



Department of Industrial, Electronic and Mechanical Engineering

Ph.D. in Applied Electronics XXXVI cycle

Quality of Experience across different application areas

Supervisor:

Prof. Marco Carli

Ph.D. Candidate:

Ilaria Mannocchi

Academic Year 2023/2024

A nonno Totto e nonno Peppe

Contents

List of Figures	v
List of Tables	viii
List of Acronyms	x
Abstract	1
1 Introduction	2
2 Quality of Experience	5
2.1 Affective Computing	5
2.2 The Body Language	7
3 HEROES: A Video-Based Human Emotion Recognition Database	14
3.1 Human Emotion Recognition Observing the Evolution of the Skeleton (HEROES) dataset creation	16
3.2 Body joint extraction	18
3.2.1 High level features	21
3.2.2 Low level features	23
3.2.3 Analysis of Annotation Quality and Inter-Rater Consistency	23
3.3 Classifications results on HEROES	24
4 Virtual Reality application	27
4.1 Device Validation	30

4.1.1	Movements along the axis	32
4.1.2	Data Analysis	33
4.1.3	Methodological Framework for Motion Accuracy Assessment	36
4.1.4	Results	37
4.2	Does the app follow the rehabilitation protocol?	44
4.2.1	Game Design and Development	45
4.2.2	The experiment	48
4.2.3	Data Collection	50
4.2.4	Results	53
4.3	Quality of Experience of the physiotherapists	59
4.3.1	Experimental Procedure	59
4.3.2	Questionnaires	61
4.3.3	Results	62
5	Quality of Experience (QoE) and the art “The Virtual Reality (VR) effect”	68
5.1	EHEM-Enhancement of Heritage Experiences: The Middle Ages.	68
5.1.1	Digitalization of Santa Maria Antiqua	72
5.1.2	The evaluation experiment	72
5.1.3	Structure of the questionnaires and method of administration	76
5.1.4	Analysis of Results	77
5.1.5	Stratified Analysis of User Experience	81
5.2	Visualizing San Saba	84
5.2.1	Program description and activities	87
5.2.2	Program Evaluation and Discussion	90
5.2.3	Knowledge before and after the summer school	90
5.2.4	Students’ comments	90
5.2.5	Guidelines	91
6	Discussion	95
6.1	Future developments, applications	108
6.2	Limits of the research	110

7 Conclusion	111
Bibliography	113

List of Figures

2.1	Interpersonal distances by Edward T. Hall's.	8
2.2	Plutchik wheel model.	10
2.3	Russell 2D model.	10
2.4	Mehrabian and Russell 3D model.	11
3.1	Example of human poses and corresponding emotions in HEROES.	15
3.2	Block diagram of the feature extraction pipeline. The fusion block is responsible for concatenating all the low-level features.	20
3.3	Confusion matrices for HEROES dataset with: (a) whole body joints information ($\mathbf{f}_{body}$), and (b) with upper body joints information only ($\mathbf{f}_{upper}$). Here, B: Boredom, D: Disgust, H: Happiness, and I: Interest.	26
4.1	Oculus Quest 2 (OQ2): from the left, the Head Mounted Display (HMD), the left controller (controller A), and right controller (controller B).	30
4.2	(a) Controller A and markers configuration used to define the (b) rigid body. (c) Controller B and the position of the synchronization marker. Controller A was allocated in a (d) custom-made support designed to be easily moved on the structures for (e) translational and (f) rotational movements.	31
4.3	Acquisition configuration of translational displacements along the (a) x-axis, (b) y-axis, and (c) z-axis, starting at a distance of about 200 mm from the reference position P_0 ; (d) representation of translational displacements P_i ($i = 1, \dots, 11$) and variations (ΔP , i.e., differences between translational displacements i^{th} and i^{th-1}).	33

4.4	Acquisition configuration of rotational displacements around the (a) x-axis, (b) y-axis, and (c) z-axis, starting at a distance of about 700 mm from the reference position R0; (d) representation of rotational displacements R_i ($i = 1, \dots, 11$) and variations (ΔR , i.e., differences between rotational displacements i^{th} and i^{th-1}).	34
4.5	Absolute error along the x-, y-, and z-axes at different distances from the HMD. . . .	39
4.6	Percentage error along the x-, y-, and z-axes at different distances from the HMD. . .	39
4.7	Error corresponding to the increment of 5 mm at the distances of 205 mm, 455 mm, and 705 mm from the HMD.	40
4.8	Bland-Altman plots comparing the translational displacements measured by the Qualisys Optical Motion Capture (OMC) system and by the Controller A of the OQ2. . . .	40
4.9	Absolute error corresponding to rotational displacements around the x-, y-, and z-axes.	42
4.10	The percentage error corresponding to rotational displacements around the x-, y-, and z- axes.	42
4.11	Bland-Altman plots comparing the rotational displacements measured by the Qualisys OMC system and by the Controller A of the OQ2.	43
4.12	Meta model of the developed immersive virtual reality application. “*” means the variable number of elements.	46
4.13	Specification of requirements of the developed immersive virtual reality application. .	47
4.14	Representation of a healthy volunteer wearing the OQ2, the Beat Saber Handle, and photo-reflective markers during flexion movement to reach the virtual target (a) and in the starting position (b) . Arrows indicated the direction of movement.	50
4.15	Representation of a healthy volunteer wearing the OQ2, the Beat Saber Handle, and photo-reflective markers during abduction movement to reach the virtual target (a) and in the starting position (b) . Arrows indicated the direction of movement.	50
4.16	Representation of a healthy volunteer wearing the OQ2, the Beat Saber Handle, and photo-reflective markers during external rotation to reach the virtual target (a) and in the starting position (b) . Arrows indicated the direction of movement.	51
4.17	Schematization of shoulder joint angles.	52

4.18	Starting position (left) and final position (right) of the elevation exercise in the sagittal plane. In the center is a representation of what the user visualizes through the HMD and the end effector circled in dotted yellow	61
4.19	Starting position (left) and final position (right) of the elevation exercise in the frontal plane. In the center is a representation of what the user visualizes through the HMD and the end effector circled in dotted yellow.	62
4.20	Starting position (left) and final position (right) of the external rotation exercise. In the center is a representation of what the user visualizes through the HMD and the end effector circled in dotted yellow	63
4.21	Physiotherapists' ratings on the acceptability questionnaire.	64
4.22	Physiotherapists' ratings on the usability questionnaire.	65
4.23	Physiotherapists' ratings on the appropriateness questionnaire.	66
5.1	Timeline guide.	73
5.2	User in the VR reconstruction of Santa Maria Antiqua.	74
5.3	Virtual reconstruction of Santa Maria Antiqua and points of interest of the apse. . . .	74
5.4	Virtual reconstruction of Santa Maria Antiqua and annotations.	75
5.5	Location of the Basilica of San Saba. Courtesy of Google Maps.	86
5.6	A detail of the virtual tour on the website.	89
5.7	Homepage of the website.. . . .	89
5.8	Computer science skills before and after the summer school.	91
5.9	Digital image skills before and after the summer school.	91
5.10	3D graphics skills before and after the summer school	92
5.11	Photogrammetry knowledge before and after the summer school.	92
5.12	Photoshop skills before and after the summer school	93
5.13	Autocad skills before and after the summer school.	93

List of Tables

2.1	Datasets in the literature and their main characteristics.	13
3.1	The emotions in the space valence-arousal.	15
3.2	Expressed emotions in the two considered scenarios.	17
3.3	Examples frames of the dataset in which non-professional actors express the emotions.	19
3.4	Video duration in seconds for emotions and actors of the HEROES dataset. . .	20
3.5	High (S_i) and low level (F_i) features for a generic i -th frame.	21
4.1	Main features of OQ2.	30
4.2	Errors for all the translational displacements along the x-, y-, and z- directions.	38
4.3	Errors for all the rotational displacements around the x-, y-, and z-axes.	41
4.4	Movements, ROM, and repetitions for each level included in the immersive virtual reality application.	49
4.5	Results for the first series of flexion for each level (Mean Absolute Error (MAE_{inf} and MAE_{sup} are the MAEs corresponding to each volunteer; N_{inf} and N_{sup} corresponded to the number of times the inferior and superior conditions were not met, respectively; $MAE_{TOT,inf}$ and $MAE_{TOT,sup}$ were the total MAE related to the inferior and superior conditions, respectively).	54

4.6	Results for the second series of flexion for each level (MAE_{inf} and MAE_{sup} are the MAEs corresponding to each volunteer for the inferior and superior conditions, respectively; N_{inf} and N_{sup} corresponded to the number of times the inferior and superior conditions were not met, respectively; $MAE_{TOT,inf}$ and $MAE_{TOT,sup}$ were the total MAE related to the inferior and superior conditions, respectively).	55
4.7	Results for abduction for each level (MAE_{inf} and MAE_{sup} are the MAEs corresponding to each volunteer for the inferior and superior conditions, respectively; N_{inf} and N_{sup} corresponded to the number of times the inferior and superior conditions were not met, respectively; $MAE_{TOT,inf}$ and $MAE_{TOT,sup}$ are the total MAE related to the inferior and superior conditions, respectively). . .	56
4.8	Results for external rotation for each level (MAE_{inf} and MAE_{sup} are the MAEs corresponding to each volunteer for the inferior and superior conditions, respectively; N_{inf} and N_{sup} corresponded to the number of times the inferior and superior conditions were not met, respectively; $MAE_{TOT,inf}$ and $MAE_{TOT,sup}$ are the total MAE related to the inferior and superior conditions, respectively). . .	57
4.9	Results for internal rotation for each level (MAE_{inf} and MAE_{sup} are the MAE corresponding to each volunteer for the inferior and superior conditions, respectively; N_{inf} and N_{sub} corresponded to the number of times the inferior and superior conditions were not met, respectively; $MAE_{TOT,inf}$ and $MAE_{TOT,sup}$ are the total MAE related to the inferior and superior conditions, respectively).	58
5.1	Results about System Usability Questionnaire (SUQ).	79
5.2	Comments and suggestions by participants.	80
5.3	Summary of the Analyzed Groups.	83
5.4	Results of the SUQ.	83
5.5	R! (R!) results of SSQ symptoms.	83
5.6	Technical program of the school.	87
5.7	Evaluation of the training period by students. Scores are in the range $[1, 5]$. . .	94
5.8	Most frequent comments from the students.	94

List of Acronyms

VR	Virtual Reality
AR	Augmented Reality
QoE	Quality of Experience
CNN	Convolutional Neural Networks
LSTM	Long-Short Time Memory
ROI	Region Of Interest
HMD	Head Mounted Display
ROM	Range of Motion
FOV	Field of View
HEROES	Human Emotion Recognition Observing the Evolution of the Skeleton
HRV	Heart Rate Variability
fNIRS	functional Near-InfraRed Spectroscopy
EEG	Electroencephalogram
k-NN	k-Nearest Neighbour
AC	Affective Computing
QoS	Quality of Service

ML	Machine Learning
ROM	Range of Movement
DOF	Degree of Freedom
OMC	Optical Motion Capture
QTM	Qualisys Track Manager
HV	Healthy Volunteer
GEMEP	Geneva Multimodal Emotion Portrayals
IEMOCAP	Interactive Emotional Dyadic Motion Capture Database
FABO	Bimodal Face and Body Gesture Database
LOA	limits of agreement
MOD	mean of difference
RC	Rotator Cuff
STEM	Science, Technology, Engineering e Mathematics
GIS	Geographic Information System
AI	Artificial Intelligence
BIM	Building Information Modeling
CAD	Computer-Aided Design
CH	Cultural Heritage
CSS	Cascading Style Sheets
GNSS	Global Navigation Satellite System
HTML	HyperText Markup Language

IR	InfraRed radiation
MR	Mixed Reality
XR	miXed Reality
SSQ	Simulator Sickness Questionnaire
MSQ	Motion Sickness Questionnaire
mAP	mean Average Precision
IoU	Intersection-over-Union
LOA	Limits of Agreement
MOD	Mean of Difference
MAE	Mean Absolute Error
EHEM	Enhancement of Heritage Experiences
SUQ	System Usability Questionnaire
OQ2	Oculus Quest 2

Abstract

Quality of Experience (QoE) is an interdisciplinary concept that assesses an individual's level of satisfaction and engagement during an interaction, based on sensory perception, subjective interpretation, and cultural influences. QoE varies among individuals due to personal factors such as cultural background and emotional state, and it is complex to evaluate as it includes sensory engagement, emotional responses, and intellectual reflection. This thesis examines how movement and emotion influence QoE, and how body movements can reveal insights into the quality of the user's experience. Existing emotion databases were examined, and a new one was developed to support an emotion recognition system based on Machine Learning (ML). Moreover, this thesis explores the potential of Virtual Reality (VR) as a tool for immersive and interactive experience, transforming the user from a passive observer to an active participant and personalizing QoE based on the user's movements and choices. The findings of this thesis open new perspectives on utilizing VR as an immersive tool to enhance user well-being and emotional communication across different fields, including rehabilitation and the arts.

Chapter 1

Introduction

Quality of Experience (QoE) is a concept that originates from the fields of psychology, philosophy, and aesthetics, often used to refer to the degree of satisfaction or achievement derived from an interaction. The QoE captures the extent to which a viewer, listener, or participant feels emotionally, intellectually, and sensorially engaged. The QoE is grounded in sensory perception. Our senses, vision, hearing, touch, and even smell act as conduits through which people engage with the world. The concept is based on several factors: subjective interpretation, sensory input, emotional resonance, and cultural or contextual influences. However individual's experience is unique, different people can have different feelings with the same experience. This variance is attributed to personal factors such as one's cultural background, prior knowledge, and emotional state. As the QoE refers to the subjective evaluation of our interactions with the world around us, the evaluation becomes even more complex, including sensory engagement, emotional responses, and intellectual reflection. Through different types of experiences, people can explore the depths of their emotional landscapes; in fact, both movement and emotion play an important role in the determination of the quality of these experiences. Examining the influence of movement and emotion on QoE, enables a deeper understanding of how various applications can serve as powerful tools for communication and emotional expression. In this area, Virtual Reality (VR) plays a fundamental role and its application varies in different fields such as education, healthcare, and entertainment. A key topic in recent research is the assessment of the quality of the user experience in virtual environments. VR can put users inside a fully immersive environment, which can significantly

amplify the QoE. It also makes QoE interactive and dynamic, depending on the viewer's movements or choices, within that virtual space. In addition, VR can allow users to physically move around in a virtual digital space while exploring the virtual environment as if it were real. This sense of embodied movement radically alters the relationship between the viewer and the virtual space, transforming the experience from passive observation to active exploration. The dynamic interplay between movement and emotion in VR, creates a feedback-loop where the viewer's actions and emotional responses continuously modify the virtual environment, resulting in a personalized and evolving QoE. This interactivity transforms the viewer from a passive observer to an active participant, deepening the emotional engagement and sensory immersion of the experience. In this context, the biggest challenge, as previously mentioned, is the variability of emotions and sensations that multiple users can experience despite experiencing the same virtual context. Based on the aforementioned, the objective of this thesis is to study the QoE of users across different application areas and through various methods.

The primary research questions based on the thesis description can be identified as follows:

- **RQ1) Can emotions be recognized through body language?:** this question investigates whether an emotion recognition system based on body movements (body language) can identify and differentiate between various emotions in real-time;
- **RQ2) What is the value of a new emotion database for emotion recognition through movement?:** the thesis addresses the need for a new database to fill existing gaps in the literature, useful for training an Artificial Intelligence (AI) system in emotion recognition;
- **RQ3) How can VR be effectively applied in rehabilitation (specifically shoulder rehabilitation) to enhance the QoE of users?:** this question explores the use of VR in a therapeutic context, assessing how emotional engagement and interactivity can improve the rehabilitation experience;
- **RQ4) How do virtual reconstruction projects of artistic sites influence the QoE of users?:** this question examines the impact of VR in the artistic and cultural context, assessing QoE through immersive experiences that allow users to virtually explore historical sites.

In Chapter 2, after a brief introduction to the concepts of Affective Computing (AC) and QoE, the concept of body language and emotions is explained. Following this, various emotion databases from existing literature are described, highlighting the need for the creation of a new one. In Chapter 3, the first study is presented, addressing the fundamental question: “Can emotions be recognized through body language? Even the lower part of the body?”. This leads to the creation of a new database that addresses the gaps identified in the literature. Using this database, a trained Machine Learning (ML) system was capable of recognizing specific emotions and distinguishing between them in real-time with a certain level of accuracy. This emotion recognition system, modified and integrated with different features, could then be applied in different contexts. In this chapter, after a brief introduction to the topic, the Human Emotion Recognition Observing the Evolution of the Skeleton (HEROES) video dataset of emotions, which has been made available to the research community, is described. Movements were collected from 16 non-professional actors expressing four different emotions (happiness, boredom, disgust, and interest) without restrictions. The dataset consisted of 256 frontal video clips, recorded using GoPro cameras. Then, in Chapter 4, the project conducted in collaboration with the University of Campus Bio-Medico of Rome is introduced. This project included three different studies that aimed to create and validate a VR application for upper limb rehabilitation. The quality of the physiotherapists’ experience, concerning this custom application, was then evaluated, and their suggestions for modifications were gathered to refine the application before patient use. The last part of the thesis is dedicated to studying QoE in a different context: the artistic one. Specifically, Chapter 5 describes the virtual reconstruction projects of the Church of Santa Maria Antiqua and the Basilica of San Saba. Along with detailing the techniques used to achieve these virtual reconstructions, the results concerning the QoE of the users are also presented. Finally, in Chapter 6, the results are discussed, with a focus on limitations and potential future work, and conclusions are drawn in Chapter 7.

Chapter 2

Quality of Experience

2.1 Affective Computing

The term AC refers to a research field dedicated to studying and developing software that recognizes human behaviors and emotions. Its goal is to design human-computer interfaces that respond to users' emotional needs and bridge the gap between humans and technology [1]. First introduced by Prof. Picard in 1997, AC has played a significant role in enabling computers to identify, express, and intelligently respond to human emotions [2]. Traditionally, “affect” was primarily studied by psychologists and rarely associated with machines; however, advancements in technology now allow affective features to be captured and processed by computers. AC builds an “affect model” based on data collected from various sensors, creating a personalized computing system capable of perceiving and interpreting human emotions and providing intelligent, sensitive, and friendly responses. This field, focused on fundamental emotions and human interactions, offers numerous applications. In the entertainment sector, for instance, modern gaming systems can capture user inputs through voice or gestures [3]. In the medical field, recent research [4] has explored the integration of AC with VR and Augmented Reality (AR) to enhance communication among professionals, patients, and healthcare providers. Additionally, AR and VR are increasingly used in medical education, allowing professionals and students to develop their skills before performing actual treatments on patients [4].

QoE reflects a user's satisfaction with a service, such as browsing a website, using an app,

or making a phone call. It is grounded in real end-user experiences and depends on both objective and subjective parameters [5]. Objective factors, linked to Quality of Service (QoS), relate to network and software performance, while subjective factors involve individual preferences, expectations, and prior video experiences. These last subjective factors (individual preferences, expectations, previous video experiences) are more challenging to classify, as they are influenced by the user's personal and sometimes physiological characteristics, which are more abstract [5]. The Qualinet White Paper on Definitions defines QoE for immersive media, such as VR, as "the degree of delight or annoyance of the user of an application or service that involves an immersive media experience. It results from how well the application or service meets the user's expectations in terms of utility and/or enjoyment, considering their personality and current state" [6]. VR technology is an advanced human-computer interface simulating a virtual environment where people can move and interact with unreal objects [7], [8]. The most recent immersive VR devices offer a three-dimensional stereoscopic vision through an Head Mounted Display (HMD) and input devices (VR controllers), in which users are immersed and can interact with the virtual environment [9], [10]. Nowadays, VR headsets are used in many different fields such as entertainment [11], education [12], [13], [14] and even in the health-care and rehabilitation sector [15], [16], [17], [18]. There are various methods for assessing the quality of user experience in VR, but no standardized systems currently exist. The most common approaches found in the literature include subjective questionnaires, which are widely used in QoE studies, and objective methods that encompass physiological, behavioral, and task performance measures. Several questionnaires have been developed to evaluate user experience in VR, with the most commonly used being Simulator Sickness Questionnaire (SSQ), Motion Sickness Questionnaire (MSQ), and System Usability Questionnaire (SUQ) [19]. At the same time, physiological parameters such as Electroencephalogram (EEG), Heart Rate Variability (HRV), functional Near-Infrared Spectroscopy (fNIRS), and galvanic skin response have been employed to assess users' emotional status while they view multimedia content [20]. However, although contact sensors provide more objective data, they often make these techniques invasive and uncomfortable for users. To address this issue, studies like [21] have evaluated tourists' emotional status and satisfaction levels during sightseeing using behavioral cues (such as eye and head/body movements) and audiovisual data (including facial and

vocal expressions). Due to their design, some medical instruments used to collect these data may affect the user experience and reduce QoE scores. Consequently, less intrusive devices, such as fitness bands and smartwatches, have also been used to gather physiological signals, as well as information on user behavior and movement. But can body language alone provide reliable insights into the emotions and quality of the user experience?

2.2 The Body Language

In this context, emotions play a crucial role, as understanding feelings is fundamental to human experience. Charles Darwin's *The Expression of the Emotions in Man and Animals*, published in 1872, is one of the earliest research works on non-verbal communication and behavior. In this work, Darwin proposed that animals, like humans, express their emotions through facial expressions. He suggested that these expressions are linked to habits and behaviors associated with specific functions throughout evolutionary history, and that both humans and animals display similar ways of expressing emotions [22]. Although facial expressions are no longer essential for their original survival purposes, they have persisted in daily life due to their significant role in communication. In the mid-1900s, more in-depth studies on body language were conducted by Birdwhistell [23], regarded as the founder of kinesics (from the Greek *kínēsis*, meaning “movement”): the study of body movement in communication. Kinesics specifically examines non-verbal communication, including body movements, gestures, facial expressions, and posture. Another important aspect of non-verbal communication is proxemics (from the Latin *proximitas*). This term, introduced by Edward Twitchell Hall in 1963 [24], refers to a branch of semiotics that studies how different cultures use space and the respective distances involved in interpersonal communication (Figure 2.1). In particular, Hall [24] described four distinct interpersonal distance zones:

- **Intimate distance** : for embracing, touching or whispering;
- **Personal distance**: for interactions among good friends or family;
- **Social distance**: for interactions among acquaintances;
- **Public distance**: used for public speaking.

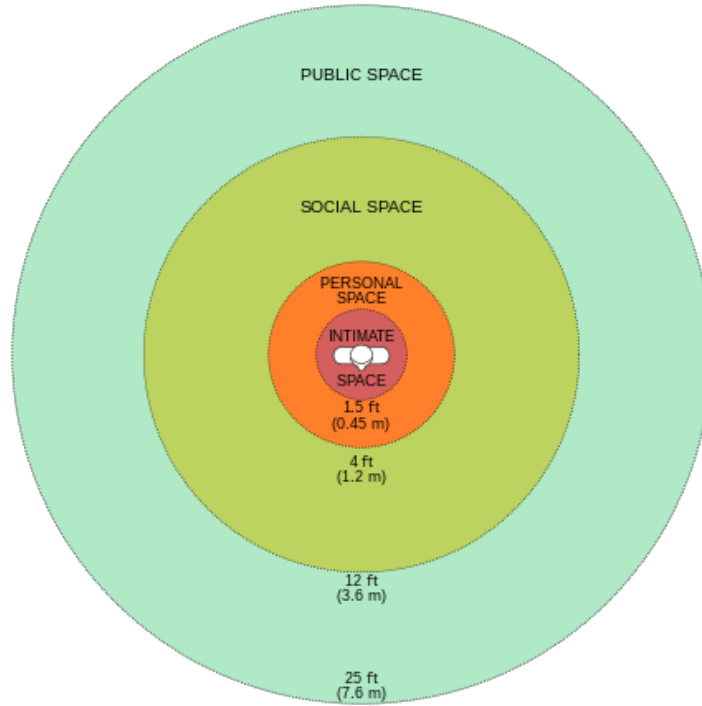


Figure 2.1: Interpersonal distances by Edward T. Hall's.

Mehrabian [25] demonstrated that in any form of communication, the impact of a message is composed of only 7% verbal speech (words), 38% vocal elements (such as pitch and inflections), and 55% non-verbal cues, including facial expressions and body language. Numerous studies support this, concluding that body language plays a crucial role in social interactions. For example, nodding can encourage conversation, while a slight tilt of the head can convey interest [22]. Non-verbal communication generally encompasses both conscious and unconscious processes of sending (encoding) and interpreting (decoding) messages, with decoding referring to the interpretation of signals received from another person. Furthermore, Mehrabian [26] described the human consistency in expressing and recognizing basic emotions: happiness, sadness, fear, anger, disgust, and surprise. These emotions, often seen as universal, are deeply embedded in human psychology and physiology, allowing individuals from diverse cultural backgrounds to recognize them with remarkable consistency. Mehrabian's model highlights that these core emotions are communicated not only through verbal expressions but also through non-verbal cues, such as facial expressions, body posture, and tone of voice, which are universally understood. This invariability suggests an evolutionary

basis, where the ability to recognize such emotions has developed as a crucial survival mechanism, enabling humans to quickly assess threats, form bonds, and navigate social interactions across different environments.

Psychological research has defined that different emotions are stimulated by different parts of the brain, and emotions are reactions to a presented stimulus which produces a specific physiological reaction [27]. Based on this understanding, emotions can be classified into two main types: discrete emotion models and multi-dimensional emotion models. The discrete model represents a more descriptive than quantitative method of analyzing emotions. Emotions are typically divided into two main categories: negative or unpleasant (including sadness, fear, and anger) and positive or pleasant (such as happiness, joy, and excitement). Ekman [28] proposed a categorial model in which the system was based on grouping emotions in different classes, such as the six basic emotions (happiness, sadness, fear, anger, disgust, and surprise), highlighting the human invariability in expressing and recognizing these universal emotions across cultures. Similarly, Plutchik [29] classified emotions into eight primary categories: fear, joy, sadness, trust, anger, surprise, anticipation, and disgust. He also developed the “wheel of emotions,” a model in which emotions are connected based on color (with lighter colors representing less intense emotions) and position (with emotions closer to the center being more intense) as shown in Figure 2.2. However, analyzing complex emotions remains challenging, and so they should also be explored using quantitative methods. In multi-dimensional space models, emotional states can be represented using either 2D or 3D models. The 2D model categorizes emotions based on two key dimensions: *valence* (how pleasant or unpleasant the emotion is) and *arousal* (the level of activation or energy experienced during the emotion) [30]. For instance, disgust is characterized by negative valence because it is a negative feeling and it has a high activation [31] while interest is identified by positive valence and low arousal [32]. The 2D model created by Russell is shown in Figure 2.3. However, the 2D model has limitations in distinguishing similar emotions that share the same quadrant, as they may have similar levels of valence and arousal. For this reason, Mehrabian and Russell proposed the 3D model considering three dimensions: valence, arousal, and dominance. Dominance, represented as the third axis, ranges from passive to dominant and reflects the level of control or power a person feels in a given emotional state (Figure 2.4). Moreover, the analysis conducted

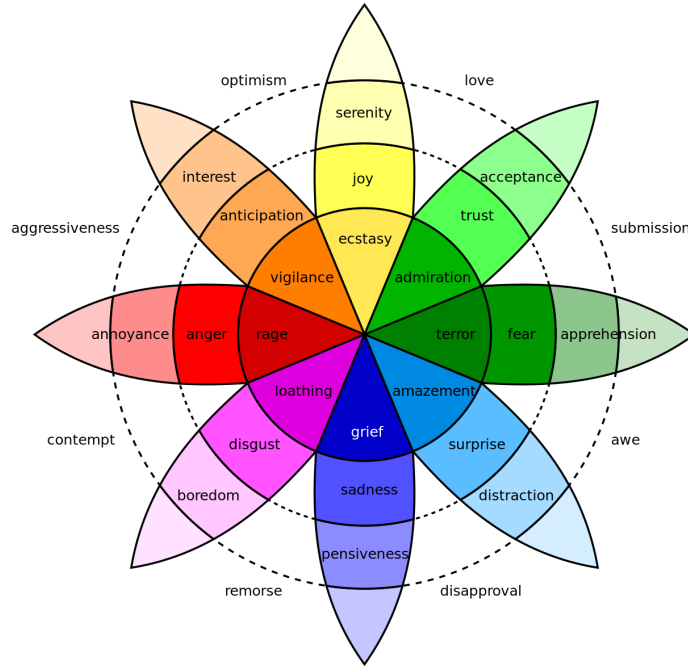


Figure 2.2: Plutchik wheel model.

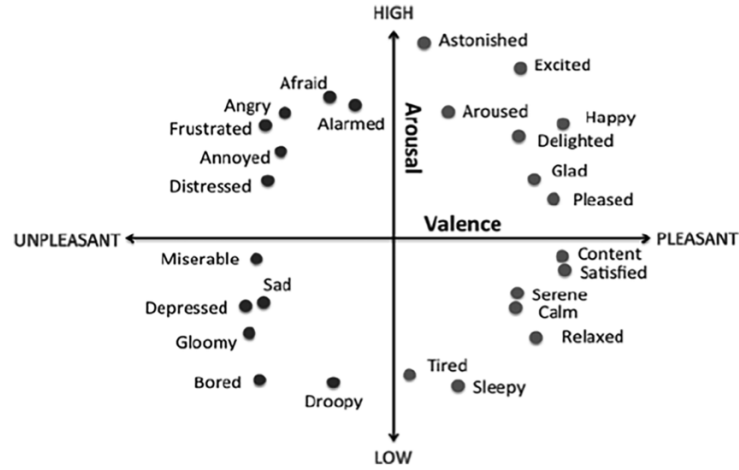


Figure 2.3: Russell 2D model.

in [33] highlighted that the frontal view provides the best results for emotion recognition and that emotions such as disgust and fear are less recognizable than happiness, anger, or sadness. The study in [34] demonstrated that human movements and posture can be associated with specific emotions. For example, boredom is characterized by limited movements, a collapsed posture, a head tilted back, and very small, slow arm movements [35]. Nonetheless, the lower

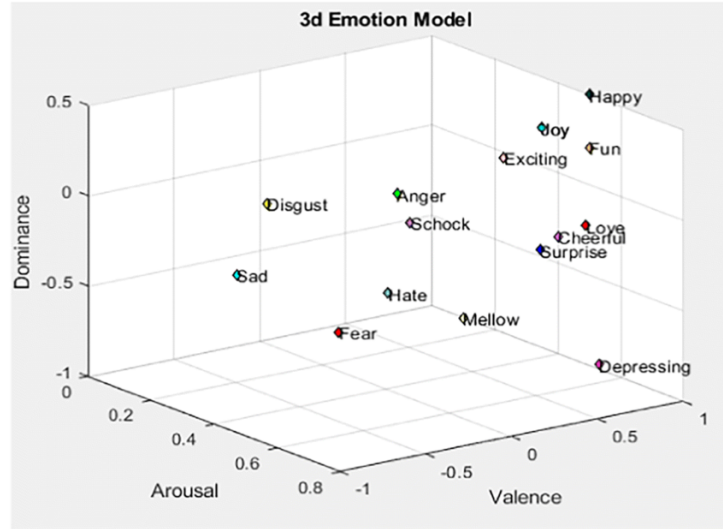


Figure 2.4: Mehrabian and Russell 3D model.

part of the body can also reveal emotional states. Some researchers suggest that the feet are the most honest part of the body, as their movements are often subconscious and harder to control. While the facial expression and the hand gestures can be manipulated to hide what one is feeling, the position and direction of the feet naturally tend to show what one is thinking or feeling. Feet pointing towards a person may indicate interest or engagement, while feet pointing away may suggest discomfort or a desire to leave the situation [36].

In this field, the automatic recognition of emotions is a challenging task that can be applied in various scenarios, such as healthcare, stress-monitoring, and security [37]. Similarly, body language may serve as a biomarker for early autism spectrum disorder diagnosis, as shown in [38]. Human movements and observation modes can also be used to assess the user's QoE and emotional state. Recent advancements in motion capture systems have enabled the collection of body joint orientations, body segment data, and facial expressions [39]. This data can be processed for automatic emotion recognition during gait [40] or other tasks, such as those studied in [41]. However, these approaches often require an optoelectronic system and a controlled environment where subjects wear reflective markers on body joints and specialized clothing, like T-shirts or gloves. This setup frequently causes discomfort and may alter natural movement and behavior. To address this issue, studies such as [42, 43] utilized a markerless motion analysis system for emotion recognition based on a deep learning approach that con-

siders facial expressions and body movements. Features like bounding volume, velocity, and acceleration [37] were employed to differentiate emotions. Moreover, research has shown that arm and upper body movements during walking, along with body space utilization, elbow angles, and spatial extension while sitting, can help predict the four basic emotions: happiness, sadness, anger, and fear [13].

The introduction of technological systems into emotion research has led to the development of databases containing various emotional expressions. In the literature, numerous emotion databases have been created, primarily composed of RGB images and videos. One widely used dataset is GEMEP [44], which includes 7,000 audio-video recordings of 18 emotions expressed by 10 professional actors. This dataset captures facial expressions and upper body movements recorded with multiple cameras. The FABO dataset [45], in contrast, comprises 206 videos focused on upper body features, combining facial expressions and body movements to capture 10 emotions. The HUMAINE database [46] includes recordings of 10 participants acting out 8 emotions—anger, despair, interest, pleasure, sadness, irritation, joy, and pride—with a camera positioned to capture movements from the front. The Theater database [47] contains two movie clips from *Death of a Salesman* by Arthur Miller, primarily focused on upper body and hand gestures. Meanwhile, the LIRIS-ACCEDE dataset [48] consists of 9,800 clips from 160 films, making it one of the largest annotated video databases. In [49], the authors introduced a dataset capturing emotional body expressions in everyday actions, featuring 11 actors expressing 8 emotions across 7 different actions, recorded using a motion capture system and four cameras. The IEMOCAP database [39] includes recordings of 10 actors with markers on their face, head, and hands, using a motion capture system to track positions during both scripted and spontaneous communication scenarios. A summary of these available datasets is provided in Table 2.1.

This literature review highlighted the need for the creation of a new database, featuring videos of non-professional actors to avoid stereotyped or exaggerated emotional expressions. Despite the range of existing datasets, this new database was designed to consider the entire body, including both upper and lower limbs, recognizing that body movements beyond facial expressions and upper body gestures are essential for a comprehensive analysis of emotions. The creation process and foundational concept of this database will be detailed in the next

chapter. This approach seeks to capture natural emotional expressions, enhancing the accuracy and reliability of emotion recognition systems and creating an automated system applicable across various contexts.

Table 2.1: Datasets in the literature and their main characteristics.

Name	Body parts	Emot	Subj	Samples	Devices	Side	Fps
GEMEP [44]	Face and upper body	18	10	>7000	Digital camera	Frontal, lateral	25
FABO [45]	Face and upper body	10	23	206	Digital camera	Frontal	15
HUMAINE [46]	Face and body	8	10	240	Camera	Frontal	25
THEATER [47]	Body	8	NA	258	Camera	NA	NA
LIRIS-ACCEDE [48]	Face and upper body	6	64	NA	Camera	Frontal	NA
EMILYA [49]	Body	8	11	7084	Motion capture system, cameras	Entire	NA
IEMOCAP [39]	Face and upper body	5	10	12 h	Motion capture system, cameras	Frontal	120
HEROES	Entire body	4	16	256	GoPro Cameras	Frontal	24

Chapter 3

HEROES: A Video-Based Human Emotion Recognition Database

Based on what has been introduced in Chapter 2, human movement and emotions are intrinsically interconnected elements of the human experience, and the relationship between physical movement and emotions has aroused great scientific interest. A central research question was formulated: “Is it possible to recognize a user’s emotional state and QoE through body language? Does the lower part of the body contribute to emotion recognition?”. As previously explained, the quality of user experience is influenced by a range of intrinsic and extrinsic factors, including individual preferences, emotional states, environmental conditions, and the specific interactions involved. These factors shape how users perceive and engage with digital systems, making it challenging to accurately assess their emotional states. To address this complexity, an initial system was developed to identify four basic emotions (happiness, boredom, disgust, and interest) using only body movements. This approach influences the understanding that emotions are often expressed through non-verbal cues, including posture, gestures, and movement patterns, which can be powerful indicators of a person’s internal state. By focusing only on body movements, the system aims to capture these emotional expressions without facial expressions or vocal intonations, providing a potentially non-intrusive and effective method for assessing user emotions across diverse contexts (Figure 3.1).

The review of the literature revealed that most databases consist of videos taken from professional actors who perform certain emotions and are focused on facial expressions and/or

upper limb movements. However, as previously mentioned, the lower part of the body can also be useful to understand the emotional human states. In this study, the goal was to create a new system capable of automatically recognizing emotions based only on limb movement and posture in real-time. To the best of current knowledge, no RGB database existed that would include the whole body (upper and lower parts), and the purpose was to create one from scratch. The HEROES database was thus proposed, consisting of RGB videos in which the entire body of the actor is visible. Moreover, while most of the existing databases were recorded using performances by professional actors guided by a director, this new database was built with the use of non-professional actors. The emotions (happiness, disgust, interest, and boredom) were selected to cover completely the high/low arousal and positive/negative valence space (see Table 3.1).

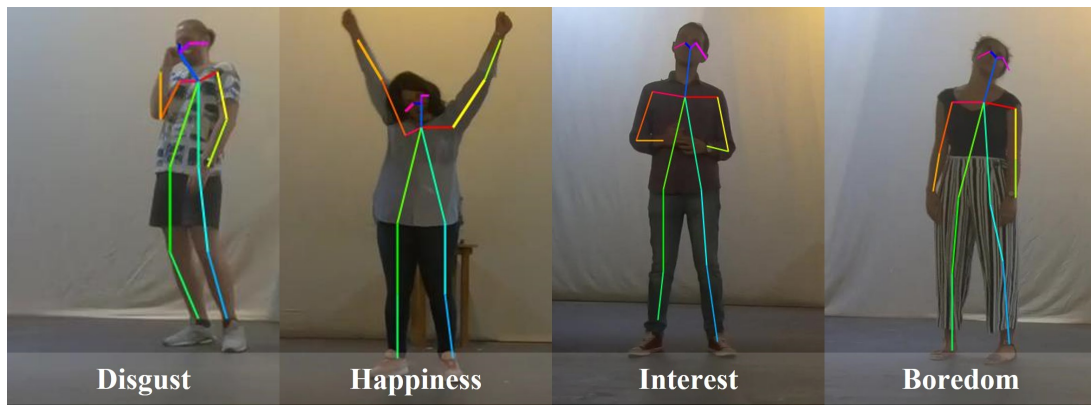


Figure 3.1: Example of human poses and corresponding emotions in HEROES.

Table 3.1: The emotions in the space valence-arousal.

	Positive valence	Negative valence
High arousal	Happiness	Disgust
Low arousal	Interest	Boredom

3.1 HEROES dataset creation

The HEROES dataset consists of videos of 16 non-professional actors (10 men and 6 women) with an average age of 27 years (± 2.8), ranging from 25 to 33. Most participants were recruited from an Italian amateur theater group and had at least 5 years of theater experience, although none had attended professional theater courses. Non-professional actors were selected to avoid stereotyped actions and to gather various types of movements associated with the same emotion [41]. All actors signed informed consent forms, where they were advised on the research objectives and agreed to the use of their recordings for research purposes. To encourage natural emotional expressions, a scenario-based approach was used, similar to those in previous studies [50] and [44]. This method was chosen for its simplicity, and it allowed for spontaneity. Each actor was provided with two scenarios representing different situations in which they were asked to express four target emotions (see Table 3.2): Scenario A represents a visit to a museum, and Scenario B represents the interaction with a person. The actors received the scenario list (Table 3.2) one day before the recording session and were instructed to act as they would in real-life situations. During the experiment, they reviewed the scenarios again before expressing each emotion, imagining the situations, and moving as if they were experiencing them. The actors were free to choose which emotion to express first and were encouraged to move and express their emotions naturally, without any restrictions or guidance. Two scenarios were selected to increase the variability of movements for the same emotion and cover different actions. Scenario A was always set in a museum context, where the actors were inclined to walk and move around the space while expressing their emotions. In Scenario B, a conversational situation was designed where participants interacted with another person, where the participants tended to maintain a static pose, but used gestures and postures to express their emotions. Then, four experts annotated the videos, assessing the perceived emotion, valence, and arousal levels. This approach aligns with existing literature on non-verbal communication and emotion recognition [45], [37]. The annotations were collected using a Matlab application that allowed the reviewers to watch each video an unlimited number of times and allowed the annotators to select the emotion (happiness, boredom, interest, or disgust), arousal, and valence for each video. The actors' faces were blurred in the videos, so the

Table 3.2: Expressed emotions in the two considered scenarios.

Emotions	Scenario A	Scenario B
Interest	I am in a museum and suddenly a painting captures my attention and I start observing it in a very interested way.	We have organized a dinner with friends. One of them tells about a journey from which he came back recently. His/her story is very interesting and I ask him to tell me more.
Boredom	My art teacher brings us to visit a museum in which the paintings are very similar. I wander among them really bored. I cannot wait for this visit to end.	I am walking in the street and I meet a friend who starts talking about the same things he talks about every time we meet. I am not interested, and I do not hide my boredom.
Happiness	Finally there will be an exhibition of my favorite artist in my city. It is incredible, I am excited! As soon as I enter in the museum my happiness is really irrepressible. I have been waiting for this moment for years.	It is incredible! I just won the lottery! I am so happy and I cannot contain the joy, I have to tell everyone! I meet a friend and tell him/her about it.
Disgust	I am in a museum in which the paintings/photos depict bloody scenes of accidents and catastrophic events which disgust me.	I meet some friends and I tell them what happened yesterday. I got home and I smelled something nauseating from the kitchen. There I found a rotting dead animal on the floor covered by ants. It was really disgusting!

annotators assessed the emotions only through body movements, with arousal and valence rated independently of facial expressions. It took an average of one hour for each observer to annotate the videos at a mean accuracy of 79%. This annotation process aimed to obtain an independent evaluation of the videos included in the dataset. The videos were recorded in a theater under fixed lighting conditions. To capture multiple perspectives, four GoPro cameras were positioned on each side of the stage. The only additional information given to the actors was on the location of the starting point and the area associated with the field of view of the cameras. They were instructed to begin and end each performance in a neutral pose (with no visible signs of muscle activation). Each actor expressed each of the four emotions five times for both scenarios, performing as they would in real-life situations without any time

constraints. With four emotions and two scenarios, 40 videos per actor were recorded from each camera. For the initial release of the HEROES dataset, a total of 256 videos were selected. More in detail, following the approach performed in [33], only videos from the frontal view camera were chosen. Four videos per actor were chosen for each emotion, with two videos corresponding to Scenario A and two to Scenario B. As previously specified, there were no time constraints for the emotional expressions, so video segments corresponding to the moments of peak emotional intensity were selected following state-of-the-art methods. The analysis by Wallbott [35] was referenced for boredom and happiness, the guidelines from [34] were followed for interest, and the study by [37] was considered for disgust. In the videos, at least one of the characteristics reported in the Table 3.3 was present. After the cropping procedure, 256 videos whose length varies between 4 and 6 seconds were obtained. The other characteristics of the dataset have been reported in Table 3.4. To assess the performance of the proposed approach, the dataset has been split into 80-20% train-test sets.

3.2 Body joint extraction

One of the primary applications for this dataset is the automatic identification and recognition of emotions across various fields, such as security, rehabilitation, and entertainment. To support this objective, a set of foundational features has been provided with the dataset, enabling the initial stages of emotion recognition and classification. This study represents a preliminary phase, designed to test the feasibility and effectiveness of using specific body movements to recognize a limited range of basic emotions. Specifically, the positions and movements of human body joints, represented as the body pose at the i -th frame $\mathbf{x}(t_i)$, were extracted using the deep learning model OpenPose, as proposed in [43]. The performance of OpenPose was evaluated by means of the mean Average Precision (mAP) metric, which considers the Intersection-over-Union (IoU) between the predicted body area and the ground truth. For 2D human pose estimation, OpenPose achieved a mAP of 79% on the MPII dataset [51], representing state-of-the-art performance. However, there remained a possibility of undetected body joints due to occlusions or noise. In this work, two strategies were proposed to mitigate the issue of missing body joints. If joints were missing in the first frame, their locations were esti-

Table 3.3: Examples frames of the dataset in which non-professional actors express the emotions.

Happiness	Disgust
 The trunk is in an upright position, and the head is tilted up. The arms are opened and perform expansive and fast movements. The legs are open or parallel.	 The body is shifted, and the trunk orientation changes side to side. One hand is up or on the mouth or covering the neck, or close to the body. Backing.
Interest	Boredom
 The forward movements are characterized by opening arms with low velocity and low dynamics. The trunk is in an upright position, and the feet point towards the object of interest.	 The head is tilted up and rotated slowly side to side while the trunk appears in a bowing posture. The movements are slow and not expansive.

mated by calculating the mean values of the image coordinates across all frames of the video. If the joints were missing in subsequent frames, the positions from the previous frame were used to estimate their locations.

The human joints considered in this study were 14 in total: nose, neck, shoulders, elbows,

Table 3.4: Video duration in seconds for emotions and actors of the HEROES dataset.

Parameter	Emotion				Actor															
	H	D	I	B	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16
Duration (s)	364	371	362	367	88	91	89	90	89	86	91	93	90	91	93	98	94	93	94	94
H: Happiness. D: Disgust. I: Interest, B: Boredom. 01 to 16: Actor numbers.																				

wrists, hips, knees, and ankles. More precisely, let $V = \{I_1, I_2, \dots, I_n\}$ represented a generic videoclip with n frames I with $H \times W$ pixels. The high-level features of the i -th frame (S_i) were defined as the set of characteristics related to human body movement through time. In this work, high-level features include the space occupied by the body during movement, the velocity and acceleration of the body joints, the curvature of arm movements, the triangular area formed by the wrists, elbows, and shoulders, and the distance between the nose and neck. Additionally, a low-level analysis was conducted using the raw image coordinates of the body joints. Statistical parameters were computed to enhance the information collected on each joint to compose the low-level features of the i -th frame (F_i). Figure 3.2 illustrates the block diagram of the feature extraction model and Table 3.5 reports the considered features.

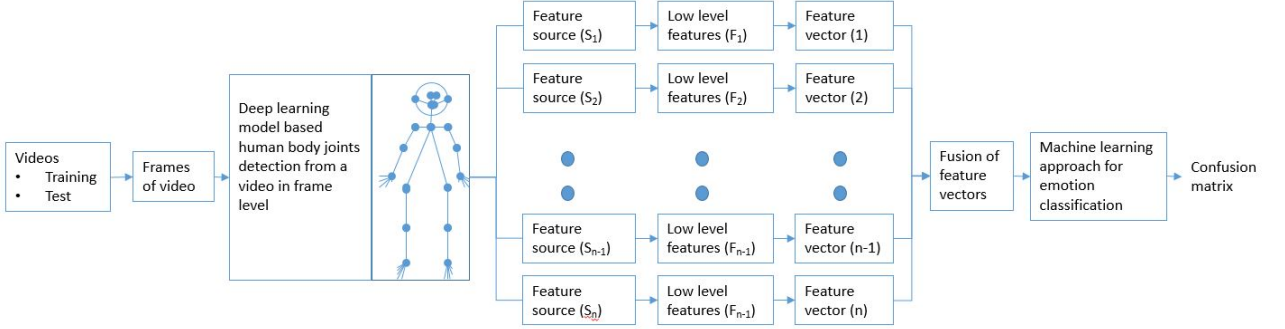


Figure 3.2: Block diagram of the feature extraction pipeline. The fusion block is responsible for concatenating all the low-level features.

A deep learning model capable of estimating real-time body poses from 2D images, as proposed in [43], inspired the extraction of different body joints from each frame of the video. Body joint features are extracted on a frame-by-frame basis. Since the duration of the input videos varies, 120 frames from each video are selected at equal intervals to collect body joint data. Those body joint coordinates are recorded for defining the feature source and eventually for extracting the required emotion-related features of a particular video.

Table 3.5: High (S_i) and low level (F_i) features for a generic i -th frame.

Features	
S_n	F_n
Surrounding space	Mean
Left arm curvature	Standard deviation
Right arm curvature	Variance
Velocity of joints	Skewness
Acceleration of joints	Kurtosis
Left hand triangular area	Third moment
Right hand triangular area	Fourth moment
Distance between nose and neck	Sensitivity factor
	Max and min surrounding space ratio

3.2.1 High level features

In the literature, different geometric and kinematic features were proposed to define the body motion for the type of emotion [13]. In the following, the list of the features considered in this work, along with their descriptions, is reported.

Space occupied by the human body

This feature measured the space occupied by a person during movement. To achieve this, a rectangular box enclosing all considered human body joints was defined. Let, x_{min} , x_{max} , y_{min} , and y_{max} be the image coordinates of the rectangular box. The surrounding space of the human body (s_A) was defined as:

$$s_A = (x_{max} - x_{min}) \cdot (y_{max} - y_{min}). \quad (3.1)$$

This value was computed for each frame I of a generic video clip V .

Velocity and acceleration of the body joints

The velocity and acceleration of a joint were associated to the speed of the motion. The formula $\mathcal{X} = \{\mathbf{x}(t_1), \mathbf{x}(t_2), \dots, \mathbf{x}(t_n)\}$ denoted the dynamic of the body joints, described as n consecutive poses $\mathbf{x}(t_n)$. As described in [52], the general expression of the velocity for both

coordinates $\mathbf{v}^k(t_i) = [v_x^k(t_i), v_y^k(t_i)]$ at time t_i can be expressed as:

$$\mathbf{v}^k(t_i) = \frac{\mathcal{X}^k(t_{i+1}) - \mathcal{X}^k(t_{i-1})}{2\Delta t}, \quad (3.2)$$

where $\mathcal{X}^k(t_i)$ was the coordinates of the k -th joint at time t_i and Δt was the time difference between consecutive poses.

Finally, the magnitude of velocity for the k -th joint was evaluated as:

$$\|\mathbf{v}^k(t_i)\| = \sqrt{(v_x^k(t_i))^2 + (v_y^k(t_i))^2}. \quad (3.3)$$

The acceleration of the body joints $\mathbf{a}^k(t_i)$ and its magnitude $\|\mathbf{a}^k(t_i)\|$ were evaluated by exploiting the same procedure but using the velocity vectors $\mathbf{v}^k(t_{i+1})$ and $\mathbf{v}^k(t_{i-1})$. The considered moving body joints were wrists, elbows, shoulders, and ankles.

Curvature of left and right arms

This feature considered the trajectory of the arms' movement. According to [13], the curvature $c^k(t_i)$ of the k -th joint at time t_i can be computed as:

$$c^k(t_i) = \frac{\|\mathbf{v}^k(t_i) \cdot \mathbf{a}^k(t_i)\|}{(\sqrt{(v_x^k(t_i))^2 + (v_y^k(t_i))^2})^3}. \quad (3.4)$$

Triangular area covered by wrist, elbow, and shoulders of left and right hands

Hand movements were very characteristic of expressing emotions. Therefore, the space occupied by a hand was calculated by evaluating the area of a triangle formed by the wrist, elbow, and shoulder as vertices.

Distance between nose and neck

Head movements changed the distance between the nose and neck, denoted as d_{nn} . It was calculated as:

$$d_{nn} = \sqrt{(x_{neck} - x_{nose})^2 + (y_{neck} - y_{nose})^2}, \quad (3.5)$$

where x_{neck} , x_{nose} , y_{neck} , and y_{nose} were the x and y coordinates of the neck and the nose.

3.2.2 Low level features

The low level features (F_i) were related to the temporal evolution of the analyzed joints. They were selected to highlight how emotions influenced the spatio-temporal evolution of human movement. The mean (μ) and standard deviation (σ) were the most commonly extracted low-level features. Similarly, variance, skewness, kurtosis, and third and fourth-order moments were calculated for each body joint. Additionally, the sensitivity factor ratio ($\frac{\sigma}{\mu}$) was determined for the minimum and maximum space occupied by the body area. The HEROES dataset included 22 distinct low-level features. Both low and high-level features were provided along with the dataset, as listed in Table 3.5.

3.2.3 Analysis of Annotation Quality and Inter-Rater Consistency

The HEROES dataset was independently annotated by four raters, who assigned both categorical emotion labels (e.g., Happiness, Boredom, Interest, Disgust) and numerical values for valence and arousal on a discrete scale. Each of the 256 video clips was evaluated by all four annotators. The agreement between the ground truth label (“GT Emotion”) and the assigned label (“Ann. Emotion”) was calculated as a binary variable (“Agreement”). Descriptive analysis shows a generally high level of consistency across annotators, with individual accuracy rates of 78.5%, 81.6%, 76.6%, and 79.3%

To assess the reliability of the manual annotations in the HEROES dataset, two complementary statistical analyses were conducted: Cohen’s Kappa to evaluate inter-rater agreement on categorical emotion labels, and one-way ANOVA to test whether arousal and valence scores vary significantly across emotion categories.

Inter-Annotator Agreement. Each of the 256 video clips in the dataset was independently annotated by four raters. Cohen’s Kappa (κ) was computed for all six possible annotator pairs to quantify agreement beyond chance:

- Annotator 1 vs Annotator 2: $\kappa = 0.73$
- Annotator 1 vs Annotator 3: $\kappa = 0.70$

- Annotator 1 vs Annotator 4: $\kappa = 0.69$
- Annotator 2 vs Annotator 3: $\kappa = 0.75$
- Annotator 2 vs Annotator 4: $\kappa = 0.72$
- Annotator 3 vs Annotator 4: $\kappa = 0.74$

According to the guidelines by Landis and Koch [53], these values indicate substantial agreement between raters, confirming a high degree of consistency in the annotation process.

ANOVA on Arousal and Valence. A one-way ANOVA was performed to determine whether arousal and valence scores differed significantly between emotion categories. The analysis was based on the average scores across annotators for each video.

The results revealed statistically significant effects for both dimensions:

- Valence: $F(3, N) = 18.45, p < 0.001$
- Arousal: $F(3, N) = 12.97, p < 0.001$

Post-hoc Tukey tests showed that valence was significantly higher for “Happiness” than for all other emotions, and arousal levels for “Disgust” and “Happiness” were significantly higher than for “Boredom” and “Interest”. These statistical findings reinforce the validity of the annotations. The high inter-rater agreement and the significant association between dimensional ratings and emotion categories suggest that the labeling process is both consistent and meaningful. This adds robustness to the dataset and strengthens its use in downstream affective computing tasks.

3.3 Classifications results on HEROES

The implementation of the proposed model and its performance evaluation were conducted in the Matlab environment, utilizing the deep learning and ML toolboxes. The body pose estimation model was developed following the reference provided by MathWorks, based on [43]. In the dataset, the video durations ranged from 4 to 6 seconds. For each video, 120 frames,

evenly spaced, were selected to extract emotion-related information from human body movements. The surrounding space occupied by the body, along with data from individual body joints, was gathered for all 256 input videos as a source of raw spatial and temporal information. Various feature sources were defined, and low-level features were derived from these sources, as described in Section 3.2. ML models are widely used for classification tasks, and in this case, the dataset included four categories of emotions. Therefore, a multiclass classification technique was required. The extracted features from the input videos were divided into training and test datasets using an 80:20 ratio and to evaluate the model's performance, the k-Nearest Neighbour (k-NN) classifier [54] was employed. The key parameters considered for defining the k-NN classifier included the number of nearest neighbors, the neighborhood search method (NS method), and the distance metric¹. To evaluate the performance of the feature set, the low-level features of the hands and feet were grouped by concatenating statistics related to velocity, hand curvatures, and acceleration. The feature sets used for the human emotion recognition task were denoted as $\mathbf{f}_{body}$ and $\mathbf{f}_{upper}$. The set $\mathbf{f}_{body}$ included features related to the velocity and acceleration of the arm and ankle joints, curvature features of both the left and right arms, the surrounding space occupied by the body, and the distance between the nose and neck. Similarly, $\mathbf{f}_{upper}$ represented a reduced version of $\mathbf{f}_{body}$ obtained by excluding the velocity and acceleration of ankle joints of the feet and the surrounding space of the body. The accuracy metric was used to evaluate the performance of the proposed feature sets. The hyperparameters used to configure the k-NN classifier were: 7 nearest neighbors and the city-block distance metric.

The confusion matrices for the HEROES dataset using the previously mentioned feature sets are presented in Figure 3.3. The accuracy of the proposed features set $\mathbf{f}_{body}$ was 94.11%, which was higher than its counterpart $\mathbf{f}_{upper}$ (84.31%). The lower classification was attributed to the exclusion of features related to the feet and the surrounding space (s_A). Moreover, the difference in the performances demonstrates the importance of lower body joints information in human emotion recognition. From the emotion classification results on HEROES dataset, it was confirmed that the information from human body joints serves as a valuable feature source for the task of human emotion recognition. This system, while currently focused on

¹<https://www.mathworks.com/help/stats/fitcknn.html>

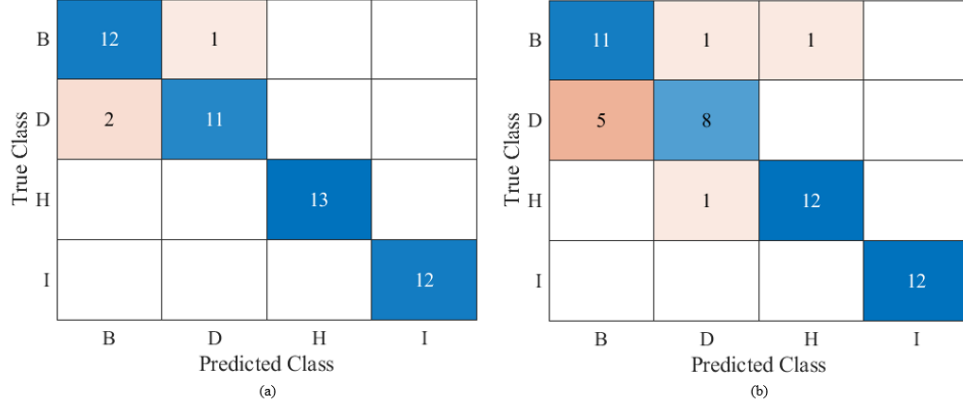


Figure 3.3: Confusion matrices for HEROES dataset with: (a) whole body joints information (f_{body}), and (b) with upper body joints information only (f_{upper}). Here, B: Boredom, D: Disgust, H: Happiness, and I: Interest.

specific emotions, offers adaptability for application in diverse contexts by adjusting the emotional categories involved and changing the features that are most effective for distinguishing particular attitudes or emotional responses. For instance, in therapeutic environments, the system could focus on recognizing stress or relaxation indicators, while in entertainment, it might prioritize identifying engagement or enjoyment. In future developments, the dataset will be expanded to include additional characteristics and a wider variety of emotions. These enhancements aim to refine the system's accuracy and its application scope, allowing it to identify more emotional states and adapt to a greater diversity of contexts and user needs.

Chapter 4

Virtual Reality application

As described in the previous chapter, recent advancements in emotion recognition technology, in particular as demonstrated in the emotion classification results from the HEROES dataset, have shown that information derived from human body joints was a valuable feature for recognizing emotions. This study could have interesting applications in various fields, including that of VR rehabilitation, where the integration of emotion recognition could enhance the adaptability and personalization of the therapy. In a rehabilitation context, emotions and psychological states are closely linked to patient motivation, engagement, and overall recovery outcomes, so it is important to consider their QoE. By incorporating systems that can recognize and classify emotions based on body movements, such as those identified in the HEROES dataset, VR applications could be adjusted in real time to respond to the emotional state of the patient to improve their experience. For example, if the system detects frustration or discouragement, it could modify the VR environment or tasks to lower difficulty levels or introduce more motivating elements, ensuring the patient remains engaged. In any case, before implementing this automatic system in a rehabilitation context, it was essential to design and validate a dedicated VR application for rehabilitation purposes. This preliminary step ensured that the VR environment could meet specific therapeutic requirements, such as providing appropriate exercises, accommodating various levels of physical ability, and offering a safe and controlled setting for users. Additionally, validation of the VR application allowed for adjustments to be made based on user feedback and clinical input, ensuring that the system could accurately capture relevant movements and deliver an engaging, effective rehabilita-

tion experience. Only after this validation process could the automatic system be confidently integrated, enabling the precise monitoring of body movements and potentially enhancing patient outcomes through feedback in real time. Based on this idea, during these years of doctoral research, a collaboration was established with the Fondazione Policlinico Universitario Campus Bio-Medico in Rome. The project required close cooperation between orthopedics, physiotherapists, and engineers to develop a VR application aimed at supporting the rehabilitation process for patients with shoulder conditions. A VR game was created with multiple levels designed to achieve specific rehabilitation objectives and improve the Range of Movement (ROM), while also providing an engaging experience for the patients. A key focus of the collaboration was to assess the QoE from both physiotherapists and patients, as well as evaluate the effectiveness of the VR game in shoulder rehabilitation. Anyway, before starting the project for rehabilitation purposes, it was necessary to validate the device, a low-cost VR headset known as Oculus Quest 2 (OQ2), typically used for entertainment. This validation was essential because the device was primarily designed for recreational use, rather than clinical or rehabilitation settings. Recent developments in VR have been applied in medical and rehabilitation fields, with various devices being used to detect and monitor human movement [55], [56]. Gaming technologies, such as Microsoft Kinect and Nintendo Wii, have been incorporated in the patient's rehabilitation protocols, showing positive outcomes [57], [58], [59], [60]. Furthermore, the use of VR systems increased rapidly as they became more affordable [61]. In VR environments, users are fully immersed in virtual spaces and could interact with virtual objects in simulated settings using controllers [61]. These technologies can also track human movements and could engage the patient to increase their adherence to the rehabilitation protocol [62]. The immersive nature of these devices allows users to experience a variety of virtual contexts, encouraging a more interactive rehabilitation experience. This approach has the potential to make long-term rehabilitation more enjoyable and engaging [61]. VR has already been used in the recovery of various clinical conditions [63], [64], [65], [66]. In the orthopedic field, rehabilitation plays a critical role after surgery or traumatic events for the recovery the compromised functionality [62]. Rotator Cuff (RC) tears are a common condition among shoulder musculoskeletal disorders [67], [68]. In [69], a new approach for the right shoulder's RC was introduced, utilizing the Oculus Rift DK2 and a motion tracking

system (Intel RealSense) to design a game that incorporates elevation and abduction movements. The system was evaluated by four experts, who determined that the application could be suitable for upper limb rehabilitation. Generally, the clinical condition of patients' shoulders is assessed using qualitative methods, such as questionnaires and clinical scales specific to the pathology [70], [71], [72]. In contrast, these new devices are able to give a quantitative evaluation of patient movements, providing real time feedback for correcting actions. Performance evaluation of VR systems for motion tracking is a fundamental step toward their application both in research and clinical purposes [73]. This assessment can be conducted by analyzing specific quality metrics, such as translational and rotational accuracy. Clinical and digital goniometers are commonly employed to measure the range of motion ROM of shoulder joints [74]. However, the Microsoft Kinect system has been identified as a potential alternative. Studies [74], [75], have compared Kinect's measurements of shoulder functional movements compared with the measure given by clinical and digital goniometers and optoelectronic system, demonstrating good reliability and validity of the device. A recent systematic review [76] highlighted moderate to good levels of intra- and inter-rater reliability in measuring shoulder ROM using various devices, such as smartphone applications, digital inclinometers/goniometers, inertial sensors, and Kinect. Other VR devices, including Oculus Rift S and Oculus Touch, have also been evaluated for rehabilitation purposes, with comparisons made to optoelectronic systems [77], [78]. These evaluations revealed significant accuracy and precision of the devices in several translations and rotations. The controllers of the OQ2 system work as peripheral accessories and act as end-effectors of the upper limb kinematic chain. The pose, consisting of position and orientation in 3D space, is fully defined by six degrees of freedom (Degree of Freedom (DOF)): three for translations and three for rotations. The aim of the first study conducted in collaboration with the Campus Bio-Medico of Rome was to assess, for the first time, the accuracy of the VR system in acquiring the translational and rotational displacements. To achieve this, a custom application was developed using Unity and the Oculus Integration Package. An optoelectronic system was used to record translational and rotational displacements to compare its measurements with those obtained from the OQ2.

4.1 Device Validation

For this purpose, an OQ2 HMD and two Oculus Touch controllers (Figure 4.1) were utilized in the study. The main technical features of the OQ2 VR system are summarized in Table 4.1. A custom application was developed using Unity (Unity, v. 2020.17) to record and store the position and orientation of the controllers at a sampling rate of 200 Hz. A QualisysTM Optical Motion Capture (OMC) system (Qualisys AB, Gothenburg, Sweden) was employed as the reference system. This system was provided with two Miquis Video cameras (sampling rate of 25 Hz) and ten Miquis M3 cameras (sampling rate of 100 Hz). According to the manufacturer, the Miquis M3 cameras offer a 3D resolution of 0.11 mm and can capture movements at distances of up to 15 meters [79]. The two Miquis video were used to have synchronized videos avail-



Figure 4.1: OQ2: from the left, the HMD, the left controller (controller A), and right controller (controller B).

Table 4.1: Main features of OQ2.

Memory	6 GB
Storage	256 GB
Display	LCD 1832 x 1920 per eye @ 72-120 Hz
Graphics	Adreno 650 (~ 1.2 TFLOPS)
Sound	2 built in speaker/3.5 mm headphone jack
Input	6DOF Inside-out tracking through 4 built-in cameras and 2 controllers with accelerometers and gyroscopes Optional: QWERTY keyboard (via Bluetooth)
Controller Input	Oculus Touch
Camera	4 infrared cameras
Mass	503g

able for use in subsequent analyses as a reference. The OMC system was utilized to track the pose (position and orientation) of the geometric center of the rigid body, which was defined using the Qualisys Track Manager (QTM) software (v. 2021.2, Build 6720). In particular, five photo-reflective markers (8 mm in diameter) were placed on controller A as shown in Figure 4.2a, and Figure 4.2b. One marker was positioned on controller B (Figure 4.2c) and was used to synchronize the signals between the OQ2 and the OMC systems. Custom 3D support was printed using polylactic acid filaments using the 3D printer Ultimaker S2+ (Figure 4.2d). This support was specifically designed and sized to hold controller A during the experimental trials. Figures 4.2e and Figure 4.2f show two structures designed for allowing translational and rotational movements. All measurements were performed by securing the left-hand controller

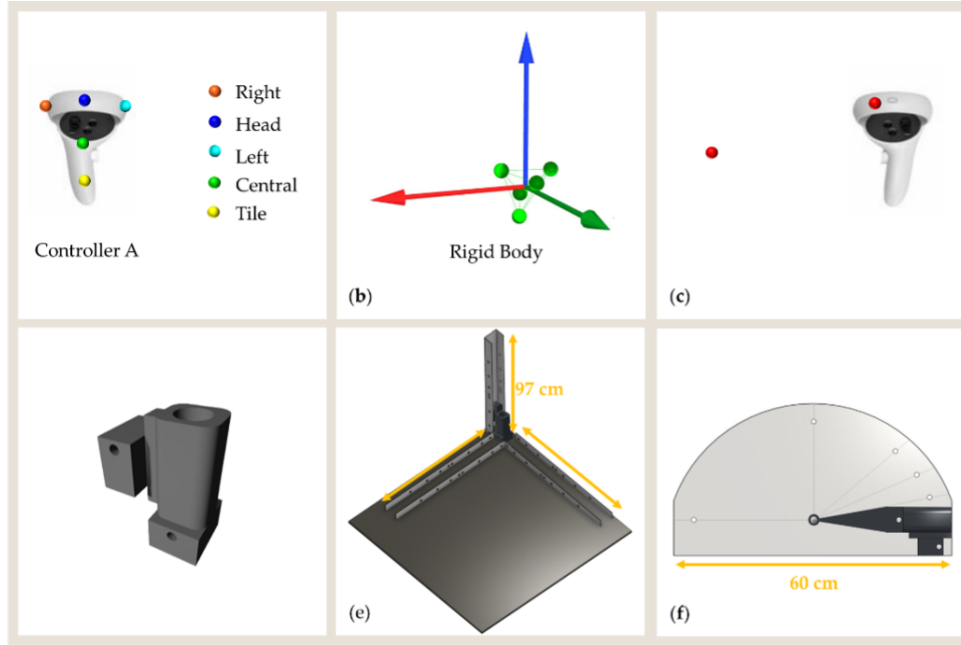


Figure 4.2: (a) Controller A and markers configuration used to define the (b) rigid body. (c) Controller B and the position of the synchronization marker. Controller A was allocated in a (d) custom-made support designed to be easily moved on the structures for (e) translational and (f) rotational movements.

(controller A) in a handmade holder, while controller B was held by an operator, referred to as Healthy Volunteer (HV), during the experiments. The HV did an initial calibration before each test to define the 3D virtual space that delineated the safe movement area, i.e., the play area. The HV was instructed to make fast movements with controller B (i.e., the one with

the synchronization marker on it) at the beginning and end of each test, to guarantee correct synchronization of the signals.

4.1.1 Movements along the axis

Traslational Movements

Measurements were conducted along each of the three 3D axes (x-width, y-depth, and z-height). A custom structure was used to move controller A along each axis (Figure 4.2e). Since the OQ2 was chosen for possible further application in shoulder rehabilitation, the minimum (approximately 200 mm) and maximum (approximately 700 mm) distances analyzed corresponded to the forearm length of a 5th percentile female and the upper limb length of a 95th percentile male, respectively, excluding the hand length [80], [81]. All the translational displacements started with the HMD at 200 mm from the origin (P_0) of the reference system until the maximum distance ($P_{11} = 705$ mm) at variable increments, as shown in Figure 4.3. At position P_0 , controller A remained stationary for approximately 20 s, and at each subsequent position P_i ($i = 1, \dots, 11$), it was held for about 10 s. At the positions P_1 (5 mm), P_6 (255 mm), and P_{11} (505 mm), an increment of 5 mm related to the previous positions ($P_0 - 0$ mm, $P_5 - 250$ mm, and $P_{10} - 500$ mm, respectively) was assessed to evaluate the OQ2's ability to detect small variations at increasing distances from the HMD. Acquisitions were repeated seven times along each axis, for a total of $N = 252$ static samples (12 positions x 7 repetitions x 3 axes).

Rotational Movements

A custom structure was designed to rotate controller A around each axis (x—roll, y—pitch, and z—yaw) as shown in Figure 4.4. The selected rotation range extended from 0° to 180° to cover the upper limb's full ROM [82]. All rotational movements were performed with the HMD positioned approximately 700 mm away from controller A during all the rotational displacements in the range 0° - 180° . Controller A executed eight rotational displacements (R_i , with $i = 0, \dots, 7$), repeating each task seven times and returning to the starting position R_0 after each repetition, for a total of 14 rotational displacements. At each R_i , controller A remained stationary for 5 s. A total of $N = 294$ (14 rotational displacements x 7 repetitions x 3 axes) static samples were

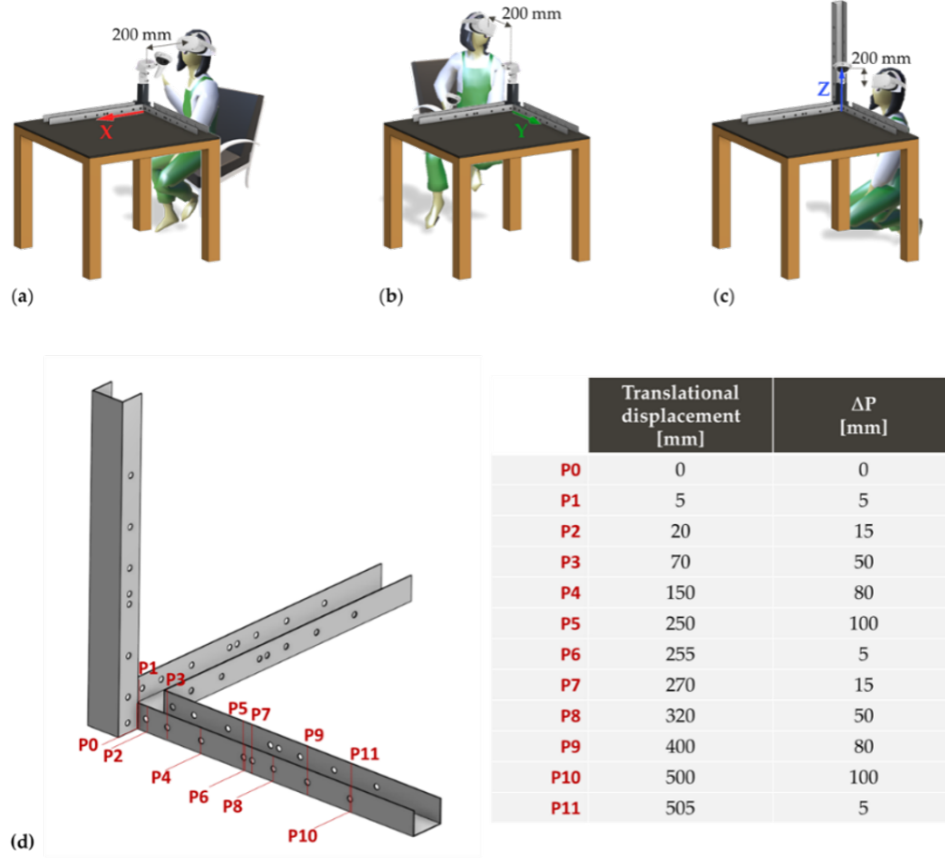


Figure 4.3: Acquisition configuration of translational displacements along the (a) x-axis, (b) y-axis, and (c) z-axis, starting at a distance of about 200 mm from the reference position P0; (d) representation of translational displacements P_i ($i = 1, \dots, 11$) and variations (ΔP , i.e., differences between translational displacements i^{th} and i^{th-1}).

acquired.

4.1.2 Data Analysis

Preprocessing (labeling and gap-filling) of marker trajectories was performed using QTM. An accurate analysis of the recorded videos allowed the identification of the start and end events of the i -th positions. As previously mentioned, the pose of controller A was tracked with the OMC by defining a rigid body in QTM, considering the evaluation of six DOF (i.e., translations along and rotations around the three axes). Data analyzed in QTM were processed using MATLAB. Since the data acquisition systems had different sampling rates (Qualisys - 100 Hz and

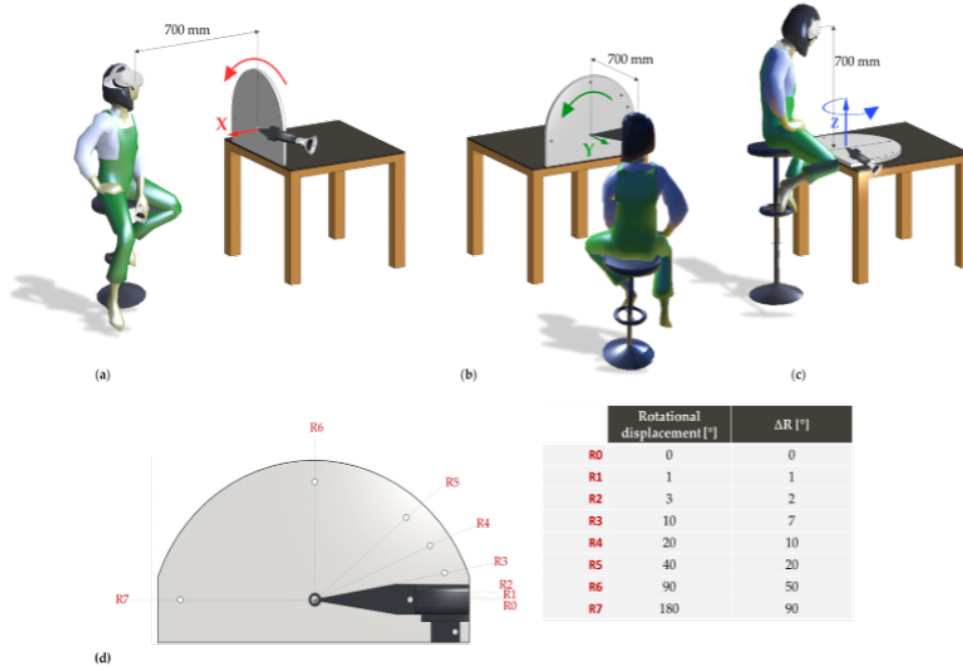


Figure 4.4: Acquisition configuration of rotational displacements around the (a) x-axis, (b) y-axis, and (c) z-axis, starting at a distance of about 700 mm from the reference position R0; (d) representation of rotational displacements R_i ($i = 1, \dots, 11$) and variations (ΔR , i.e., differences between rotational displacements i^{th} and i^{th-1}).

OQ2 - 200 Hz), resampling at the minimum sampling rate was performed. Additionally, because Unity uses a left-handed coordinate system and Qualisys a right-handed one, controller data were converted to match the OMC coordinate system. Moreover, the orientation of the controller A was reported as a quaternion in Unity, so a conversion from quaternion to rotation matrix was performed. A quick upper limb elevation movement was used as a common event to synchronize the test system (i.e., OQ2) and the reference one (i.e., Qualisys™). Synchronization points were determined by identifying the maximum values of the trajectories of the Sync marker at the beginning and the end of all the trials. These peaks corresponded to the highest positions reached by controller B during the rapid arm elevation performed by the HV. After signal synchronization, both translation and rotation signals were windowed to select 3 s between two consecutive events. The selected windows of each position and rotation were defined assuming a steady-state condition reached by controller A, free from artifacts caused by movement required to reposition the controller.

Translational Accuracy

The data were averaged over the 3 seconds of recording to determine the mean positions measured at each translational displacement performed. The absolute positional error e_P were calculated as follows:

$$e_P = | p^Q - p^C |, \quad (4.1)$$

where p^Q was the position measured by the Qualisys OMC system and p^C was the position measured by the controller. The positional percentage error $e_{P,\%}$, was evaluated as follows:

$$e_{P,\%} = \frac{e_P}{p^Q} \cdot 100. \quad (4.2)$$

The error e^Δ corresponding to the smaller step size (i.e., 5 mm) performed at P0 (0 mm), P5 (250 mm), and P10 (500 mm) was computed as follows:

$$e^\Delta = \Delta Q - \Delta C, \quad (4.3)$$

where ΔQ and ΔC were the steps measured by the Qualisys OMC system and the OQ2, respectively, between the position P_{i+1} and P_i .

Rotational Accuracy

The mean rotation matrix was calculated for each rotational displacement achieved by both the OQ2 and OMC systems. The rotation matrix M , relating the starting ($R_{0,i}$) and ending (R_i^T) rotation matrices, was computed as follows [77]:

$$M = R_{0,i} R_i^T, \quad i = 1, \dots, 7. \quad (4.4)$$

The final rotation angle Θ was computed as follows:

$$\text{tr} M = 1 + 2 \cos \Theta \quad (4.5)$$

then,

$$\Theta = \arccos \left(\frac{\text{tr} M - 1}{2} \right). \quad (4.6)$$

The absolute rotational error e_R was calculated as follows:

$$e_R = | \Theta^C - \Theta^Q |, \quad (4.7)$$

where Θ^Q was the rotation measured by the Qualisys OMC system and Θ^C was the rotation angle measured by controller A. The rotational percentage error $e_{R,\%}$ was evaluated as follows:

$$e_{R,\%} = \frac{e_R}{\Theta^Q} \cdot 100. \quad (4.8)$$

A Bland-Altman analysis was conducted for both translational and rotational displacements to evaluate the level of agreement between the two measurement systems. The Bland-Altman plots display the average of the two compared measures on the x-axis and their difference on the y-axis. This visualization allows us to assess agreement and potential systematic bias. Limits of agreement ($\text{mean} \pm 1.96 \cdot \text{SD}$) are included to identify outliers or inconsistencies. The analysis defined the 95% Limits of Agreement (LOA), representing the range within which 95% of the differences between the two systems are expected to fall [83]. LOA were expressed as Mean of Difference (MOD) $\pm 1.96 \cdot \text{SD}$ where MOD represents the mean of the difference and SD the standard deviation of the difference between the Qualisys OMC and OQ2. In a Bland-Altman plot, the x-axis represents the mean of the two measurements being compared (in this case, the values obtained from the Oculus Quest 2 and the optoelectronic reference system), while the y-axis shows the difference between the two measurements for each data point. This graphical method allows for the identification of systematic bias and the evaluation of agreement between the two systems.

4.1.3 Methodological Framework for Motion Accuracy Assessment

To evaluate the tracking accuracy of the OQ2, a comparative analysis was performed against a gold-standard optoelectronic motion capture system. The aim was to assess the system's precision in detecting translational and rotational movements under controlled conditions.

Experimental Setup. A rigid 3D-printed support was used to perform incremental movements along the three Cartesian axes (X, Y, Z) and to apply controlled angular rotations around each axis. The OQ2 headset was securely mounted on the support, and its internal tracking data were recorded at each step. Simultaneously, the reference system captured the actual displacement with millimeter and sub-degree precision.

Data Acquisition. Movements were performed in steps for translations and for rotations. For each condition, both systems registered the spatial coordinates and/or orientation angles. At each step, the experiment was repeated three times to account for variability and improve statistical robustness.

Error Computation. For each measurement, the absolute error was computed as:

$$\text{Error} = |x_{\text{OQ2}} - x_{\text{ref}}|$$

for translations, and

$$\text{Error}_\theta = |\theta_{\text{OQ2}} - \theta_{\text{ref}}|$$

for rotations. Errors were calculated per axis and aggregated across trials. Additionally, percentage error and mean absolute error (Mean Absolute Error (MAE)) were computed.

Data Analysis. The collected data were grouped by motion step and axis. For each group, the mean, standard deviation, and confidence intervals were calculated. Trends in tracking accuracy were visualized through line plots and Bland-Altman diagrams. Differences between axes were also analyzed to identify potential anisotropies in tracking performance.

This structured methodology provides a reliable and replicable framework for assessing the motion accuracy of consumer-grade VR systems such as the OQ2.

4.1.4 Results

Translational Accuracy

The absolute and percentage errors, corresponding to all the translational displacements performed along the x-, y-, and z- directions are reported in Table 4.2. The MAE, expressed as mean $\pm$ standard deviation, was 13.52 ± 6.57 mm along the x-axis at a distance of approximately 700 mm from the HMD, corresponding to a translational displacement of 500 mm from controller A's initial position (i.e., at about 200 mm from the HMD). Along x-, y-, and z-directions, the absolute errors increased as the distance of the controller A from the HMD rose (Figure 4.5). The Highest percentage errors were found to be 29.5%, 11.8%, and 16.7% for the x-, y-, and z-axes, respectively (Figure 4.6). These results corresponded to the smallest

Table 4.2: Errors for all the translational displacements along the x-, y-, and z- directions.

Directional axis	Translational displacement P [mm]		Absolute Error Mean (SD) [mm]	Percentage Error Mean (SD) [%]
X	P1	5	1.38 (0.85)	29.5 (18.5)
	P2	20	2.13 (0.88)	11.0 (4.4)
	P3	70	2.22 (1.47)	3.2 (2.1)
	P4	150	4.04 (3.22)	2.7 (2.1)
	P5	250	2.41 (1.92)	1.0 (0.8)
	P6	255	2.03 (1.72)	0.8 (0.7)
	P7	270	3.21 (1.96)	1.2 (0.7)
	P8	320	5.11 (4.39)	1.6 (1.4)
	P9	400	7.06 (3.99)	1.8 (1.0)
	P10	500	13.52 (6.57)	2.7 (1.3)
	P11	505	11.67 (4.21)	2.3 (0.8)
Y	P1	5	0.55 (0.47)	11.8 (10.4)
	P2	20	1.35 (0.94)	6.0 (3.7)
	P3	70	1.68 (1.36)	2.4 (1.9)
	P4	150	2.68 (1.47)	1.8 (1.0)
	P5	250	4.22 (2.35)	1.7 (0.9)
	P6	255	4.43 (2.84)	1.7 (1.1)
	P7	270	4.07 (2.51)	1.5 (0.9)
	P8	320	5.02 (3.32)	1.6 (1.0)
	P9	400	7.11 (3.02)	1.8 (0.8)
	P10	500	8.63 (3.74)	1.7 (0.8)
	P11	505	9.74 (5.65)	1.9 (1.1)
Z	P1	5	0.81 (0.73)	16.7 (16.4)
	P2	20	1.16 (0.52)	8.5 (3.9)
	P3	70	1.16 (0.57)	1.8 (0.9)
	P4	150	1.85 (1.48)	1.3 (1.0)
	P5	250	3.59 (2.12)	1.5 (0.9)
	P6	255	3.21 (2.51)	1.3 (1.0)
	P7	270	4.17 (3.72)	1.6 (1.4)
	P8	320	4.93 (4.60)	1.6 (1.5)
	P9	400	5.50 (4.43)	1.4 (1.1)
	P10	500	7.44 (4.78)	1.5 (1.0)
	P11	505	7.60 (5.56)	1.5 (1.1)

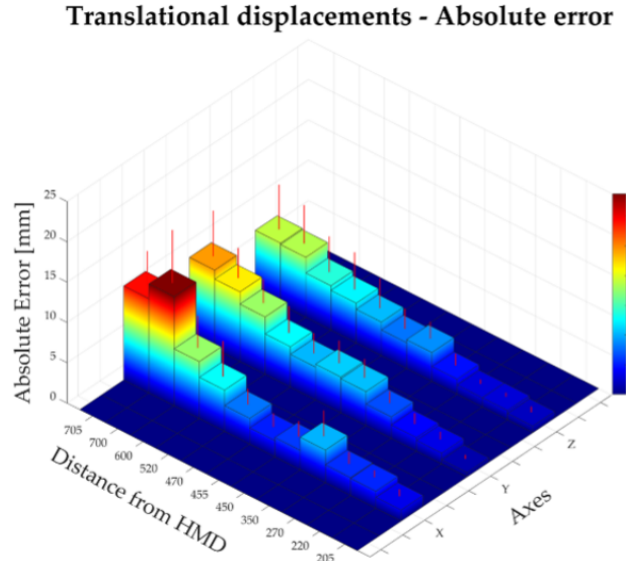


Figure 4.5: Absolute error along the x-, y-, and z-axes at different distances from the HMD.

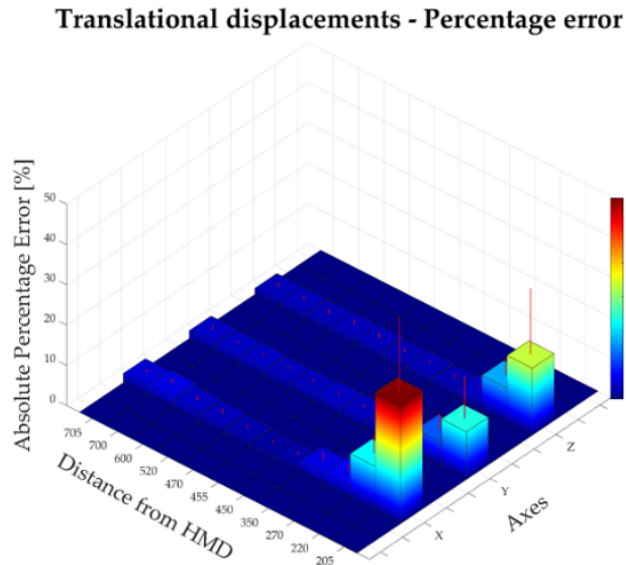


Figure 4.6: Percentage error along the x-, y-, and z-axes at different distances from the HMD.

increments (i.e., 5 mm) from the starting position, at which an absolute error of 1.38 mm, 0.55 mm, and 0.81 mm was found for the x-, y-, and z-axes, respectively. The results showed that the OQ2 was able to discriminate the smallest step size of 5 mm at the distances of 205 mm, 455 mm, and 705 mm from the HMD, although the performance was better at shorter distances (Figure 4.7). The Bland-Altman analysis for translational displacements indicated a decline in

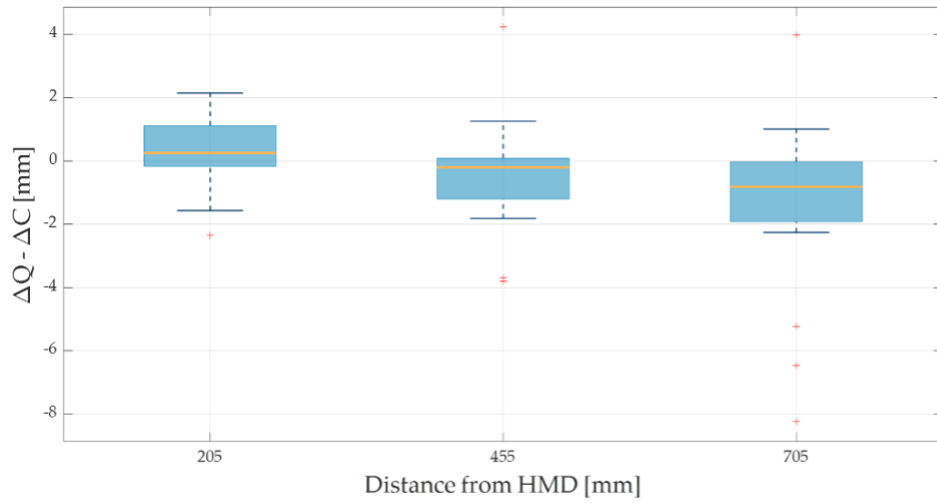


Figure 4.7: Error corresponding to the increment of 5 mm at the distances of 205 mm, 455 mm, and 705 mm from the HMD.

the performance of the OQ2 at greater distances. As shown in Figure 4.8 by the increasing dispersion of the differences between the two measurement systems as the distance of controller A from the HMD increases.

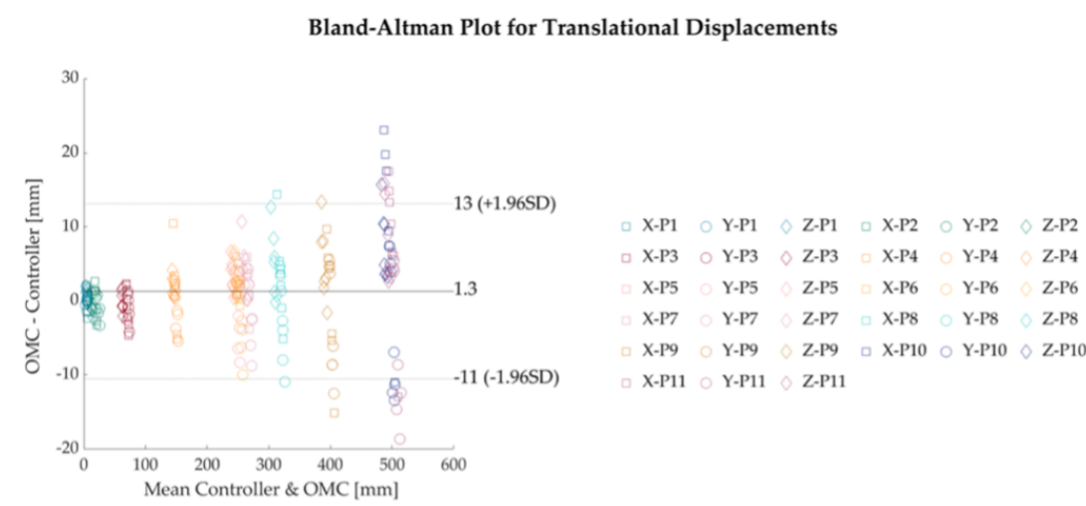


Figure 4.8: Bland-Altman plots comparing the translational displacements measured by the Qualisys OMC system and by the Controller A of the OQ2.

Table 4.3: Errors for all the rotational displacements around the x-, y-, and z-axes.

Directional axis	Rotational displacement R		Absolute Error Mean (SD) [°]	Percentage Error Mean (SD) [%]
X	R1	1	0.10 (0.05)	5.9 (3.1)
	R2	3	0.07 (0.05)	3.3 (2.4)
	R3	10	0.15 (0.12)	1.6 (1.2)
	R4	20	0.09 (0.06)	0.5 (0.3)
	R5	40	0.13 (0.07)	0.3 (0.2)
	R6	90	0.30 (0.12)	0.3 (0.1)
	R7	180	0.63 (0.13)	0.4 (0.1)
Y	R1	1	0.31 (0.25)	42.9 (44.6)
	R2	3	0.13 (0.10)	4.2 (3.0)
	R3	10	0.08 (0.04)	0.9 (0.4)
	R4	20	0.12 (0.06)	0.6 (0.3)
	R5	40	0.07 (0.07)	0.2 (0.2)
	R6	90	0.50 (0.11)	0.6 (0.1)
	R7	180	0.30 (0.13)	0.2 (0.1)
Z	R1	1	0.43 (0.19)	42.6 (21.5)
	R2	2	0.09 (0.08)	3.2 (2.6)
	R3	10	0.18 (0.24)	1.8 (2.4)
	R4	20	0.31 (0.22)	1.6 (1.1)
	R5	40	1.11 (0.37)	2.8 (0.9)
	R6	90	0.75 (0.48)	0.8 (0.5)
	R7	180	0.44 (0.27)	0.2 (0.2)

Rotational Accuracy

The absolute and percentage errors corresponding to all the rotational displacements performed around the three axes are reported in Table 4.3. The MAE, expressed as mean $\pm$ standard deviation, was $1.11 \pm 0.37^\circ$ for a rotation of 40° around the z-axis (Figure 4.9). The highest percentage errors were observed at 5.9%, 42.9%, and 42.6% for a rotation of 1° around the x-, y-, and z-axis, respectively (Figure 4.10). The rotational displacements of 1° around all axes corresponded to a small absolute error of 0.10 ± 0.05 mm, 0.31 ± 0.25 mm, and 0.43 ± 0.19 mm for the x-, y-, and z-axes, respectively. Bland-Altman analysis showed acceptable agreement between the two measurement systems with LOAs within 1.7° and MOD equal to -0.08° (Figure 4.11).

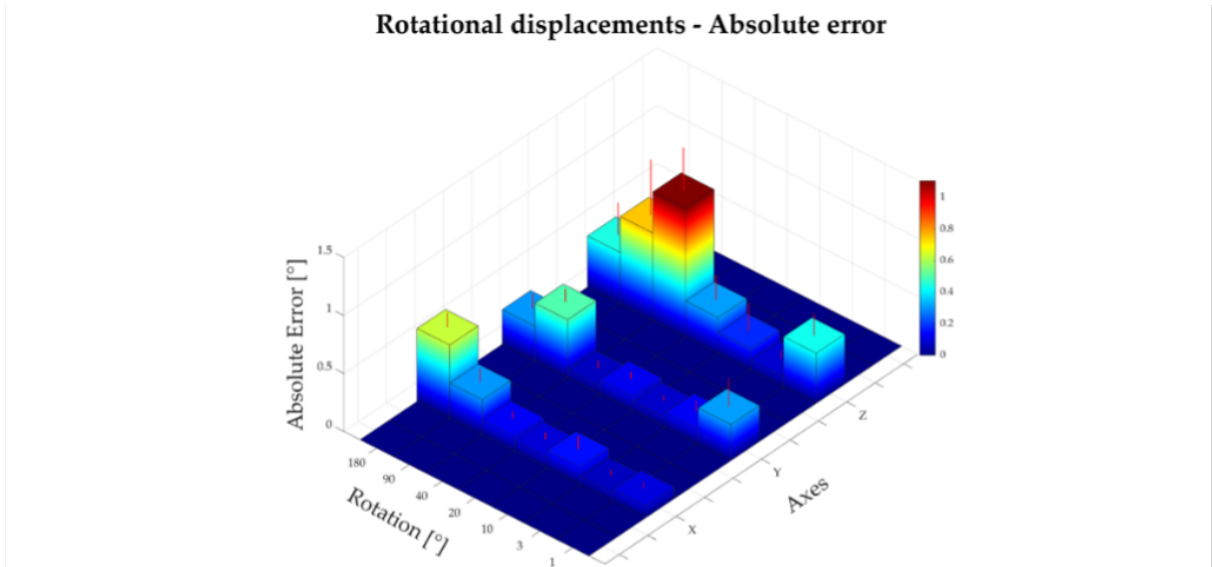


Figure 4.9: Absolute error corresponding to rotational displacements around the x-, y-, and z-axes.

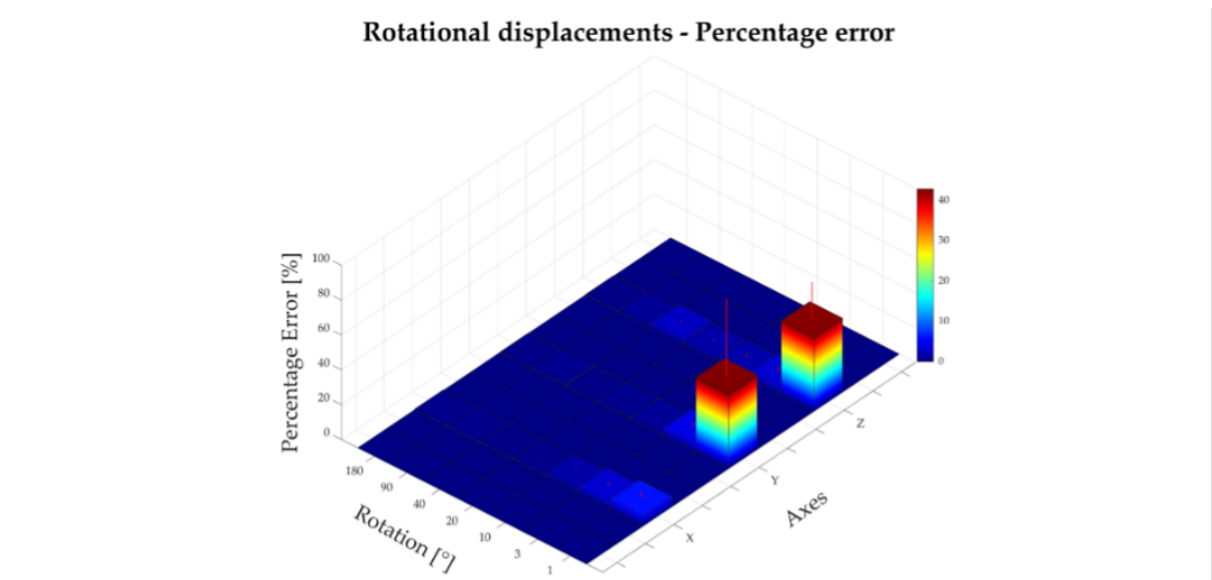


Figure 4.10: The percentage error corresponding to rotational displacements around the x-, y-, and z-axes.

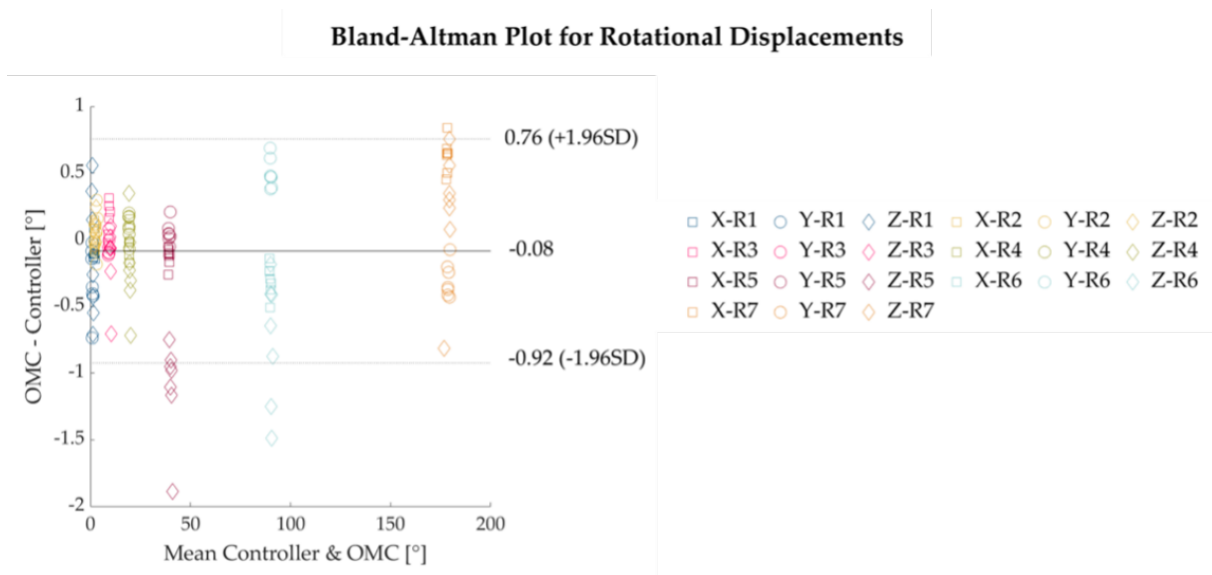


Figure 4.11: Bland-Altman plots comparing the rotational displacements measured by the Qualisys OMC system and by the Controller A of the OQ2.

4.2 Does the app follow the rehabilitation protocol?

Following the validation of the OQ2 device, the next step was to create and subsequently investigate the performance of the customized VR app. The main goal was to develop an application specifically suited for patients with RC tears (following recommendations developed by the American Society of Shoulder and Elbow Therapists [84]) and then verify if, during the VR application, the achievement of essential ROM targets outlined in a standard rehabilitation protocol, were achieved. The custom VR application was divided into 4 levels depending on the degree of difficulty, exercises, and ROM targets. This aspect was crucial to ensure that the application not only met therapeutic requirements but also accurately replicated the movements necessary for effective physical recovery. Specifically, the application needed to enable patients to perform targeted exercises with an appropriate ROM, guaranteeing that each proposed difficulty level matched the standards of the rehabilitation protocol. This verification was essential, as insufficient or excessive movement ranges could interfere with recovery or, worse, cause injury. Additionally, the application had to allow for the progressive adaptation of exercises, allowing a gradual increase in difficulty aligned with the patient's physical improvements. Since VR technology focuses users' attention on the immersive experience, the design of VR applications must be shaped to the specific needs of the target population and carefully adapted to the intended context of use. This means that the application should be designed by taking into account the needs of the final users, so they can intuitively and appropriately operate with their physical and cognitive abilities. RC diseases have a high incidence in the working population, leading to symptoms such as pain, reduced ROM, and work absences [85], [86], [87]. Therefore, creating a VR app that helps in rehabilitation by promoting movement in a controlled and engaging virtual environment could be useful for improving patient outcomes. Encouraging patients to do therapeutic exercises in a VR environment could improve their motivation and can increase their adherence to their rehabilitation protocol. Shoulder rehabilitation after RC repair focuses on obtaining full painless shoulder ROM gradually introducing muscle strengthening exercises under the supervision of expert physiotherapists who adjust the difficulty of the exercises based on the patient's recovery progress [88]. In the field of orthopedic rehabilitation, few studies have evaluated the effectiveness of kinematic

interventions in virtual environments for patients [89], [90], [91], [92]. One of the few studies that was in line with the methodology proposed for this new research was conducted by Baldominos et al., who tested a custom application designed to reproduce abduction and adduction movements with four physiotherapy professionals following RC repair. This study used Oculus Rift DK2 VR in association with the motion tracking system Inter RealSense [69].

4.2.1 Game Design and Development

The application, initially employed for validating the performance of the OQ2 instrument, was further enhanced and expanded using Unity3D (version 2020.17) in combination with the Oculus Integration Package. This development allowed the creation of a fully immersive VR environment, improving graphics, user interaction, and real-time feedback. By integrating the Oculus package, the app could fully utilize VR headsets and motion tracking systems, enabling more precise and responsive user experiences. This expansion not only optimized the app's functionality but also facilitated a more engaging and realistic approach to physical therapy exercises. As mentioned previously, the game was designed considering the recommended rehabilitation exercises following arthroscopic RC repair [84]. A set of C# scripts and free Unity components have been used and implemented to develop different functions in the VR app. The design process, to create the VR rehabilitation scenario, included the specification of: requirements, context, and objectives. Figure 4.12 illustrated the meta model that described the specifications of the developed application, defining the collection of objects and their relationship. Each box in Figure 4.12 represented an object with its associated properties. The boxes highlighted with a more intense color (Context, Objective, and Requirement), represented the main objects around which the entire application was created. The lines connecting the boxes illustrated the relationships between the elements of context, objectives, and requirements. Finally, each layout consisted of a collection of many objects indicated by a number next to the boxes.

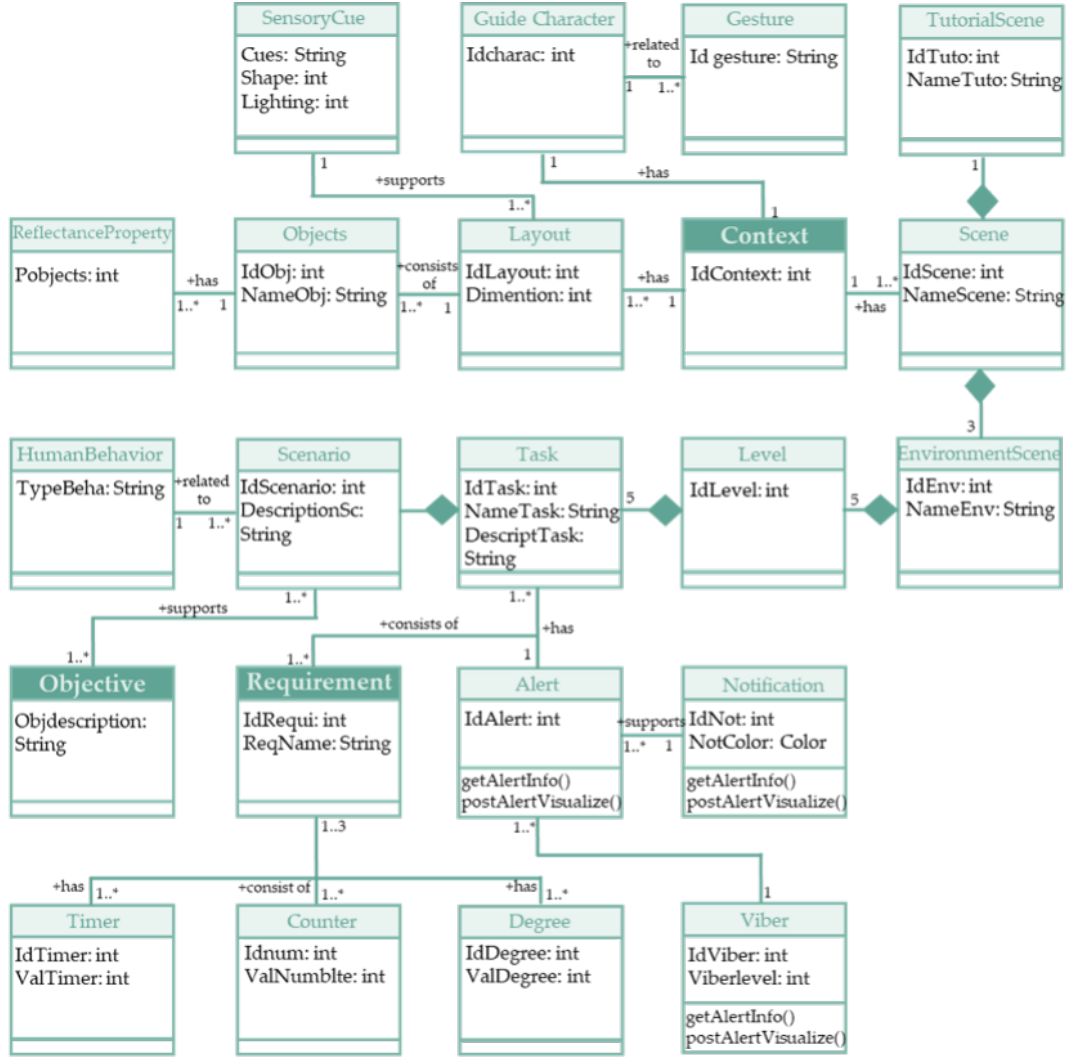


Figure 4.12: Meta model of the developed immersive virtual reality application. “*” means the variable number of elements.

Specification of Requirements, Context and Objectives

The following requirements were implemented in the developed VR app: alertness, arm length, height, field of vision, fault tolerance, timer, and counter (Figure 4.13). In this prototype, an alert system was incorporated to notify users of correct hand positioning, guiding the appropriate movement of the arm. Upper limb length and height could be adjusted by pressing the trigger on the controller, making the app accessible to a diverse population. These parameters were crucial for accurately positioning objects within the scene, improving the Field of View (FOV) for each user. Additionally, a controller vibration and an arrow of different colors

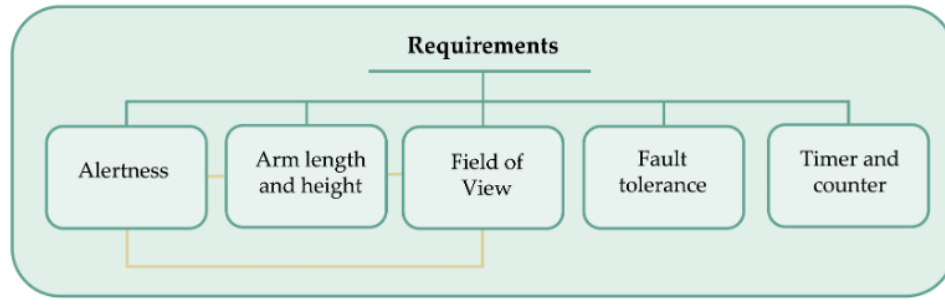


Figure 4.13: Specification of requirements of the developed immersive virtual reality application.

(adapted to the degree of shoulder movement, user height, and arm length) can alert patients to right-hand movement and object position. The arrow indicated angular degrees and was dynamically adjusted based on the user's body dimensions. Fault tolerance was designed to maintain continuous system operation by ensuring the reliability of various components, allowing the app to function even if FOV or task/scene changes occurred during interactions. Thus, an algorithm was developed to ensure that the controller of the VR device will work during the whole process. Furthermore, a timer and counter were set for each scene to monitor the duration of each training session and track the number of movements performed. To improve the patient experience, appropriate scenes, layouts, and characters were incorporated into the virtual environment. The layout included all of which influenced the user's perception of the environment, such as: auditory feedback, visual configuration of the surroundings, and lighting. The shapes, colors, and lighting intensity were considered critical variables, as they could significantly affect the user's attitude and QoE. Specific sensory cues were selected in the developed VR app to optimize the user experience. In this customized VR app, rehabilitation movements were simulated within a naturalistic setting, under suitable lighting conditions. These design choices aimed to improve task performance, enhancing the appearance of the performance area, and positively affecting the patients' perceptions and overall experience. Additionally, the VR app showed a virtual physiotherapist at the beginning and end of each exercise to provide directions and motivational feedback to the patient. An immersive and naturalistic scenario was developed, allowing users to perform rehabilitation exercises within a virtual environment full of natural world with objects like trees, apples, animals, and vegetables. Before task execution, patients may want to be trained on the use of the VR app.

Consequently, a tutorial was incorporated at the beginning of the VR experience to assist users in managing the device and understanding the functionality of the VR application.

4.2.2 The experiment

Population and Equipment

Five volunteers (male/females: 2/3; mean age $\pm$ standard deviation: 24.4 ± 3.0 years old; body mass: 67.6 ± 16.9 kg; height: 1.70 ± 0.10 m) were enrolled in this study. Eligibility for participation was determined by the following criteria: age $age \geq 18$ years, no musculoskeletal disorders of the shoulder, and no history of shoulder surgery. All but one of the participants were right-handed. The experimental sessions took place in the Laboratory of Motion Analysis at the Fondazione Policlinico Universitario Campus Bio-Medico in Rome. Volunteers wore the OQ2 and interacted with the virtual environment using two controllers attached to the ends of a double-handled stick (Beat Saber Handle), allowing them to move and perform tasks within the customized virtual environment. The VR system tracked the movements of the head and hands without the need for external sensors, replicating these movements in the virtual space. Shoulder kinematics was recorded as always recorded by the QualisysTM motion capture system. In each session, photo-reflective markers (8 mm in diameter) were placed on key anatomical landmarks: incisura jugularis, processus xiphoideus, C7 vertebra, T8 vertebra, and bilaterally on the acromioclavicular joint, coracoid process, trigonum spinae scapulae, angulus inferior, angulus acromialis, medial and lateral epicondyles, and radial and ulnar styloid [93]. The center of glenohumeral rotation was estimated using the linear regression method [94]. Additionally, five rectangular marker clusters, each containing four markers, were placed on the thorax, upper arms, and forearms to track segmental movements during dynamic trials [95]. Ten Miquis M3 cameras recorded markers' trajectories at 100 Hz.

Movements Protocol

After configuring the play area according to the user's arm length and height, all participants completed four levels simulating rehabilitative exercises following RC repair. In the first phase of the experiment, participants calibrated the system by measuring their arm length and

Table 4.4: Movements, ROM, and repetitions for each level included in the immersive virtual reality application.

	Level	Flexion 1	Flexion 2	Abduction	External rotation	Internal Rotation
L1	Repetitions	10	10	10	8	-
	ROM	60°-90°	60°-90°	≤ 45°	≤ 20°	-
L2	Repetitions	10	10	10	8	8
	ROM	90°-120°	90°-120°	45°-80°	20°-30°	≤ 20°
L3	Repetitions	10	10	10	10	10
	ROM	130°-155°	130°-155°	80°-120°	30°-45°	20°-50°
L4	Repetitions	12	12	12	12	12
	ROM	≥ 140°	≥ 140°	≥ 120°	≥ 45°	≥ 50°

height. These measurements were used as input for the scripts designed to place objects in the virtual environment, ensuring they could be reached at each level within the predefined ROM. During the development of the VR application in Unity, developers used a common goniometer to fix the position of the virtual objects in the virtual environment. Table 4.4 reports the shoulder movements, ROM, and number of repetitions for each level. The first level included two sets of ten repetitions each of flexion movements in the range of 60°–90°, one set of ten repetitions of abduction movements up to 45°, and one set of eight repetitions of extra-rotations up to 20°. The second level consisted of two sets of ten repetitions each of flexion movements in the range 90°–120°, one set of ten repetitions of abduction movements in the range 45°–80°, one set of eight repetitions of extra-rotations in the range 20°–30°, and one set of eight repetitions of intra-rotations up to 20°. The third level included two sets of ten repetitions each of flexion movements within the range of 130°–155°, one set of ten repetitions of abduction movements between 80°–120°, one set of ten repetitions of extra-rotations ranging from 30°–45°, and one set of ten repetitions of intra-rotation in the range of 20°–50°. The fourth level consisted of two sets of twelve repetitions for flexion movements within the range of 140°–ROM_{MAX}, one set of twelve repetitions for abduction movements between 120°–ROM_{MAX}, one set of twelve repetitions for external rotations ranging from 45°–ROM_{MAX}, and one set of twelve repetitions of intra-rotation in the range of 50°–ROM_{MAX}.

Figures 4.14, 4.15, 4.16 present a HV provided with the OQ2, Beat Saber handles, and photo-

reflective markers while performing flexion, abduction, and external rotation movements, respectively. Internal rotation is not shown because it was the same as external rotation, but in the opposite direction.

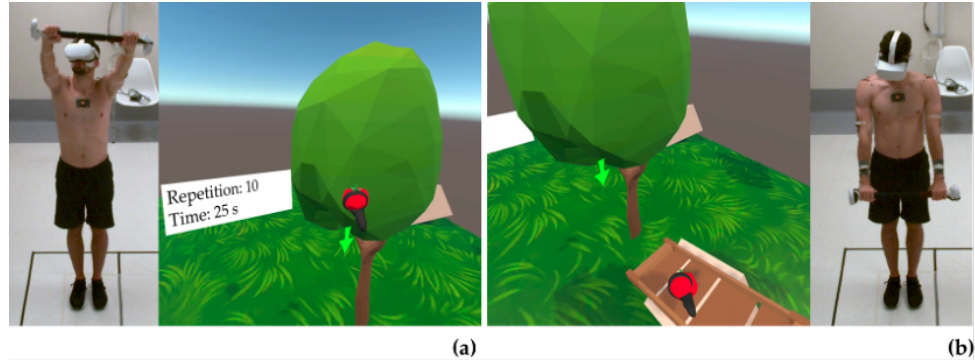


Figure 4.14: Representation of a healthy volunteer wearing the OQ2, the Beat Saber Handle, and photo-reflective markers during flexion movement to reach the virtual target (a) and in the starting position (b). Arrows indicated the direction of movement.

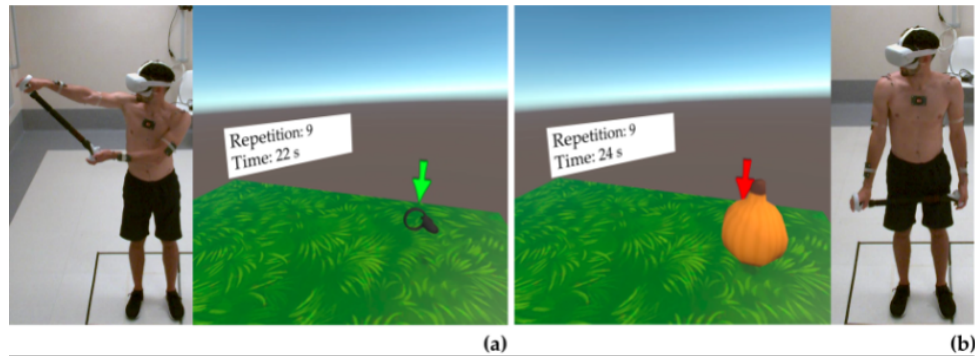


Figure 4.15: Representation of a healthy volunteer wearing the OQ2, the Beat Saber Handle, and photo-reflective markers during abduction movement to reach the virtual target (a) and in the starting position (b). Arrows indicated the direction of movement.

4.2.3 Data Collection

All participants were instructed to hold a static N-pose for 3 seconds to define the kinematic model based on the trajectories of the anatomical photo-reflective markers. Marker trajectories were collected and pre-processed (e.g., gap filling, labeling) with QTM software

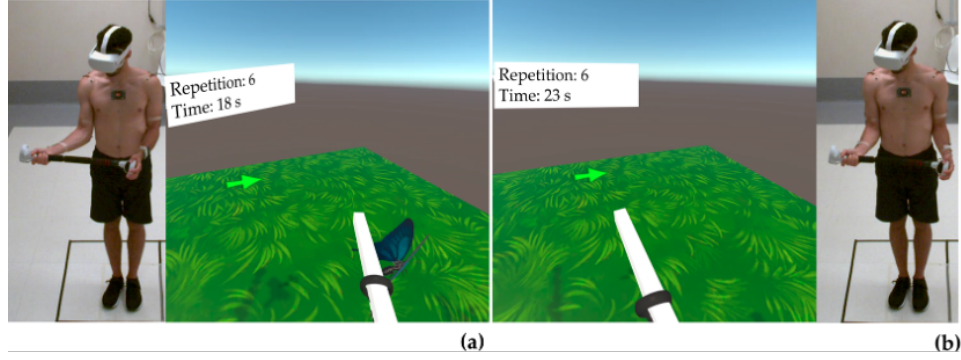


Figure 4.16: Representation of a healthy volunteer wearing the OQ2, the Beat Saber Handle, and photo-reflective markers during external rotation to reach the virtual target **(a)** and in the starting position **(b)**. Arrows indicated the direction of movement.

(v. 2023.1, build 7985). The processed trajectories were then imported into Visual 3D (C-Motion, Inc., Germantown, WD, USA) for kinematic analysis through a custom pipeline. In Visual 3D, the trajectories of anatomical markers during the static N-pose were used to establish the local coordinate system and orientation of the thorax, humeri, and forearms [93]. Only data from the right side were further analyzed as the VR app was designed only for the right side at that time. The humerus 3D orientation was expressed relative to the thorax. In particular the rotation sequences used to evaluate humerothoracic angles were abduction–adduction, flexion–extension, and internal–external rotation (YXZ sequence in Visual 3D) for movements in the frontal and transverse planes, and flexion–extension, abduction–adduction, and internal–external rotation (XYZ sequence in Visual 3D) for movements in the sagittal plane [96], [97] (Figure 4.17). In Figure 4.17, a schematic representation of shoulder joint angles starting from the kinematic model definition (all markers are displayed except for trigonum spinae scapulae, angulus inferior, and angulus acromialis) is presented. The local coordinate systems of the thorax, humeri, and forearms were based on the ISB recommendations [93]. Movements of the humerus relative to the thorax (humerothoracic joint angles) were measured using specific rotation sequences: the XYZ sequence was used for flexion (rotation around the X-axis), and the YXZ sequence for abduction (rotation around the Y-axis) and axial rotation (rotation around the Z-axis) [96], [97]. The kinematic data obtained from Visual3D were exported in .mat format for further analysis in MATLAB® (version R2022b). The

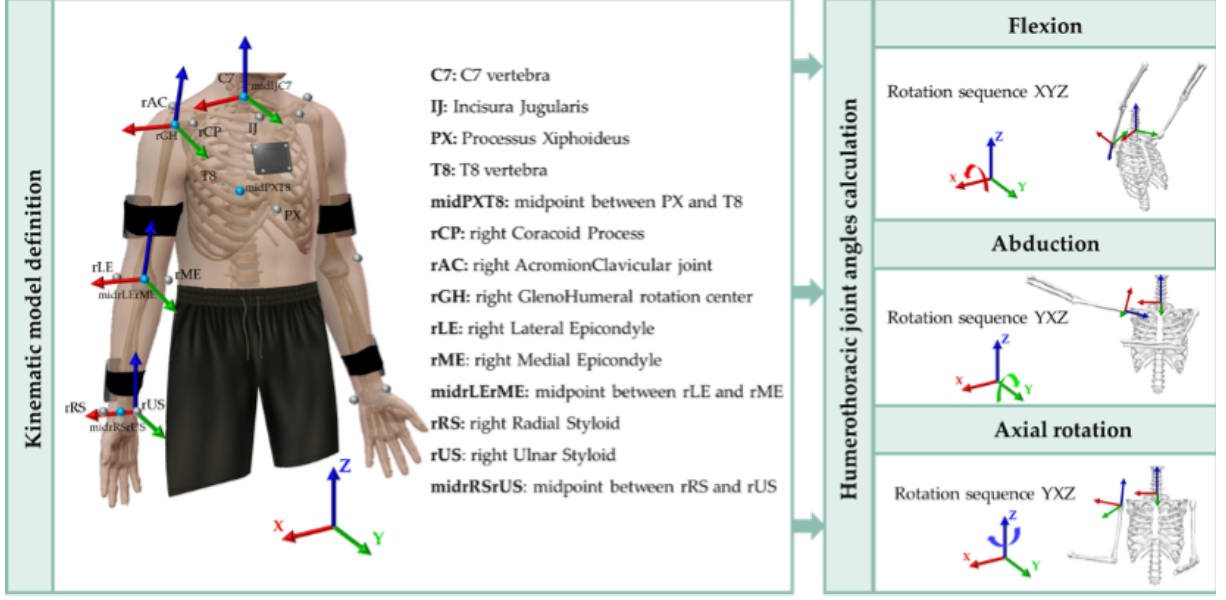


Figure 4.17: Schematization of shoulder joint angles.

waveforms corresponding to flexion, abduction, internal, and external rotation angles were analyzed to detect the peak values reached during each repetition. For each identified repetition, the peak value was evaluated to verify if it met the ROM criteria defined in the VR application, as outlined below:

$$InfCond = Peaks_i < InferiorLimit? \begin{cases} err_{i,inf} = |x_{GS,i} - x_{hyp,inf}| & \text{if } InfCond == 1 \\ err_{i,inf} = [] & \text{if } InfCond == 0 \end{cases}, \quad (4.9)$$

$$SupCond = Peaks_i > SuperiorLimit? \begin{cases} err_{i,sup} = |x_{GS,i} - x_{hyp,sup}| & \text{if } SupCond == 1 \\ err_{i,sup} = [] & \text{if } SupCond == 0 \end{cases}, \quad (4.10)$$

where $InfCond$ and $SupCond$ corresponded to the inferior and superior bounds, respectively; $err_{i,inf}$ and $err_{i,sup}$ corresponded to the absolute error at the i -th repetition for the lower and upper bounds of the established ROM, respectively; $x_{GS,i}$ was the angular value measured by the gold standard at the i -th repetition; $x_{hyp,inf}$ and $x_{hyp,sup}$ corresponded to the lower and upper hypothesized values of the ROM, respectively. For each subject, the MAE for the inferior (MAE_{inf}) and superior (MAE_{sup}) conditions were computed as follows:

$$MAE_{inf} = \text{mean}(err_{inf}), \quad (4.11)$$

$$MAE_{sup} = \text{mean}(err_{sup}) . \quad (4.12)$$

Let N_{inf} , and N_{sup} be the total number of times the inferior and superior conditions were not met for all participants, respectively; then, the total MAE for the inferior ($MAE_{tot,inf}$) and superior ($MAE_{tot,sup}$) conditions were computed as follows:

$$MAE_{tot,inf} = \text{mean}(err_{N,inf}) , \quad (4.13)$$

$$MAE_{tot,sup} = \text{mean}(err_{N,sup}) . \quad (4.14)$$

where $err_{N,inf}$ and $err_{N,sup}$ were vectors $1 \times N_{inf}$ and $1 \times N_{sup}$ reporting the concatenated vector of errors made by all volunteers.

4.2.4 Results

Results for the first and second rounds of flexion, seen in Tables 4.5 and 4.6, respectively, and for abduction (Table 4.7), demonstrated low values for MAE and for the number of times conditions were not met overall. In the case of external rotation (Table 4.8) and internal rotation (Table 4.9), the total MAE values were similar to those observed for flexion and abduction, but the number of times the conditions were not met was significantly higher. This discrepancy was attributed to the volunteers' difficulty in performing isolated shoulder rotations, conducting them to compensating with trunk movements.

Flexion

The results of the first series of flexion are illustrated in Table 4.5. At level 1, the total inferior MAE was 6.4, with this value progressively decreasing with each level. Inferior conditions were not satisfied on 8 occasions by participants. Total superior MAE was found to be 2.9, and superior conditions were not met 5 times. In level 2, the total inferior MAE was 5.1, and there were 15 times where inferior conditions were not met (the highest number of unmet conditions across both the first and second rounds of flexion). No errors or unmet conditions were recorded for the superior limit. At level 3, the total inferior MAE was 1.3, with 4 unmet conditions. The total superior MAE was 8.6, with superior conditions unmet 3 times. In level 4, volunteers only had one recorded instance of an unmet condition for inferior limit and a

Table 4.5: Results for the first series of flexion for each level (MAE_{inf} and MAE_{sup} are the MAEs corresponding to each volunteer; N_{inf} and N_{sup} corresponded to the number of times the inferior and superior conditions were not met, respectively; $MAE_{TOT,inf}$ and $MAE_{TOT,sup}$ were the total MAE related to the inferior and superior conditions, respectively).

Level	Volunteer	MAE_{inf}	N_{inf}	$MAE_{TOT,inf}$	MAE_{sup}	N_{sup}	$MAE_{TOT,sup}$
L1	V1	1.0	8	6.4	✓	5	2.9
	V2	3.3			✓		
	V3	2.7			3.4		
	V4	✓			2.6		
	V5	10.2			✓		
L2	V1	2.3	15	5.1	✓	0	✓
	V2	5.5			✓		
	V3	2.6			✓		
	V4	2.8			✓		
	V5	7.9			✓		
L3	V1	✓	4	1.3	✓	3	8.6
	V2	✓			✓		
	V3	0.9			✓		
	V4	1.7			✓		
	V5	✓			8.6		
L4	V1	✓	1	0.8	–	–	–
	V2	✓			–		
	V3	0.8			–		
	V4	✓			–		
	V5	✓			–		

Table 4.6: Results for the second series of flexion for each level (MAE_{inf} and MAE_{sup} are the MAEs corresponding to each volunteer for the inferior and superior conditions, respectively; N_{inf} and N_{sup} corresponded to the number of times the inferior and superior conditions were not met, respectively; $MAE_{TOT,inf}$ and $MAE_{TOT,sup}$ were the total MAE related to the inferior and superior conditions, respectively).

Level	Volunteer	MAE_{inf}	N_{inf}	$MAE_{TOT,inf}$	MAE_{sup}	N_{sup}	$MAE_{TOT,sup}$
L1	V1	6.8	6	4.6	✓	6	10.5
	V2	2.8			✓		
	V3	✓			✓		
	V4	✓			10.5		
	V5	2.3			✓		
L2	V1	8.0	8	12.3	✓	7	9.2
	V2	15.4			✓		
	V3	✓			✓		
	V4	✓			9.2		
	V5	13			✓		
L3	V1	✓	2	4.8	✓	2	2.1
	V2	7.9			✓		
	V3	✓			✓		
	V4	1.7			✓		
	V5	✓			2.1		
L4	V1	✓	8	2.6	–	–	–
	V2	2.6			–		
	V3	✓			–		
	V4	✓			–		
	V5	✓			–		

total inferior MAE equivalent to 0.8, which is the lowest error value for inferior flexion in both the first and second rounds.

For the second series of flexion, as shown in Table 4.6 the total inferior MAE at level 1 was 4.6, with inferior conditions unmet 6 times. The total superior MAE reached 10.5, the highest for both flexion rounds, with 6 unmet superior conditions. In level 2, volunteers had a total inferior MAE equivalent to 12.3, the highest value of mean error for inferior flexion overall. Participants did not meet inferior conditions 8 times. The total superior MAE was found to be 9.2, and only one volunteer (V4) contributed to this error value. In total, conditions were unmet 7 times. At level 3, the total inferior and superior MAE values were 4.8 and 2.1, respectively, with 2 unmet conditions for each. In level 4, there were 8 unmet conditions, and the total inferior MAE was 2.6.

Table 4.7: Results for abduction for each level (MAE_{inf} and MAE_{sup} are the MAEs corresponding to each volunteer for the inferior and superior conditions, respectively; N_{inf} and N_{sup} corresponded to the number of times the inferior and superior conditions were not met, respectively; $MAE_{TOT,inf}$ and $MAE_{TOT,sup}$ are the total MAE related to the inferior and superior conditions, respectively).

Level	Volunteer	MAE_{inf}	N_{inf}	$MAE_{TOT,inf}$	MAE_{sup}	N_{sup}	$MAE_{TOT,sup}$
L1	V1	–	–	–	5.9	14	6.1
	V2	–			7.8		
	V3	–			✓		
	V4	–			✓		
	V5	–			1.2		
L2	V1	✓	0	✓	✓	4	4.7
	V2	✓			4.7		
	V3	✓			✓		
	V4	✓			✓		
	V5	✓			✓		
L3	V1	4.0	15	8.6	✓	0	✓
	V2	✓			✓		
	V3	11.4			✓		
	V4	8.2			✓		
	V5	8.0			✓		
L4	V1	6.6	22	6.6	–	–	–
	V2	5.8			–		
	V3	5.9			–		
	V4	10.8			–		
	V5	7.0			–		

Abduction

Results for abduction are reported in Table 4.7. At level 1, abduction showed the highest total superior MAE and the largest number of instances where superior conditions were unmet, with values of 6.1 and 14, respectively. In level 2, no errors were recorded for the inferior condition during abduction, and volunteers consistently met the required conditions. For the superior condition, the total superior MAE was 4.7, with 4 instances where the conditions were unmet. At level 3, the total inferior MAE for abduction was 8.6, the highest across all levels, with 15 instances of unmet conditions. No errors or unmet conditions were recorded for the superior condition. At level 4, abduction had a total inferior MAE of 6.6, with 22 instances where conditions were not met, representing the highest number of unmet conditions recorded for both superior and inferior limits in abduction.

Table 4.8: Results for external rotation for each level (MAE_{inf} and MAE_{sup} are the MAEs corresponding to each volunteer for the inferior and superior conditions, respectively; N_{inf} and N_{sup} corresponded to the number of times the inferior and superior conditions were not met, respectively; $MAE_{TOT,inf}$ and $MAE_{TOT,sup}$ are the total MAE related to the inferior and superior conditions, respectively).

Level	Volunteer	MAE_{inf}	N_{inf}	$MAE_{TOT,inf}$	MAE_{sup}	N_{sup}	$MAE_{TOT,sup}$
L1	V1	–	–	–	11.5	32	6.5
	V2	–			3.4		
	V3	–			1.9		
	V4	–			5.7		
	V5	–			7.9		
L2	V1	✓	1	2.3	2.9	12	2.6
	V2	✓			1.7		
	V3	✓			✓		
	V4	2.3			✓		
	V5	✓			2.6		
L3	V1	✓	16	6.3	2.6	5	2.6
	V2	✓			✓		
	V3	1.3			✓		
	V4	10.3			✓		
	V5	0.1			✓		
L4	V1	1.1	26	19.6	–	–	–
	V2	✓			–		
	V3	15.8			–		
	V4	26.6			–		
	V5	✓			–		

External Rotation

Results for external rotation are reported in Table 4.8. In level 1, external rotation had a total superior MAE value of 6.5, while conditions were not met 32 times. In level 2, external rotation had a total inferior MAE equivalent to 2.3, and volunteers did not meet conditions only once. The total superior MAE was found to be 2.6, and conditions were not met 12 times. In level 3, external rotation had a total inferior MAE equivalent to 6.3, and volunteers did not meet conditions 16 times. The total superior MAE was found to be 2.6, and conditions were not met 5 times. In level 4, external rotation had a total inferior MAE of 19.6, and conditions were not met 26 times; 19.6 represents the highest total MAE value for all recorded movements.

Table 4.9: Results for internal rotation for each level (MAE_{inf} and MAE_{sup} are the MAE corresponding to each volunteer for the inferior and superior conditions, respectively; N_{inf} and N_{sup} corresponded to the number of times the inferior and superior conditions were not met, respectively; $MAE_{TOT,inf}$ and $MAE_{TOT,sup}$ are the total MAE related to the inferior and superior conditions, respectively).

Level	Volunteer	MAE_{inf}	N_{inf}	$MAE_{TOT,inf}$	MAE_{sup}	N_{sup}	$MAE_{TOT,sup}$
L2	V1	–	–	–	17.3	33	9.8
	V2	–			2.8		
	V3	–			8.0		
	V4	–			0.2		
	V5	–			12.4		
L3	V1	✓	6	2.0	✓	0	✓
	V2	1.9			✓		
	V3	✓			✓		
	V4	2.2			✓		
	V5	✓			✓		
L4	V1	✓	30	8.0	–	–	–
	V2	4.3			–		
	V3	7.6			–		
	V4	10.7			–		
	V5	3.1			–		

Internal Rotation

Results for internal rotation are reported in Table 4.9. At level 2, internal rotation showed a total superior MAE of 9.8, with conditions unmet 33 times, the highest number of unmet conditions recorded for all recorded movements. At level 3, the total inferior MAE for internal rotation was 2.0, with 6 unmet conditions. For the superior condition, no errors or unmet conditions were recorded. In level 4, internal rotation has a total inferior MAE equivalent to 8.0, and volunteers did not meet conditions 30 times. The data showed encouraging results, with volunteers successfully meeting the target conditions most of the time. Although the outcomes were positive, additional studies comparing VR applications with established clinical standards are still needed. However, these findings added to the growing body of research that supports the use of VR in shoulder rehabilitation programs.

4.3 Quality of Experience of the physiotherapists

Once the device was validated and it was confirmed that the range of movement performed within the VR app aligned with rehabilitation guidelines, the next critical phase involved evaluating the usability, acceptability, and suitability of the custom VR application. The assessment process starts by obtaining feedback from experts in the field, particularly experienced physical therapists, before the application is introduced in clinical trials with patients. These professionals, with their knowledge of rehabilitation practices, provided suggestions to identify potential improvements in the device. Their expertise was the key to determining if the VR system could be integrated into rehabilitation settings and complement existing methods [98]. The feedback from therapists provides a critical foundation for refining the device and the app. Their recommendations may lead to modifications in the user interface, exercise parameters, or even the hardware itself to make the system more accessible and effective. Examining factors such as: wearability, safety, stability, comfort, size, functionality, and mode of use in terms of acceptability, usability, and appropriateness was useful for both developers and end users. These assessments offered important perspectives for optimizing the design and ensuring that the device met the specific requirements of its intended application.

4.3.1 Experimental Procedure

A qualitative study was conducted to assess overall satisfaction and the feasibility of the custom VR application using OQ2 for shoulder rehabilitation. The study involved eleven physiotherapists (5 males and 6 females, with a mean age of 37.9 years) from the Fondazione Policlinico Universitario Campus Bio-Medico of Rome. Participants were required to meet the following criteria:

- work experience of at least two years in the rehabilitation of shoulder pathologies;
- willingness to respond to online questionnaires;
- signature of the informed consent form.

On the contrary the exclusion criteria for the use of the OQ2 were:

- being tired;
- needing sleep;
- having digestive problems;
- being under emotional stress or anxiety;
- suffering from flu, headaches, migraines, or earaches;
- being pregnant.

The gaming experience of physiotherapists was assessed by asking about how frequently they used video games in general (often, sometimes, never), as well as their use of VR devices and specifically the OQ2. None of the participants had visual impairments that could interfere with the VR experience. The VR application was set in a naturalistic environment, allowing users to interact with virtual objects through the controllers. The application was the same as used in the previous study, with four levels of difficulty corresponding to different stages of rehabilitation. Given that shoulder musculoskeletal rehabilitation often includes cane-assisted exercises (for semi-passive movements), all tasks were designed to be performed using both hands, with a stick handle grip and controllers positioned at each end. The movements available in the application, as previously mentioned, include flexion, abduction, internal rotation, and external rotation, as illustrated in Figures 4.18, 4.19, and 4.20.

The physiotherapists tested the VR rehabilitation program in a single session. Each session, lasting approximately 45 minutes, involved completing levels 1 and 4, which represent, respectively, the easiest and most challenging levels. The study protocol for each participant included the following steps: a) an introduction to the use of OQ2; b) calibration of the OQ2 system; c) start of the game; d) training and familiarization with the game; e) subject height setting; f) level 1 exercises, including sagittal plane elevations, frontal plane elevations, and external rotation exercises; g) level 4 exercises, including sagittal plane elevations, frontal plane elevations, and both external and internal rotation exercises. At each level, the exercises were performed consecutively, starting immediately after the previous one. For sagittal plane elevations, participants were tasked to take an apple from a tree and put it inside a basket, requiring a top-to-bottom movement (Figure 4.18). The height of the apple increased progressively

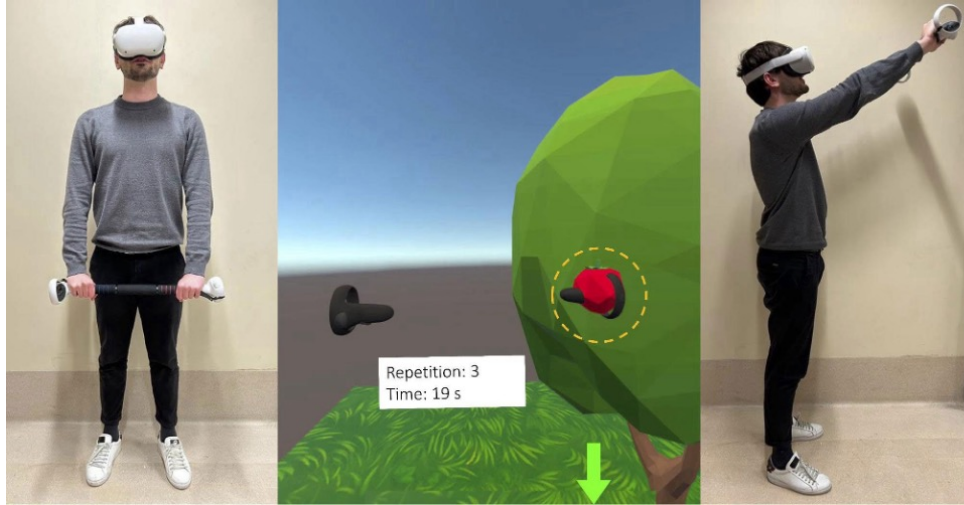


Figure 4.18: Starting position (left) and final position (right) of the elevation exercise in the sagittal plane. In the center is a representation of what the user visualizes through the HMD and the end effector circled in dotted yellow

through the levels, simulating a higher target as synonymous of better ROM ability. Similarly, the second series of elevation in the sagittal plane required performing the movement from bottom to top, and with increasing difficulty through the levels. For the frontal plane elevations, participants began with their arms along their body and moved them laterally to touch a virtual fruit or vegetable, which disappeared after the contact. This exercise involved performing frontal plane elevations, with the targets positioned progressively higher across the levels (Figure 4.19). In the external rotation exercises, players were required to start with their elbows flexed to 90° and their shoulders adducted. In this position, players were asked to grasp a rod to be rotated externally successively until hitting a butterfly (Figure 4.20). For the internal rotation exercises, participants followed similar starting conditions, but the internal rotation movements were performed in the opposite direction.

4.3.2 Questionnaires

After the experiment, physiotherapists completed online questionnaires consisting of both closed and open-ended questions. Two User Experience questionnaires evaluating VR system acceptability and usability were proposed [99], [100]. The acceptability questionnaire contained 13 questions focused on various aspects of the device's acceptability, such as wearability,

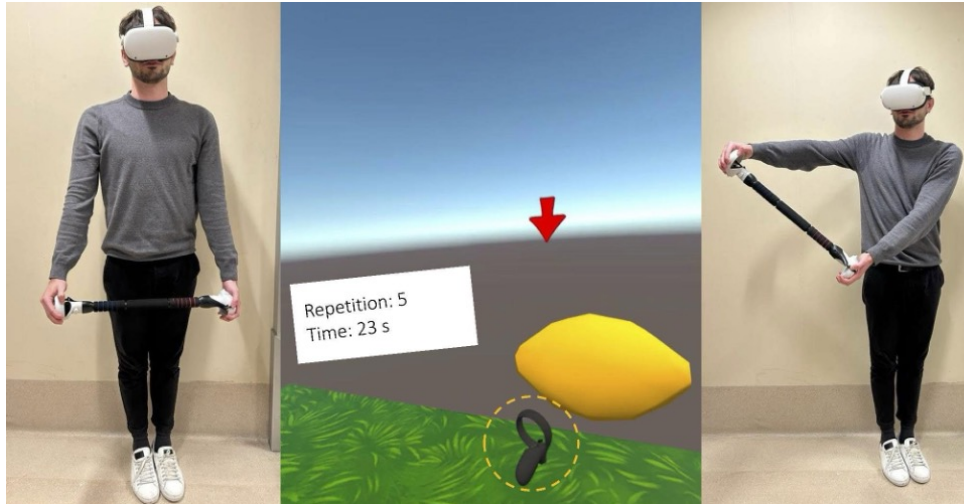


Figure 4.19: Starting position (left) and final position (right) of the elevation exercise in the frontal plane. In the center is a representation of what the user visualizes through the HMD and the end effector circled in dotted yellow.

safety, stability, ease of control, comfort, size, and overall acceptability of the device. Physiotherapists rated statements on a 7-point Likert scale (7=strongly agree, 4=neutral, 1=strongly disagree) [99]. The usability questionnaire included 15 questions divided into three categories: utility, playability, and use mode [100]. Participants rated statements on a 5-point Likert scale (5=strongly agree, 3=neutral, 1=strongly disagree). Additionally, the research group developed a custom appropriateness questionnaire, which included one open-ended question and three multiple-choice questions. These questions aimed to collect the physiotherapists' insights and opinions on the appropriateness of OQ2 and the VR rehabilitation program for patients with shoulder musculoskeletal disorders. Specifically, the questionnaire aimed to explore three additional aspects related to the use of the VR system: which type of shoulder musculoskeletal pathology and in what period of the rehabilitation was more indicated, and if the supervision of a physiotherapist was required.

4.3.3 Results

The physiotherapists involved in the study had an average of 15 years of professional experience (range: 3–25 years, median: 16 years). The majority of the participants (n=10) were familiar with the use of video games, a smaller number had experience with VR devices (n=3),

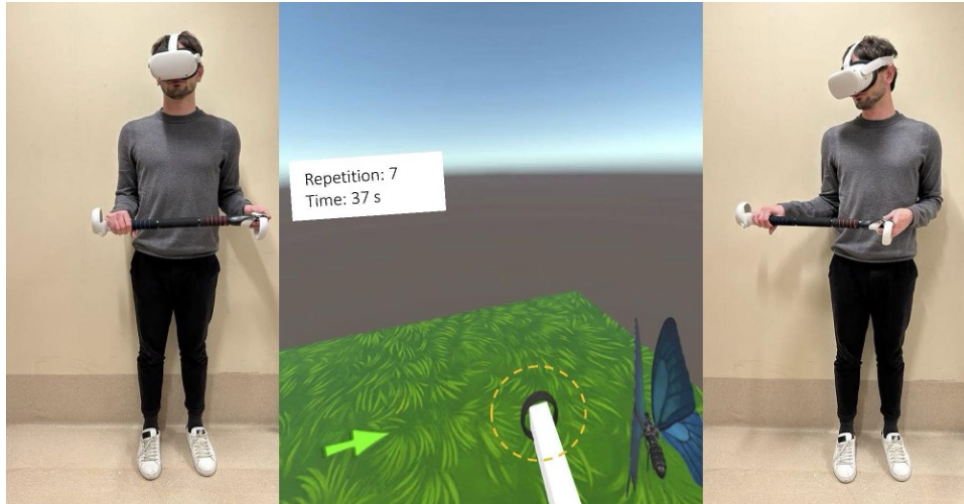


Figure 4.20: Starting position (left) and final position (right) of the external rotation exercise. In the center is a representation of what the user visualizes through the HMD and the end effector circled in dotted yellow

and only one had prior experience with the QQ2 system.

Acceptability questionnaire

Results from the acceptability questionnaire ($n=11$), a 1–7 scale, indicated an overall positive report from the physiotherapists. Figure 4.21 shows more details about each item's responses. All the physiotherapists found QQ2 easy to wear ($Q1 - \text{mean} \pm \text{std} = 6.2 \pm 0.9$) and easy to control ($Q4 - \text{mean} \pm \text{std} = 6.0 \pm 0.9$). Additionally, they felt confident ($Q2 - \text{mean} \pm \text{std} = 6.1 \pm 0.9$) and safe using the device ($Q3 - \text{mean} \pm \text{std} = 6.0 \pm 0.9$). The majority (73%) enjoyed the experience with QQ2 ($Q6 - \text{mean} \pm \text{std} = 5.2 \pm 1.7$), and most (82%) agreed, somewhat agreed, or strongly agreed that they found it comfortable ($Q5 - \text{mean} \pm \text{std} = 5.6 \pm 1.1$). Although 36% of physiotherapists showed disagreement, most (64%) would like to use QQ2 weekly ($Q7 - \text{mean} \pm \text{std} = 4.5 \pm 2.5$), and they would recommend QQ2 to a colleague ($Q8 - \text{mean} \pm \text{std} = 4.6 \pm 2.1$). Only 18% reported feeling a sense of well-being after using the device ($Q9 - \text{mean} \pm \text{std} = 3.3 \pm 1.4$), while most (82%) had a positive opinion regarding its size ($Q11 - \text{mean} \pm \text{std} = 5.6 \pm 2.2$). There was an equal split (45%) in agreement and disagreement regarding expectations of the device ($Q10 - \text{mean} \pm \text{std} = 4.0 \pm 1.8$). A slight majority (55%) believed that regular use of QQ2 could provide benefits ($Q12 - \text{mean} \pm \text{std} = 4.4 \pm 1.8$), and most (73%) agreed, somewhat agreed, or strongly

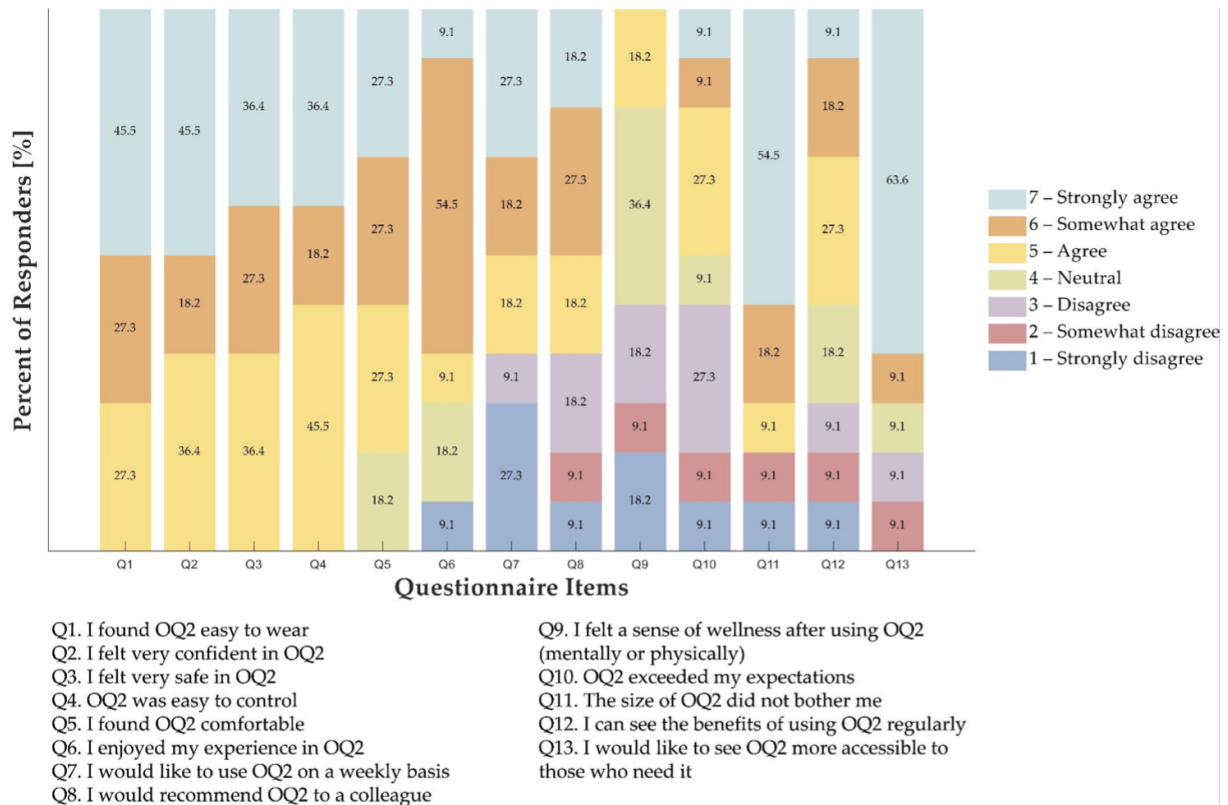


Figure 4.21: Physiotherapists' ratings on the acceptability questionnaire.

agreed that OQ2 should be more widely accessible to those who need it (Q13 – mean±std=5.8±1.9).

Usability questionnaire

Results from the usability questionnaire (n=11), a 1–5 scale, are reported in Figure 4.22. The majority of physiotherapists (73%) agreed or strongly agreed that sessions involving video games were more engaging (Q1 – mean±std=4.0±1.2). Additionally, 63.7% found the developed VR application interesting (Q2 – mean±std=3.5±1.1) and expressed a desire to continue using the games if they could (Q4 – mean±std=3.4±1.6). More than half (55%) agreed or strongly agreed that the physical therapy protocol created in the VR environment met a real need for patients with shoulder diseases (Q3 – mean±std=3.2±1.5). However, only 45% of physiotherapists indicated that they would use the VR system for home rehabilitation (Q5 – mean±std=3.3±1.6). All physiotherapists agreed or strongly agreed that assistance was crucial for using OQ2 and engaging with the VR rehabilitation app (Q8 – mean±std=4.3±0.5). The majority (82%) felt that the games were intuitive and easy to understand (Q6 – mean±std=4.3±0.8). However, only 45%

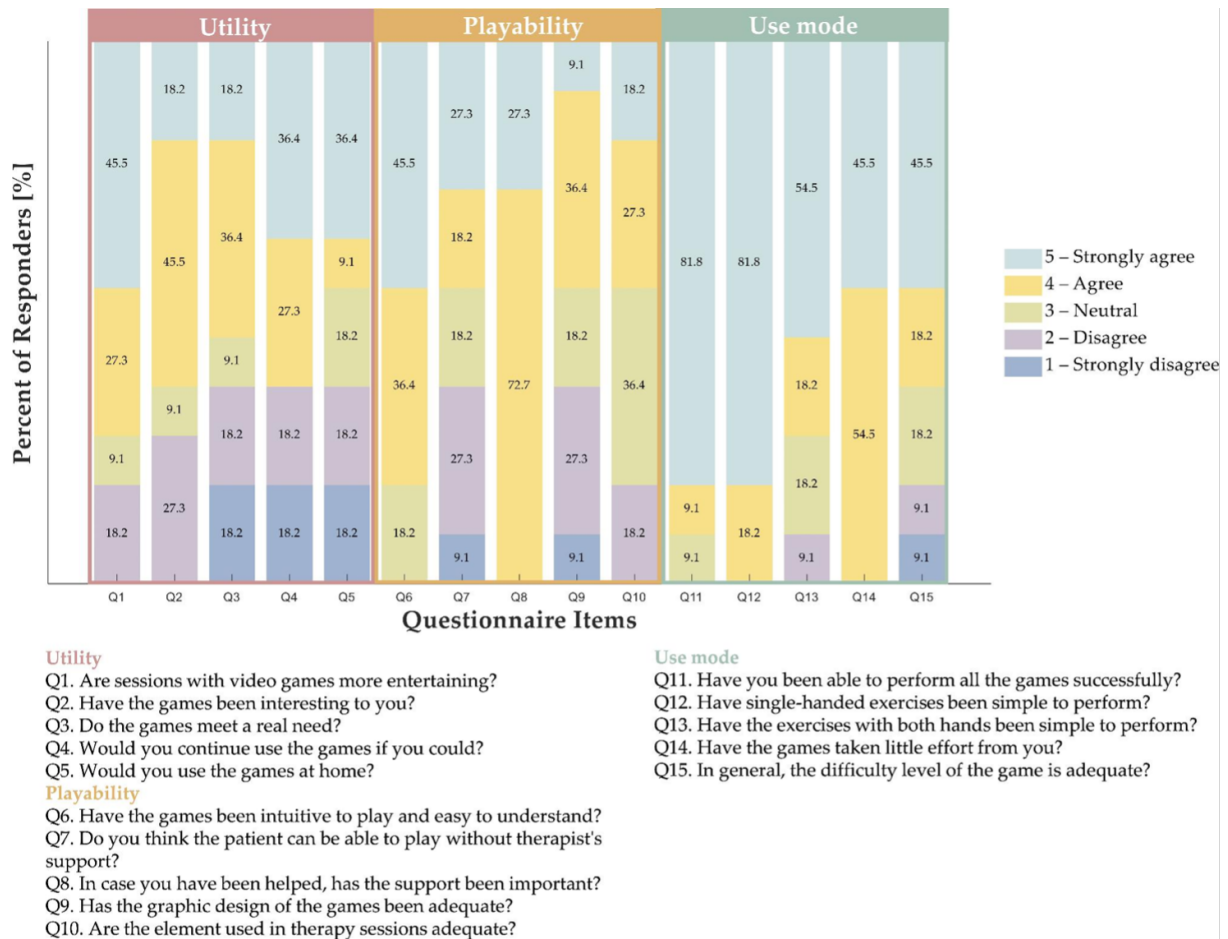


Figure 4.22: Physiotherapists' ratings on the usability questionnaire.

of physiotherapists agreed and strongly agreed that patients can be able to play without the therapist's support (Q7 – mean±std=3.3±1.4), the graphic design of the games was adequate (Q9 – mean±std=3.1±1.2) and the elements used in the therapy sessions were adequate (Q10 – mean±std=3.4±1.0). All the physiotherapists found that the games required little effort from their (Q14 – mean±std=4.4±0.5), and the single-handed exercises were easy to perform (Q12 – mean±std=4.8±0.4). The exercises with both hands were considered simple by 73% of physiotherapists (Q13 – mean±std=4.2±1.1). A majority (82%) successfully completed the physical therapy program (Q11 – mean±std=4.7±0.6), and 64% felt that the difficulty level of the game was appropriate (Q15 – mean±std=3.8±1.4).

Appropriateness questionnaire

Figure 4.23 shows the appropriateness questionnaire results for the multiple-choice questions (Question 1: For which patients would you use OQ2; Question 2: When would you use the OQ2 in the treatment? Question 3: In which environment would you use the OQ2?).

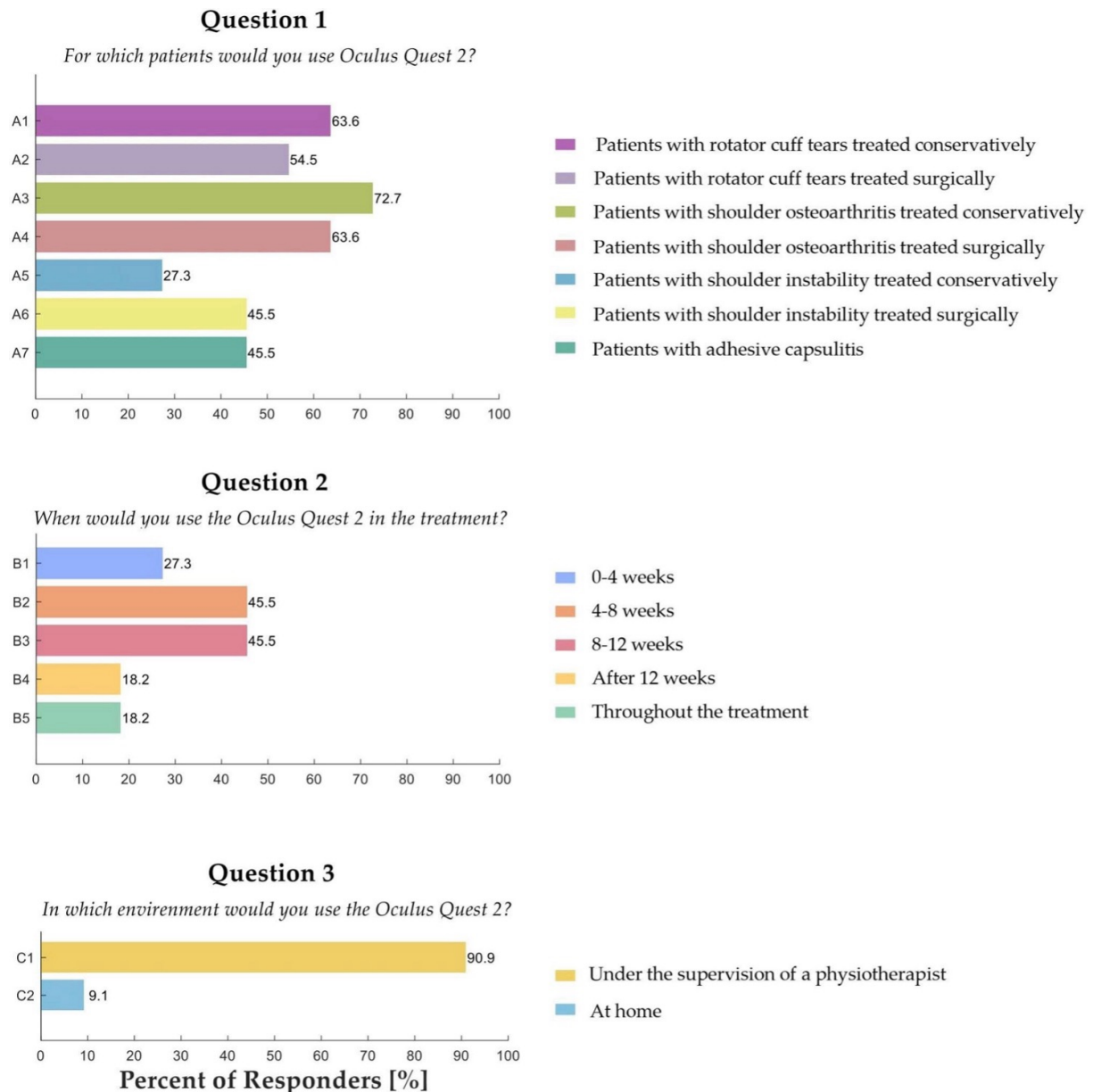


Figure 4.23: Physiotherapists' ratings on the appropriateness questionnaire.

The results indicated that 63.6% of the physiotherapists would prefer to use the OQ2 system for patients with conservatively treated RC tears or for those with shoulder osteoarthritis fol-

lowing surgical treatment. Regarding the treatment phase, 45.5% of physiotherapists expressed interest in using the VR application during the mid-rehabilitation phase (4–8 weeks) and the late rehabilitation phase (8–12 weeks). Additionally, 91% of the physiotherapists believed that patients should use OQ2 under the supervision of a therapist rather than independently at home. In response to the open-ended question, “*What do you think about your experience with OQ2? Do you have any suggestions?*”, the experimenters systematically collected and analyzed the physiotherapists’ feedback, which focused on categories such as namely exercise correctness, graphic design, user autonomy, and engagement. About 63% of physiotherapists expressed doubts regarding the accuracy of the exercises, pointing out that head and trunk rotations were necessary to complete the movements, which did not isolate shoulder motions like flexion-extension, adduction-abduction, and internal-external rotation. Some physiotherapists also noted that performing the exercises, exclusively, in a standing position could lead to patient fatigue. Overall, physiotherapists rated the graphics of the developed VR app as elementary but pleasant. The physiotherapists suggested improvements in the app design, including the addition of an avatar that mirrors the patient’s movements and helps guide them through the exercises. According to the physiotherapists’ feedback, the OQ2 system showed potential as a valuable tool for shoulder musculoskeletal rehabilitation, offering both comfort and an engaging experience for patients. They believed that the system could enhance patient involvement in their rehabilitation process, making therapy more interactive and motivating. However, the physiotherapists also pointed out that some improvements were necessary to enable patients to use the system independently, without constant supervision from a therapist. In particular, increasing user independence would allow the extension of the use of the system for home rehabilitation. In addition to these suggested improvements, further research was needed to carefully evaluate the clinical effectiveness of the OQ2 system in patients with shoulder musculoskeletal disorders.

Chapter 5

QoE and the art “The VR effect”

5.1 EHEM-Enhancement of Heritage Experiences: The Middle Ages.

In response to the feedback received during the revision of the thesis, this chapter has been enriched with a methodological analysis of the questionnaires, a stratified analysis of the collected data and further considerations on usability and user experience.

User QoE in VR art environments is defined by the depth of engagement, interactivity, and emotional connection that users feel when exploring digital recreations of artworks or historical sites. VR art transforms passive observation into active participation, allowing users to immerse themselves in artistic spaces, examine details up close, and interact with virtual objects. This immersive nature amplifies sensory experiences and encourages deeper emotional involvement, making art exploration more interactive, educational, and impactful compared to traditional methods. The intersection of VR and art has opened new possibilities for exploring and preserving cultural heritage, offering users immersive experiences that enable them to engage with historical and artistic works in ways that were previously inaccessible. One such project embodying this concept is Enhancement of Heritage Experiences (EHEM), which was an important part of my doctoral research and was developed in collaboration with the University of Barcelona, the University of Tuscia, Roma Tre University, and CYENS – Centre of Excellence in Cyprus. The primary aim of the EHEM project was to create virtual recon-

structions of medieval artistic heritage, with a particular focus on architecture adorned with mural paintings. These reconstructions aimed to be as accurate and faithful as possible, reflecting the original appearance of these sites at different historical moments. The lifecycle of these buildings was digitally reproduced across all their temporal phases. Additionally, the 3D model included detailed information on the different materials, techniques, and restoration interventions that shaped the monuments over the centuries. This approach not only facilitated the visualization of the evolution of cultural heritage but also provided a deeper understanding of how these spaces might have looked across various historical periods. This was achieved through a technological tool that ensured their preservation, allowed for regular updates by incorporating new research, and disseminated information tailored to various levels of interest and expertise. The three monuments that served as test cases for the project share a rich history of architecture and paintings were: Santa Maria Antiqua (Rome, Italy) which presented as a major challenge up to ten layers of pictorial decorations, the Enkleistra of Agios Neophytos (Cyprus), also exhibited successive architectural and decorative phases from the twelfth to the sixteenth centuries, and Sant Quirze de Pedret (Catalonia, Spain), two layers of frescoes was challenging for visitors, since most of the paintings were detached and are now displayed in different museums. Therefore, all these monuments involved multiple architectural phases and superimposed layers of pictorial decorations that should be displayed in a chronological sequence to facilitate a better understanding of them. At the start of the project, consortium members were asked to define user stories describing desired functionalities within the applications. User stories are brief descriptions of features, written from the end-user's perspective, with the primary goal of focusing on user needs and the value the software provides. Users were divided into two categories: civilians (general public) and experts/professionals, including restorers and curators. Using these stories as a basis, the following user requirements were identified for the various applications:

- **Free Exploration:** users should be able to navigate freely within the monument's 3D virtual representation environment;
- **Real-Time Interaction:** the application should facilitate real-time interaction with the monument's elements;

- **Painting Information Access:** users would need tools to access information about specific paintings within the monument, including details such as subject matter, creation date, and historical context;
- **Localization:** the application should provide a user localization feature within the monument's 3D environment;
- **Mural Details:** users should be able to access information regarding murals, including their representation, dating, and original state;
- **Temporal Navigation:** users should have the ability to virtually change the time and date within the monument's 3D environment;
- **Temporal Visualization:** experts would need the ability to view the monument at different points in time;
- **Image Retrieval:** the application would allow experts to retrieve relevant photographs associated with specific points of interest within the monument;
- **Interactive Editing:** experts would require tools to select specific areas of interest within the monument and edit certain features, such as applying virtual lights;
- **Advanced Painting Search:** the application would facilitate complex searches for paintings, including functionalities like displaying multiple points of interest simultaneously based on a user-defined question.

These user requirements provided a foundation for developing the 3D monument exploration applications, ensuring they catered to the needs of both civilian visitors and expert users. However, a raw 3D model alone, even if it represented the complete temporal evolution of architecture and decoration, was insufficient for researchers to analyze, interpret, or compare hypotheses about a monument. It also did not enable the general public to fully understand the monument and its significance. To address this, models needed to be enriched with additional assets, enhancing their value for study, documentation, and dissemination. This project involved enriching the model with the following digital assets:

- **Graphic Materials:** these included photographs, drawings, maps, and documents for historical, artistic, educational, and research purposes. They provided visual documentation of architectural and decorative phases and were often geo- and time-referenced to specific parts of the monument;
- **Annotations:** essential for cultural heritage, annotations consisted of text, images, and multimedia attached to monument images (usually photos), aiding in understanding and preservation, particularly for monuments with layered artworks and complex iconography;
- **Storytelling:** designed for various users, from art historians to the general public, storytelling involves a structured tour through different monument locations and historical phases. It included multimedia elements like text, audio, and video, allowing for interactive and immersive exploration.

Furthermore, navigating a 3D model, especially of monuments, is challenging. Users may get lost, struggle to reach specific parts of the monument, or face difficulties with understanding where the main areas are and how to access them. Guided tours and predefined points of view are essential tools to simplify this navigation. These tools require content creators, such as art historians, to navigate the 3D model and save camera settings (location, orientation) for future use. Predefined camera positions can be linked to points of interest (so that clicking on a point of interest automatically moves the camera to that location) or integrated into various parades that structure a guided tour.

In this framework, the Department of my University, in collaboration with the Department of Archaeology and Digital Heritage, was responsible for the digitalization of the church of Santa Maria Antiqua. The main objectives of the project included:

- Evaluate **Usability:** understanding how easily users interacted with the application, navigated its features, and accomplished tasks; navigate its features, and accomplish tasks;
- Assess **User Satisfaction:** gathering feedback on user satisfaction, preferences, and any health issues encountered during usage;

- Validate **Design Choices**: determining whether the application design aligned with user expectations and needs;
- Identify **Areas for Improvement**: identifying aspects that require refinement or improvement.

5.1.1 Digitalization of Santa Maria Antiqua

Projects like EHEM not only preserve historical art for future generations but also democratize access to cultural heritage, making it possible for people around the world to explore medieval architecture and art without the constraints of location or physical problems. These types of reconstructions integrate historical and artistic knowledge along with a diachronic perspective of heritage, serving as a tool for researchers, restorers, and heritage curators. For Santa Maria Antiqua, an immersive virtual experience was developed for use with Oculus Quest 2 headsets¹. The project was created using Unreal Engine 5, with code written in C. The exploration of the 3D monument started on a specially designed web landing page, featuring a timeline that guided users through the experience (Figure 5.1). The virtual reconstruction setup allowed users to independently explore the environment, focus on points of interest, and switch between viewing the current state of the environment and a reconstructed version of the apsidal wall (Figure 5.2 and Figure 5.3). A reconstructive panel was also included, featuring descriptions of the various pictorial phases to provide information on the stratigraphy of the walls (Figure 5.4).

5.1.2 The evaluation experiment

In addition to creating historically accurate representations, a key focus of the project was to assess the QoE of the users, taking into account the various influencing factors and dimensions that may impact the overall experience, with particular attention to presence, immersion, and discomfort [19]. Presence refers to the feeling of being physically located in the virtual environment. A high level of presence enhances the connection between the user and the virtual reconstruction, making the experience more meaningful and impactful. In art-related VR

¹EHEM: Santa Maria Antiqua 3D model on Sketchfab.

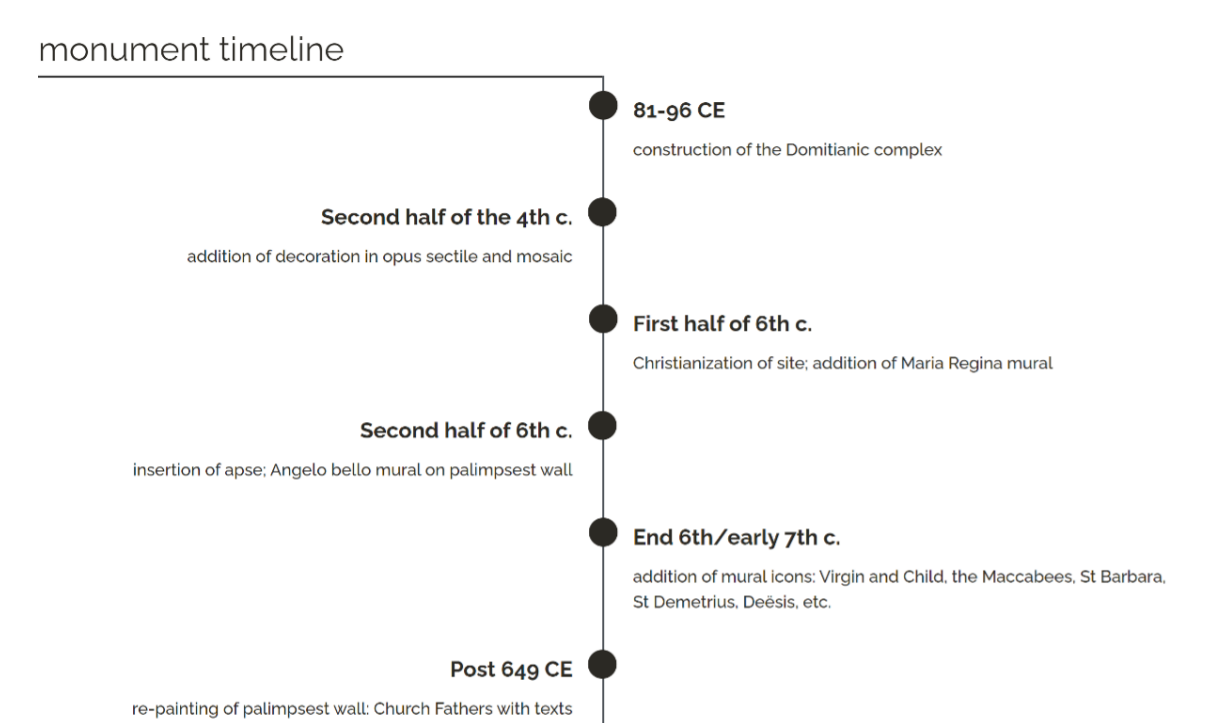


Figure 5.1: Timeline guide.

projects, the sense of presence allows users to feel as if they are standing within the historical space, experiencing the architecture and murals as though they were there. Immersion is related to the depth of engagement a user feels with the virtual world. In the context of medieval art and architecture, immersion means more than just visual fidelity; it involves creating an experience that fully captures the user's senses and attention. This can include using detailed textures in the mural paintings, lighting that accurately reflects the historical environment, and interactive features that let users explore the art. These elements help users better understand the cultural importance of the artwork. Discomfort in VR can happen due to things like feeling sick from motion, difficulty using the system, or confusing design. In art projects like EHEM, it's important to reduce discomfort so that users can fully enjoy the experience without being distracted, particularly because, in the VR app of Santa Maria Antiqua, users could explore the virtual scene through head and eye movements. Since the goal is to engage people with the art and history, any discomfort could take away from their ability to appreciate the details and beauty of the work. When these elements are balanced, the result is a rich, immersive exploration of cultural heritage that allows users to step into the past and experience



Figure 5.2: User in the VR reconstruction of Santa Maria Antiqua.



Figure 5.3: Virtual reconstruction of Santa Maria Antiqua and points of interest of the apse.

art in its historical context. Moreover, the evaluation of these elements through user feedback allows researchers to continually improve the virtual experience, ensuring that it is both educational and engaging. The subjective evaluation involved a group of participants interacting



Figure 5.4: Virtual reconstruction of Santa Maria Antiqua and annotations.

with the Santa Maria Antiqua VR application. The experiment took place at the Department of Humanities (Roma Tre University), and participants were chosen to cover a broad demographic spectrum in terms of age and education levels, and backgrounds. Two OQ2 devices were utilized, each set up in a separate navigation spaces. Participants were recruited through a call shared via the department's social media channels. The navigation in the application was controlled through head and eye movements. By moving their heads, users were able to control a pointer that enabled them to reach and zoom in on specific points of interest. The experiment had no fixed duration, allowing users to stop whenever they wished. On average, the time spent interacting with the VR application was approximately 10 minutes. Before starting, participants had a brief training phase where they were taught how to use the VR equipment, activate the teleport feature, and navigate within the virtual environment. They were also informed about the purpose of the experiment and provided with an overview of the application's content. Once familiarized with the system, participants were encouraged to explore the virtual version of Santa Maria Antiqua and interact with the virtual environment. After completing the tasks, participants provided detailed feedback through interviews. They were asked to complete three questionnaires: a general questionnaire assessing the over-

all experience, the SSQ [101], and the SUQ [102]. Additionally, comments, suggestions, and potential improvements were also collected.

5.1.3 Structure of the questionnaires and method of administration

To assess the QoE of users within the VR application dedicated to the reconstruction of Santa Maria Antiqua, three questionnaires were employed: a general demographic questionnaire, the SSQ, and the SUQ. These tools were selected due to their relevance in scientific literature for evaluating immersive virtual environments in terms of both effectiveness and usability. A moderately noisy environment was selected to simulate the potential in-situ use of the application, recreating conditions where visitors might explore Santa Maria Antiqua while others are using the VR app. The general questionnaire aimed to collect demographic data (age, gender, educational background) and preliminary information about participants' previous experience with virtual reality and with the archaeological site featured in the simulation. Additionally, participants were asked to rate their overall satisfaction with the experience and their interest in visiting the real site, using a 5-point Likert scale. The questionnaires were administered with Google Forms immediately after the VR experience. The SSQ, consists of 16 symptoms commonly associated with cybersickness, grouped into three subscales: oculomotor symptoms, disorientation, and general discomfort. Participants were asked to rate each symptom on a four-point scale: none, slight, moderate, or severe. This tool was essential in identifying any discomfort caused by the use of the OQ2 headset, particularly given the navigation system based on head and eye movements. The SUQ was used to evaluate the perceived usability of the application. The 10 questions of the SUQ were rated on a 7-point scale (1 = strongly disagree, 7 = strongly agree), and addressed aspects such as ease of use, functional integration, need for technical support, system consistency, and ease of learning. A final usability score was calculated following standard procedures, providing an overall usability index useful for future comparisons. The combination of subjective assessments (questionnaires) and behavioral observations allowed for a more comprehensive and reliable evaluation of QoE, integrating personal perceptions with qualitative data regarding user-system interaction.

5.1.4 Analysis of Results

The questionnaires proposed to users, along with the observation of participants' behavior during the experiment, were designed to evaluate how effectively users could access and interact with the cultural content. The analysis conducted provided insight into how well the application presented and engaged users with multimedia content related to the site. Furthermore, by observing participants' interactions and gathering their feedback, valuable insights were obtained regarding the application's strengths and areas for improvement.

The collected data were analyzed to derive meaningful conclusions for:

- **User Satisfaction:** participants' subjective feedback was categorized to identify common themes. Positive feedback indicated user satisfaction, while negative comments highlighted areas requiring improvement;
- **Qualitative Insights:** participants' qualitative feedback offered valuable perspectives on their preferences, challenges, and suggestions.

The participant group consisted of 19 individuals (10 females and 9 males), with an average age of 28.21 ± 8.05 years. The 36.8% of the population had never used VR before, while among the remaining 63.2%, 35.7% felt comfortable using VR (on a scale from 1 to 5), and 21.4% did not feel confident in their abilities. Nearly half of the population (47.4%) had never been to Santa Maria Antiqua but 100% of the users expressed a desire to visit the real site after this VR experience. On a scale of 1 to 5, 94.8% rated the overall experience as 4 and 5, only one person rated it with the score 3.

Simulator Sickness Questionnaire

The SSQ is a tool used to assess symptoms of cybersickness, a form of discomfort or illness experienced by users of VR systems. Participants rate each symptom on a scale ranging from None, Slight, Moderate, to Severe. The symptoms evaluated include general discomfort, fatigue, headache, eye strain, difficulty focusing, increased salivation, sweating, nausea, difficulty concentrating, fullness of the head, blurred vision, dizziness (eyes open), dizziness (eyes closed), stomach awareness, and burping. Regarding the simulator sickness questionnaire,

most participants (84.2%) reported no general discomfort, while 10.5% experienced slight discomfort and 5.3% reported moderate discomfort. Fatigue was not commonly reported, with only 5.3% experiencing slight fatigue. Headaches affected 10.6% of participants, split between slight (5.3%) and moderate (5.3%), while the majority reported no headaches. Eye strain was a more noticeable issue, with 26.3% of participants reporting this symptom. Difficulty focusing was the most common symptom, with 47.4% experiencing slight difficulty and 5.3% reporting a moderate level. Increased salivation affected 10.5% of participants, and 15.8% experienced nausea. Regarding sweating, 10.6% reported symptoms (5.3% slight and 5.3% moderate). Concentration issues were noted by 15.8% as slight and by 5.3% as moderate. For the fullness of head symptom, 15.8% experienced slight discomfort. Blurred vision was reported by 26.3% as slight and 5.3% as moderate. Dizziness with eyes open affected 15.8%, while 5.3% experienced the same issue with eyes closed. No significant symptoms were reported regarding stomach awareness or burping.

System Usability Questionnaire

The questionnaire aimed to assess the usability of the product, determining how effectively, efficiently, and satisfactorily it could be used by specific users to achieve certain objectives within a particular context of use. Answers were given on a scale from 1 to 7 (1 = Strongly disagree, 7 = Strongly agree), providing a single index of perceived usability. The scores and questions are shown in the Table 5.1. After this experiment most participants (12 out of 19) expressed that they would like to use the system frequently, and an equal number found the system easy to use, compared to just one participant who found it unnecessarily complex. For the third question, 68.4% agreed or strongly agreed that the system was simple to use, while 15.3% disagreed, and 15.8% gave a neutral response. Around half of the participants answered that they did not need the support of a technical person to operate the system, while six participants declared the opposite. Almost all users (94.7%) agreed that the system's various functions were well integrated, while 79% of participants believed there was some inconsistency, with only one user was completely agree with this statement. Most participants (16 out of 19) agreed or strongly agreed that the system was easy to learn, while only one participant found the system cumbersome. A large majority (94.7%) disagreed or strongly

disagreed with the notion that the system was difficult to use. Thirteen users felt confident using the system, two were not confident, and three gave a neutral response. Lastly, 73.7% of participants reported that they did not need to learn a lot before using the system, highlighting how user-friendly the system was and how it could be used without requiring specialized courses or training. Finally, the usability evaluation, gave a final score of 80.65. This score reflected the overall ease of use, intuitiveness, and user satisfaction with the application. An 80.65 rating suggested a generally positive user experience, indicating that most participants found the interface accessible and the features functional. Such a score also highlighted areas of strength in the design while leaving room for potential improvements based on specific user feedback. In Table 5.2 some comments and suggestions provided by participants after the experiment are reported.

Table 5.1: Results about SUQ.

SUQ	Average	SD	Variance
1) I think that I would like to use this system frequently	4,53	0,69	0,48
2) I found the system unnecessarily complex	1,31	0,47	0,92
3) I thought the system was easy to use	3,73	1,28	1,64
4) I think that I would need the support of a technical person to be able to use this system	2,63	1,38	1,91
5) I found the various functions in this system were well integrated	4,47	0,61	0,37
6) I thought there was too much inconsistency in this system	1,78	1,18	1,39
7) I would imagine that most people would learn to use this system very quickly	4,47	0,77	0,59
8) I found the system very cumbersome to use	1,63	0,95	0,91
9) I felt very confident using the system	3,89	1,10	1,21
10) I felt very confident using the system	1,63	1,30	1,69

Table 5.2: Comments and suggestions by participants.

Comments	Suggestions
The device is useful and interesting. It allows you to immerse yourself, firsthand, looking at details that you normally don't have the opportunity to observe.	Nothing
Very interesting and fun	A more precise order of consultation would be useful
Interesting experience	Propose its use again at the university with other reconstructions
The reproduction of Santa Maria Antiqua is an excellent teaching tool both for those who study art and for those who have never approached the subject.	it would be interesting to improve the resolution quality of the inscriptions
It was a truly evocative experience: a sort of journey through time and materiality.	For the first part: the altitude and the sensation of feeling pressed against the priest in the apse are a bit annoying. For the rest there is nothing to say, it is impeccable!!!
Crazy	Being able to see the various decorative phases in the periods
Beautiful experience that allows you to see details and details invisible in real life	expand the project into other spaces
Super interesting	I have no suggestions to make
Excellent experience	Nothing

Table 5.2: (continued)

Comments	Suggestions
The graphics are one hundred percent immersive, and this allows even those who have not visited the church to be able to enjoy the unique experience that was experienced first and foremost by those who worked on the palimpsest wall and then also by those who carried out the project!	Integrate both experiences
A fun and very interesting experience to complete the use of the monument by fully understanding it.	I would like to see the church rebuilt in its entirety
Involvement	It would have been more pleasant with background music
To improve	Perfect like this
Great app, I felt really inside the experience	I have no suggestions to make.
Surprising and fascinating	-
Unique experience	-
Good experience.	-

5.1.5 Stratified Analysis of User Experience

To better understand the influence of individual factors on the QoE, a stratified analysis was conducted based on users' age and their prior experience with VR. The total number of participants ($n = 19$) was divided into subgroups as follows: 13 users aged 25 or younger, and 6 over 25 years old; 7 were first-time VR users, while 12 had used VR before (Table 5.3). This exploratory analysis aimed to identify trends and differences in usability perception and discomfort symptoms across distinct user profiles.

Grouping Criteria Age: Participants were split into two age groups:

- Group A: Users aged 25 or younger (n = X)
- Group B: Users over 25 years old (n = Y)

Grouping Criteria VR Experience:

- Group C: First-time VR users (n = 7; 36.8%)
- Group D: Users with prior VR experience (n = 12; 63.2%)

Preliminary observations suggest that users with prior VR experience (Group D) reported higher confidence levels in using the system, as reflected in their SUQ responses. Specifically, users in Group D were more likely to:

- Agree that the system was easy to use;
- Feel confident during interaction;
- Require less technical support.

On the other hand, first-time users (Group C) showed slightly higher levels of cybersickness symptoms, particularly in categories such as eye strain and difficulty focusing—consistent with findings in previous VR usability studies. Regarding age differences, younger participants (Group A) appeared more comfortable with the navigation system based on head and eye movement. Older users (Group B) occasionally expressed initial hesitation during the training phase but showed no major usability issues once familiarized with the interface. However, discomfort levels (as measured by the SSQ) did not vary significantly between the two age groups. Despite the limited sample size, this stratified analysis offers preliminary evidence that prior exposure to VR technologies can enhance both usability and user comfort. These findings support the inclusion of brief training or familiarization sessions in future applications to reduce entry barriers, especially for first-time users. These results are shown in Tables 5.4 and 5.5.

Usability perception SUQ scores showed clear differences: participants with previous VR experience reported a higher average usability score (83.2) compared to first-time users (75.6). This suggests that familiarity with the medium improves confidence and efficiency

Table 5.3: Summary of the Analyzed Groups.

Variable	Group	Number of users
Age	≤ 25 years	13
	>25 years	6
VR Experience	First experience	7
	Already used	12

Table 5.4: Results of the SUQ.

Group	SUQ Mean	Main observations
≤ 25 years	82.1	Generally higher ratings and more confidence
>25 years	76.4	More variability in responses
First experience in VR	75.6	Greater uncertainty in some questions
Users with VR experience	83.2	More confident and positive responses on ease of use

in interacting with the virtual environment. Similarly, younger users (≤ 25) scored slightly higher on average (82.1) compared to older users (> 25), who showed greater variability in their responses.

Cybersickness symptoms SSQ were more frequent among first-time users, particularly in areas like eye strain, nausea, and disorientation. In contrast, experienced users reported only mild discomfort, with general fatigue being the most common issue. Among age groups, older participants showed slightly higher reports of visual discomfort (e.g., difficulty focusing, dizziness), although differences were not statistically significant due to the small sample size.

These results suggest that both age and prior exposure to VR can influence the subjective

Table 5.5: **R!** (R!)results of SSQ symptoms.

Group	Frequent symptoms	Notes
≤ 25 years	Eye strain, blurred vision	In line with the average
>25 years	Dizziness, difficulty focusing	More discomfort, especially ocular
First experience in VR	Eye strain, nausea, discomfort	More frequent than experienced users
Users with VR experience	General discomfort, minor fatigue	Mild and less frequent symptoms

experience of immersive applications. Brief training sessions before use, or adaptive interfaces for less experienced users, could mitigate discomfort and improve overall QoE.

5.2 Visualizing San Saba

The cultural heritage sector is becoming increasingly competitive and specialized, with employers looking for candidates who have a unique blend of archaeological, art historical, and technological skills. As new technologies emerged, topics related to art and design became more interdisciplinary. Professionals in the field were expected not only to study and promote cultural heritage but also to understand and apply the latest technologies to enhance and communicate their work [103]. For example, even in serious games, there is cross-contamination between the world of gaming and the world of cultural outreach. In gaming, many scenarios were inspired by archaeological and cultural themes, while gaming software and visualization techniques were increasingly used to develop multimedia products aimed at disseminating cultural content. As a result, cultural heritage professionals who aimed to create such multimedia content needed technological expertise, including knowledge of game engines, 3D modeling, and texturing [104], [103]. In this context, the European Union funded projects focused on digital cultural heritage [105]. Among these initiatives, the European platform [106] provided access to more than 50 million books, paintings, films, museum objects, and archival records from across Europe. The 3D-ICONS project [107] focused on creating 3D models and digitizing significant European monuments and buildings. Additionally, the ARCHES project [108] aimed to develop and assess the use of multimedia technologies to promote inclusion for museum visitors with disabilities. These initiatives highlight the growing importance of the intersection between art and technology in the future. Students of archaeology and art history may be inclined to enroll in engineering-based courses for various reasons. First of all, such courses provide a deeper understanding of the technologies and tools involved in creating digital models, which are valuable for studying and disseminating cultural heritage [109], [110]. By learning different technologies such as 3D printing, VR, and computer-aided design, humanities students expand their career opportunities [111], [112]. In addition, the teaching methods associated with technical disciplines help develop essential skills like

problem-solving, critical thinking, and project management [113], [114]. Furthermore, during courses, humanities students have the opportunity to collaborate with engineers and other students from Science, Technology, Engineering e Mathematics (STEM) degrees, encouraging interdisciplinary projects that can inspire new forms of academic research and multimedia installations. Finally, an engineering-oriented school can also enhance the overall portfolio of humanities students. They will learn new techniques and technologies that can lead to more innovative digital products of scientific investigations [115]. However, there was a lack of specific studies addressing the integration of technology into cultural heritage training programs.

In response to this gap, the summer school “Visualizing San Saba: Geo-Technology Applied to Cultural Heritage” was developed. This section is focused on describing the summer school and the outcomes of the experiment conducted to introduce art students to new technologies, highlighting how the interaction between different disciplines provided added value. The program provided cultural heritage students with the opportunity to discuss and learn about new technologies that are increasingly applied in the artistic field. At the end of the summer school, the students created a website featuring the digitalization of the Basilica of San Saba in Rome as the final product. The course was developed based on the idea that short programs can play a crucial role in introducing students to the languages, tools, and programs of the digital world. Such courses help students learn how to work within multidisciplinary teams, which is especially important given the current lack of digital humanities professionals and researchers who can bridge the gap between the humanities and STEM disciplines in the cultural heritage sector. The research questions of the study carried out during the school were:

- RQ1: What are the best practices for designing and implementing an engineering school program for art students?
- RQ2: What specific skills and knowledge do humanities students gain?
- RQ3: How do humanities students’ perceptions and attitudes toward engineering change after attending a STEM oriented school?

Additionally, there is little research on the long-term effects of engineering-based education on the digitization of cultural heritage and the career development of archaeology and art

history students. One of the main contributions of this work was to define best practices for the design and implementation of an engineering-based curriculum for humanities students, as well as an analysis of the impact such programs have on the arts community and the art field in general.

The Basilica of San Saba

The Basilica of San Saba is a medieval church in Rome, Italy, dating back to the 7th century and known for its well-preserved murals. Located on the Little Aventine Hill (see Figure 5.5), it is one of the oldest surviving examples of an Early Medieval Greek church in the city. The summer school *Visualizing San Saba* aimed to promote the site through digital tools, focusing particularly on the 13th-century frescoes in the outer left aisle. A team of graduate and doctoral students virtually reconstructed and restored the panel of Gregory the Great with St. Saba and St. Andrew using advanced visualization technologies.



Figure 5.5: Location of the Basilica of San Saba. Courtesy of Google Maps.

5.2.1 Program description and activities

Implementation

The week-long Summer School introduced M.S. and Ph.D. students to interdisciplinary research bridging humanities, science, and technology. The initiative focused on the archaeological, architectural, and artistic reconstruction of the Monastery of San Saba, aiming to develop digital tools for its dissemination and preservation. Nineteen students from various humanities departments across Italy received training in data acquisition and cultural heritage digitization, ultimately producing outputs such as a virtual tour and a 3D reconstruction of the Basilica. The project-based format was supervised by a multidisciplinary team of academics and experts from institutions including Roma Tre, Tuscia, and Palermo Universities, with experience in 3D, VR, AR, Mixed Reality (MR), and miXed Reality (XR) technologies [116], [117]. The immersive, hands-on setting inside the Basilica itself facilitated experiential learning, consistent with educational research emphasizing the effectiveness of real-world, active engagement [118]. Theoretical activities were carried out in rotating working groups. The full weekly schedule is provided in Table 5.6.

Table 5.6: Technical program of the school.

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
09:00-11:00	School introduction	Grouped activities	Theoretical lecture	Grouped activities	Theoretical lecture	Website development	Group discussion
11:00-13:00	Visiting of the Basilica						Website presentation
14:00-16:30	Basilica's history lecture	Grouped activities	Post-production	Theoretical lecture	Website development		End of the school
16:30-18:00	Theoretical lecture		Data management				
18:00-19:00	Daily meeting	Daily meeting	Daily meeting	Daily meeting	Daily meeting	Daily meeting	

Acquisition tools

The goal of this topic was to introduce art students to one of the primary methods for extracting information from physical objects and environments. Specifically, various tools and technologies were utilized for this purpose:

- **Photogrammetry:** enabled students to create detailed 3D models of artworks and architectural features, supporting both documentation and creative reinterpretation [119].

- **Geographic Information System (GIS):** was used to collect and analyze spatial data via Global Navigation Satellite System (GNSS), helping students integrate geographic information into cultural heritage interpretation [120].
- **Total Station:** allowed for high-precision topographic mapping of the site, introducing students to professional surveying techniques.
- **Omnidirectional camera:** captured 360° imagery for the development of an immersive virtual tour of the Basilica.
- **Thermal imaging camera:** identified moisture and structural issues not visible to the naked eye, contributing to preservation assessment.

Rendering multimedia tools

To process and render the collected data, various software programs are available. In this case, students were introduced to two widely used tools in the field of cultural heritage conservation and restoration. Regarding the rendering step of multimedia data, the students studied the following tools:

- **Photoshop and AutoCAD:** Photoshop was used for editing raster graphics, image manipulation, and photo editing. AutoCAD, commonly used in architecture, engineering, and design, was utilized for drawing, modeling, and labeling. It allows for the creation of precise 2D and 3D blueprints, plans, and technical drawings [121];
- **Web development tools:** to demonstrate how the work could be shared online, students were introduced to the basics of HTML, CSS, and JavaScript. These tools enabled them to contribute to the official website of the Basilica by integrating their work, including the virtual tour and 3D models of the Basilica. The website can be accessed to explore the virtual tour (Figure 5.6) of the Basilica². In Figure 5.7 the opening page is shown.

One lecture focused on how art-related 3D multimedia can be applied to immersive technologies such as VR, AR, and XR, collectively referred to as MR. The goal of this lecture was to

²<https://dhfactory.it/sansaba/>

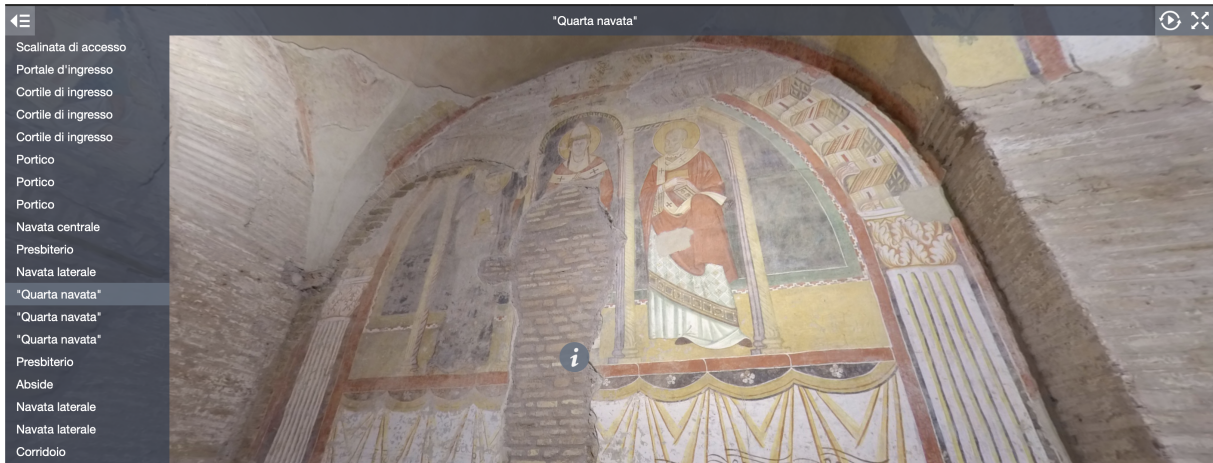


Figure 5.6: A detail of the virtual tour on the website.



Figure 5.7: Homepage of the website..

demonstrate that, unlike asynchronous media such as photos and videos, MR provides a level of engagement closely resembling real-life experiences [117]. However, due to time and hardware limitations, it was not possible to apply these concepts immediately. Future editions of the summer school will emphasize these aspects.

Final project and work presentation

Through a combination of theoretical and practical lessons, students worked on a final project, which was presented to a panel of instructors and industry experts. This project served as an assessment of the students' ability to apply the skills and knowledge gained throughout the course to a real-world scenario.

5.2.2 Program Evaluation and Discussion

To assess the effectiveness of the summer school, a student-oriented self-assessment study was conducted. Specifically, students completed questionnaires at the beginning and end of the course, which included questions about their knowledge and skills regarding the technologies and tools introduced, as well as their overall satisfaction with the experience. The school quality was rated on a scale from 1 to 5, where 1 corresponded to “*unsatisfactory*” and 5 to “*satisfactory*”. Differently, technical knowledge was rated on a scale from 0 to 3, where 0 indicated “*no knowledge*” and 3 indicated “*good knowledge*” of the subject. This section evaluates the self-assessment questionnaires filled out by participants before and after the training. Additionally, students’ feedback was analyzed to provide insights for the next edition of the summer school.

5.2.3 Knowledge before and after the summer school

To address RQ2, students’ self-assessed skills were compared before and after the summer school using pie charts. Most participants had limited prior experience, except in areas like computer science and digital imaging. As shown in Figures 5.8–5.9, the adopted learning model led to a marked increase in perceived competence and confidence. Significant gains were also observed in less familiar domains such as 3D graphics and photogrammetry (Figures 5.10–5.11), despite some students revising their initial self-assessments downward post-training. Improvements in software use, notably Photoshop and AutoCAD, further confirmed the program’s impact (Figures 5.12, 5.13).

5.2.4 Students’ comments

Student evaluations confirmed the overall success of the summer school, with an average satisfaction score of 4.63 out of 5 (Table 5.7). However, feedback (Table 5.8) highlighted a recurring concern about limited time, which received a lower score (4.11). Participants expressed interest in extending the program and in accessing further learning opportunities. Suggestions included offering online resources, workshops, or internships to support continued skill development. Despite this, the training model proved effective in enhancing students’ technical

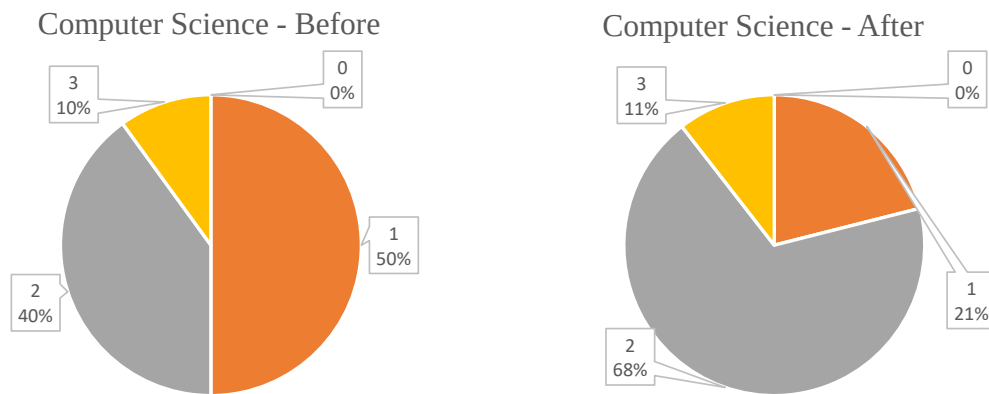


Figure 5.8: Computer science skills before and after the summer school.

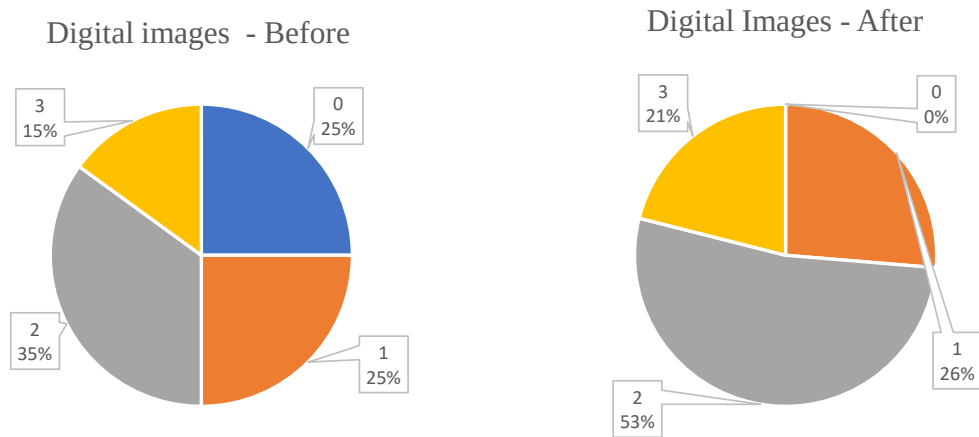


Figure 5.9: Digital image skills before and after the summer school.

competencies and interest in the application of technology to the arts, laying the groundwork for future growth in the field.

5.2.5 Guidelines

Based on the experience, several guidelines can be proposed to support the design of similar training programs (addressing RQ1):

- **Define clear objectives** to guide curriculum design and align activities with intended learning outcomes. Early assessment can help adapt content to student needs.

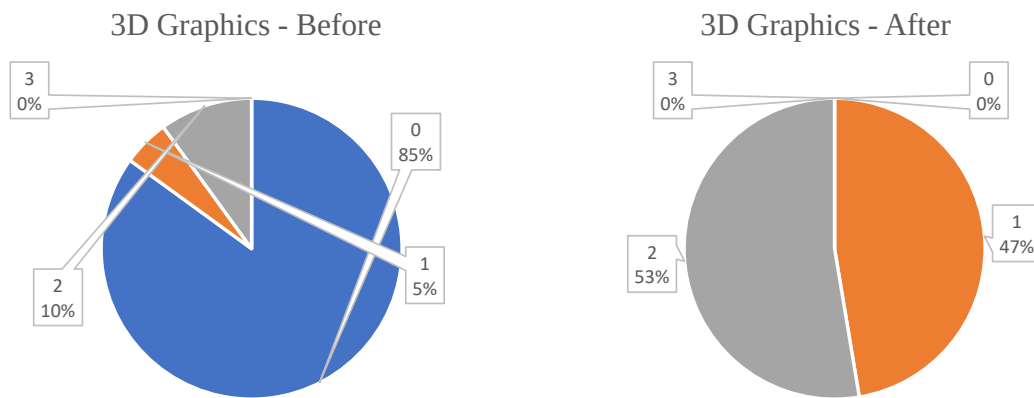


Figure 5.10: 3D graphics skills before and after the summer school

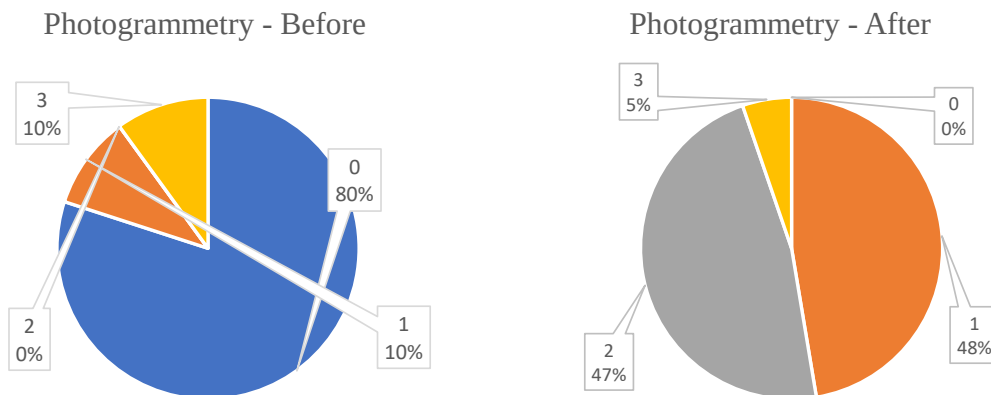


Figure 5.11: Photogrammetry knowledge before and after the summer school.

- **Adopt project-based learning** to connect theory with practice and strengthen technical and collaborative skills.
- **Ensure access to up-to-date technology** to facilitate hands-on training with relevant tools.
- **Promote interdisciplinary collaboration** to foster teamwork and broaden perspectives across domains.
- **Extend or supplement the program** through follow-up activities like online resources or workshops, addressing students' desire for more learning time.

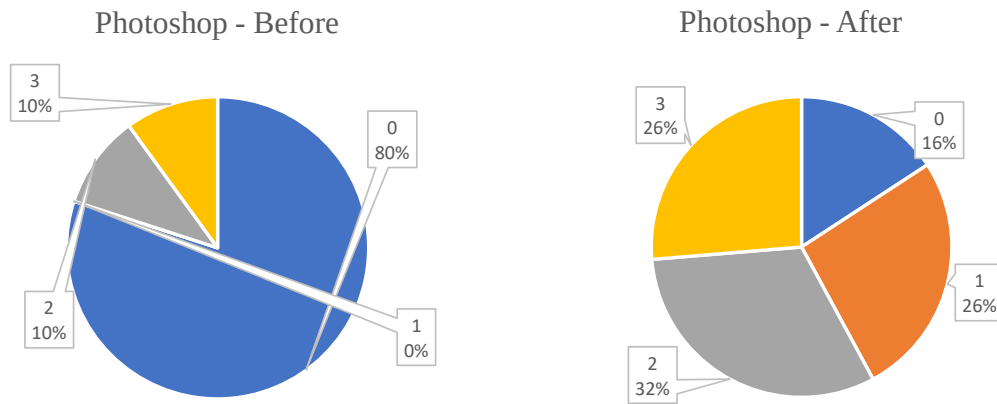


Figure 5.12: Photoshop skills before and after the summer school

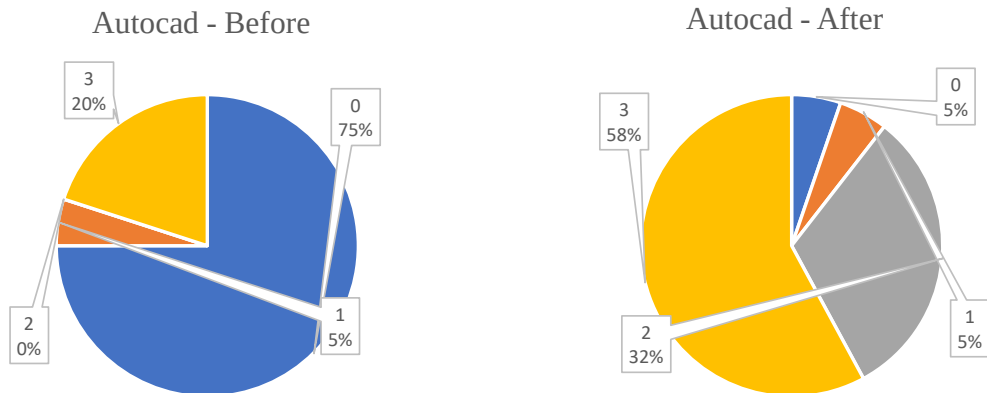


Figure 5.13: Autocad skills before and after the summer school.

- **Regularly evaluate outcomes** to refine the program and enhance future editions.

Table 5.7: Evaluation of the training period by students. Scores are in the range $[1, 5]$.

	1	2	3	4	5	Mean
Achievements of objectives	0	0	1	4	14	4.684
Topics covered	0	0	4	3	12	4.421
Available time	0	1	3	8	7	4.105
Training methods	0	0	1	5	13	4.632
Comfort of the school	0	0	0	7	12	4.632
Organisation of the school	0	0	0	4	15	4.790
Expert contributions	0	0	0	3	16	4.842
Overall satisfactory	0	0	0	1	18	4.947
Average scores						4.632

Table 5.8: Most frequent comments from the students.

Comments
<i>“Deepen the topics by doing more practice.”</i>
<i>“Spend more time on each topic.”</i>
<i>“More time.”</i>
<i>“Provide the material before the start of the school.”</i>
<i>“Stay overnight on-site to maximize the experience.”</i>
<i>“More spacious classrooms.”</i>
<i>“Change the time of year.”</i>
<i>“Less theory and more practice.”</i>

Chapter 6

Discussion

QoE refers to the subjective perception of the users when they interact with digital products, services, or environments. It is a critical metric that refers to the overall satisfaction or dissatisfaction of the users. Unlike traditional QoS measures, focused on latency, bandwidth, or processing power, QoE includes a broader, more subjective assessment of how technology impacts users on a personal and emotional level. Basically, QoE captures how well a system meets the expectations and needs of users, taking into account their comfort, ease of use, immersion, and emotional engagement. QoE has acquired increasing importance in various fields due to its holistic nature. In sectors like telecommunications, media streaming, health-care, and interactive technologies (such as VR, AI, and ML), understanding and optimizing QoE has become essential, as these technologies become more integrated into daily life. This focus is the key to creating systems that are not only functional and efficient but also enjoyable and engaging for users. Evaluating QoE often requires a combination of subjective and objective measures. Objective metrics might include technical data such as frame rates, latency, or resolution, which contribute to the overall performance of a system. However, subjective feedback, often collected through user surveys, interviews, or emotional response measurements, is often essential to understanding the human experience behind the technology. In this context, the research question “**RQ1) Can emotions be recognized through body language?**” was addressed in Chapter 3, which presented a new approach to emotion recognition by analyzing full body movements, focusing on upper and lower body gestures. Traditionally, emotion recognition systems have centered on facial expressions and speech, but the HEROES

dataset explored how emotions can be identified through body postures and motions, an area less studied in AC. The HEROES database is a valuable contribution to the field of AC, which tries to understand and model human emotions using computational methods. By recognizing emotions from body movements, this database offers a new perspective on how emotions can be interpreted through non-verbal cues. These approaches originate from the idea that emotions are best related through the face and voice. As suggested by Mehrabian’s model of communication [26], only 7% of communication is verbal, with 55% of it being non-verbal behaviors, including facial expressions and gestures. While facial expressions have been studied for many years, the role of the entire body in showing emotions is often ignored, even though it is very important. Researchers now recognize that body posture and movement, including both the upper and lower limbs, play a crucial role in expressing emotions [37]. To answer to the **RQ2) “What is the value of a new emotion database for emotion recognition through movement?”** HEROES database is a significant contribution to this growing area of research, focusing on whole body movements to analyze and classify emotional states. This video-based dataset, developed using 16 non-professional actors, aimed to enhance emotion recognition by considering both upper and lower body movements, providing a more comprehensive approach than previous datasets that primarily focused on the face and upper body. The emotions in the dataset were chosen to represent both positive and negative valences, as well as high and low arousal levels, a concept derived from dimensional models of emotion proposed by Russel [122]. The chosen emotions were: happiness, disgust, interest, and boredom. The actors were recorded in two distinct scenarios: an observational setting, such as a museum visit, and an interactive setting involving conversations with another person. This approach was designed to reflect real-life contexts where different emotional responses may be triggered, securing a more natural and authentic representation of emotions. A key feature of the HEROES database is its emphasis on the whole body, including the lower limbs, which were often ignored in emotion recognition studies. For instance, movements of the feet and legs, such as the orientation of the feet towards or away from an object of interest, provide valuable insights into the subject’s emotional state. This aspect remains largely unexplored in existing datasets like GEMEP [44] and FABO [45], which focus more on upper body movements and facial expressions. Another critical aspect of the HEROES database is

the use of non-professional actors. This decision was made to avoid stereotyped or exaggerated expressions often observed in professional performances [47], to aim for more natural and spontaneous movements that better represent real-life emotions. Additionally, the videos were recorded using multiple GoPro cameras from different angles, providing diverse perspectives that enhance the future applications of the dataset for ML and motion analysis. The dataset was also manually annotated by experts to ensure the accuracy of the emotions represented. Four experts annotated each video, focusing only on body movements, as the actors' faces were intentionally blurred to prevent facial expressions from influencing the annotation and to preserve privacy. The resulting accuracy of emotion recognition, based on these annotations, was found to be 79%, which underscores the reliability of HEROES. The dataset was subsequently split into training and testing sets to evaluate the accuracy of emotion recognition models. One of the key innovations introduced is the use of body joints as features for emotion recognition. Through the deep learning model OpenPose, the HEROES dataset tracks 14 body joints, including the nose, neck, shoulders, elbows, wrists, hips, knees, and ankles. By analyzing the velocity, acceleration, and space occupied by these joints, the dataset allows the development of ML models that can accurately classify emotional states based on full-body movements. For instance, the study highlights how emotions like happiness can be detected through expansive, fast movements of the arms and an upright posture, while disgust is associated with defensive gestures such as covering the mouth or neck. Similarly, interest is marked by movements with low velocity, while boredom appears through slow, limited actions and a collapsed posture. The results of the study underscore the importance of considering lower body movements in emotion recognition. The accuracy of the classifier when using the whole body joint information was 94.11%, significantly higher than the 84.31% accuracy achieved when only upper body joints were considered. This demonstrates the critical role that lower body movements play in conveying emotional states and suggests that future emotion recognition systems would benefit from incorporating these movements into their analysis. Furthermore, the dataset's focus on non-professional actors allows a degree of spontaneity and authenticity in the recorded emotional expressions, making it more applicable to real-world scenarios. Unlike professional actors, who may over exaggerate emotional responses, non-professional actors are more likely to demonstrate natural and delicate

emotional cues, enhancing the dataset’s relevance for practical applications. Concerning the inter-rater agreement analysis for the annotation, using Cohen’s Kappa further supports the robustness of the annotation process, the computed κ values ranged from 0.69 to 0.75 across all annotator pairs, indicating a substantial level of agreement according to the standard interpretation scale proposed by Landis and Koch [53]. Such consistency is particularly relevant given the subjective nature of emotion recognition tasks. The results suggest that, despite possible ambiguity in interpreting subtle emotional cues from non-verbal behavior, annotators converged on similar interpretations for most clips. This implies that the emotion categories used were well understood and sufficiently discrete to allow reliable labeling. The substantial agreement achieved across four independent raters strengthens the validity of the dataset and ensures that the emotion labels are not the result of random or inconsistent judgments. Consequently, these annotations can be considered a reliable ground truth for evaluating affective computing models. The statistical validation confirms that the dimensional annotations (valence and arousal) are meaningfully associated with the categorical emotion labels. The one-way ANOVA revealed that both valence and arousal varied significantly across emotional classes, with $F(3, N) = 18.45$ for valence and $F(3, N) = 12.97$ for arousal ($p < 0.001$ in both cases). These results suggest that annotators assigned dimensional values in a consistent manner aligned with the semantics of the emotional states. Specifically, emotions such as *Happiness* received higher valence scores, while *Disgust* and *Happiness* were associated with higher arousal. Conversely, *Boredom* was characterized by low scores in both dimensions. This pattern aligns with existing emotion theory (e.g., Russell’s Circumplex Model), where valence and arousal form a two-dimensional affective space. The coherence of the numerical annotations with these theoretical expectations reinforces the reliability of the data and justifies their use in training affective models based on dimensional representations.

The HEROES dataset has a range of potential applications across various fields, from healthcare to security. In fields like healthcare, entertainment, and education, QoE is critical for determining the success and adoption of new technologies. In healthcare, emotion recognition systems could monitor patients’ emotional states in real time, offering valuable insights into their mental health and well-being. Specifically in healthcare, technologies must not only perform tasks efficiently but also provide patients with an experience that promotes comfort,

confidence, and engagement [62], [123]. Collecting quantitative data on the kinematics of patients with shoulder musculoskeletal impairments is highly beneficial for both clinicians and patients. Interest in wearable devices that can monitor joint movements non-intrusively has increased in recent years, as they offer valuable insights into rehabilitation progress and athlete performance in fields ranging from rehabilitation to sports medicine. For rehabilitation technologies, VR can provide real-time feedback, making exercises more interactive and enjoyable. In virtual rehabilitation settings, QoE is enhanced by providing real-time feedback and creating a motivating and immersive experience that can improve adherence to therapy [90]. For patients recovering from injuries or surgeries, high QoE means easier and more enjoyable rehabilitation, which can be connected to better health outcomes. The immersive experience provided by VR devices can incorporate visual, vibrotactile, or auditory feedback, enhancing rehabilitation by guiding patients through specific tasks [61]. The difficulty level of rehabilitation exercises must be adjusted based on the patient's joint function to ensure a virtual experience that is both gratifying and appropriate for his condition. Regarding the **"RQ3) How can VR be effectively applied in rehabilitation (specifically shoulder rehabilitation) to enhance the QoE of users?"** the studies conducted in collaboration with the Campus Bio-Medico of Rome aimed to develop a VR application designed specifically for patients with musculoskeletal disorders of the shoulder. For this purpose, the OQ2 VR system was used because it does not require external devices to track the HMD and controllers, and it is a wireless system. This ensures high mobility without the need for a restricted workspace. As the OQ2 is a VR tool originally designed for video games, it is essential to evaluate its performance in terms of translation and rotation precision before using it in clinical settings (Chapter 4 of the thesis presents this evaluation in detail). The measurements performed aimed to determine if the OQ2, using an inside-out tracking, is accurate enough to measure position and orientation, as the VR system is planned to be applied in future rehabilitation sessions for patients with shoulder musculoskeletal disorders. To confirm the validity of the results, this analysis was compared with previous studies in the literature that evaluated the accuracy of rotational and translational tracking of VR device controllers by setting up experimental configurations where the position and orientation of the controller can be varied using established methods [77], [78], [124]. The common goal of these studies was the potential application of VR

devices in industrial research or clinical contexts in fact this preliminary evaluation of the device's performance, such as accuracy and precision, was essential before it can be applied in real-world settings. A study evaluating the performance of the Oculus Rift S, which utilizes the same inside-out tracking technology as the OQ2, reported a translational accuracy of 4.36 ± 2.91 mm and a rotational accuracy of $1.13^\circ \pm 1.23^\circ$ for the controller [77]. These results are similar to those from our study, although a larger absolute error in translational accuracy was recorded [77]. Specifically, a higher absolute error of 13.52 ± 6.57 mm was observed at a distance of 500 mm from the HMD along the x-axis, approximately corresponding to the length of a fully extended elbow. This error equates to a percentage error of 2.7%. Given that the OQ2 controllers depend on the camera sensors within the HMD, such results were anticipated. The authors in [77] indicated that optimal tracking performance was achieved when the controller was kept close to the HMD, unlike in our study, where the distance between the HMD and the controller varied. The translational increments were chosen to cover a wide range of anthropometric measurements of the target population, from a minimum distance representing the forearm length of a 5th percentile female to a maximum distance corresponding to the upper limb length (excluding the hand) of a 95th percentile male [81], [80]. In our study, the maximum absolute error for rotational displacements was $1.11^\circ \pm 0.37^\circ$ for a 40° rotation around the z-axis. To evaluate the rotational accuracy, the controller was held about 700 mm away from the HMD, simulating the worst case scenario (the maximum distance) to prevent shorter distances from affecting the results. Furthermore, the maximum error for the OQ2 controller's translational accuracy (13.52 ± 6.57 mm) appeared acceptable for monitoring shoulder joint angles. For instance, with a fully extended arm of a 95th percentile male (700 mm), a displacement of about 20 mm at the hand corresponded to an angular deviation of 1.64° . This is comparable to the maximum MAE for rotational accuracy in this study and is smaller than the errors reported for the Oculus Rift S controller or shoulder joint angle measurements using Kinect-based motion capture. The results of our study are suitable for the intended application: rehabilitation exercises for patients with shoulder joint issues (e.g., RC injuries, frozen shoulder, scapular dyskinesis). The strong accuracy observed in both translational and rotational movements under static conditions provides a solid foundation for future studies evaluating VR devices under dynamic conditions during upper limb movements. So the technological

trade-off between the convenience of wireless VR devices and the requirement for high accuracy has been effectively addressed. Previous research has also explored the positional and translational accuracy of other VR systems, such as the HTC Vive and Oculus Touch controllers, which use an inside-in tracking system. Compared to the results of our study, inside-in systems demonstrate greater accuracy. For example, in [124] HTC Vive controller accuracy below 0.4° for rotation and less than 3 mm for translation is reported. Although these results are better than those of the OQ2 controller, it is still a good option for tracking upper limb movements in clinical settings. After validating the instrument, in this case the OQ2, it was necessary to evaluate the custom VR application designed to guide patients through exercises following RC surgery. As described in Chapter 4, the app's ability to help users during rehabilitation protocol was tested by comparing it to a high-precision stereophotogrammetric system. With the OQ2 Touch controllers, patients' hand movements and gestures can be integrated into the virtual environment, allowing their actions to be represented in an intuitive and immersive way. As already mentioned previously, the use of gaming elements and diverse virtual scenarios has the potential to increase patient compliance, enhancing the overall effectiveness of rehabilitation programs [125], [91], [126]. Furthermore, the adaptability of this tool, combined with ML algorithms, can allow healthcare professionals to personalize rehabilitation programs for patients [126]. This personalized approach can improve rehabilitation results by adapting tasks to each patient's specific needs [91]. Along with this, these devices can also create digital records of rehabilitation sessions. The healthcare experts, having easy access to these records, could focus more attention on patient care, facilitating the evaluation and adjustment of treatment plans. In literature, a double-blind pilot study revealed that manipulated virtual real-time presentation of a subject's movements during shoulder exercises promoted greater active flexion ROM [89]. Another study investigated the influence of game features and practice type, task-oriented or imitation-oriented, on shoulder muscle activity [127]. While VR offers many potential benefits for shoulder rehabilitation, only a few studies have looked into the validity and effectiveness of using virtual environments for orthopedic rehabilitation [89], [90], [128]. None of these studies have assessed the performance of VR applications with established criteria like clinical guidelines or gradual ROM targets. For this reason, the study conducted compared the movements performed at various levels of

the VR app with gold-standard values using a motion capture system. The methodology of the study carried out with Campus Biomedico of Rome, was inspired by a similar approach used in a 2015 study, where the authors combined motion tracking and VR, using the Oculus Rift DK2, to develop a video game [69]. The game guided patients through specific movements that they had to replicate, and their movements were simultaneously monitored to assess if they were performed correctly [69]. However, that study did not involve clinical trials with patients, but experts tested the product and provided feedback. Therefore, while the methodologies are similar, the results are not directly comparable. Although the current results stand alone, as similar findings had not yet been reported in the literature to our knowledge, the VR app demonstrated strong performance when compared to gold-standard values measured through motion tracking. For the first and second series of flexion, the total lower MAE did not exceed 6.4, except one instance where the value reached 12.3. As for the number of times the inferior flexion conditions were not met by the five volunteers, this only occurred once, with a value over 8. For superior flexion, volunteers showed a margin of error in just 6 out of 30 trials, while in the remaining 24 trials, no individual MAE values were recorded. Regarding abduction, there were no errors in the lower condition for level 2 and the upper condition for level 3. However, in levels 3 and 4 for the lower condition, the number of unmet conditions slightly increased, reaching a value of 22, though the total lower MAE did not surpass 8.6. Overall, the results for both flexion and abduction appeared promising, with volunteers frequently achieving the gold-standard target. Internal and external rotations showed similar total MAE values to those observed for flexion and abduction. However, the number of times patients failed to meet the required conditions was higher in these trials. This difference was likely since patients were unable to perform isolated shoulder rotations and instead used their trunk to assist with the internal and external movements. Despite these limitations, given the novelty of the topic and the limited research available, the current results provided a valuable foundation for larger-scale investigations in the future. Factors such as age, specific conditions, pain levels, and emotional state may influence the use of VR applications, but exploring these effects was beyond the scope of this study. However, future research would benefit from examining these factors more closely, as stratifying for them could provide insights into how they impact the use of VR devices in rehabilitation. This would further validate the use of VR

applications for shoulder rehabilitation, particularly following arthroscopic RC repair. In general, more evidence is needed to definitively confirm that VR rehabilitation programs can help patients achieve the same gold standard of movements as traditional rehabilitation programs. However, the current results suggest that this may be possible soon, and further investigation is required to better assess the effectiveness of these applications and their development into highly effective therapeutic tools. The growing impact of VR on shoulder rehabilitation is evident, as shown by the promising results of this study and the increasing number of experimental studies focused on developing these VR applications. Once the device has been validated, the performance of a customized VR app for shoulder rehabilitation has been evaluated by comparing the movements performed by volunteers at different levels of the program with a stereophotogrammetric system considered the gold standard, the last step of this series of studies was to investigate the QoE of some specialized physiotherapists on the use of this custom VR application as described in Chapter 4. The approach followed the standards proposed in ISO 9241-210:2019 (Ergonomics of human-system interaction Part 210: Human-centered design for interactive systems) and followed key steps such as identifying the end-users and context of use, determining user needs and system requirements, designing and developing the solution, and evaluating its effectiveness. The primary contribution of this research was to provide valuable design recommendations for future developments, drawing from the insights of experienced physiotherapists specializing in shoulder rehabilitation. These clinicians identified the challenges that the patients could experience during physical therapy and the potential benefits of using VR in their recovery. The recommendations focus on refining the VR game to eliminate unnecessary movements of the shoulder, improving more autonomous use, and offering feedback to improve user engagement and accuracy during rehabilitation exercises. The focus behind this study was to explore various aspects such as utility, playability, usability, ease of use, ease of understanding, safety, comfort, size, and fit, taking into account the opinions of physiotherapy experts specializing in orthopedic shoulder conditions. In terms of acceptability, the study results showed that physiotherapists had positive views on the use of the VR system for shoulder rehabilitation. Specifically, they found it comfortable, safe, and easy to use. Regarding usability, the physiotherapists noted that the VR system was engaging, intuitive, and easy to understand. However, in terms of playability, not all physiotherapists

believed that the VR application achieved the end users' requirements. In particular, the main difficulty identified was the need to perform additional movements (such as head and trunk rotation, or cervical extension), which did not allow for proper isolation of shoulder movements and potentially could negatively impact on the rehabilitation treatment. This problem revealed another complication: many physiotherapists thought that patients would not be able to use the system independently without external support to prevent incorrect movements unless feedback on exercise accuracy was provided. This observation highlighted the need for a clearer description of the protocol execution and, most importantly, feedback for the therapist to help the patient make corrections. The data provided by the VR system could be used to assess shoulder movements during exercises and establish performance indicators to evaluate how well the protocol was being followed. Furthermore, adding a patient avatar could enhance engagement, improve autonomy in use, and ensure the correct execution of exercises. Regarding appropriateness, physiotherapists considered the OQ2 and the VR rehabilitation program suitable for treating shoulder conditions, particularly for patients with shoulder osteoarthritis who are undergoing conservative treatment. This was because the VR system required active or assisted active movements, while post-surgical patients typically start with passive movements, with active movements being introduced later in the recovery process. Other studies had also assessed the usability of VR systems for rehabilitation, though mostly in different clinical areas, such as post-surgery elbow recovery, stroke rehabilitation, and Parkinson's disease management [100], [129], [130]. Similar to our research, previous studies had evaluated usability aspects such as the game's intuitiveness, clarity, ease of use, comfort, enjoyment, practicality, and modes of interaction. In recent years, only a few studies had explored the use of VR systems to improve joint function in patients with shoulder musculoskeletal disorders. Although these studies are not directly comparable due to differences in the types of VR systems and the use of custom-designed applications, the majority have highlighted the positive impact of VR on shoulder rehabilitation [131], [132], [130], [133]. Moreover, various studies had demonstrated the effectiveness of VR in the rehabilitation of patients with conditions such as spinal cord injuries, chronic stroke, and neurological disorders [99], [134], [135]. These studies had shown that VR can significantly enhance the efficiency of physical therapy and promote motor learning following trauma.

Although used in different fields, VR has a shared goal in both healthcare and cultural heritage: increasing user engagement. It helps patients' motivation to follow the rehabilitation exercises, while also allowing students to explore and understand historical sites. In both cases, VR creates an interactive environment that encourages greater participation and learning. Similarly, VR's application in cultural heritage has opened new opportunities for education and preservation. User QoE plays a fundamental role in defining how users interact with art in VR. In the context of digital and immersive art, QoE is the combination of sensory, emotional, and cognitive factors that collectively define how satisfying, engaging, and impactful an experience is. When applied to VR art environments, these factors are more evident due to the immersive and interactive nature of the environment. VR provides a platform for users to not only view art but to experience it in a way that is deeply personal, investigative, and dynamic. Digitalization of artworks involves the process of converting physical works of art into digital formats, which has obtained significant importance in recent years. Overall, digitalization not only ensures the protection of cultural treasures but also expands their impact by making them available to a much larger audience. In particular, this digitalization offers several advantages as: preservation, accessibility, education, Cultural Heritage, and art restoration. Furthermore, immersion is about how much users feel like they are part of the virtual environment. This happens by creating a sense of presence, where users feel as if they are physically inside the virtual world. In traditional art galleries or museums, the viewer is often confined to passive observation, limited by physical distance, display cases, or the barriers created by space and time. VR, however, breaks these limitations by placing users directly inside the artistic space. VR can transport users to places and times they have never been, recreating historical moments or spaces. Interactivity is another critical factor in the QoE in VR art environments. In a traditional museum setting, interaction with art is usually restricted and the users can look but not touch, with little opportunity for interaction beyond visual appreciation. In contrast, VR allows for multiple forms of interaction, transforming the art. Users can manipulate virtual objects, change perspectives, or even alter the artwork itself, depending on how the VR environment is designed. This level of interactivity not only makes the experience more engaging but also adds an educational dimension. Such interactivity encourages exploration and curiosity, leading to a more active form of learning and

appreciation. Moreover, by interacting directly with the artwork, users can develop a personal connection to it, making the experience feel more intimate and meaningful. The last research question **RQ4) “How do virtual reconstruction projects of artistic sites influence the QoE of users?”** is addressed in Chapter 5. The EHEM project focuses on creating digital reconstructions of medieval architecture and mural paintings, and its approach was based exactly on what had been discussed so far. By integrating 3D models with historical data, it allows users to explore these cultural sites through time, revealing changes in structure, colors, and lighting. The project improves research and public interaction by utilizing VR, providing a stratified understanding of medieval heritage and its evolution. The EHEM project aimed to virtually reconstruct a part of the church of Santa Maria Antiqua, which is usually not very accessible. Furthermore, through this VR application, the user could move in the virtual space approaching specific points of the apse that would not have been possible to appreciate in a normal visit to the church. One of the most useful aspects of VR art is its potential to make art more accessible to a wider audience. Traditional art galleries and museums are often bound by geographical, financial, and physical limitations, preventing many people from engaging with art in person. VR can overcome these barriers by offering virtual tours and exhibitions that can be accessed from anywhere in the world. This allows individuals who might not have the opportunity to visit famous museums or galleries to experience the same works of art in a digital format. In addition to geographical accessibility, VR art can also accommodate individuals with physical disabilities, allowing them to explore spaces that might otherwise be challenging to navigate. For example, individuals who use wheelchairs may find it difficult to access certain galleries or exhibits, but in a VR environment, they can move freely and explore without physical constraints. In addition, in the project concerning Santa Maria Antiqua, the user could interact with the system by reading the captions or information relating to some specific points of the art work made the experience more engaging and instructive. The study aimed to collect user QoE through their comments and through two questionnaires: SSQ and SUQ. In particular, the experiment reported general users’ satisfaction, 94.8% of the participants rated the experience as 4 or 5 out of 5. For the SUQ, most participants found the system easy to use (12 out of 19), with 68.4% agreeing that the system was simple. Additionally, 94.7% felt that the system’s functions were well integrated. A large majority (73.7%) reported that

they didn't need much prior training to use the system, highlighting its simplicity to use. The SSQ evaluated that: most participants (84.2%) reported no general discomfort, while 10.5% experienced slight discomfort and 5.3% had moderate discomfort. The most common symptoms were difficulty focusing (52.7%) and eye strain (26.3%). The stratified analysis of usability and discomfort based on age and prior VR experience offers valuable insights into how individual factors influence user experience in immersive environments. The results suggest that prior exposure to VR plays a significant role in enhancing usability, as users with previous experience (Group D) reported greater ease of use, confidence, and autonomy during interaction with the system. These observations are consistent with the notion that familiarity with interactive technologies can reduce the cognitive and operational load associated with first-time use. Conversely, first-time users (Group C) exhibited slightly higher levels of cybersickness symptoms, particularly visual fatigue and difficulty focusing. This finding aligns with prior literature indicating that initial exposure to VR can trigger stronger discomfort responses due to a lack of habituation. Although the sample size limits the statistical generalizability of the results, the observed trends reinforce the importance of considering user background in the design and deployment of VR applications. In terms of age-related differences, younger users (Group A) generally showed greater adaptability to the head- and eye-based navigation system, while older users (Group B) initially required more time to become familiar with the interface. However, once acclimated, both groups reported comparable levels of comfort, as measured by the SSQ. This suggests that initial usability barriers can be mitigated through brief familiarization or training phases. In summary, these findings highlight the importance of adjusting the onboarding process to match the user's level of experience. They also point out the potential benefits of using flexible interfaces or personalized tutorials, especially in educational or clinical settings where users may have very different backgrounds.

The "Visualizing San Saba" project shows how VR allows students to engage with historical artifacts and sites in an immersive way, creating virtual reconstructions that offer greater understanding of historical preservation. In particular, the students followed interactive lessons to learn how to use different tools that would allow them to digitize a cultural heritage. The students initially had limited knowledge of the technologies being taught, such as 3D graphics, photogrammetry, and software like Photoshop and AutoCAD. After the program, their techni-

cal skills significantly improved, as demonstrated by the comparison between self-assessment questionnaire results before and after the course. A pie chart analysis showed noticeable improvement in students' abilities, particularly in using AutoCAD and Photoshop, with the percentage of students reporting "good knowledge" increasing after the summer school. Students were largely satisfied with the program, with an average satisfaction score of 4.63 out of 5. However, some feedback suggested the need for more practice time, additional material, and the extension of the program duration. They also expressed a desire for more practical activities compared to the theory provided. Comments included requests for providing resources ahead of time, increasing practice opportunities, and changing the year when the course takes place. Time was one of the main limitations noted by students. Despite this, the feedback on the program's design and implementation was positive, indicating that the interdisciplinary approach was valuable and effectively helped bridge the gap between the humanities and engineering disciplines. The discussions highlighted that, although the summer school effectively enhanced the students' skills and knowledge, improvements could be made in managing time and finding a better balance between theoretical learning and practical application. Finally, artistic VR experiences with historical references can be effectively applied in educational settings. This approach can enhance student engagement and provide immersive learning opportunities, enabling students to explore and interact with historical sites and artworks in ways that deepen their understanding and connection to the material. Through the integration of VR, schools can offer a more interactive and enriching approach to history and art education, making cultural heritage more accessible and engaging for learners of all ages.

6.1 Future developments, applications

Future research should focus on expanding the use of HEROES in rehabilitation and cultural heritage contexts. One direction would be to incorporate videos recorded from multiple camera angles, generating a richer and more comprehensive dataset. This approach could significantly enhance automatic emotion recognition by analyzing human behavior and body movements from various perspectives. Multi-angle data can capture subtle movements or behaviors that may not be visible from a single viewpoint, allowing for a more precise under-

standing of the user's physical and emotional state. As the research progresses, more emotions or behaviour could be integrated and analyzed, identifying new features that can help as discriminating factors for these emotions, allowing the application in VR context for a deeper understanding of users' experience in virtual environments. By incorporating additional emotional states such as frustration, joy, engagement, or anxiety and analyzing these alongside behavioral patterns, VR systems could offer a more comprehensive vision of user experiences. Moreover, automatic emotion recognition through body movement can significantly enhance (VR) rehabilitation by providing real-time feedback on patients' emotional states. This integration allows for the development of adaptive VR rehabilitation programs that respond to the emotional needs of patients, increasing their motivation and engagement. By analyzing body movements to detect emotions, therapists can adjust exercises not only to improve physical recovery but also to address psychological well-being. This holistic approach ensures that patients receive a more personalized and effective rehabilitation experience, ultimately leading to better outcomes in both physical and emotional health. In the context of rehabilitation, the VR app for shoulder rehabilitation could be tested on patients with a variety of shoulder conditions, thus expanding its range of applications and assessing its overall effectiveness. Testing the app on individuals with different pathologies, such as RC tears, shoulder impingement, frozen shoulder, or arthritis, would allow researchers to evaluate the app's adaptability to diverse patient needs. This would also allow a detailed assessment of the patients' QoE and the potential benefits of integrating VR into personalized rehabilitation protocols. Moreover, the effectiveness of the VR rehabilitation program could be compared with conventional therapy to explore how it measures up in terms of outcomes, patient engagement, and recovery speed. In the field of cultural heritage, the digitization of artworks could be extended to less accessible monuments of historical and artistic interest. Integrating an automatic emotion recognition system could further enhance the understanding of QoE in this context, avoiding the use of devices that can annoy the users. By analyzing how users emotionally respond to specific artistic and historical elements, it allows for more personalized and impactful cultural experiences in VR. Additionally, artistic VR experiences with historical references can be effectively applied in educational settings to enhance student engagement and provide immersive learning opportunities.

6.2 Limits of the research

Although the research was complex, it nevertheless has some limitations:

- QoE is inherently subjective and varies from person to person, influenced by cultural factors, age, and familiarity with technology. This limits the ability to obtain objective, consistent QoE measurements;
- the research carried out for emotion recognition has considered videos recorded of non-professional actors, which may not be generalizable to broader populations. In fact, they could express the emotions in different way despite of in real-world contexts;
- the same research is limited to four basic emotions, considering that other emotional states could be more challenging and complex. Furthermore, interpreting body language without considering environmental context and individual factors can lead to inaccuracies and misinterpretations;
- the rehabilitation studies were conducted on a small sample of HV in a controlled environment; however, to fully validate the effectiveness of VR applications in orthopedic rehabilitation, further large-scale clinical studies are needed. These future studies should focus on patients with a range of clinical conditions to assess the broader applicability and effectiveness of VR in real-world therapeutic settings;
- in the artistic field, a notable limitation is the sample size. Moreover, considering different scenarios can indeed be valuable and helpful, as it allows for a more comprehensive understanding of potential outcomes and challenges.

Chapter 7

Conclusion

The central theme of this thesis is the QoE, a concept that extends across multiple application domains, from VR for physical rehabilitation to the enjoyment of art and cultural heritage. Across the chapters, it has become evident that QoE is not merely a technical topic linked to the quality of a service or device, but a subjective experience that engages sensory, emotional, and cognitive aspects, influenced by the interaction between the user and the technology. The study demonstrates that emotion recognition should not focus solely on facial expressions or hand gestures but also consider movements of the lower body. Leg and foot movements have proven to enhance the accuracy of emotion interpretation, as they provide subtle, significant cues about a person's emotional state, including body orientation. With further research, body language could evolve as a primary tool for emotional monitoring in VR contexts, enhancing applications in rehabilitation, education, and cultural heritage experiences. In the rehabilitation field, the use of VR has shown significant prospective in enhancing the QoE of the patients. The VR application developed for shoulder rehabilitation demonstrated that immersion in virtual environments not only motivates patients to adhere to their therapeutic protocol but also allows for the collection of precise quantitative data on the movements performed. This translates into an objective and real-time assessment of the patient's condition, improving the interaction between therapist and patient and leading to more effective treatment. In this context, QoE manifests as a combination of satisfaction, engagement, and the improvement of motor abilities. Using VR in the arts and cultural sectors has shown that experiencing art through virtual reconstructions helps people to better understand and appreciate

historical heritage. The “Visualizing San Saba” and EHEM projects confirmed that immersion in virtual environments enables users to interact with historical artifacts and sites in new and engaging ways, offering a personalized and interactive experience. Here, the QoE is not only in the accuracy of the reconstructions but also in the user’s ability to explore, learn, and emotionally connect with the historical context. The variability in emotional responses and sensations experienced by users presents a challenge but also represents a valuable opportunity for continuously improving the solutions offered. In conclusion, this thesis has contributed to demonstrating how VR and QoE can be successfully applied in a variety of contexts, enhancing both emotional engagement and the effectiveness of interactions. The ongoing evolution of VR technologies presents extraordinary opportunities to expand these research fields, making the user experience more dynamic, interactive, and personalized. The future points toward greater personalization of experiences, with the integration of new QoE assessment techniques, ensuring that the end user enjoys a more complete experience in different fields as therapeutic, educational, or cultural settings.

Bibliography

- [1] Geoffrey Gaudi, Bill Kapralos, Kc Collins, and Alvaro Uribe Quevedo, *Affective Computing: An Introduction to the Detection, Measurement, and Current Applications*, pp. 25–43, 01 2022.
- [2] R.W. Picard, E. Vyzas, and J. Healey, “Toward machine emotional intelligence: analysis of affective physiological state,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 10, pp. 1175–1191, 2001.
- [3] Tarun Vashishth, Bhupendra Kumar, Rajneesh Panwar, Kewal Sharma, Sachin Chaudhary, and Mr. Vikas Sharma, *Virtual Reality (VR) and Augmented Reality (AR) Transforming Medical Applications*, 10 2023.
- [4] Suma Dawn, “Virtual reality and augmented reality based affective computing applications in healthcare, challenges, and its future direction,” in *Affective Computing in Healthcare: Applications based on biosignals and artificial intelligence*, pp. 10–1. IOP Publishing Bristol, UK, 2023.
- [5] Sahin Coskun, Gokce Yilmaz, Federica Battisti, Musaed Alhussein, and Saiful Islam, “Measuring 3d video quality of experience (qoe) using a hybrid metric based on spatial resolution and depth cues,” *Journal of Imaging*, vol. 9, pp. 281, 12 2023.
- [6] Andrew Perkis, Christian Timmerer, Sabina Baraković, Jasmina Baraković Husić, Søren Bech, Sebastian Bosse, Jean Botev, Kjell Brunnström, Luis Cruz, Katrien De Moor, et al., “Qualinet white paper on definitions of immersive media experience (imex),” *arXiv preprint arXiv:2007.07032*, 2020.

- [7] Myeongseok Park, Yunsik Cho, Giri Na, and Kim Jinmo, “Application of virtual avatar using motion capture in immersive virtual environment,” *International Journal of Human-Computer Interaction*, pp. 1–15, 09 2023.
- [8] Jie Hao, Zhengting He, Xin Yu, and Andréas Remis, “Comparison of immersive and non-immersive virtual reality for upper extremity functional recovery in patients with stroke: A systematic review and network meta-analysis,” *Neurological Sciences*, vol. 44, no. 8, pp. 2679–2697, 2023.
- [9] Peng-ran Liu, Lin Lu, Jia-yao Zhang, Tong-tong Huo, Song-xiang Liu, and Zhe-wei Ye, “Application of artificial intelligence in medicine: an overview,” *Current medical science*, vol. 41, no. 6, pp. 1105–1115, 2021.
- [10] Salih Beyaz, “A brief history of artificial intelligence and robotic surgery in orthopedics & traumatology and future expectations,” *Joint Diseases and Related Surgery*, vol. 31, no. 3, pp. 653, 2020.
- [11] Christina Tsita, Maya Satratzemi, Alexandros Pedefoudas, Charalabos Georgiadis, Maria Zampeti, Evi Papavergou, Syrago Tsiara, Eleni Sismanidou, Petros Kyriakidis, Dionysios Kehagias, et al., “A virtual reality museum to reinforce the interpretation of contemporary art and increase the educational value of user experience,” *Heritage*, vol. 6, no. 5, pp. 4134–4172, 2023.
- [12] Marina Paolanti, Mariapaola Puggioni, Emanuele Frontoni, Lorella Giannandrea, and Roberto Pierdicca, “Evaluating learning outcomes of virtual reality applications in education: a proposal for digital cultural heritage,” *ACM Journal on Computing and Cultural Heritage*, vol. 16, no. 2, pp. 1–25, 2023.
- [13] F. Ahmed, A. S. M. Bari Hossain, and M. L. Gavrilova, “Emotion recognition from body movement,” *IEEE Access*, vol. 8, pp. 11761–11781, 2020.
- [14] Mario A Rojas-Sánchez, Pedro R Palos-Sánchez, and José A Folgado-Fernández, “Systematic literature review and bibliometric analysis on virtual reality and education,” *Education and Information Technologies*, vol. 28, no. 1, pp. 155–192, 2023.

- [15] Chenli Lin, Yuanyuan Ren, and Aming Lu, “The effectiveness of virtual reality games in improving cognition, mobility, and emotion in elderly post-stroke patients: a systematic review and meta-analysis,” *Neurosurgical Review*, vol. 46, no. 1, pp. 167, 2023.
- [16] Polyana Cristina Alves Brepohl and Higor Leite, “Virtual reality applied to physiotherapy: a review of current knowledge,” *Virtual Reality*, vol. 27, no. 1, pp. 71–95, 2023.
- [17] Andrea Demeco, Laura Zola, Antonio Frizziero, Chiara Martini, Arrigo Palumbo, Ruben Foresti, Giovanni Buccino, and Cosimo Costantino, “Immersive virtual reality in post-stroke rehabilitation: a systematic review,” *Sensors*, vol. 23, no. 3, pp. 1712, 2023.
- [18] Elizabeth Chaplin, Christos Karatzios, and Charles Benaim, “Clinical applications of virtual reality in musculoskeletal rehabilitation: A scoping review,” in *Healthcare*. MDPI, 2023, vol. 11, p. 3178.
- [19] Sara Vlahovic, Mirko Suznjevic, and Lea Skorin-Kapov, “A survey of challenges and methods for quality of experience assessment of interactive vr applications,” *Journal on Multimodal User Interfaces*, vol. 16, no. 3, pp. 257–291, 2022.
- [20] J. Marín-Morales, C. Llinares, J. Guixeres, and M. Alcañiz, “Emotion recognition in immersive virtual reality: From statistics to affective computing,” *Sensors*, vol. 20, no. 18, pp. 5163, 2020.
- [21] Y. Matsuda, D. Fedotov, Y. Takahashi, Y. Arakawa, K. Yasumoto, and W. Minker, “Emotour: Estimating emotion and satisfaction of users based on behavioral cues and audio-visual data,” *Sensors*, vol. 18, pp. 3978, 2018.
- [22] C. Darwin and P. Prodger, *The expression of the emotions in man and animals*, Oxford University Press, USA, 1998.
- [23] Ray L Birdwhistell, *Kinesics and context: Essays on body motion communication*, University of Pennsylvania press, 2010.
- [24] Edmund T Hall and Edward T Hall, *The hidden dimension*, vol. 609, Anchor, 1966.
- [25] A. Mehrabian, *Nonverbal Communication*, Routledge: UK, 2017.

- [26] Albert Mehrabian, "Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament," *Current Psychology*, vol. 14, pp. 261–292, 1996.
- [27] Taciana Saad Rached and Angelo Perkusich, "Emotion recognition based on brain-computer interface systems," *Brain-computer interface systems-Recent progress and future prospects*, pp. 253–270, 2013.
- [28] P. Ekman, "Strong evidence for universals in facial expressions: A reply to russell's mistaken critique," *Psychol. Bull.*, vol. 115, no. 2, pp. 268–287, 1994.
- [29] Robert Plutchik and H Kellerman, *Theories of emotion*, vol. 1, Academic Press, 1980.
- [30] Glenn F Wilson and Christopher A Russell, "Real-time assessment of mental workload using psychophysiological measures and artificial neural networks," *Human factors*, vol. 45, no. 4, pp. 635–644, 2003.
- [31] L. F Russell, J. A. & Barrett, "Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant," *Journal of Personality and Social Psychology*, vol. 76, no. 5, pp. 805–819, 1999.
- [32] D. Glowinski, N. Dael, A. Camurri, G. Volpe, M. Mortillaro, and K. Scherer, "Toward a minimal representation of affective gestures," *IEEE Trans. on Affective Computing*, vol. 2, no. 2, pp. 106–118, 2011.
- [33] M. Coulson, "Attributing emotion to static body postures: recognition accuracy, confusions, and viewpoint dependence," *Jour. Nonverbal Behavior*, vol. 28, no. 2, pp. 117–139, 2004.
- [34] M. de Meijer, "The contribution of general features of body movement to the attribution of emotion," *Journal of Nonverbal Behavior*, vol. 13, no. 4, pp. 247–268, 1989.
- [35] H. G. Wallbott, "Bodily expression of emotion," *European Journal of Social Psychology*, vol. 28, no. 6, pp. 879–896, 1998.
- [36] A. Pease and B. Pease, *The definitive book of body language*, Peace International, 2004.

- [37] F. Noroozi, C. A. Corneanu, D. Kamińska, T. Sapiński, S. Escalera, and G. Anbarjafari, “Survey on emotional body gesture recognition,” *IEEE Trans. on Affective Computing*, vol. 12, no. 2, pp. 505–523, 2021.
- [38] M. Alcañiz Raya, J. Marín-Morales, M. E. Minissi, G. Teruel Garcia, L. Abad, and I. A. Chicchi Giglioli, “Machine learning and virtual reality on body movements’ behaviors to classify children with autism spectrum disorder,” *J Clin Med*, vol. 9, no. 5, pp. 1260, 2020.
- [39] C. Busso, M. Bulut, and CC. Lee et al., “Iemocap: interactive emotional dyadic motion capture database,” *Lang Resources & Evaluation*, vol. 42, no. 4, pp. 335–359, 2008.
- [40] G. Venture, H. Kadone, T. Zhang, J. Grèzes, A. Berthoz, and H. Hicheur, “Recognizing emotions conveyed by human gait,” *Int. J. Soc. Robot*, vol. 6, pp. 621–632, 2014.
- [41] Y. Ma, H.M. Paterson, and F.E Pollick, “A motion capture library for the study of identity, gender, and emotion perception from biological motion,” *Behavior Research Methods*, vol. 38, no. 4, pp. 134–141, 2006.
- [42] P. P. Filntisis, N. Efthymiou, P. Koutras, G. Potamianos, and P. Maragos, “Fusing body posture with facial expressions for joint recognition of affect in child–robot interaction,” *IEEE Robotics and Automation Letters*, vol. 4, pp. 4011–4018, 2019.
- [43] Z. Cao, G. Hidalgo, T. Simon, S. Wei, and Y. Sheikh, “Openpose: Realtime multi-person 2d pose estimation using part affinity fields,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 1, pp. 172–186, 2021.
- [44] D. Glowinski, A. Camurri, G. Volpe, N. Dael, and K. Scherer, “Technique for automatic emotion recognition by body gesture analysis,” in *IEEE CVPRw*.
- [45] H. Gunes and M. Piccardi, “A bimodal face and body gesture database for automatic analysis of human nonverbal affective behavior,” in *Int. Conf. on Pattern Recognition (ICPR’06)*, 2006.

- [46] G. Castellano, S. D. Villalba, and A. Camurri, “Recognising Human Emotions from Body Movement and Gesture Dynamics,” in *ACII*, 2007.
- [47] M. Kipp and J.-C. Martin, “Gesture and emotion: Can basic gestural form features discriminate emotions?,” in *Int. Conf. on Affective Computing and Intelligent Interaction and Workshops 2009*.
- [48] Yoann Baveye, Emmanuel Dellandréa, Christel Chamaret, and Liming Chen, “Liris-accede: A video database for affective content analysis,” *IEEE Trans. on Affective Computing*, vol. 6, no. 1, pp. 43–55, 2015.
- [49] N. Fourati and C. Pelachaud, ““emilya: Emotional body expression in daily actions database,”,” in *Int. Conf. Lang. Resources Eval*, 2014.
- [50] H. G. Wallbott and K. R. Scherer, “Cues and channels in emotion recognition,” *J. Personal. Soc. Psychol*, vol. 51, no. 4, pp. 690–699, 1986.
- [51] Mykhaylo Andriluka, Leonid Pishchulin, Peter Gehler, and Bernt Schiele, “2d human pose estimation: New benchmark and state of the art analysis,” in *IEEE CVPR*, 2014.
- [52] C. Larboulette and S. Gibet, “A review of computable expressive descriptors of human motion,” in *Proc. 2nd Int. Workshop on Movement and Computing*, 2015.
- [53] J. Richard Landis and Gary G. Koch, “The measurement of observer agreement for categorical data,” *Biometrics*, vol. 33, no. 1, pp. 159–174, 1977.
- [54] N. S. Altman, “An introduction to kernel and nearest-neighbor nonparametric regression,” *The American Statistician*, vol. 46, no. 3, pp. 175–185, 1992.
- [55] Lucas Medeiros Souza do Nascimento, Lucas Vacilotto Bonfati, Melissa La Banca Freitas, José Jair Alves Mendes Junior, Hugo Valadares Siqueira, and Sergio Luiz Stevan Jr, “Sensors and systems for physical rehabilitation and health monitoring—a review,” *Sensors*, vol. 20, no. 15, pp. 4063, 2020.
- [56] Huiyu Zhou and Huosheng Hu, “Human motion tracking for rehabilitation—a survey,” *Biomedical signal processing and control*, vol. 3, no. 1, pp. 1–18, 2008.

- [57] Sohrab Almasi, Hossein Ahmadi, Farkhondeh Asadi, Leila Shahmoradi, Goli Arji, Mojtaba Alizadeh, and Hoshang Kolivand, “Kinect-based rehabilitation systems for stroke patients: A scoping review,” *BioMed research international*, vol. 2022, no. 1, pp. 4339054, 2022.
- [58] Bojan Milosevic, Alberto Leardini, and Elisabetta Farella, “Kinect and wearable inertial sensors for motor rehabilitation programs at home: State of the art and an experimental comparison,” *Biomedical engineering online*, vol. 19, pp. 1–26, 2020.
- [59] Gustavo Saposnik, Robert Teasell, Muhammad Mamdani, Judith Hall, William McIlroy, Donna Cheung, Kevin E Thorpe, Leonardo G Cohen, and Mark Bayley, “Effectiveness of virtual reality using wii gaming technology in stroke rehabilitation: a pilot randomized clinical trial and proof of principle,” *Stroke*, vol. 41, no. 7, pp. 1477–1484, 2010.
- [60] José-Antonio Gil-Gómez, Roberto Lloréns, Mariano Alcañiz, and Carolina Colomer, “Effectiveness of a wii balance board-based system (ebavir) for balance rehabilitation: a pilot randomized clinical trial in patients with acquired brain injury,” *Journal of neuro-engineering and rehabilitation*, vol. 8, pp. 1–10, 2011.
- [61] Gordon Tao, Bernie Garrett, Tarnia Taverner, Elliott Cordingley, and Crystal Sun, “Immersive virtual reality health games: a narrative review of game design,” *Journal of NeuroEngineering and Rehabilitation*, vol. 18, pp. 1–21, 2021.
- [62] Alessandra Berton, Umile Giuseppe Longo, Vincenzo Candela, Sara Fioravanti, Lucia Giannone, Valeria Arcangeli, Viviana Alciati, Claudia Berton, Gabriella Facchinetti, Anna Marchetti, et al., “Virtual reality, augmented reality, gamification, and telerehabilitation: psychological impact on orthopedic patients’ rehabilitation,” *Journal of clinical medicine*, vol. 9, no. 8, pp. 2567, 2020.
- [63] Marco Iosa, Merve Aydin, Carolina Candelise, Natascia Coda, Giovanni Morone, Gabriella Antonucci, Franco Marinozzi, Fabiano Bini, Stefano Paolucci, and Gaetano Tieri, “The michelangelo effect: art improves the performance in a virtual reality task developed for upper limb neurorehabilitation,” *Frontiers in psychology*, vol. 11, pp. 611956, 2021.

- [64] Patricia Sánchez-Herrera-Baeza, Roberto Cano-de-la Cuerda, Edwin Daniel Oña-Simbaña, Domingo Palacios-Ceña, Jorge Pérez-Corrales, Juan Nicolas Cuenca-Zaldivar, Javier Gueita-Rodriguez, Carlos Balaguer-Bernaldo de Quirós, Alberto Jardón-Huete, and Alicia Cuesta-Gomez, “The impact of a novel immersive virtual reality technology associated with serious games in parkinson’s disease patients on upper limb rehabilitation: a mixed methods intervention study,” *Sensors*, vol. 20, no. 8, pp. 2168, 2020.
- [65] Belinda Kipping, Sylvia Rodger, Kate Miller, and Roy M Kimble, “Virtual reality for acute pain reduction in adolescents undergoing burn wound care: a prospective randomized controlled trial,” *Burns*, vol. 38, no. 5, pp. 650–657, 2012.
- [66] Carla Winter, Florian Kern, Dominik Gall, Marc Erich Latoschik, Paul Pauli, and Ivo Käthner, “Immersive virtual reality during gait rehabilitation increases walking speed and motivation: a usability evaluation with healthy participants and patients with multiple sclerosis and stroke,” *Journal of neuroengineering and rehabilitation*, vol. 18, no. 1, pp. 68, 2021.
- [67] Umile Giuseppe Longo, Laura Risi Ambrogioni, Alessandra Berton, Vincenzo Candela, Arianna Carnevale, Emiliano Schena, Eugenio Gugliemelli, and Vincenzo Denaro, “Physical therapy and precision rehabilitation in shoulder rotator cuff disease,” *International Orthopaedics*, vol. 44, pp. 893–903, 2020.
- [68] Umile Giuseppe Longo, Arianna Carnevale, Carlo Massaroni, Daniela Lo Presti, Alessandra Berton, Vincenzo Candela, Emiliano Schena, and Vincenzo Denaro, “Personalized, predictive, participatory, precision, and preventive (p5) medicine in rotator cuff tears,” *Journal of personalized medicine*, vol. 11, no. 4, pp. 255, 2021.
- [69] Alejandro Baldominos, Yago Saez, and Cristina Garcia Del Pozo, “An approach to physical rehabilitation using state-of-the-art virtual reality and motion tracking technologies,” *Procedia Computer Science*, vol. 64, pp. 10–16, 2015.
- [70] Alessandra Berton, Sergio De Salvatore, Vincenzo Candela, Gabriele Cortina, Daniela Lo Presti, Carlo Massaroni, Stefano Petrillo, and Vincenzo Denaro, “Delayed rehabilitation protocol after rotator cuff repair,” *Osteology*, vol. 1, no. 1, pp. 29–38, 2020.

- [71] Umile Giuseppe Longo, Sebastiano Vasta, Nicola Maffulli, and Vincenzo Denaro, “Scoring systems for the functional assessment of patients with rotator cuff pathology,” *Sports medicine and arthroscopy review*, vol. 19, no. 3, pp. 310–320, 2011.
- [72] Umile Giuseppe Longo, Alessandra Berton, Philip Michael Ahrens, Nicola Maffulli, and Vincenzo Denaro, “Clinical tests for the diagnosis of rotator cuff disease,” *Sports Medicine and Arthroscopy Review*, vol. 19, no. 3, pp. 266–278, 2011.
- [73] Bruno Bonnechère, Bart Jansen, and S Van Sint Jan, “Cost-effective (gaming) motion and balance devices for functional assessment: need or hype?,” *Journal of biomechanics*, vol. 49, no. 13, pp. 2561–2565, 2016.
- [74] Burakhan Çubukçu, Uğur Yüzgeç, Raif Zileli, and Ahu Zileli, “Reliability and validity analyzes of kinect v2 based measurement system for shoulder motions,” *Medical engineering & physics*, vol. 76, pp. 20–31, 2020.
- [75] Bin Chen, Wenli Wang, Guojiong Hu, Rongzhou Zhong, Xiaoyun Su, Hongping Zhi, and Wenxin Niu, “Concurrent validity of a markerless motion capture system for the assessment of shoulder functional movement,” *Medicine in Novel Technology and Devices*, vol. 15, pp. 100131, 2022.
- [76] Peter Beshara, David B Anderson, Matthew Pelletier, and William R Walsh, “The reliability of the microsoft kinect and ambulatory sensor-based motion tracking devices to measure shoulder range-of-motion: A systematic review and meta-analysis,” *Sensors*, vol. 21, no. 24, pp. 8186, 2021.
- [77] Tyler A Jost, Bradley Nelson, and Jonathan Rylander, “Quantitative analysis of the oculus rift s in controlled movement,” *Disability and Rehabilitation: Assistive Technology*, vol. 16, no. 6, pp. 632–636, 2021.
- [78] Leia C Shum, Bulmaro A Valdés, and HF Machiel Van der Loos, “Determining the accuracy of oculus touch controllers for motor rehabilitation applications using quantifiable upper limb kinematics: Validation study,” *JMIR Biomedical Engineering*, vol. 4, no. 1, pp. e12291, 2019.

- [79] Qualisys, ” .
- [80] John W Garrett, “The adult human hand: some anthropometric and biomechanical considerations,” *Human factors*, vol. 13, no. 2, pp. 117–131, 1971.
- [81] Ronald Huston, *Principles of biomechanics*, CRC press, 2008.
- [82] Deanna H Gates, Lisa Smurr Walters, Jeffrey Cowley, Jason M Wilken, and Linda Resnik, “Range of motion requirements for upper-limb activities of daily living,” *The American journal of occupational therapy*, vol. 70, no. 1, pp. 7001350010p1–7001350010p10, 2016.
- [83] Paul S Myles and James Cui, “I. using the bland–altman method to measure agreement with repeated measures,” 2007.
- [84] Charles A Thigpen, Michael A Shaffer, Bryce W Gaunt, Brian G Leggin, Gerald R Williams, and Reg B Wilcox III, “The american society of shoulder and elbow therapists’ consensus statement on rehabilitation following arthroscopic rotator cuff repair,” *Journal of shoulder and elbow surgery*, vol. 25, no. 4, pp. 521–535, 2016.
- [85] Kostas J Economopoulos and Stephen F Brockmeier, “Rotator cuff tears in overhead athletes,” *Clinics in sports medicine*, vol. 31, no. 4, pp. 675–692, 2012.
- [86] Kevin D Plancher, Jaya Shanmugam, Karen Briggs, and Stephanie C Petterson, “Diagnosis and management of partial thickness rotator cuff tears: a comprehensive review,” *JAAOS-Journal of the American Academy of Orthopaedic Surgeons*, vol. 29, no. 24, pp. 1031–1043, 2021.
- [87] Alexander L Lazarides, Eduard Alentorn-Geli, JH James Choi, Joseph J Stuart, Ian KY Lo, Grant E Garrigues, and Dean C Taylor, “Rotator cuff tears in young patients: a different disease than rotator cuff tears in elderly patients,” *Journal of shoulder and elbow surgery*, vol. 24, no. 11, pp. 1834–1843, 2015.
- [88] Jeffrey D Osborne, Ashok L Gowda, Brett Wiater, and J Michael Wiater, “Rotator cuff rehabilitation: current theories and practice,” *The Physician and sportsmedicine*, vol. 44, no. 1, pp. 85–92, 2016.

- [89] Isabella Schwartz, Ori Safran, Naama Karniel, Michal Abel, Adina Berko, Martin Seyres, Tamir Tsoar, and Sigal Portnoy, “Positive effect of manipulated virtual kinematic intervention in individuals with traumatic stiff shoulder: a pilot study,” *Journal of Clinical Medicine*, vol. 11, no. 13, pp. 3919, 2022.
- [90] Nihan Ozunlu Pekiavas and Nevin Ergun, “Comparison of virtual reality exergaming and home exercise programs in patients with subacromial impingement syndrome and scapular dyskinesis: Short term effect,” *Acta orthopaedica et traumatologica turcica*, vol. 51, no. 3, pp. 238–242, 2017.
- [91] Won Kee Chang, Jong In Lee, Ji Hye Hwang, and Jae-Young Lim, “Post-operative rehabilitation using a digital healthcare system in patients who had undergone rotator cuff repair: protocol for a single-center randomized controlled trial,” *Trials*, vol. 23, no. 1, pp. 667, 2022.
- [92] Monique Wochatz, Sophie Rabe, Tilman Engel, Steffen Mueller, and Frank Mayer, “Scapular kinematics during unloaded and maximal loaded isokinetic concentric and eccentric shoulder flexion and extension movements,” *Journal of Electromyography and Kinesiology*, vol. 57, pp. 102517, 2021.
- [93] Ge Wu, Frans CT Van der Helm, HEJ DirkJan Veeger, Mohsen Makhsous, Peter Van Roy, Carolyn Anglin, Jochem Nagels, Andrew R Karduna, Kevin McQuade, Xuguang Wang, et al., “ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion—Part II: shoulder, elbow, wrist and hand,” *Journal of biomechanics*, vol. 38, no. 5, pp. 981–992, 2005.
- [94] CGM Meskers, Frans CT Van der Helm, LA Rozendaal, and PM Rozing, “In vivo estimation of the glenohumeral joint rotation center from scapular bony landmarks by linear regression,” *Journal of biomechanics*, vol. 31, no. 1, pp. 93–96, 1997.
- [95] Quinn A Boser, Aida M Valevicius, Ewen B Lavoie, Craig S Chapman, Patrick M Pilarski, Jacqueline S Hebert, and Albert H Vette, “Cluster-based upper body marker models for three-dimensional kinematic analysis: Comparison with an anatomical model and reliability analysis,” *Journal of biomechanics*, vol. 72, pp. 228–234, 2018.

- [96] Andreas Kontaxis, Andrea Giovanni Cutti, Garth R Johnson, and HEJ Veeger, “A framework for the definition of standardized protocols for measuring upper-extremity kinematics,” *Clinical Biomechanics*, vol. 24, no. 3, pp. 246–253, 2009.
- [97] AG Cutti, ILARIA Parel, M Raggi, E Petracci, A Pellegrini, AP Accardo, R Sacchetti, and G Porcellini, “Prediction bands and intervals for the scapulo-humeral coordination based on the bootstrap and two gaussian methods,” *Journal of biomechanics*, vol. 47, no. 5, pp. 1035–1044, 2014.
- [98] Brandon Birckhead, Carine Khalil, Xiaoyu Liu, Samuel Conovitz, Albert Rizzo, Itai Danovitch, Kim Bullock, Brennan Spiegel, et al., “Recommendations for methodology of virtual reality clinical trials in health care by an international working group: iterative study,” *JMIR mental health*, vol. 6, no. 1, pp. e11973, 2019.
- [99] Nick Birch, Jon Graham, Tom Priestley, Chris Heywood, Mohamed Sakel, Angela Gall, Andrew Nunn, and Nada Signal, “Results of the first interim analysis of the rapper ii trial in patients with spinal cord injury: ambulation and functional exercise programs in the rex powered walking aid,” *Journal of neuroengineering and rehabilitation*, vol. 14, pp. 1–10, 2017.
- [100] Edwin Daniel Oña, Carlos Balaguer, Roberto Cano-de la Cuerda, Susana Collado-Vázquez, and Alberto Jardón, “Effectiveness of serious games for leap motion on the functionality of the upper limb in parkinson’s disease: a feasibility study,” *Computational intelligence and neuroscience*, vol. 2018, no. 1, pp. 7148427, 2018.
- [101] Shelly L Ames, James S Wolffsohn, and Neville A McBrien, “The development of a symptom questionnaire for assessing virtual reality viewing using a head-mounted display,” *Optometry and Vision Science*, vol. 82, no. 3, pp. 168–176, 2005.
- [102] J Brooke, “Sus-a quick and dirty usability scale. usability eval ind 189: 4–7,” 1996.
- [103] Trine Bille, “Creative labor: Who are they? what do they do? where do they work? a discussion based on a quantitative study from denmark,” in *Careers in creative industries*, pp. 36–65. Routledge, 2011.

- [104] Lawrence Argent, Bill Depper, Rafael Fajardo, Sarah Gjertson, Scott T Leutenegger, Mario A Lopez, and Jeff Rutenbeck, “Building a game development program,” *Computer*, vol. 39, no. 6, pp. 52–60, 2006.
- [105] Vassilis Pouloupoulos and Manolis Wallace, “Digital technologies and the role of data in cultural heritage: The past, the present, and the future,” *Big Data and Cognitive Computing*, vol. 6, no. 3, pp. 73, 2022.
- [106] “The Europeana platform,” <https://digital-strategy.ec.europa.eu/en/policies/europeana>, Accessed: 2023-01-29.
- [107] “3D-ICONS: 3D Digitisation of Icons of European Architectural and Archaeological Heritage,” <http://3dicons-project.eu/activities>, Accessed: 2023-01-29.
- [108] “ARCHES: Accessible Resources for Cultural Heritage EcoSystems,” <https://www.arches-project.eu>, Accessed: 2023-01-29.
- [109] S Reid Warren, “The status of the humanities in engineering education,” *IEEE Transactions on Education*, , no. 2/3, pp. 69–71, 1964.
- [110] Ashwin Mohan, Dominike Merle, Christa Jackson, John Lannin, and Satish S Nair, “Professional skills in the engineering curriculum,” *IEEE Transactions on Education*, vol. 53, no. 4, pp. 562–571, 2009.
- [111] Lucio Tommaso De Paolis, Federica Faggiano, Carola Gatto, Maria Cristina Barba, and Valerio De Luca, “Immersive virtual reality for the fruition of ancient contexts: The case of the archaeological and naturalistic park of santa maria d’agnano in ostuni,” *Digital Applications in Archaeology and Cultural Heritage*, vol. 27, pp. e00243, 2022.
- [112] Augusto Ciuffoletti, “Teaching networks to digital humanists,” *IEEE Transactions on Education*, vol. 64, no. 3, pp. 253–260, 2020.
- [113] Juraj Petrović and Predrag Pale, “Achieving scalability and interactivity in a communication skills course for undergraduate engineering students,” *IEEE transactions on education*, vol. 64, no. 4, pp. 413–422, 2021.

- [114] Marcel M Robles, “Executive perceptions of the top 10 soft skills needed in today’s workplace,” *Business communication quarterly*, vol. 75, no. 4, pp. 453–465, 2012.
- [115] Francesca Matrone, Andrea Lingua, Roberto Pierdicca, Eva Savina Malinverni, Marina Paolanti, Eleonora Grilli, Farbio Remondino, Arnadi Murtiyoso, and Tania Landes, “A benchmark for large-scale heritage point cloud semantic segmentation,” *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 43, pp. 1419–1426, 2020.
- [116] K. Gary, “Project-Based Learning,” *Computer*, vol. 48, no. 9, pp. 98–100, 2015.
- [117] S. Muenster, “Digital 3D Technologies for Humanities Research and Education: An Overview,” *Applied Sciences*, vol. 12, no. 5, pp. 2426, 2022.
- [118] J. S. Krajcik and P. C. Blumenfeld, *Project-Based Learning*, p. 317–334, Cambridge Handbooks in Psychology. Cambridge University Press, 2005.
- [119] J. S. Aber, I. Marzolf, and J. B. Ries, “Photogrammetry,” in *Small-Format Aerial Photography*, James S. Aber, Irene Marzolf, and Johannes B. Ries, Eds., pp. 23–39. Elsevier, Amsterdam, 2010.
- [120] A. Mościcka and A. Zwirowicz-Rutkowska, “On the use of geographic information in humanities research infrastructure: A case study on cultural heritage,” *ISPRS International Journal of Geo-Information*, vol. 7, no. 3, pp. 106, 2018.
- [121] V. Nikolić, N. Smičiklas, and A. Ilić, “Mobile platform usage in creating conservation and restoration documentation,” in *Proceedings of the IV Advanced Ceramics and Applications Conference*, 2017, pp. 257–269.
- [122] A. Mehrabian J. Russell, “Evidence for a three-factor theory of emotions,” *J. Research in Personality*, vol. 11, no. 2, pp. 273–294, 1977.
- [123] Won-Seok Kim, Sungmin Cho, Jeonghun Ku, Yuhee Kim, Kiwon Lee, Han-Jeong Hwang, and Nam-Jong Paik, “Clinical application of virtual reality for upper limb motor rehabilitation in stroke: review of technologies and clinical evidence,” *Journal of clinical medicine*, vol. 9, no. 10, pp. 3369, 2020.

- [124] Kate A Spitzley and Andrew R Karduna, “Feasibility of using a fully immersive virtual reality system for kinematic data collection,” *Journal of biomechanics*, vol. 87, pp. 172–176, 2019.
- [125] Arianna Carnevale, Ilaria Mannocchi, Mohamed Saifeddine Hadj Sassi, Marco Carli, Giovanna De Luca, Umile Giuseppe Longo, Vincenzo Denaro, and Emiliano Schena, “Virtual reality for shoulder rehabilitation: Accuracy evaluation of oculus quest 2,” *Sensors*, vol. 22, no. 15, pp. 5511, 2022.
- [126] Bruno Cunha, Ricardo Ferreira, and Andreia SP Sousa, “Home-based rehabilitation of the shoulder using auxiliary systems and artificial intelligence: an overview,” *Sensors*, vol. 23, no. 16, pp. 7100, 2023.
- [127] Ken Chen and Karen B Chen, “Task-oriented and imitation-oriented movements in virtual reality exercise performance and design,” *Human factors*, vol. 65, no. 1, pp. 125–136, 2023.
- [128] Burak Menek, Devrim Tarakci, Ela Tarakci, and Merve Yilmaz Menek, “Investigation on the efficiency of the closed kinetic chain and video-based game exercise programs in the rotator cuff rupture: A randomized trial,” *Games for Health Journal*, vol. 11, no. 5, pp. 298–306, 2022.
- [129] Asmus Andresen, Jonas Valvik, Mikkel Rosholm, Theo Khumsan, Niels Christian Nilsen, and Ali Adjorlu, “Elbovr: Iterative development of a vr application for post-surgery elbow rehabilitation,” in *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 2023, pp. 138–142.
- [130] Lauren Baron, Vuthea Chheang, Amit Chaudhari, Arooj Liaqat, Aishwarya Chandrasekaran, Yufan Wang, Joshua Cashaback, Erik Thostenson, and Roghayeh Leila Barmaki, “Virtual therapy exergame for upper extremity rehabilitation using smart wearable sensors,” in *Proceedings of the 8th ACM/IEEE International Conference on Connected Health: Applications, Systems and Engineering Technologies*, 2023, pp. 92–101.

- [131] Niamh Brady, Jeremy Lewis, Karen McCreesh, Beate DeJaco, and Joseph G McVeigh, “Physiotherapist beliefs and perspectives on virtual reality–supported rehabilitation for the assessment and management of musculoskeletal shoulder pain: A focus group study protocol,” *HRB Open Research*, vol. 4, 2021.
- [132] M Ora Powell, Aviv Elor, Mircea Teodorescu, and Sri Kurniawan, “Openbutterfly: Multimodal rehabilitation analysis of immersive virtual reality for physical therapy,” *American Journal of Sports Science and Medicine*, vol. 8, no. 1, pp. 23–35, 2020.
- [133] Pei-Jung Lin, Han Yu Chen, Stephen Hung, Chun-Chen Lin, and Yu-Chen Wang, “An upper extremity rehabilitation system using virtual reality technology,” in *2018 15th International Symposium on Pervasive Systems, Algorithms and Networks (I-SPAN)*. IEEE, 2018, pp. 253–256.
- [134] Ángela Aguilera-Rubio, Alicia Cuesta-Gómez, Ana Mallo-López, Alberto Jardón-Huete, Edwin Daniel Oña-Simbaña, and Isabel M^a Alguacil-Diego, “Feasibility and efficacy of a virtual reality game-based upper extremity motor function rehabilitation therapy in patients with chronic stroke: A pilot study,” *International journal of environmental research and public health*, vol. 19, no. 6, pp. 3381, 2022.
- [135] Aviv Elor, Mircea Teodorescu, and Sri Kurniawan, “Project star catcher: A novel immersive virtual reality experience for upper limb rehabilitation,” *ACM Transactions on Accessible Computing (TACCESS)*, vol. 11, no. 4, pp. 1–25, 2018.