guidance.



(b) From left to right: Patient's Biosignals Viewer page; Camera Viewer page; Off / See-Through Mode; Instructions page

Figure 4.3: ARSurgery displayed pages, with sequential guidance for the setup phase.

ARSurgery was developed in such way that it can be used in both portrait and landscape modes. Thus, it allows the user to explore the app in portrait mode, if desired. However, if the application is mirrored on the AR headset right at the beginning, as it should, it can be fully used in landscape mode, maximizing the headset screen usable area.

Finally, the app - particularly the calibration phase - was all-around protected with a blocking / notification system that prevents the user from progressing in an undesired order or missing steps, e.g., in case the user selects "Connect BITalino EMG" without having the smartphone Bluetooth enabled, the app warns the user that he should turn it on beforehand. While the app is gathering the EMG signal characteristics, it notifies the user with a "Gathering rest stats" popup, as well as a "Stats completed!" once the process is concluded. Examples of these messages are shown in Figure 4.5.

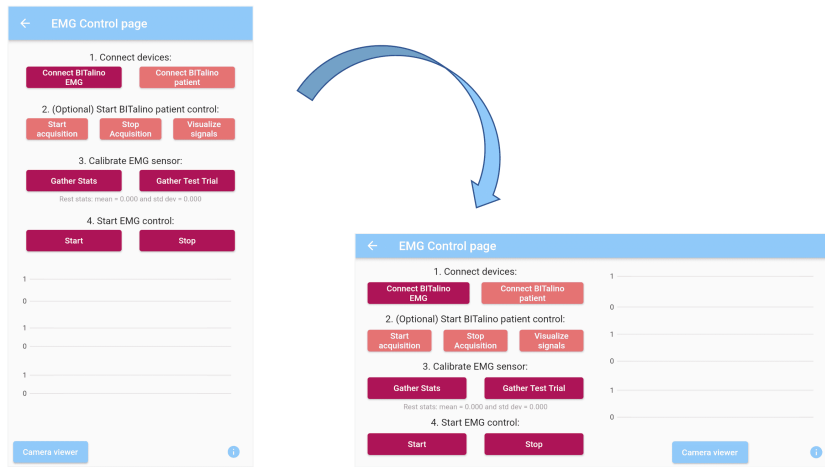


Figure 4.4: Illustration of ARSurgery portrait and landscape modes.

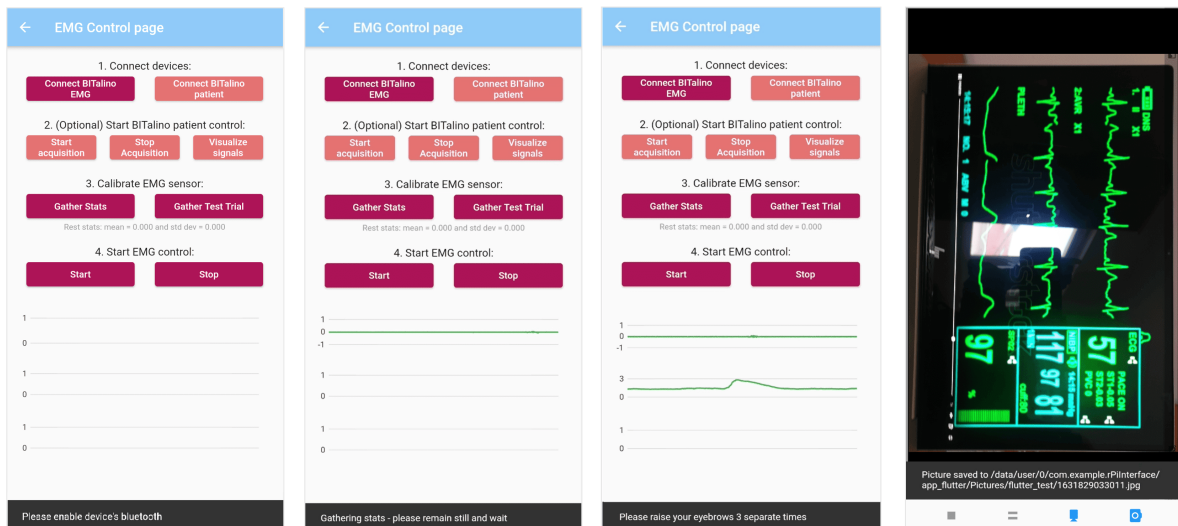


Figure 4.5: Examples of ARSurgery displayed notifications and warnings.

4.2.2 Smartphone - AR Headset connection

The connection between the smartphone where ARSurgery is installed, and the AR head-mounted display, is made resorting to the *screen mirroring* feature, which establishes a direct wireless connection between the sender (smartphone) and the receiver (AR headset). Screen mirroring can be achieved with the default application most smartphones include, such as *Smart View* for Android devices or *AirPlay* for iOS devices, or with any third party application developed for that purpose, e.g. *Chrome Cast*, available on the App Store. With screen mirroring, the content being displayed on the smartphone, in this case, ARSurgery, is presented on the AR headset in real-time. This interface is very suitable for this type of systems, since it is user friendly and practical to operate, in addition to being wireless, which is desirable

in this context.

4.2.3 Signal processing and onset detector

To identify the facial commands done by the user, the system must include an onset detector, capable of processing the EMG signal and distinguishing muscle activations from periods of myoelectric silence. Furthermore, the onset detector should be computationally light, since it is to be implemented on a mobile application, and should perform in real-time, i.e., with low latency, to efficiently interpret user's commands. The detector used to identify onsets and offsets on the EMG signal was based on an amplitude threshold, computed beforehand, in the setup phase. In order to enhance the quality of the EMG analysed by the threshold detector, the signal was pre-processed and filtered in real-time. Experimental tests were carried out to compare the performance of different filters, both with Finite Impulse Response (FIR) and Infinite Impulse Response (IIR), for these specific application. Based on these experimental tests, further presented and detailed on Section 5.1.2, it was decided to process the input signal in real-time using a Weighted Moving Average (WMA) filter, with a window size of 100 data points.

Figure 4.6: EMG raw signal - without applying noise attenuating filters (left). EMG signal, filtered with a Weighted Moving Average (WMA) FIR filter, with a window size of 100 samples (right).

4.2.4 Methodology validation

To evaluate the system's overall performance, and the success of the hands-free control mechanism of the AR headset, an experimental study was designed and conducted. The experiment was composed by two distinct tests: a first one, whose goal was to measure the response level of the app's algorithm to the input movements of the user, and a second, to evaluate the prototype's overall performance, with the electrodes integrated in the AR headset.

For the first test, the participants only interacted with the electrodes and the developed app, disregarding the headset. The electrodes were placed on the participants' forehead with the help of a headband, built for that purpose. As for the second test, the participants were asked to place the AR headset, attached with the support substrate for the electrodes, on their forehead. Both montages are shown in Figure 4.7.

In both tests, the participants were asked to use the ARSurgery app, following a defined protocol, with a set of at least two trials, each composed by a specific sequence of the forehead movements described above – brief eyebrow raises, and extended eyebrow raises, designated by BER and EER respectively. The protocol was designed to let the user run across all pages. The first test evaluates the quality of the recognising mechanism and generated response. The second test, on the other hand,

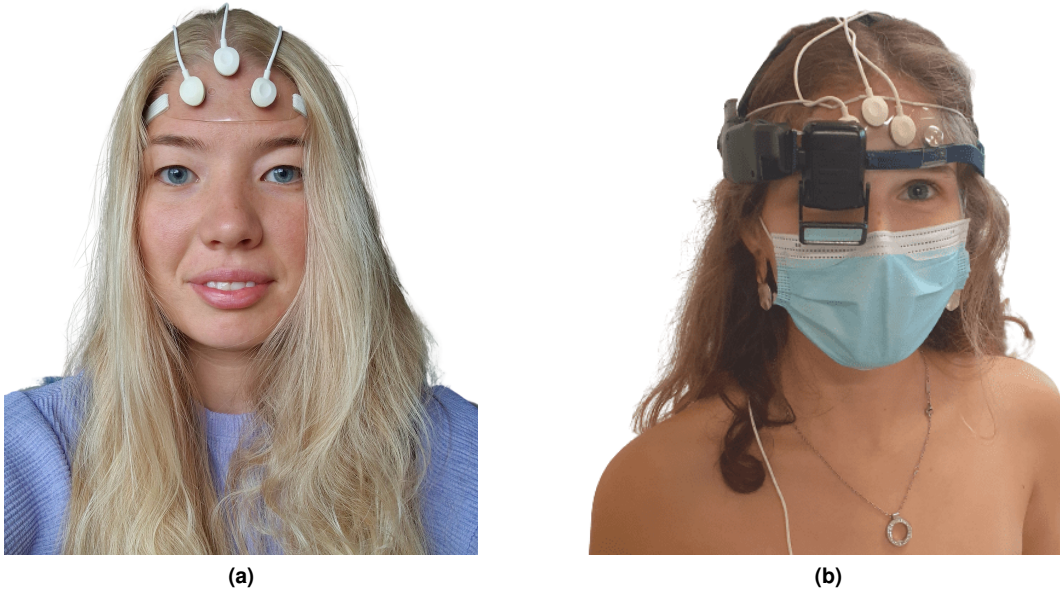


Figure 4.7: Montages used in Test 1 (left) and Test 2 (right).

evaluates the prototype as a system, validating its general usability, and if factors such as the headset's weight and ergonomics do not interfere with its use.

Additionally, and since surgeon doctors represent the main interest group for this project, it was also asked for a group of surgeons to answer a two-part survey regarding the developed system's suitability and usability.

The results for both tests as well as the answers to the survey are presented in Section 5.1.5.

5

Results

Contents

5.1	Characterization and optimization of the EMG signal and hardware	45
5.2	Study with the interest group (surgeon doctors)	62

5.1 Characterization and optimization of the EMG signal and hardware

As introduced in Sections 3.2.2, 3.2.3 and 4.2.3, a primary experimental study was conducted, to gather information about the facial EMG signal, and further decide about the optimal conditions for its recording. This study's goal was to ascertain which electrode type should be used in the signal acquisition, as well as to define the most suitable pipeline to process the signal, including which type of digital filter to use. Three participants took part in this study. The inclusion criteria for participation was being of legal age (> 18 years) and the absence of impairments that could compromise the execution of the required tasks. The characteristics of the participants are depicted in the Table 5.1.

Table 5.1: Summary of participants characteristics

Subjects	Subject 1	Subject 2	Subject 3
Gender	Female	Female	Male
Age (years)	23	57	54
Average (Age)	43.7 ± 15,4		

The acquisition of the electromyographic signals was carried out following an experimental protocol composed by four different facial gestures - brief (BER) and extended (EER) (≈ 3 seconds) eyebrows' raising, single-eye winking (W), and Eye Blink (EB). The movements of brief eyebrow's raising, single eye winking and extended eyebrow's raising are the movements being tested as potential inputs for the control of the augmented reality headset. The eye-blinking is included in the protocol for further testing regarding the control of false positives since blinking is a physiological action that occurs naturally. The defined sequence of the facial gestures for each trial is depicted in Figure 5.1. For simplification purposes, in this trial representation, BER, W and EB movements were considered transient in time, although that may slightly vary depending on the individual performing them.

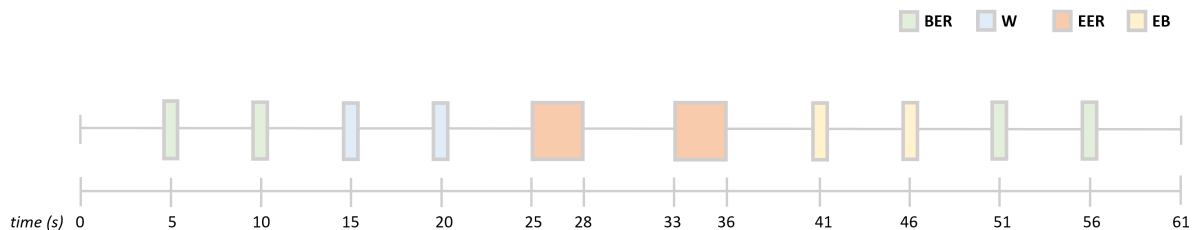


Figure 5.1: Illustration of the experimental protocol for the acquisition of the electromyographic signals. BER = Brief Eyebrow Raise; W = Single-eye wink; EER = Extended (≈ 3 seconds) Eyebrow Raise; EB = Eye-Blink.

5.1.1 Electromyographic signal pre-processing

To analyze the acquired electromyographic signals, namely, to evaluate their quality, a set of pre-processing steps was performed. Firstly, the time series values (Analog to Digital Converter (ADC)) sampled from the EMG sensor's channel were converted to their physiological unit of measurement, mV, according to the transfer function represented in Equation (5.1)

VCC is the operating voltage, 3300 mV.

G_{EMG} is the sensor gain, 1009.

n is the number of bits of the channel, 10.

$$signal_{EMG}(mV) = \frac{\left(\frac{ADC}{2^n} - \frac{1}{2}\right) * VCC}{G_{EMG}} * 1000 \quad (5.1)$$

The second step applied to the time series was the subtraction of the baseline average along the signal, to center the time series around 0 mV, and the third step was to perform a full wave rectification, to enhance the performance of the algorithms applied later on.

Lastly, the time frames corresponding to each muscle activation onset and offset were manually annotated - signal ground truth - resorting to the SignalBit software [50, 51], to ensure that the regions corresponding to the different movements were precisely delimited.

An illustration of the pre-processing applied to the EMG signals, along with the onset annotations, is presented in Figure 5.2.

5.1.2 Electromyographic signal quality

Since muscle activity is being measured with surface electromyographic (sEMG) sensors, the acquired signals are contaminated with noise, reducing the quality of the signal, i.e., reducing its Signal to Noise Ratio (SNR). In addition, this noise also depends on the type of electrodes used during the signals' acquisition, as well as the positioning of the electrodes' montage in relation to the muscle being analyzed, in this case, the forehead of the subjects. Thus, in order to determine which type of electrodes results in a higher quality signal, as well as their optimal positioning in the forehead, several tests were performed while varying these two parameters. To assess the sEMG signal quality, different metrics were computed, for both the baseline of the signal, as well as for the time frames of muscle activation. The SNR, defined in the Equation (5.2) was also computed for the entirety of the signal. The metrics computed for each region are designated in the Table 5.2. The mean, standard deviation, and maximum amplitude were calculated for the regions of muscle activity and for the signal's baseline. The signal's integral was also calculated for the facial movement time periods.

$$SNR(dB) = 20 \times \log_{10} \frac{Power(activation)}{Power(baseline)} \quad (5.2)$$

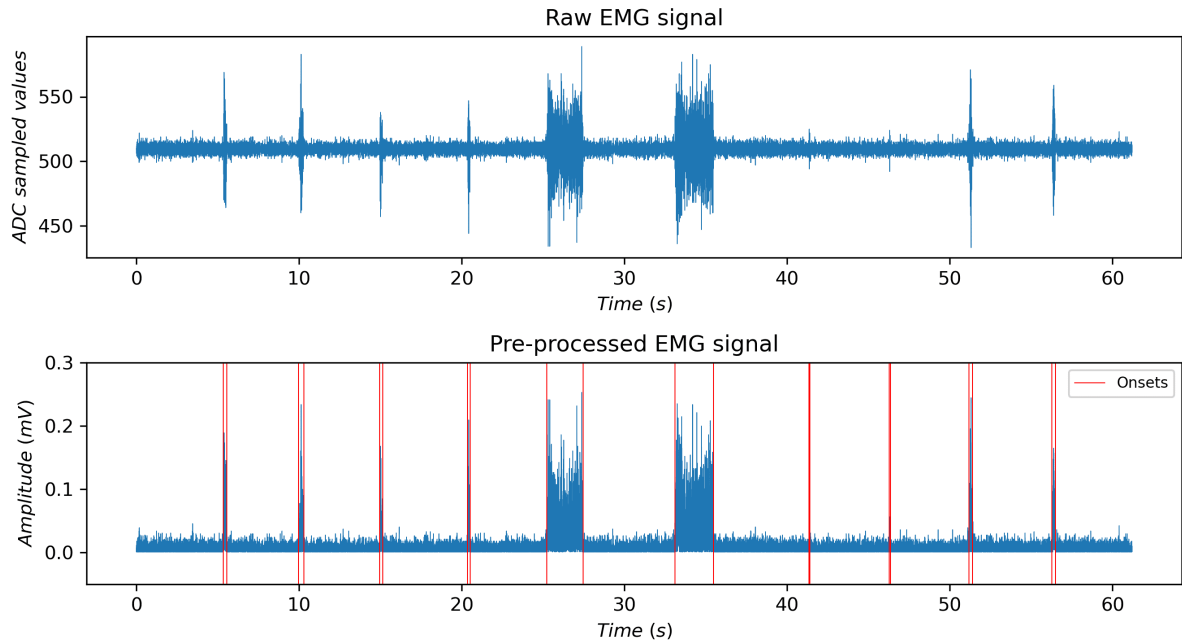


Figure 5.2: Illustration of the EMG signal before and after the pre-processing: conversion to mV units, subtraction of the baseline mean and full wave rectification. Red vertical lines indicating the manually annotated onset and offset for each muscle activation.

Table 5.2: Metrics computed from the signal, depending on the signal's region

Signal Region	Metrics Calculated				
	Mean (μ)	Standard deviation (σ)	Maximum amplitude	Integral of region (\int)	SNR (dB)
Baseline	X	X	X		
Activation region	X	X	X	X	
Total signal					X

Electrode model selection

Regarding the electrodes used in this experiment, two types of non-gelled Ag/AgCl electrodes were analysed, due to the nature of their application being an augmented reality headset. The first set of tests were performed between the standard electrodes, Eg, and the raised profile dry electrodes, Er - presented on Table 3.2. These tests were performed with both types of electrodes applied simultaneously on the skin of the subjects, proximal to each other, to ensure the comparison between the two of them was the most accurate possible. In these tests, two movements were tested - eyes' blinking (natural physiological movement) and single-eye winking (one of the voluntary movements to test as input for the AR system), EB and W, respectively.

Figure 5.3 shows a portion of the simultaneously acquired signals, in this case, a sequence of inter-

calated EB and W (EB - W - EB - W). Please note that, as expected, the signal obtained when using the standard Eg electrodes shows a higher resolution, having more defined activation regions during both movements being executed. The probability histograms on Figure 5.4 show the different features calculated for the two signals. As expected, the Eg signals present a higher Signal-to-Noise ratio than the Er ones (35,1 dB vs 5,25 dB). The Eg baseline results are also more reliable, being less distributed across values, and the standard deviation of the signal is also lower for the Eg, meaning lower noise acquired. In terms of the muscle activation regions, in this case, the single-eye winking, the Eg electrodes provide signals with higher integral (32,75 mV vs 5,47 mV) and maximum amplitude (0,7 mV vs 0,1 mV), i.e., a higher intensity signal. Nonetheless, although the results for the Eg electrodes are, as anticipated, generally better, even with the Er dry electrodes, it is still possible to clearly detect the peaks corresponding to the single-eye winking movements, which is the main goal, and the statistics are generally on the same order of magnitude as the Eg. The low resolution of the EB movements recorded with the latter is not concerning at all, since the goal is to actually discard physiological eye-blinking that may happen during the usage of the system.

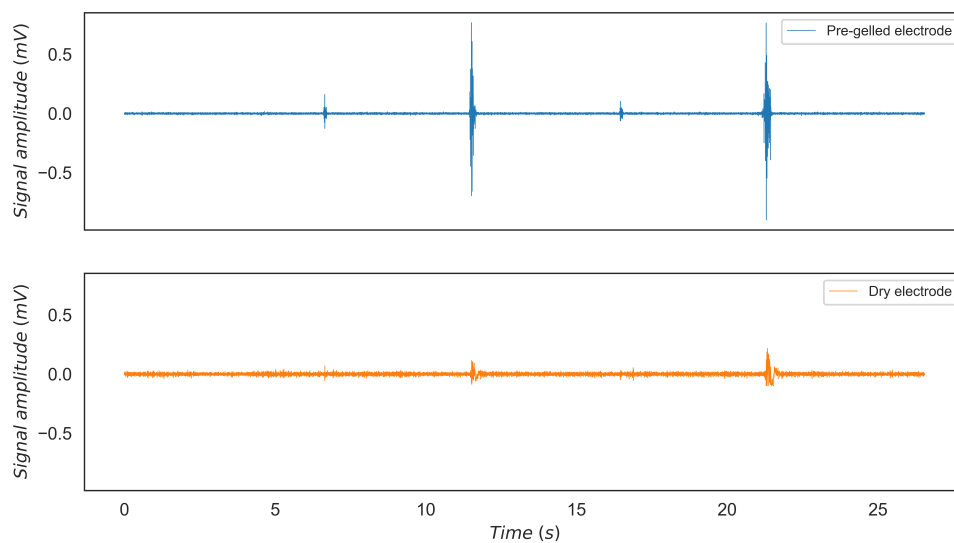
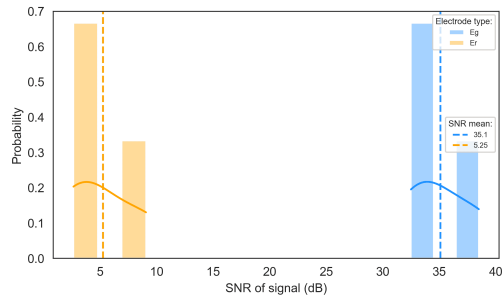
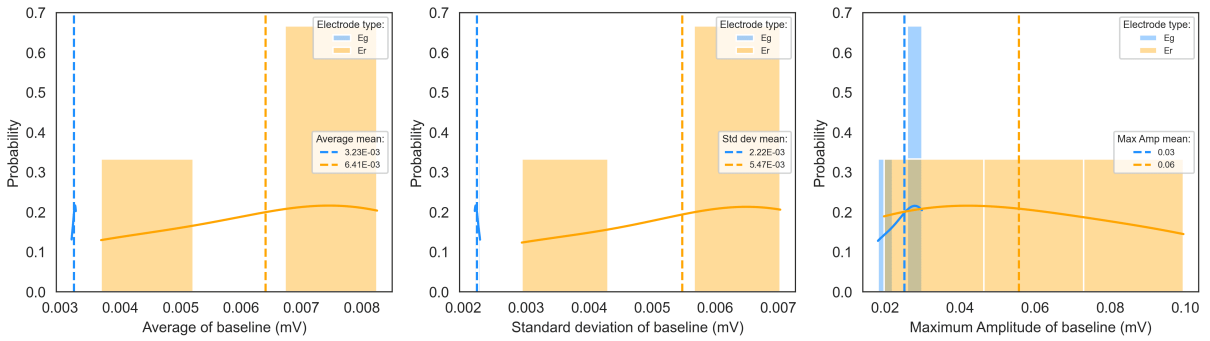


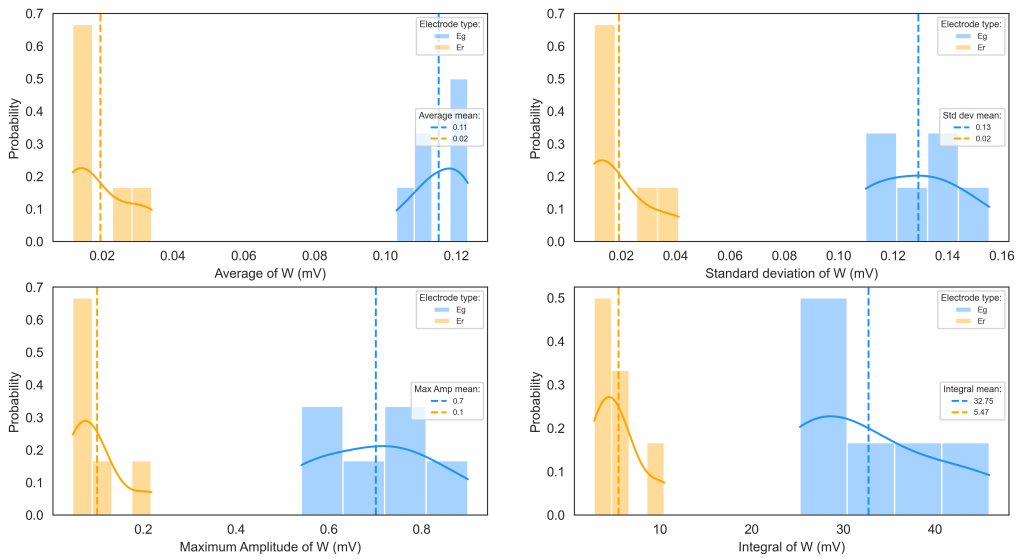
Figure 5.3: Illustration of the EMG signal acquisition using two types of electrodes simultaneously - Ag/AgCl pre-gelled electrodes - Eg (top), and Ag/AgCl raised profile dry electrodes - Er (bottom).



(a) Signal-to-Noise Ratio (SNR)



(b) Baseline average, standard deviation and maximum amplitude



(c) Single-eye Winking (W) average, standard deviation, maximum amplitude and integral

Figure 5.4: Statistical comparison of EMG test trials using Ag/AgCl pre-gelled electrodes - Eg (blue), and Ag/AgCl raised profile dry electrodes - Er (yellow). Top graph: Signal-to-Noise Ratio (SNR). Middle graphs: Average, Standard Deviation, and Maximum Amplitude of signal baseline. Bottom graphs: Average, Standard Deviation, Maximum Amplitude, and Integral of signal for single-eye winking (W) movement.

Following this primary experiment, a second set of tests was done, to compare the two types of dry electrodes, Ef and Er. However, for these, each electrode type was tested separately, with different trials recorded for each one of the electrodes. The number of trials each subject performed with the different electrodes is shown in Table 5.3.

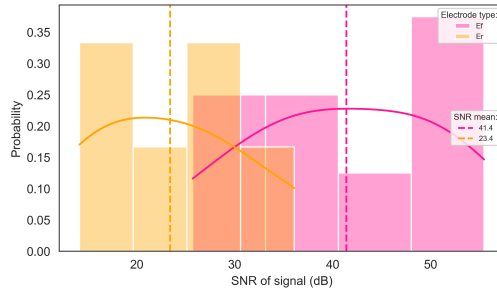
Table 5.3: Number of trials performed by each subject, with each of the three types of electrodes - Er, Ef and Eg.

Electrode type	Subject 1	Subject 2	Subject 3	Total:
Ef	3	5	0	8
Er	3	0	3	6
Eg	7	0	0	7
Total:	13	5	3	21

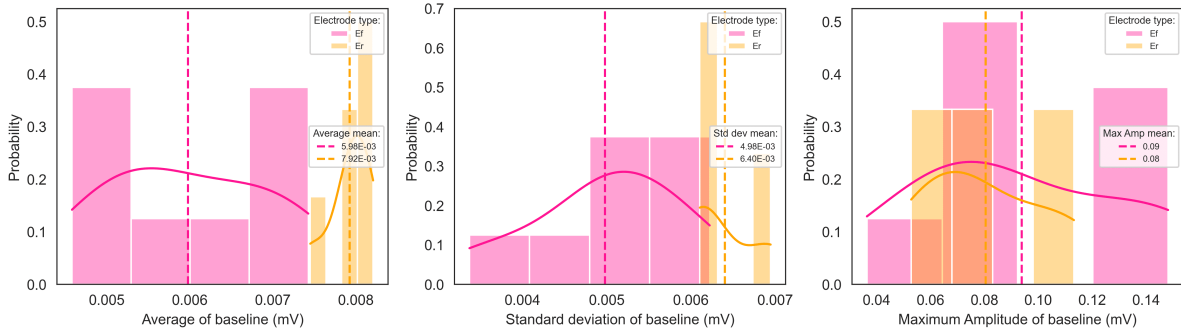
The metrics for the signal's baseline, for each signal segment - Brief Eyebrow Raise (BER), Single-eye Wink (W) and Extended Eyebrow Raise (EER), and for the total signal, were computed according to Table 5.2, and the results were aggregated in histograms, to better illustrate the comparison between the two types of electrodes. Each segment of the signal has histograms for its corresponding metrics, and each histogram contains the comparison between the two Ag/AgCl dry electrodes models studied – Ef (represented in pink) and Er (represented in yellow). Additionally, each histogram shows probability density estimation curves for the different electrode models, as well as vertical lines indicating the mean result of that metric. The results are presented in terms of relative frequency, i.e., the bars height sums up to 1, to consider the imbalance between number of trials acquired with each electrode type, and to normalize the data.

The histograms were divided according to two criteria: (1) number of included subjects, and (2) filtering of the signal. In a first phase, trials from all subjects were included for the calculation of the metrics. After that, to investigate the impact of inter-subject variability and detect possible outliers, the trials were split by subject, and computed for each of them, individually. On the other hand, the metrics were computed both for the “raw signal”, i.e., without any type of noise attenuating filter, and also after applying a filter to it. In this case, the filter selected to present was the “adaptive filter error”, which consists in calculating the difference vector between the original signal and the signal resulting from applying an adaptive Least Mean Square (LMS) filter. For abridgement purposes, the first set of histograms, in Figure 5.5, show the results while considering all subjects, and without filtering the signal. The second set of histograms, in Figure 5.6, show the metrics considering only one of the subjects, and passing the adaptive filter on the signal. Additionally, to help in the interpretation and comparison between electrode models, the relative difference between their metrics' mean values was obtained, according to Equation (5.3). These results are summarized in Table 5.4 and Table 5.5.

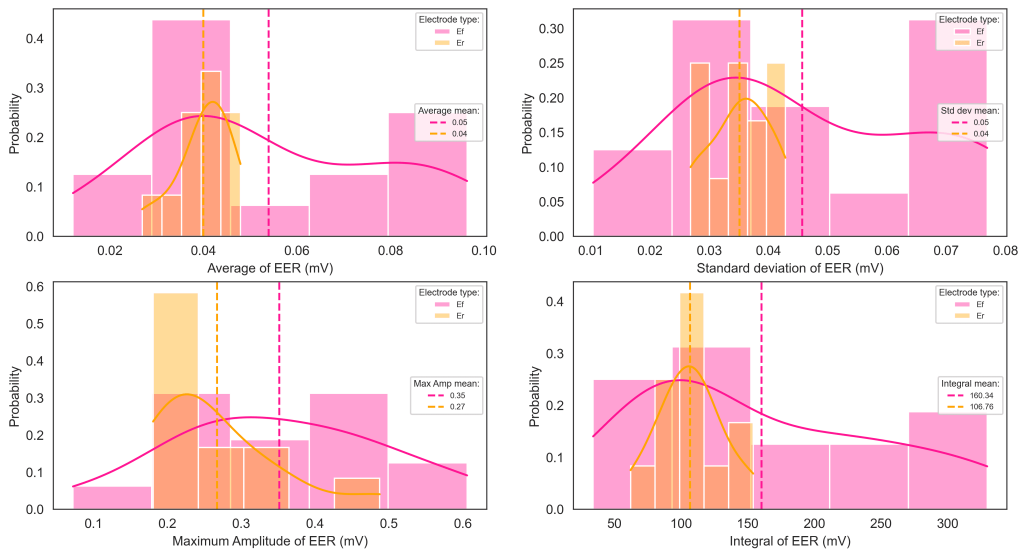
$$Diff(\%) = \frac{signal_{Ef} - signal_{Er}}{signal_{Ef}} \times 100 \quad (5.3)$$



(a) Signal-to-Noise Ratio (SNR)

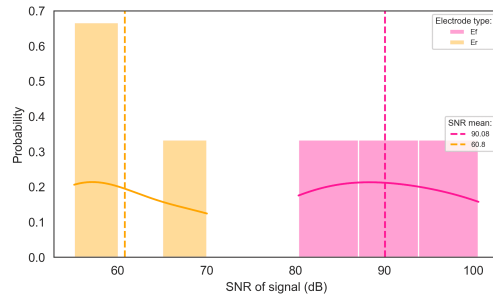


(b) Baseline Average, Standard Deviation and Maximum Amplitude

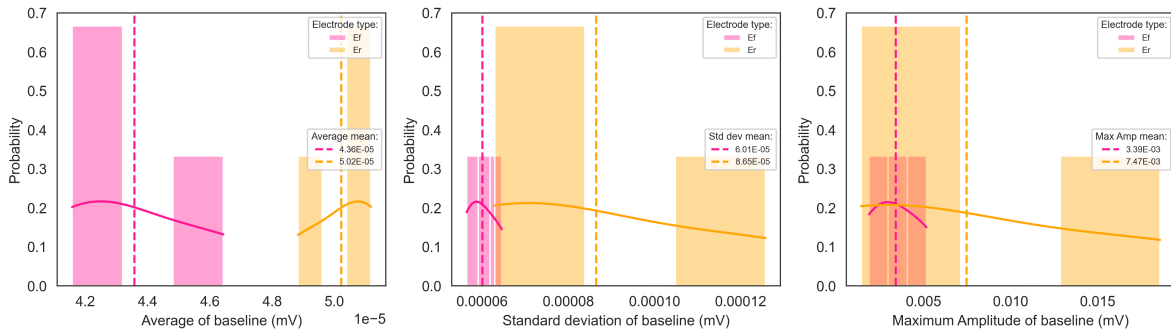


(c) Extended Eyebrows' Raising (EER) Average, Standard Deviation, Maximum Amplitude and Integral

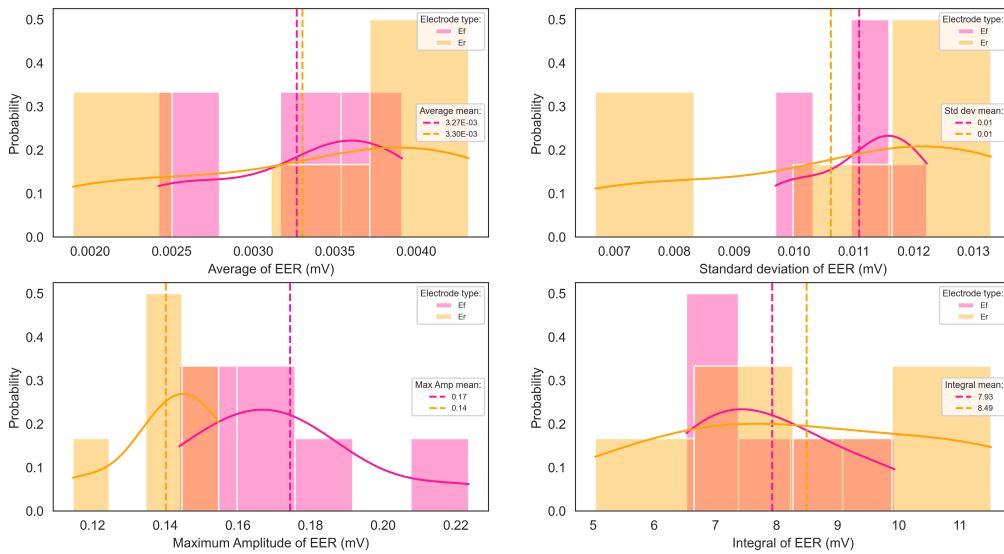
Figure 5.5: Statistical comparison of EMG test trials using raised profile (Er - yellow) and flat profile (Ef - pink) Ag/AgCl dry electrodes. Statistics include all subjects without signal filtering. Top graph: Signal-to-Noise Ratio (SNR). Middle graphs: Average, Standard Deviation, and Maximum Amplitude of signal baseline. Bottom graphs: Average, Standard Deviation, Maximum Amplitude, and Integral of signal for extended eyebrows' raising movement.



(a) Signal-to-Noise Ratio (SNR)



(b) Baseline average, standard deviation and maximum amplitude



(c) Extended Eyebrows' Raising (EER) average, standard deviation, maximum amplitude and integral

Figure 5.6: Statistical comparison between EMG test trials obtained with Ag/AgCl dry electrodes - raised profile (Er - yellow), and flat profile (Ef - pink). Statistics are based on Subject 1, with signal filtering using an LMS adaptive filter. Top graph: Signal-to-Noise Ratio (SNR). Middle graphs: Average, Standard Deviation, and Maximum Amplitude of signal baseline. Bottom graphs: Average, Standard Deviation, Maximum Amplitude, and Integral of EER of signal for extended eyebrows' raising movement.

Table 5.4: Relative difference (%) between metrics' mean values obtained for Ef and Er electrode models, while considering all subjects, and without applying any filter to the signals.

Signal Region	Metrics Calculated				
	Mean (μ)	Standard deviation (σ)	Maximum amplitude	Integral of region (\int)	SNR (dB)
Baseline	-32,44%	-28,51%	11,11%		
W	50,00%	50,00%	58,82%	61,17%	
BER	33,33%	20,00%	26,67%	-5,81%	
EER	20,00%	20,00%	22,86%	33,42%	
Total signal					43,48%

Table 5.5: Relative difference (%) between metrics' mean values obtained for Ef and Er electrode models, while considering one subject, and applying an adaptive filter on the signals.

Signal Region	Metrics Calculated				
	Mean (μ)	Standard deviation (σ)	Maximum amplitude	Integral of region (\int)	SNR (dB)
Baseline	-15,14%	-43,93%	-120,35%		
W	63,11%	57,11%	37,50%	65,88%	
BER	6,13%	0,00%	0,00%	30,86%	
EER	-0,92%	0,00%	17,65%	-7,06%	
Total signal					32,50%

Some conclusions can be withdrawn when analyzing the histograms (and corresponding tables) with the obtained statistics for the signals, both pre- and post-filtering. Looking at the SNR histograms, the Ef electrodes reach greater values than the Er - 41,40 dB vs 23,4 dB (all subjects, pre-filtering) and 90,08 dB vs 60,8 dB (one subject, post-filtering). This corresponds to an SNR increase of 43,48 % / 32,50 % for the Ef model. Furthermore, the baseline statistics also favor the usage of the Ef electrodes - both the average and the standard deviation of the baseline are significantly lower for these electrodes, which is the desirable, since the baseline ideally corresponds to myoelectric silence. For the mean and standard deviation of the signals, there is a respective decrease of 32,44% and 28,51% (all subjects, pre-filtering), and 15,14% and 43,93% (one subject, post-filtering). Finally, it is worth analysing the statistics for the input facial movements being studied, particularly the Extended Eyebrows' raising (EER), which is the one represented in the histograms. Before filtering the signal segments corresponding to facial movements, their mean, maximum amplitude and integral are significantly higher for the Ef electrodes. This is according to the expected, as these metrics are interconnected. A signal with generally higher amplitude tends to result in a higher integral and is more likely to reach a greater maximum value. Additionally, the standard deviation is also expected to increase as it is derived from a signal with higher amplitude, implying larger variations in magnitude. On the other hand, when looking at the filtered signals, the difference between Er and Ef electrodes is severely decreased, demonstrating the effect of applying a smoothing filter on the signal, to reduce noise contamination.

With these results was possible to decide on the Ef Ag/AgCl dry electrodes, more proximal to the skin – shown in the second row of Table 3.3, to conduct further testing, and integrate the Augmented Reality system.

Filtering of the acquired signal

The next step was to study if applying a digital filter to the acquired signals would enhance them, and if so, which type of filter should be applied. Because the goal is to detect electromyographic onsets and offsets (temporal features), the aim of applying a filter is to reduce noise the electrodes may capture, enhancing the signal's SNR the most possible. That way, the most important features to retain are temporal, meaning that filters which attenuate components with certain frequencies can be applied, for example moving averages. Several filters were selected as candidates to be analysed. Although there are various types of complex filters that have a high level performance, considering this signal processing should be done in real-time, and on a mobile (Android/iOS) Operating System, the filters were selected considering also their computational complexity. Both Infinite Impulse Response (IIR) filters and Finite Impulse Response (FIR) filters were tested:

- IIR Butterworth filters:
 - 2nd order, with low cutoff frequency $f_{c,low} = 20\text{Hz}$ and high cutoff frequency $f_{c,high} = 450\text{Hz}$;
 - 10th order, with low cutoff frequency $f_{c,low} = 20\text{Hz}$ and high cutoff frequency $f_{c,high} = 40\text{Hz}$;
- FIR Moving Average filters:
 - Simple Moving Average (SMA), with window size = 10 samples;
 - Exponential Moving Average (EMA), with window size = 10 samples;
 - Gaussian Moving Average (GMA), with window size = 10 samples;
- FIR Adaptive Least Mean Square (LMS) Error filter, with step size = 5 samples.

The performance of the filters on the EMG signal enhancement was evaluated through the outcome from an onset detector applied to the filtered signal. The selected onset detection algorithm was based on the approach proposed by Hodges and Bui (1996) [52], available on the *BioSPPy* Python toolbox¹. This detector receives as inputs: (1) the filtered EMG signal; (2) a sample segment corresponding to a rest period (electromyographic silence); (3) the sampling rate; (4) the detection threshold, i.e., the value from which the signal is considered a muscle activation; and (5) the detection window size, i.e., the number of samples for which the mean must exceed the defined threshold (in seconds).

¹ Available at <https://github.com/PIA-Group/BioSPPy/blob/master/biosppy/signals/emg.py>.

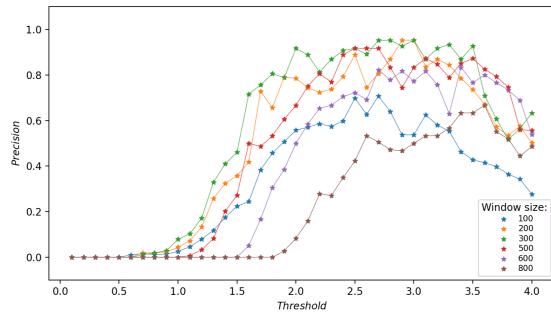
The performance of the detector while varying the filter was evaluated through Precision-Recall curves. These curves summarize the trade-off between the true positive rate – recall – and the positive predictive value – precision. The mathematical expressions for these concepts are shown in Equation (5.4) and Equation (5.5).

$$Recall = \frac{\#TruePositives}{(\#TruePositives + \#FalseNegatives)} \quad (5.4)$$

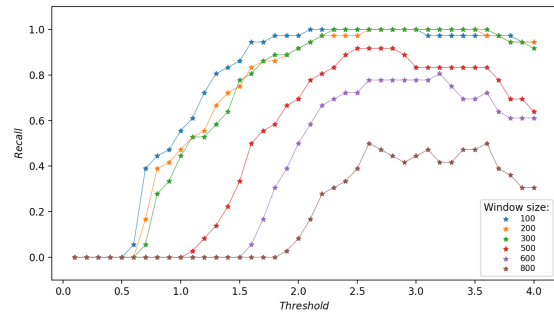
$$Precision = \frac{\#TruePositives}{(\#TruePositives + \#FalsePositives)} \quad (5.5)$$

These performance metrics are widely used in machine learning contexts, because they allow to evaluate, for a certain classifier, which set of hyper-parameters optimize it, defining a balance between both False Negatives and False Positives rates. To establish the definition of the different variables - True Positives (TPs), False Positives (FPs) and False Negatives (FNs) - in this specific context, several steps were performed. Firstly, all the segments (pairs of onset-offset) corresponding to muscle activations were annotated, for each EMG signal acquisition performed following the defined protocol. This annotations - ground truth - were registered, once again, resorting to the SignalBit software. As for the detections given by the detector, each one was classified as a True Positive (TP) or False Positive (FP), depending on its temporal location. If the detection coincided with a real onset, within a certain time tolerance, the detection would be considered as a TP. Otherwise, it would be a FP. In the case of multiple detections in the region of the real onset, the closest to it would be the TP, and the remaining would be classified as FP. The tolerance was calculated considering the average latency between the real onsets and the obtained detections, provided it does not exceed a maximum latency of 150 samples (150 ms), considered as the acceptable for the purpose of its application. In the case of muscular activity not being detected by the detector, they were considered as False Negatives (FN).

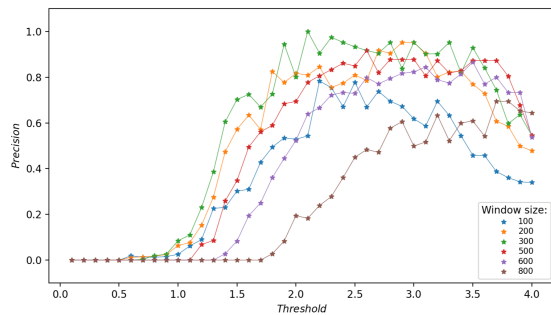
The detector was tested with detection threshold values varying between 0,1 and 4,0 (with increments of 0,1), and detection window sizes of 100 ms, 200 ms, 300 ms, 500 ms, 600 ms and 800 ms. The performance of the detector, depending on the filter used on the input EMG signal, was analyzed through the graphs of Precision and Recall presented in Figure 5.7. Although all the mentioned filters were analyzed, it was decided to present only one filter of each filter type.



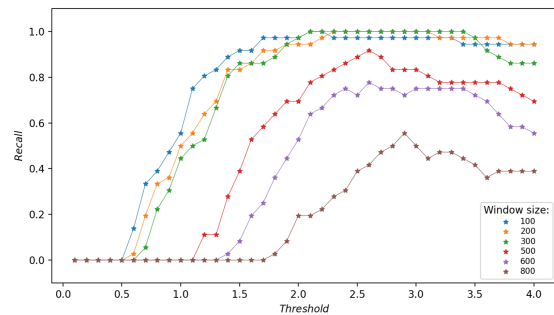
(a) Precision - Raw Signal



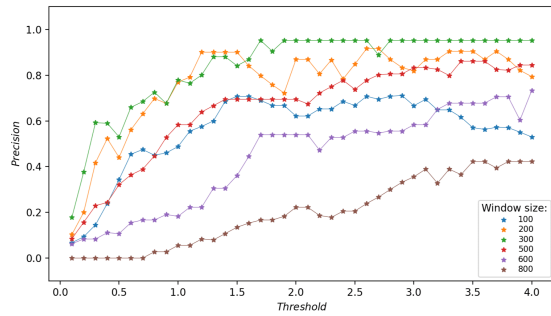
(b) Recall - Raw Signal



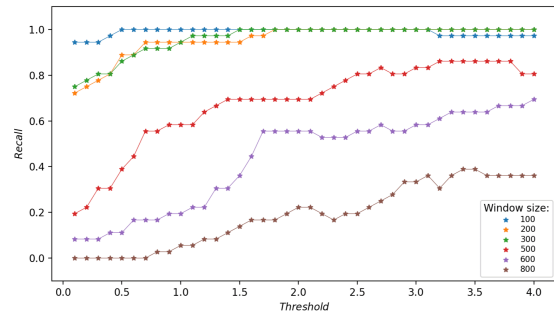
(c) Precision - Butterworth Filter (2nd order)



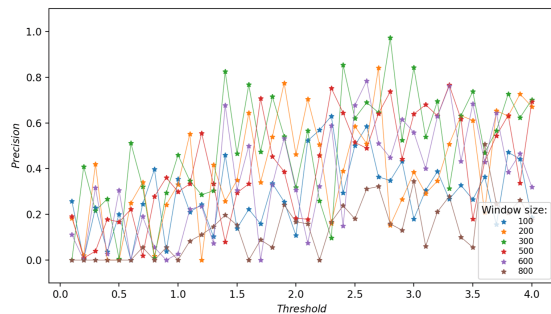
(d) Recall - Butterworth Filter (2nd order)



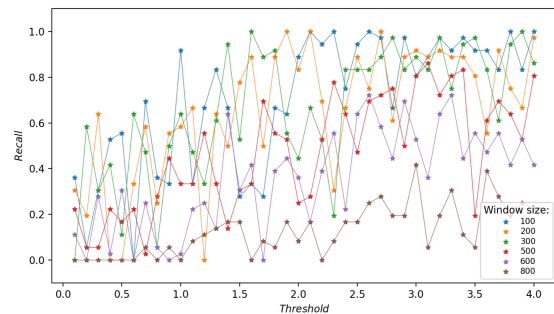
(e) Precision - Simple Moving Average (W = 10)



(f) Recall - Simple Moving Average (W = 10)



(g) Precision - LMS Adaptive Filter (L = 5, step-size = 1)



(h) Recall - LMS Adaptive Filter (L = 5, step-size = 1)

Figure 5.7: Precision and Recall curves for Hodges and Bui EMG onset detector, with respect to the detection threshold and detection window size, considering different filters applied to input EMG signal. (a, b) No filtering (Raw signal), (c, d) Butterworth 2nd order filter with cut-off frequencies of 20Hz and 450Hz, (e, f) Simple Moving Average filter with window size of 10 samples, and (g, h) Least Mean Squares Adaptive filter with length of 5 and step-size of 1.

Based on the Precision and Recall curves obtained for the different filters, the Moving Average filters result in overall higher precision and recall values, across the range of detection thresholds and window sizes. Looking at the SMA (with a 10 sample window) filter graphs, they are more stable with respect to the threshold used, turning the detector more robust. For that reason, it was decided to use a Moving Average filter to process the input EMG signal being acquired from the user.

5.1.3 Detector's hyperparameters

Looking once more at the graphs on Figure 5.7, some conclusions can be drawn regarding the most suitable hyperparameters (detection threshold and detection window size) values to use on the onset detector. Firstly, in terms of the detection threshold, the graphs generally show that using lower thresholds result in decreased precision and recall values. This occurs because when using very low thresholds, while the actual onsets (TP) are detected, there is also a significantly higher rate of false positives (FP), leading to reduced Precision. Conversely, applying excessively high thresholds will result in a higher FN rate, consequently reducing the Recall. In this particular context and with the analysed signal processing, the ideal threshold value falls approximately within the range of 2.1 to 3.3.

Regarding the detection window, this parameter determines the minimum duration during which the EMG signal needs to exceed the defined detection threshold. The graphs show that applying wider windows on the detector - 500 ms (red curves), 600 ms (purple curves) and 800 ms (brown curves) - decrease the maximum obtainable precision and recall values, regardless of the chosen threshold. This is to be expected, because larger windows, while preventing false detections (FP), can also decrease the number of real onsets (TP) detected, for weaker or shorter EMG signal onsets, consequently increasing the FN rate and resulting in lower Recall values. Conversely, applying a very short window - 100 ms (blue curves) - while it prevents missing shorter onsets, it also greatly increases the number of false onsets detected (FP), resulting in a lower Precision, which is demonstrated by the blue curves presented in the precision graphs. In this case, the window size that optimizes the Hodges and Bui detector's performance is 200 ms.

5.1.4 Onset detector implementation

The previous analysis was based on the detector that Hodges and Bui proposed, computed in the BioSPPy toolbox. The detector implemented in the ARSurgery Mobile application was adapted from this, but has slight differences due to the context constraints of this application, and being computed for real-time usage.

Firstly, the EMG electrodes being used aren't fixed to the skin, being instead fixed on the head-mounted AR display itself. This increases the inter-subject and inter-trial variability between acquired signals, since the EMG electrodes may coincide with slightly different points of the forehead, which influ-

ences the amplitude of the signal recorded. As an attempt to overcome this, the implemented detector uses a dynamic threshold instead of a fixed value, to enhance the detector's adaptability to the input signal. The threshold in the implemented detector is calculated during a test trial performed beforehand, where the user performs 3 separate brief eyebrow raises (BER). The test trial is pre-processed, rectified, and smoothed with a Weighted Moving Average of 100 samples. The baseline average and standard deviation are calculated, and a test function is computed following Equation (5.6). The threshold is defined as 1σ above the test function average [52]. Figure 5.8 illustrates a segment of an EMG signal capturing two instances of BER and two instances of EER movements, and its resulting test function, as well as the corresponding computed threshold.

$$testfunction(x) = \frac{1}{\sigma_{baseline}} * (x - \mu_{baseline}) \quad (5.6)$$

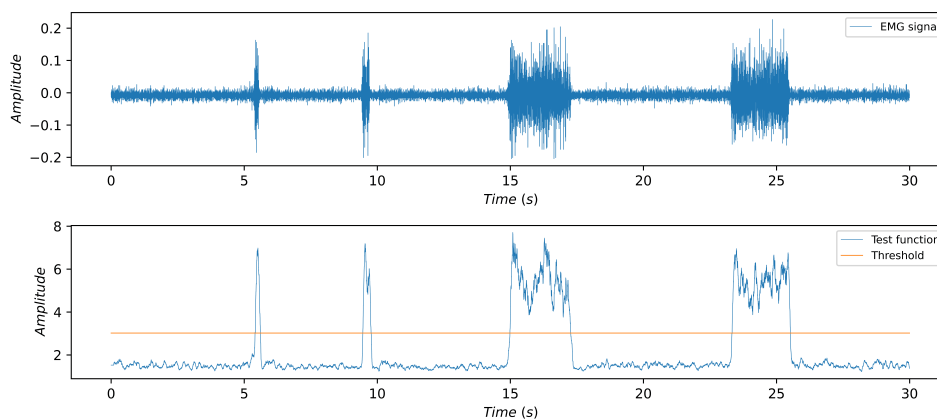


Figure 5.8: Illustration of an EMG signal segment and its resulting test function, as well as the corresponding computed threshold.

It is important to note that along these experimental tests, the Single-Eye Winking movement (W) was eventually excluded as an input command for the AR head-mounted display, due to its highly variable intensity, even among repetitions performed by the same subject. The input commands that were implemented in the final prototype were the BER and EER movements. The EER movement controls the navigation from / to the Calibration Page, allowing the user to exit it after pressing "Start Acquisition", and directing them to the Camera Viewer. To switch between this latter, the Patient's Biosignals Viewer, and the Off-Mode / See-Through Screen, the user can perform a BER movement.

5.1.5 AR system prototype validation

As described in Section 4.2.4, two different tests were performed to validate both the onset detector and the EMG-controlled AR headmounted display as a system. In both tests, the participants were asked to use the ARSurgery app, following the sequence of forehead movements shown in Figure 5.9, designed to guide the participants through the different pages of the app, ensuring the hands-free control

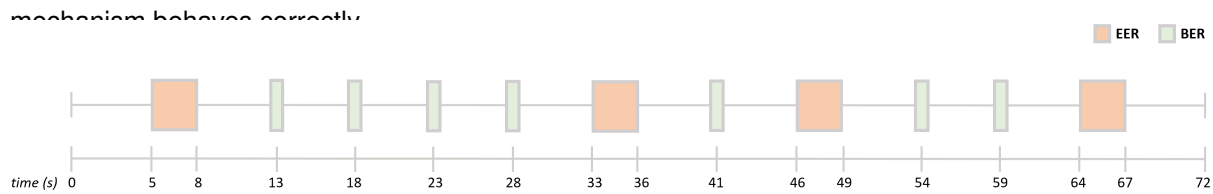


Figure 5.9: Illustration of the movements' sequence followed by participants in Test 1 and Test 2. Sequence includes alternating Brief Eyebrow Raises (BER) and Extended Eyebrow Raises (EER) lasting ≈ 3 s.

For each of the movements executed, the corresponding triggered action on the app was registered: When the app was in accordance with the movement made, it was considered as a True Positive (TP). The cases where the movement did not elicit any response from the app were considered as False Negatives (FN). On the other hand, if the app perceived a trigger in the absence of any deliberate movements from the participant, it was annotated as a False Positive (FP). Furthermore, because there are two distinct events being analyzed in these trials, and their differentiating parameter is the duration of the onset, if the app did not distinguish between the BER and the EER movements - perceiving the BER as EER, or vice versa - that was also registered.

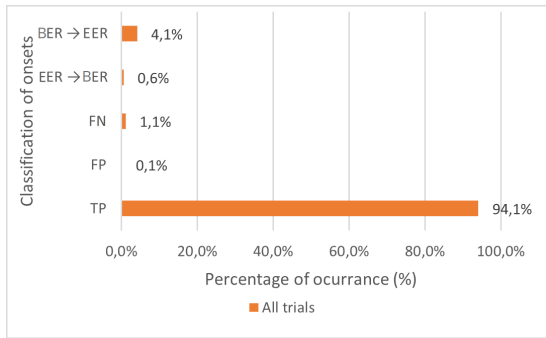
Since the goal of this project is to be applied in the surgical field, this experimental study was performed on two distinct groups of participants: a control group, composed by regular subjects, and a test group, composed by surgeon doctors, with different specializations. This study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Instituto Superior Técnico (Statement nº 19/2021). Written consent was obtained for the participants prior to their enrollment, and all data were treated anonymously. The consent document is present in Appendix A. Table 5.6 summarizes the data about the participants. The trials, both inter and intra-subject, were all performed independently – for each trial the app was reset to guarantee that all the steps were performed each time, including the defining of the EMG signal statistics and threshold. Table 5.7 shows the total number of trials that were performed, for both the first and second tests. The results of the experimental tests are presented in Figure 5.10.

Table 5.6: Summary of participants data - number of participants, sex and age - discriminated for both study groups - control group and test group

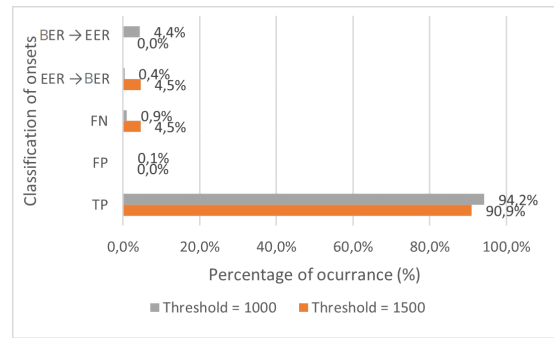
Trial group	Nb of participants (#)	Sex (Female or Male)	Age (years)
Control	18	15F, 9M	30 \pm 16
Test	2	1F, 1M	46 \pm 21

Table 5.7: Total number of experimental trials performed by the participants, discriminated by the test type - Test 1 and Test 2

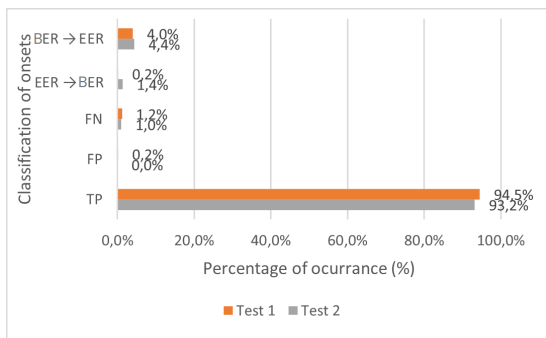
	Test 1	Test 2	Total
Nb of trials	60	36	96



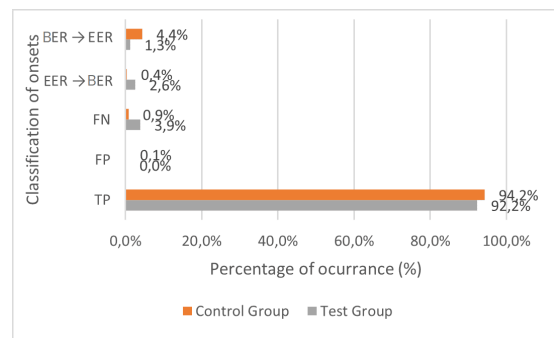
(a) All trials



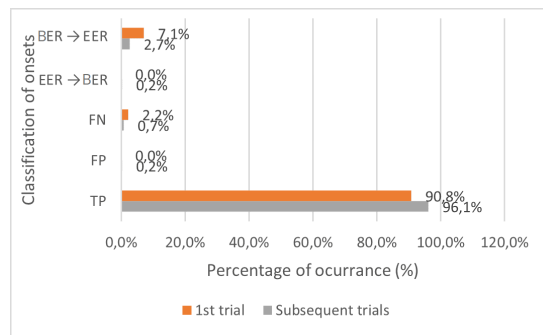
(b) Threshold value



(c) Test type



(d) Target Group



(e) Learning Curve

Figure 5.10: Performance metrics results of experimental tests including True Positive Rate (TP), False Positive Rate (FP), False Negative Rate (FN), EER Movement mismatch Rate (EER -> BER), and BER Movement mismatch Rate (BER -> EER). Emphasis on comparing test type (Test 1 vs Test 2), threshold value (threshold = 1000 vs threshold = 1500), target group (Control Group vs Test Group), and learning rate (first trial vs subsequent trials).

As depicted in Figure 5.10a, when considering the total amount of trials performed, for both tests, without any type of discrimination, the results from the experimental tests are very satisfactory. Overall, the percentage of onsets perceived correctly by the app (TP rate), of 94,1%, is very promising, considering the straightforwardness of the criteria used for the onset distinction and classification. The second most important aspect to assess in this type of control mechanism is the absence of response from the

app when it is not required by the participant, i.e., the FP rate. The FP rate determined, of 0,1%, was also very satisfactory, meaning that almost no occurrences of this type happened. However, it is worth mentioning that the optimistic results for this particular metric can be related, to a certain extent, with the fact that the trials were performed in a controlled environment, where the participants were focused on the solicited task, and therefore less prone to make spontaneous facial movements that lead to false positives, as it can occur in real life scenarios.

In terms of the FN rate, meaning the percentage of times the participants performed one of the two desired movements, without eliciting any response from the app, a percentage of 1,1% was obtained. In this context, this value is sufficiently low, to not compromise the correct functioning of the application.

Now, regarding the capacity of the algorithm to distinguish correctly between the two eyebrow movements being analyzed, which only differ in the duration of the stimulus: From Figure 5.10a, it is possible to observe that the brief eyebrow raise (BER) is perceived as a extended eyebrow raise (EER) much more frequently than the opposite – 4,1% and 0,6% respectively. However, these percentages are highly related to the temporal parameter that was defined in the algorithm as the value that separates a brief from a extended eyebrow raise. This value must be selected considering two aspects – on the one hand, because this app is to be used by different subjects, and must be the most intuitive possible, the value should be high enough to clearly separate the extended ER movements from the brief ER ones, this way respecting the fact that different subjects may have slightly different notions and/or executions of the brief eyebrow raise. On the other hand, the value should be low enough for the EER movement to be feasible, and not uncomfortable to execute. Hence, during the experimental study, two different thresholds were tested, to analyze how that affects the behavior of the algorithm, and to explore which threshold would better compromise these two conditions. Looking at Figure 5.10b, it is noticeable that varying this threshold does impact the distinction of the brief and extended eyebrow raises – with a threshold of 1500 milliseconds, there is a higher percentage of EER being perceived as ER, and a null percentage of the opposite case – 4,5% and 0,0% respectively; However, if the threshold is decreased to 1000 milliseconds, the reverse occurs – the majority of the EER are well identified (0,4% of mismatch cases), but some of the BER are mistaken by EER (4,4% of mismatch cases). It is worth mentioning that only the first subject – one of the surgeon doctors – performed the trials with the threshold of 1500 milliseconds, and the subject reported that this value was too high for the LER movement to be comfortable to execute, especially during surgery, therefore the threshold was adjusted to 1000 milliseconds, for all the subsequent subjects who performed the trials. Hence, the results discussed above regarding this parameter, are partly reflected by the unbalanced number of participants in the two groups, which does not compromise the conclusion that both thresholds provide acceptable results, and a value between 1000 and 1500 milliseconds would be the most adequate.

Another aspect of the results that is important to analyze is the difference between the two tests

requested to be carried out by the participants. As can be seen in Figure 5.10c, the results obtained for the Test 1, which only involved the sEMG sensors and the app, disregarding the AR headset, are generally more optimistic than the ones obtained for the Test 2, which assessed the entire AR system with integrated sEMG sensors – a 1,3% higher TP rate and a summed 1,6% lower rate of mismatches.

Additionally, during the experimental study, there were trials that were considered invalid, when one of three cases occurred:

1. The statistics of the resting EMG signal, acquired by the app as a first step, necessary to its proper functioning, were registered and taken into consideration. If the standard deviation of the resting EMG signal was superior to 0,2, which indicates an inadequate positioning of the electrodes on the forehead, as described in Section 4.2.1, the trial was considered invalid.
2. If the trial progressed without incidents, but after a certain moment, the app started to show an abnormal behavior – for example a constant switching between the displayed pages – and the EMG signal being acquired, which is shown on the app, is in accordance with that, i.e., the EMG signal itself is very heterogeneous and does not have the typical shape, then the trial was also considered invalid. The reason for this decision is that this change in behavior most probably results from a displacement of the electrodes on the skin, due to movement from the subject, or a sudden lack of contact with the skin.
3. If, during the trial, more than six out of the eleven movements required to perform by the participant, were incorrectly classified by the app, or not detected, the trial was also considered invalid, because the reason for this occurrence is probably related with the step of acquiring the initial parameters, namely, the threshold necessary to distinguish between an onset and a resting period. In that step, it is important that the three test onsets the subject performs to define the threshold are similar in intensity. However, although rarely, it was observed that some participants tended to execute the first test onset with a much higher intensity than the other two, which compromises the calibration. This incorrect threshold defining jeopardizes the correct testing of the algorithm.

5.2 Study with the interest group (surgeon doctors)

Since the surgeon group represents the main interest group for this project, it was also asked for its participants to answer a survey regarding the developed system's usability. This survey comprised two parts: a more generic one, adapted from the System Usability Scale (SUS) [53], which is commonly used to assess the usability of new technology systems. This scale was chosen because it is very simple to administer to participants, and can be used on small sample sizes with reliable results, which is the case of the surgeons' group. The SUS is composed by 10 different affirmations, and the subject must choose

between 5 levels of agreement for each of those affirmations. The final score is attributed based on the answers of the subject to the different items: for the odd affirmations, the score corresponding to the chosen level of agreement is subtracted by 1; for the even affirmations, the level score is subtracted to 5. Then, those results are added together, and multiplied by 2,5. The final score obtained must be superior to 68 to be considered satisfactory/above average. Also, because the SUS is a somewhat generic questionnaire, a second part with questions regarding this specific developed system, developed for the surgery context, was added. This second set of questions followed a semi-structured interview format. The two-part survey applied to the surgeons' group is detailed below:

Table 5.8: Two-part survey applied to the test group

Survey	
	System Usability Scale
Part 1	1 I think I would like to use this tool frequently.
	2 I found the tool unnecessarily complex.
	3 I thought the tool was easy to use.
	4 I think that I would need the support of a technical person to be able to use this system.
	5 I found the various functions in this tool were well integrated.
	6 I thought there was too much inconsistency in this tool.
	7 I would imagine that most people would learn to use this tool very quickly.
	8 I found the tool very cumbersome to use.
	9 I felt very confident using the tool.
	10 I needed to learn a lot of things before I could get going with this tool.
	Specific Questionnaire
Part 2	1 Would the use of this system during surgery provide an effective improvement of the surgeon's vision field?
	2 Is this system comfortable to use during a surgery?
	3 Are there any spatial constraints between the AR headset and the medical equipment necessary to put on the surgeon's head (e.g.: surgical magnifying glasses)?

The mean score obtained by the surgeons' group on the first part of the survey - the SUS questionnaire - was 82,5 points. This scale does not correspond to percentages. Instead, to evaluate the obtained result, the latter should be converted into a percentile rank, in order to assess how it compares to others. The percentile values according to the SUS score obtained are presented in Figure 5.11.

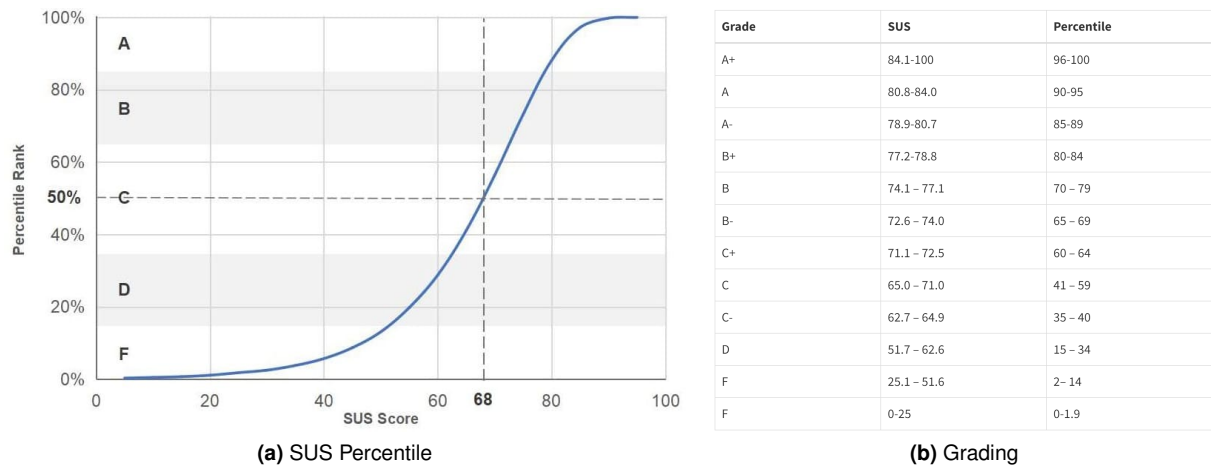


Figure 5.11: SUS scores, respective percentiles and grades.

From Figure 5.11, the mean 82,5 points obtained on the questionnaires from the surgeon doctors' group, is equivalent to the percentile 90% - 95%. This means the designed prototype had a very satisfying reception in terms of usability, including its user-friendliness. Regarding the second part of the questionnaire applied to the surgeon doctors, there was a very positive feedback on their behalf, and constructive suggestions were made in terms of possible applications for the prototype, describing specific scenarios in the surgery context where the system would be useful: Preoperative imaging (e.g., CT-Scan) were considered very relevant to present in the AR display, on demand. Furthermore, screen casting of the monitors displaying real-time intraoperative imaging were also considered very useful. For example, the real-time imaging from fluorescence cholangiography for bile duct visualization during laparoscopic cholecystectomy, is presented on a screen which is usually situated next to the surgeon. In the interviews with the surgeons, feedback was received towards screen casting those monitors displaying image intensifiers (to enhance blood vessels, guidewires), directly on the AR display, in order to reduce the number of times the surgeon needs to look away, and enhance their focus on the primary task. Another case study found useful was for the anesthesiologist. At times, during surgery, the latter leaves the Operating Room, being only called if / when necessary, according to the patient's vital signals as well. Screen casting the monitor displaying the patient's vital signals on the anesthesiologist's AR headset, would enable them to monitor the signals on demand, assisting them in that task. In terms of the control mechanism, the feedback received was consistently positive, with both surgeon doctors finding it intuitive and the two commands easy to execute.

6

Conclusion

Contents

6.1	Conclusions	67
6.2	Future Work	67

6.1 Conclusions

The increasing complexity of variables and information in the operating room has created challenges for surgical teams, which at the worst level, may potentially impact intervention outcomes. When it comes to the information flow in the OR during surgery, currently, the latter is still mostly static, with several monitors and displays being scattered along the room, and with the surgeon and auxiliary team having to actively seek the information. Improving the access, visualization, and integration of information in the operating room is a crucial challenge to enhance the flow of surgery and minimize disruptions that can impact patient safety.

Surgeon doctors and auxiliary team members using head-mounted displays to screen relevant patient information is a viable option, provided their control is user-friendly, and does not interfere with the surgery flow. Furthermore, the head-mounted display should be ergonomic, and do not physically interfere with the surgical apparatus the surgeon is required to use (e.g., surgical mask, magnifying glasses, and surgical headlamp).

Considering these requirements and constraints, a hands-free Augmented Reality headset controlled through forehead movements using electromyography (EMG) was successfully developed. Hardware was chosen according to experimental tests carried out. The developed solution was utilized and evaluated by two groups of subjects, including two surgeon doctors. Overall, the results on the control mechanism accuracy were very promising. Additionally, the surgeon doctors' interviews, questionnaires and usage feedback demonstrated a highly positive reception of the designed prototype in terms of usability and user-friendliness. The feedback received from the surgeon doctors was not only positive but also provided valuable insights into potential applications of the prototype in different surgical contexts.

6.2 Future Work

In terms of future developments for this system, there are several aspects that may be tackled to enhance the solution. On the one hand, a wider range of input movements may be considered, in order to increase the degrees of freedom in terms of app navigation and selection options. This would also enable displaying different contents broadcasted on the headset screen.

On the other hand, connecting the monitors in the OR to ARSurgery via Wi-Fi instead of pointing the smartphone's camera to them, would enhance the screen casting quality and allow switching between multiple monitors.

In terms of the onset detector, allowing ARSurgery to save the calibration parameters for each user, would simplify and shorten the setup phase. However, for this to be an effective option, it is important to either enhance the stability of electrode placement on the user's forehead, or enhance the robustness of

the algorithm for processing the acquired EMG signal. Furthermore, ARSurgery's calibration phase still involves using the smartphone's touchscreen to select the appropriate options. A notable enhancement would be the automation of these steps, eliminating the need for manual hand actions, thereby making the interface fully hands-free.

Lastly, the solution would benefit from incorporating robust data protection measures, giving users their own profile, secured with personalized passwords, ensuring the privacy and security of each user's information.

Bibliography

- [1] C. F. Rodrigues, “Técnicas avançadas em engenharia biomédica: Improving surgeon’s view: Augmented reality in the operating room,” Instituto Superior Técnico, Tech. Rep., 2021.
- [2] J. A. Wahr, R. L. Prager, J. Abernathy Iii, E. A. Martinez, E. Salas, P. C. Seifert, R. C. Groom, B. D. Spiess, B. E. Searles, T. M. Sundt III *et al.*, “Patient safety in the cardiac operating room: human factors and teamwork: a scientific statement from the american heart association,” *Circulation*, vol. 128, no. 10, pp. 1139–1169, 2013.
- [3] E. A. Martinez, A. Shore, E. Colantuoni, K. Herzer, D. A. Thompson, A. P. Gurses, J. A. Marsteller, L. Bauer, C. A. Goeschel, K. Cleary *et al.*, “Cardiac surgery errors: results from the uk national reporting and learning system,” *International journal for quality in health care*, vol. 23, no. 2, pp. 151–158, 2011.
- [4] G. Palmer, J. H. Abernathy III, G. Swinton, D. Allison, J. Greenstein, S. Shappell, K. Juang, and S. T. Reeves, “Realizing improved patient care through human-centered operating room design: a human factors methodology for observing flow disruptions in the cardiothoracic operating room,” *Anesthesiology*, vol. 119, no. 5, pp. 1066–1077, 2013.
- [5] U. Matern and S. Koneczny, “Safety, hazards and ergonomics in the operating room,” *Surgical endoscopy*, vol. 21, no. 11, pp. 1965–1969, 2007.
- [6] D. Silver, A. D. Kaye, and D. Slakey, “Surgical flow disruptions, a pilot survey with significant clinical outcome implications,” *Current pain and headache reports*, vol. 24, no. 10, pp. 1–4, 2020.
- [7] A. Joseph, A. Khoshkenar, K. M. Taaffe, K. Catchpole, H. Machry, and S. Bayramzadeh, “Minor flow disruptions, traffic-related factors and their effect on major flow disruptions in the operating room,” *BMJ quality & safety*, vol. 28, no. 4, pp. 276–283, 2019.
- [8] L. Al-Hakim, “The impact of preventable disruption on the operative time for minimally invasive surgery,” *Surgical endoscopy*, vol. 25, no. 10, pp. 3385–3392, 2011.

- [9] K. R. Catchpole, A. E. Giddings, M. Wilkinson, G. Hirst, T. Dale, and M. R. de Leval, "Improving patient safety by identifying latent failures in successful operations," *Surgery*, vol. 142, no. 1, pp. 102–110, 2007.
- [10] D. A. Wiegmann, A. W. ElBardissi, J. A. Dearani, R. C. Daly, and T. M. Sundt III, "Disruptions in surgical flow and their relationship to surgical errors: an exploratory investigation," *Surgery*, vol. 142, no. 5, pp. 658–665, 2007.
- [11] M. R. de Leval, J. Carthey, D. J. Wright, V. T. Farewell, J. T. Reason *et al.*, "Human factors and cardiac surgery: a multicenter study," *The Journal of thoracic and cardiovascular surgery*, vol. 119, no. 4, pp. 661–672, 2000.
- [12] N. Sevdalis, D. Forrest, S. Undre, A. Darzi, and C. Vincent, "Annoyances, disruptions, and interruptions in surgery: the disruptions in surgery index (disi)," *World journal of surgery*, vol. 32, no. 8, pp. 1643–1650, 2008.
- [13] T. Zhan, K. Yin, J. Xiong, Z. He, and S.-T. Wu, "Augmented reality and virtual reality displays: perspectives and challenges," *Iscience*, vol. 23, no. 8, p. 101397, 2020.
- [14] M. Eckert, J. S. Volmery, C. M. Friedrich *et al.*, "Augmented reality in medicine: systematic and bibliographic review," *JMIR mHealth and uHealth*, vol. 7, no. 4, p. e10967, 2019.
- [15] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino, "Augmented reality: A class of displays on the reality-virtuality continuum," in *Telemanipulator and telepresence technologies*, vol. 2351. Spie, 1995, pp. 282–292.
- [16] R. Skarbez, M. Smith, and M. C. Whitton, "Revisiting milgram and kishino's reality-virtuality continuum," *Frontiers in Virtual Reality*, vol. 2, p. 647997, 2021.
- [17] J. Carmigniani, B. Furht, M. Anisetti, P. Ceravolo, E. Damiani, and M. Ivkovic, "Augmented reality technologies, systems and applications," *Multimedia tools and applications*, vol. 51, no. 1, pp. 341–377, 2011.
- [18] G. L. Calhoun and G. R. McMillan, "Hands-free input devices for wearable computers," in *Proceedings Fourth Annual Symposium on Human Interaction with Complex Systems*. IEEE, 1998, pp. 118–123.
- [19] J. Lai and J. Vergo, "Medspeak: Report creation with continuous speech recognition," in *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*, 1997, pp. 431–438.
- [20] S. M. Faizan, "Applying speech recognition and language processing methods to transcribe and structure physicians' audio notes to a standardized clinical report format," 2020.

- [21] S. Das, R. Bakis, A. Nádas, D. Nahamoo, and M. Picheny, "Influence of background noise and microphone on the performance of the ibm tangora speech recognition system," in *1993 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2. IEEE, 1993, pp. 71–74.
- [22] K. Takemura, K. Takahashi, J. Takamatsu, and T. Ogasawara, "Estimating 3-d point-of-regard in a real environment using a head-mounted eye-tracking system," *IEEE Transactions on Human-Machine Systems*, vol. 44, no. 4, pp. 531–536, 2014.
- [23] L. Massin, V. Nourrit, C. Lahuec, F. Seguin, L. Adam, E. Daniel *et al.*, "Development of a new scleral contact lens with encapsulated photodetectors for eye tracking," *Optics Express*, vol. 28, no. 19, pp. 28 635–28 647, 2020.
- [24] J.-J. Yang, G. W. Gang, and T. S. Kim, "Development of eog-based human computer interface (hci) system using piecewise linear approximation (pla) and support vector regression (svr)," *Electronics*, vol. 7, no. 3, p. 38, 2018.
- [25] F. D. Perez Reynoso, P. A. Niño Suarez, O. F. Aviles Sanchez, M. B. Calva Yañez, E. Vega Alvarado, and E. A. Portilla Flores, "A custom eog-based hmi using neural network modeling to real-time for the trajectory tracking of a manipulator robot," *Frontiers in Neurorobotics*, vol. 14, Sep 2020.
- [26] J. R. Wolpaw, "Chapter 6 - brain-computer interfaces," in *Neurological Rehabilitation*, ser. Handbook of Clinical Neurology, M. P. Barnes and D. C. Good, Eds. Elsevier, 2013, vol. 110, pp. 67–74. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B978044452901500006X>
- [27] E. López-Larraz, "What is bci? an introduction to brain-computer interface using eeg," Dec 2022. [Online]. Available: <https://www.bitbrain.com/blog/brain-computer-interface-using-eeg-signals>
- [28] A. Barreto, S. Scargle, and M. Adjouadi, "A practical emg-based human-computer interface for users with motor disabilities," 2000.
- [29] M. Reis, C. Almeida, and R. M. Rocha, "On the performance of surface electromyography-based onset detection methods with real data in assistive technologies: Comparative analysis and enhancements via sensor fusion," *Multimedia Tools and Applications*, vol. 77, pp. 11 491–11 520, 2018.
- [30] A. Kaur, "Wheelchair control for disabled patients using emg/eog based human machine interface: a review," *Journal of medical engineering & technology*, vol. 45, no. 1, pp. 61–74, 2021.

- [31] B. Zhu, D. Zhang, Y. Chu, X. Zhao, L. Zhang, and L. Zhao, "Face-computer interface (fci): Intent recognition based on facial electromyography (femg) and online human-computer interface with audiovisual feedback," *Frontiers in Neurorobotics*, vol. 15, p. 692562, 2021.
- [32] M.-S. Song, S.-G. Kang, K.-T. Lee, and J. Kim, "Wireless, skin-mountable emg sensor for human-machine interface application," *Micromachines*, vol. 10, no. 12, p. 879, 2019.
- [33] M. Schliebener, L. Kraft, and M. Dufner, "An emg-based approach toward the assessment of implicit self-esteem," *Acta Psychologica*, vol. 234, p. 103868, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0001691823000446>
- [34] S. Orguc, H. S. Khurana, K. M. Stankovic, H. Leel, and A. Chandrakasan, "Emg-based real time facial gesture recognition for stress monitoring," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018, pp. 2651–2654.
- [35] B. Rodríguez-Tapia, I. Soto, D. M. Martínez, and N. C. Arballo, "Myoelectric interfaces and related applications: Current state of emg signal processing—a systematic review," *IEEE Access*, vol. 8, pp. 7792–7805, 2020.
- [36] I. Rodríguez-Carreño, L. Gila-Useros, and A. Malanda-Trigueros, "Motor unit action potential duration: measurement and significance," in *Advances in clinical neurophysiology*. IntechOpen, 2012.
- [37] A. Jaramillo-Yáñez, M. E. Benalcázar, and E. Mena-Maldonado, "Real-time hand gesture recognition using surface electromyography and machine learning: A systematic literature review," *Sensors*, vol. 20, no. 9, p. 2467, 2020.
- [38] K. R. Mills, "The basics of electromyography," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 76, no. suppl 2, pp. ii32–ii35, 2005. [Online]. Available: <https://jnnp.bmj.com/content/76/suppl.2/ii32>
- [39] A. Péter, E. Andersson, A. Hegyi, T. Finni, O. Tarassova, N. Cronin, H. Grundström, and A. Arndt, "Comparing surface and fine-wire electromyography activity of lower leg muscles at different walking speeds," *Frontiers in physiology*, p. 1283, 2019.
- [40] H. Baars, T. Jöllenbeck, H. Humburg, and J. Schröder, "Surface-electromyography: skin and subcutaneous fat tissue attenuate amplitude and frequency parameters," in *ISBS-Conference Proceedings Archive*, 2006.
- [41] S. Day, "Important factors in surface emg measurement," *Bortec Biomedical Ltd publishers*, pp. 1–17, 2002.

- [42] E. A. Clancy, E. L. Morin, and R. Merletti, "Sampling, noise-reduction and amplitude estimation issues in surface electromyography," *Journal of electromyography and kinesiology*, vol. 12, no. 1, pp. 1–16, 2002.
- [43] M. B. I. Reaz, M. S. Hussain, and F. Mohd-Yasin, "Techniques of emg signal analysis: detection, processing, classification and applications," *Biological procedures online*, vol. 8, pp. 11–35, 2006.
- [44] P. Konrad, "The abc of emg," *A practical introduction to kinesiological electromyography*, vol. 1, no. 2005, pp. 30–5, 2005.
- [45] D. Batista, H. Plácido da Silva, A. Fred, C. Moreira, M. Reis, and H. A. Ferreira, "Benchmarking of the bitalino biomedical toolkit against an established gold standard," *Healthcare Technology Letters*, vol. 6, no. 2, p. 32–36, 2019.
- [46] H. Silva, R. Scherer, J. Sousa, and A. Londral, "Towards improving the usability of electromyographic interfaces," *Biosystems amp; Biorobotics*, p. 437–441, 2013.
- [47] Z. Lu and P. Zhou, "Hands-free human-computer interface based on facial myoelectric pattern recognition," *Frontiers in Neurology*, vol. 10, 2019.
- [48] P. Yupapin, Hamed, Salleh, Tan, Ismail, Ali, Dee-Uam, and Pavaganun, "Human facial neural activities and gesture recognition for machine-interfacing applications," *International Journal of Nanomedicine*, p. 3461, 2011.
- [49] A. S. Carmo, M. Abreu, A. L. N. Fred, and H. P. da Silva, "Epibox: An automated platform for long-term biosignal collection," *Frontiers in Neuroinformatics*, vol. 16, 2022. [Online]. Available: <https://www.frontiersin.org/articles/10.3389/fninf.2022.837278>
- [50] A. P. Alves, H. Plácido da Silva, A. Lourenco, and A. Fred, "Signalbit: A web-based platform for real-time biosignal visualization and recording," 01 2013.
- [51] A. Lourenço, H. P. da Silva, C. Carreiras, A. Priscila Alves, and A. L. Fred, "A web-based platform for biosignal visualization and annotation," *Multimedia Tools and Applications*, vol. 70, no. 1, p. 433–460, Mar 2013.
- [52] P. W. Hodges and B. H. Bui, "A comparison of computer-based methods for the determination of onset of muscle contraction using electromyography," *Electroencephalography and Clinical Neurophysiology/Electromyography and Motor Control*, vol. 101, no. 6, p. 511–519, 1996.
- [53] R. A. Grier, A. Bangor, P. Kortum, and S. C. Peres, "The system usability scale," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 57, no. 1, p. 187–191, 2013.



Informed Consent and Request for Data Collection

CONSENTIMENTO INFORMADO

Título do Estudo

Analysis of facial surface electromyography (sEMG) and Electroencephalography (EEG) signals for hands free control of augmented reality headsets

Enquadramento

Obrigado por se voluntariar para esta experiência para avaliação de comandos fisiológicos para interação homem-máquina, desenvolvida pelo IT – Instituto de Telecomunicações. Esta experiência vai ser realizada por Margarida Ramalho e Paulo Rodrigues, estudantes de Mestrado do Instituto Superior Técnico – Universidade de Lisboa. O investigador responsável será a Prof. Ana Fred (IT-IST, UL), com o apoio do Prof. Hugo Plácido da Silva (IT-IST, UL), do Prof. Paulo Correia (IT-IST, UL), da Dra. Carolina Rodrigues (Hospital de Santa Marta), e do Mestre Miguel Martins (Rotacional, Lda.).

Propósito

Avaliar a qualidade de diferentes comandos fisiológicos – gestos faciais – tendo em vista a sua aplicação no controlo hands-free de um headset de realidade aumentada (RA) em contexto ecológico, com recurso a sensores do tipo wearable.

Procedimento

Cada experiência envolve um participante disposto numa sala, respeitando as medidas sanitárias e recomendações da Direção Geral da Saúde (DGS). Prevê-se a utilização da sala de reuniões do Instituto de Telecomunicações (IST Torre Norte – Piso 11), ainda que possa ser necessário recorrer a outras localizações que reúnam condições adequadas. O participante terá um ou mais sensores *wearable* aplicados no seu corpo, requerendo a aplicação de eléctrodos na região da testa, por forma a captar os seus sinais de Eletromiografia de Superfície (sEMG), Eletroencefalografia (EEG) e movimento (IMU). Após a aplicação do sensor, o nível de conforto será avaliado utilizando escalas qualitativas como a System Usability Scale (SUS). Durante a experiência, o participante realizará uma sequência de gestos faciais com amplitudes temporais diferentes, tendo simultaneamente os sinais captados e enviados em tempo real para o computador pessoal do investigador, onde são armazenados localmente. Após a experiência, ser-lhe-á pedido para preencher um questionário de auto-reporte do nível de interesse nos conteúdos exibidos.

Registo de Dados

Além dos elementos registados no presente formulário, durante a experiência serão registados os sinais dos sensores *wearable* e as respostas ao questionário de auto-reporte. Todos os dados registados (EMG,

Página 1 de 2

EEG, IMU e questionário de auto-reporte) estarão desassociados do seu nome. A associação é feita unicamente através do código.

Privacidade

Os dados são protegidos por palavra-passe e guardados para estudos futuros por um prazo de 5 anos (ou até solicitar a eliminação). Se autorizar será mantida uma versão anonimizada dos dados numa base de dados pública, para fins de investigação. Em qualquer momento poderá ter acesso aos dados, pedir a sua correção ou até a eliminação, enviando um email para os responsáveis pelo tratamento dos dados (Margarida Ramalho <margarida.v.b.l.ramalho@tecnico.ulisboa.pt> e Paulo Rodrigues <paulo.f.rodrigues@tecnico.ulisboa.pt>). Os dados poderão ser partilhados com terceiros se concordar.

Riscos

Os riscos de participação neste estudo não são superiores aos da vida quotidiana, havendo apenas recolha de dados.

Participação

A participação nesta experiência é voluntária, devendo ser feita de forma esclarecida e livre. Poderá questionar a sua participação e desistir em qualquer momento sem consequências, caso em que os seus dados serão destruídos.

Esclarecimento de Dúvidas

Poderá esclarecer as suas dúvidas em qualquer altura com o executante da experiência, ou por e-mail contactando a Prof. Ana Fred <afred@lx.it.pt>, o Prof. Hugo Plácido da Silva <hsilva@lx.it.pt>, ou o encarregado de proteção de dados do IT, Marcelino Pousa <mpousa@av.it.pt>. Em último caso tem o direito de apresentar reclamação à Comissão Nacional de Proteção de Dados (CNPD).

Acordo de Participação

Eu li e entendi os detalhes deste estudo, tendo esclarecido eventuais questões complementares com o executante. Ao assinar abaixo, concordo em participar nesta experiência.

Código: Nome completo: Data:

Género (deixar em branco caso não pretenda indicar): Idade:

Email: Assinatura:

AUTORIZAÇÃO PARA UTILIZAÇÃO DE DADOS

Título do Estudo

Analysis of facial surface electromyography (sEMG) and Electroencephalography (EEG) signals for hands free control of augmented reality headsets

Acordo para Utilização de Dados

Obrigado por participar na nossa experiência. Com o propósito de estudar os diferentes modos de interação homem-máquina e, no futuro, implementar comandos fisiológicos no controlo hands-free de um headset de realidade aumentada (RA), os seus dados são muito importantes para o IT.

Os dados que recolhemos (indicados no consentimento informado) não serão associados ao seu nome ou a outras informações que o identifiquem. Com este formulário, você pode consentir explicitamente a finalidade para a qual podemos utilizar os seus dados. Se, por algum motivo, desejar que os seus dados sejam excluídos imediatamente e não sejam utilizados para análise, por favor escreva a inicial do seu nome abaixo:

Desejo que os meus dados sejam excluídos e não sejam usados em qualquer análise

Se permitir que utilizemos os seus dados, assinale, com a sua inicial, os propósitos específicos:

a. A equipa pode visualizar e analisar os meus dados

b. A equipa pode mostrar os meus dados em reuniões confidenciais com parceiros

c. A equipa pode partilhar os meus dados com parceiros para análise posterior

d. A equipa pode mostrar os meus dados em publicações académicas

e. A equipa pode mostrar os meus dados em conferências

f. A equipa pode mostrar os meus dados em meios de comunicação (notícias, revistas, etc.)

g. A equipa pode partilhar uma versão anonimizada dos meus dados para a criação de uma base de dados pública a disponibilizar à comunidade científica

Confirmo que li todas as declarações acima, e que a equipa só pode usar meus dados para as finalidades que eu assinalei.

Nome completo: Data:

Email: Assinatura:

Página 1 de 1

