Remote Visual Observation of Real Places through Virtual Reality Headsets

ALESSIO REGALBUTO

PRINCIPAL SUPERVISOR: DR. SALVATORE LIVATINO
SECONDARY SUPERVISOR: DR. IOSIF MPORAS

Thesis submitted to the University of Hertfordshire in partial fulfilment of the requirements of the degree of Doctor of Philosophy (PhD)

United Kingdom

October 2018

ABSTRACT

Virtual Reality has always represented a fascinating yet powerful opportunity that has attracted studies and technology developments, especially since the latest release on the market of powerful high-resolution and wide field-of-view VR headsets. While the great potential of such VR systems is common and accepted knowledge, issues remain related to how to design systems and setups capable of fully exploiting the latest hardware advances.

The aim of the proposed research is to study and understand how to increase the perceived level of realism and sense of presence when remotely observing real places through VR headset displays. Hence, to produce a set of guidelines that give directions to system designers about how to optimize the display-camera setup to enhance performance, focusing on remote visual observation of real places. The outcome of this investigation represents unique knowledge that is believed to be very beneficial for better VR headset designs towards improved remote observation systems.

To achieve the proposed goal, this thesis presents a thorough investigation of existing literature and previous researches, which is carried out systematically to identify the most important factors ruling *realism*, *depth perception*, *comfort*, and *sense of presence* in VR headset observation. Once identified, these factors are further discussed and assessed through a series of experiments and usability studies, based on a predefined set of research questions.

More specifically, the role of *familiarity* with the observed place, the role of the *environment* characteristics shown to the viewer, and the role of the *display* used for the remote observation of the virtual environment are further investigated. To gain more insights, two usability studies are proposed with the aim of defining guidelines and best practices.

The main outcomes from the two studies demonstrate that test users can experience an enhanced realistic observation when natural features, higher resolution displays, natural illumination, and high image contrast are used in Mobile VR. In terms of comfort, simple scene layouts and relaxing environments are considered ideal to reduce visual fatigue and eye strain. Furthermore, sense of presence increases when observed environments induce strong emotions, and depth perception improves in VR when several monocular cues such as lights and shadows are combined with binocular depth cues.

Based on these results, this investigation then presents a focused evaluation on the outcomes and introduces an innovative eye-adapted High Dynamic Range (HDR) approach, which the author believes to be of great improvement in the context of remote observation when combined with eye-tracked VR headsets. Within this purpose, a third user study is proposed to compare static HDR and eye-adapted HDR observation in VR, to assess that the latter can improve *realism*, *depth perception*, sense of *presence*, and in certain cases even *comfort*. Results from this last study confirmed the author expectations, proving that eye-adapted HDR and eye tracking should be used to achieve best visual performances for remote observation in modern VR systems.

Keywords: Virtual Reality, Remote Visual Observation, Realistic Visual Reproduction, Stereoscopic 3D Visualization, VR Displays, 360 Camera View, Usability Evaluations, Eye-Tracking, HDR imaging, Eye-Adapted HDR.

DECLARATION STATEMENT

I certify that the work submitted is my own and that any material derived or quoted from the published or unpublished work of other persons has been duly acknowledged (ref. UPR AS/C/6.1, Appendix I, Section 2 – Section on cheating and plagiarism)

Student Full Name: Alessio Regalbuto

Student Registration Number: 14184995

Signed: Alessio Regalbuto

Date: 31/10/2018

ACKNOWLEDGEMENTS

I dedicate this thesis to all the people that have supported me with their love and have never stopped believing in me since the very beginning of this long challenging journey.

My special thanks go to my family, which managed to be present every day of the last 3 years despite the long physical distance, morally helping me endure all the obstacles of my new adventure in the United Kingdom. I would like to thank my mother and Attilio for their daily priceless support, my brother Mirko for his great ideas and comforting presence, my father and Rinalda, my aunties and uncles, and my grandparents, who I am sure will keep on guiding me and my future towards great profitable and amazing experiences.

Furthermore, I feel extremely lucky to have met so many inspiring people during my stay at University of Hertfordshire, first of which my Supervisor Dr Salvatore Livatino, who taught me not only how to be a good PhD student, but also how to improve my social and professional skills through team working and fortitude in every circumstance. I truly believe his supervision has been remarkable, and his passion for his work exemplary. I also express my gratitude to my Supervisor Dr Iosif Mporas, who provided me with precious suggestions and gave me the honour to be part of the local organizing committee for the SPECOM & ICR 2017 conferences.

I would like also to mention the fantastic flatmates I have been living with during these years: Gloria for our crazy conversations and movie marathons, Devansh for his hilarious humour, Zainab for our endless board games, Josh for bringing always the good mood, Rachel for her funny stories, Andy & Co for our cool gaming parties and procrastination moments.

A special mention goes also to my Italian colleagues Vincenzo Rinaldi and Ippazio Alessio, Dario Guastella, Francesco Mergiotti, Pasquale Bacone, and Valerio Scarlato with whom I have spent daily unforgettable (sometimes even crazy) moments, and who helped me a lot to better develop my professional career. Furthermore, I am thankful to my colleague Giuseppe Murabito for his kindness and distinctive dedication to work, and to my friend Greg for his help during my last phases of studies. A nice thought also goes to all my friends in Italy who remotely kept me company despite the distance, including Nicoletta, Michele, Fulvia, Frank, Giorgio, and Paolo.

Finally, I want to particularly thank my girlfriend Nathalie, who is probably the person who supported me the most throughout my last sprint of this PhD, providing me with all her efforts, energies, and love to help me realize this dream.

I can say with no doubt that without you all this life experience would not have been the same. You have all my gratitude, and my love.

Hoping in a prosperous future to all of us,

Alessio Regalbuto

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. INTRODUCTION

1.1 Virtual Reality

Virtual Reality (VR) has been defined as "*a computer-simulated environment with which a person can interact almost as if it were part of the real world*" [1]. It is typically considered a medium [2], which stands between the human being and a virtual world, "*having no noticeable interface, so that the boundaries between users and virtual worlds are seemingly non-existent*" [2]. Despite VR has been proposed and investigated already from a few decades, it still is more than ever seen as a revolutionary new media [3], because of the VR great potential and the new advances in related technologies. To reduce the boundaries between users and virtual worlds means technologies that provoke the human sensory system the same way the real world does. This would induce observers believe that the observed world is real.

An important milestone of VR technology was reached in 1950 by Morton Heilig with the development of the *Sensorama simulator* [4]. This device represented one of the first attempts to innovate the perception of digital video footage by introducing 3D images, touch feedback, and sense of smell simultaneously. As a result, the viewer had the illusion to be teleported inside a digital believable reproduction of reality, purely virtual. Great has then been the interest and the evolution of VR hardware, with further milestones achieved through the years, including the development of systems such as the first Head Mounted Display (HMD) [5] and the CAVE [6].

The focus has mostly been on VR displays because of the dominant role played by the *Human Visual System (HVS)* in human perception [7]. Indeed, the richness of information the visual world offers (e.g. the perception of distant, noxious, or fragile objects [8]), coupled to the lack of accuracy of the other senses (e.g. audition has severe accuracy limitations in perceiving directions and distances), and the level of reliability humans put over the sense of sight (studies prove that it can bias information coming from the other senses [9]), show that human vision has the highest potential to achieve accurate and reliable space awareness of surrounding existing places.

The research work described in this thesis focuses on understanding, developing and assessing VR system that mainly involve the HVS. Furthermore, it looks at the specific aspect of replicating real places observation through the use of VR technology, with a specific attention to VR headset technology.

Figure 1 *- Oculus Rift CV1, virtual reality headset and controllers.*

Modern Virtual Reality applications make use of a Head Mounted Display (HMD) more than ever before. The **Figure 1** shows an example of a popular VR HMD (the Oculus Rift [10]). Other examples are the HTC Vive [11], the StarVR [12], and the Samsung Gear VR [13].

The great interest and potential towards VR, which have fuelled the development of new hardware and software, are the many applications. Among them we find:

- **Training and education.** VR delivers learning experiences with unprecedented levels of interactivity;
- **Architecture and Planning.** VR makes it possible for engineers and manufacturers to experience their creations before these are built. This also reduces production costs and development times;
- **Entertainment and gaming.** Games in VR can offer real-life simulated environments and intuitive interfaces that allow players to participate more naturally with higher levels of realism;
- **Simulations.** VR makes complex situations easier by offering a safe environment for training purposes with higher accuracy and interactivity (e.g. flight simulators);
- **Conferencing.** VR brings digital workers together in digital meetings and conferences, removing the barrier of distance and enabling higher levels of telecommunication;
- **Help and healing.** VR can help people with phobias or pathological issues to better tolerate their fears and recover from traumas.
- **E-tourism.** VR offers the possibility to visualize places around the world realistically, regardless of the viewer’s actual location. This has a great potential, especially for places not easily accessible to humans.

1.2 VR applied to remote observation of real places

The possibility to remotely explore existing places through 3D 360 panoramic photographs is of great interest for a wide range of applications, e.g. dangerous environments exploration, archaeological cultural patrimony preservation. The interest has further grown since the introduction of *Mobile VR* headsets, which allow consumers to use their smartphones as VR displays with no need to buy expensive devices (see **Figure 2**).

Figure 2 *- Mobile VR headset, which can host a smartphone as VR display.*

Compared to standard virtual reality devices, Mobile VR offers similar functionalities at reduced costs with better portability. Most of the time this can still deliver believable remote observation to users, who can feel teleported in other environments through accurate 3D perception and natural body movements, with no need for expensive laptops by taking advantage of the power of modern smartphones.

Furthermore, some types of Mobile VR headsets offer adjustable lenses and variable distances from the smartphone display. This enables more comfortable visualization even to viewers affected by visual problems (e.g. astigmatism, myopia, etc.). To improve even more the quality of remote observations, another chased goal has been the enhancement of *telepresence*, which "*refers to the set of technologies that induce people to feel as if they are actually present in a different place or time"* [14], making use of visual, audio, and sometimes haptic stimuli. Therefore, spatial audio capabilities, high resolution visual content, and tactile feedbacks are additional features that applied to Mobile VR can offer improved performances, making virtual remote observation even more credible.

1.3 Existing issues

Despite the great potential of Mobile VR some issues and drawbacks are still present, the variety of smartphones and their very different display specifications represent an obstacle to comprehensive optimal setups and configurations of both graphics and hardware for Mobile VR headsets. Thus, there is room for several improvements on the visualization of 3D 360 panoramas, including discomfort reduction (e.g. visual fatigue [15], motion sickness), depth distortions and visual artefacts removal, and enhanced level of realism. Such improvements would benefit remote observation by limiting the gap between direct observation of real places and remote visualization of their digital reproduction.

Additional issues come from the actual generation of 3D 360 panoramas when taking photographs of existing places: unconventional and expensive cameras are usually required, and the lack of awareness on effects of the used setup can result in poorly generated 3D 360 panoramas.

Furthermore, eyes differences and dissimilarities among VR users introduce additional variables that can affect the final performances of the remote visual observation. Among them, psychological factors and subjective experiences can induce users to perceive the same virtual environment differently. Therefore, users that try for the first time VR need to achieve good quality visual experiences as the risk of rejecting this technology is high, due to possible bias caused by visualized contents.

Among the main elements affecting the VR headset users' experience:

- **Discomfort**: several users report that many Virtual Reality experiences cause visual fatigue, eye strain, or motion sickness. Most of the time, this is due to erroneous HMD configurations, wrong optical lenses (which can cause blurred vision), visual delays (which cause conflicting sensory information, thus motion sickness and headache).

- **Depth distortions and visual artefacts**: the use of HMDs with too large interpupillary distance (IPD) values (which prevent the correct depth perception of the visualized virtual environment), wrong camera stereoscopic configuration when capturing depth information from the real scene (e.g. excessive camera baseline), and different graphical and device fields of views can introduce distance estimation errors and incorrect depth representations of the virtual environment.

- **Provided level of realism**: a poor unrealistic reproduction of the virtual environment deteriorates the usefulness of the remote observation and can represent a serious problem when applied to critical applications such as robotic teleoperations and simulators.

- **Cost**: the use of Virtual Reality to access remote visual observation of existing places has been represented by the prohibitive costs of Desktop VR HMDs: the need of a powerful computer or laptop to allow the use of such headsets has hindered the use of this technology.

There is also lack of guidelines on how to best setup acquisition and visualization systems for applications focusing on the use of VR headsets, e.g. to increase the level of realism, sense of presence and comfort.

1.4 Aims and main objectives

It is believed that guidelines related to optimizing VR headset system setup can be very beneficial. These include to provide good design principles toward system elements such as: camera and display specifications. Such recommendations and directives are believed could be gathered and proposed to VR system designers, based on a deep analysis of previous findings (also related to other fields and type of VR technology) and on focused assessment trials. The ultimate goal would be to deliver higher quality remote observation systems, which can be of benefit for all applications that make use of Mobile VR. High quality VR is deemed reachable with today technology, but it is hindered by the lack of guidelines about how to best use this really powerful technology, which today more than ever is focusing on the specific use of VR headsets.

This research investigation aims at understanding the role of technology setup and camera-display parameters in providing convincing visual realism, for reducing the perceptual gap between direct and indirect observation of real places when using Mobile VR headsets.

To achieve this goal, a thorough search in the literature is proposed by this thesis to identify the most important parameters ruling *realism*, *depth perception*, *comfort*, and *sense of presence* in VR headset observation. Furthermore, it is decided to apply the identified findings and assess them through a series of experiments and usability studies, focused on a number of set research questions.

More specifically, a systematic investigation on the role of *familiarity* with the visualized place, on the role of the *environment* characteristics shown to the viewer, and on the role of the *display* used for the remote observation of the virtual environment is carried out. Once that an insight on the proposed research questions has been gained, it is decided to proceed on a focused evaluation, which would involve the use of eye-tracked VR observation and the use of High Dynamic Range (HDR) images. These are further assessed to investigate the impact and potential benefit of eye-adapted HDR observation when viewing real places remotely in VR.

The ultimate objective is to produce a research that represents meaningful knowledge, which I plan to make available through this thesis to future system designers in form of a set of guidelines, towards a much-improved remote observation of real environments through Virtual Reality devices.

1.5 Outline of Dissertation

This thesis is organized as follows. Chapter 2 presents definitions and background knowledge that can help the reader to better understand topics of this research investigation.

Chapter 3 presents the proposed investigation, specifying aims and motivations (see section 3.1), and introducing three proposed research phases: the first is relative to a systematic review for learning more information on the role of camera-display parameters (see section 3.2); the second related to building knowledge from assessments by making use of user studies on the role of familiarity, environment, and display, and their impact on realism, comfort, presence, and depth perception (see section 3.3); the third phase presenting a focused user study designed after reviewing previous results, to investigate on eye-adapted HDR visualization (see section 3.4).

Chapter 4 presents information from the presented literature review, discussing the approach that has been used to collect research papers (see section 4.1), guidelines and best practices resulting from previous research outcomes (see section 4.2), a model that has been developed to better organize parameters that were found during this investigation (see section 4.3), and an extended model focused on further parameters having a relevant impact on sense of presence and distance estimation in VR (see section 4.4).

Chapter 5 presents more details on the first proposed user study, to investigate on the effect of users having different familiarities with the presented environment in VR, and on the possible influence of the environment over realism, comfort, presence, and depth perception. Furthermore, it discusses the evaluation design of the proposed user study (see section 5.1), implementation details (see section 5.2), an extended pilot test that was useful to optimize questionnaires and procedure designs (see section 5.3), results of the proposed user study (see section 5.4), a section analysing results in more details (see section 5.5), and a section summarising research outcomes in simple guidelines (see section 5.6).

Chapter 6 presents the second user study, which investigated the role of display and environment on the visual perception of viewers in virtual reality. This introduces the chosen evaluation design (see section 6.1), implementation details (see section 6.2), information on procedure and extended pilot test (see section 6.3), results of the proposed user study (see section 6.4), result analysis (see section 6.5), and a summary of outcomes through simple guidelines (see section 6.6).

Chapter 7 introduces the third usability evaluation, which based on previous results was focused on the possible advantages of eye-adapted HDR visualization of 3D panoramic environments for remote observation. This presents the chosen evaluation design (see section 7.1), implementation details (see section 7.2), further information on the procedure used for this user study (see section 7.3), results of the proposed user study (see section 7.4), result analysis (see section 7.5), and a summary of outcomes through simple guidelines (see section 7.6).

Finally, Chapter 8 presents a summary of the investigation (see section 8.1), highlights on the main contributions of this research outcomes (see section 8.2), and gives suggestions for future works that might continue the work of this thesis (see section 8.3).

CHAPTER 2. BACKGROUND KNOWLEDGE

This chapter introduces concepts and technologies that will be discussed by this thesis, to help the reader understand the proposed investigation and research outcomes.

Virtual Reality is a medium between the human being and virtual/replicated worlds, thus this chapter presents relevant subjects on both sides, starting with the human visual system to then proceed to computer vision and visualization technologies with regards to remote virtual observation of real scenes. Eventually, the most popular VR headsets are presented.

The *Human Visual System (HVS)* is considered the most dominant human sense for perceiving surrounding environments in comparison to the other four (hearing, taste, smell, touch) [7]. Indeed, the richness of information the visual world offers (e.g. vision allows the perception of distant, noxious, or fragile objects [8]), the lack of accuracy of the other senses (e.g. audition has severe accuracy limitations in perceiving directions and distances), and the level of reliability humans put over the sense of sight (studies prove that it can bias information coming from the other senses [9]) show that human vision has the highest potential to achieve accurate and reliable space awareness of surrounding existing places.

In order to understand how the *HVS* works and how it can be emulated by modern visual technologies to remotely observe existing places, the structure and behaviour of human eyes are analysed by reporting definitions for stereoscopy and depth cues (see section 2.1). Furthermore, the differences between direct and indirect observation systems are presented (see section 2.2). Then, noteworthy definitions for remote visual observation of real places are discussed, focusing on the concepts of realism, sense of presence, immersion, comfort, and real-time / non-real-time visual observation (see section 2.3).

In the context of approaches to achieve enhanced remote observation of existing places, the problem of acquiring depth information of depicted environments using 3D cameras is presented. This includes the definition of the relevant internal and external parameters ruling the acquisition of 3D photos (see section 2.4). Furthermore, panoramic visualization widening the amount of information of the reproduced remote environment is discussed (see section 2.5), presenting panoramic paintings, panoramic photography, and stitching issues. Approaches to acquire mono and stereo panoramas are also investigated (see section 2.6).

3D visualization systems are presented to provide a base to understand the most effective techniques that enable visual perception of the *HVS* through the provision of specific monocular and binocular depth cues (see section 2.7). Within this context, I acquired a greater understanding of 3D visualization systems through hands-on development of Computer Generated 3D scenes shown on 3D displays, by making use of Unity 3D [16], OptiTrack motion cameras [17], and Nvidia 3D glasses [18]. This has been very beneficial to focus my investigation on the most relevant aspects and parameters that have an influence over the HVS depth perception and space awareness. Some examples of developed 3D scenes, including 3D prototypes produced to showcase Dr Livatino's filed invention titled "Coherent touchless interaction with stereoscopic 3D images" [19], are shown in Appendix C.

The potential of Virtual Reality devices is discussed and presented thoroughly. VR observers interact through eye and head movements. Therefore, the capability of the visualization system to adapt the shown 3D content according to observer's perspective is crucial to better reduce the gap between direct and indirect visual observation of real environments. For this reason, the use of Virtual Reality as a mean to interactively visualize stereoscopic panoramic photographs is being considered: thanks to the capabilities of *Head Mounted Displays* (*HMDs*) observer's head movements can be recorded. This allows the system to suitably combine viewer's head rotations with the rotation of the visualized virtual panorama. As a result, the viewer is given the illusion to naturally explore the surrounding space by simply looking around it through head rotations (see section 2.8).

The emerging of portable virtual reality devices that rely on smartphones is presented. These devices define a subset of virtual reality acknowledged as "Mobile VR", which addresses technologies that by means of modern smartphones inside portable plastic *HMDs* is able to show virtual content such as 3D panoramic pictures without the need for Desktop computers (see sub-section 2.8.1).

Furthermore, the use of HDR images and the reproduction of human eye adaptation behaviour in virtual reality scenes have been considered, as these are relevant to improve quality of the indirect observations by showing a wider range of colours and light information (see section 2.9). The use of eye-tracking devices applied to virtual reality is also investigated (see section 2.10).

Finally, this chapter introduces VR usability evaluations (see section 2.11), to help the reader understand what an user study is and how it can be useful when conducting a research investigation.

2.1 The Human Visual System and Direct Observation of Real Places

2.1.1 Stereoscopy and depth perception on human eyes

Depth perception is relevant to humans to enable more accurate spatial awareness of surrounding environments. As shown in **Figure 3**, the advantage resides on the opportunity to faithfully estimate objects' distance and better interpret real life situations.

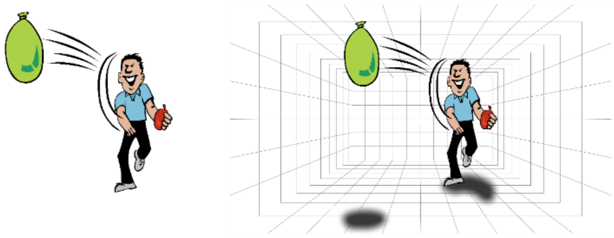


Figure 3 *- The advantage of 3D vision over 2D vision.*

On the human eye, depth perception is accomplished by means of *monocular* and *binocular cues*. The former conveys some sensation of depth even in standard two-dimensional images, and consists of *position* (see paragraph 2.1.2.1), *perspective* (see paragraph 2.1.2.2), *relative size* (see paragraph 2.1.2.3), *shading and shadows* (see paragraph 2.1.2.4), *depth from motion* (see paragraph 2.1.2.5), *occlusions* (see paragraph 2.1.2.6), *depth-of-field and defocus blur* (see paragraph 2.1.2.7), *texture gradients* (see paragraph 2.1.2.8), *motion parallax* (see paragraph 2.1.2.9), *accommodation* (see paragraph 2.1.2.10); the latter mainly relies on *binocular convergence* (see paragraph 2.1.3.1), and *binocular parallax* (see paragraph 2.1.3.2). Within this context, the following space definitions are given [20] and shown in **Figure 4**.

- Personal space refers to the zone immediately surrounding the observer's head, generally within arm's reach and slightly beyond, not exceeding 2 meters of distance.
- Action space refers to the zone within which actions are performed by the human body, and it ranges between 2 meters and 30 meters from the viewer.
- Vista space refers to the zone that is unperturbed by the motions of the observer and that goes beyond 30 meters from the viewer.

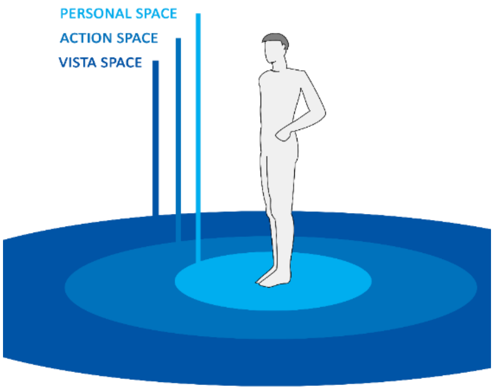


Figure 4 *- Scheme showing personal, action, and vista spaces.*

Among above definitions, viewer's surrounding space is classified into *egocentric* and *exocentric*. The *egocentric space* defines the distance measurement between observer and a target object within the environment, whilst the *exocentric space* defines the distance between targets other than the observer. In both cases, the observer is considered the centre of reference [21] [22].

Some depth cues are independent of distance, such as occlusion or relative size, whereas others are distance-dependent, such as disparity or vergence [23].

2.1.2 Monocular cues

2.1.2.1 Position

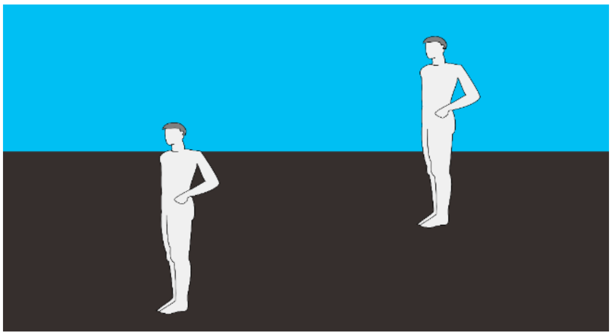


Figure 5 *- Example of Position - Monocular cue. The man on the right appears farther than the man on the left.*

Subjects that lay on a ground plane and are located near the bottom of a picture are perceived closer to the viewer (see example in **Figure 5**). This implies a priori knowledge and is considered as one of the oldest known depth cues.

2.1.2.2 Perspective

Figure 6 *- Example of Perspective - Monocular cue. Photo taken from Pixabay [24].*

Perspective is a monocular cue that shows parallel lines converging in distance, at infinity. This allows the viewer to reconstruct relative distance of two parts of an object or features of a landscape (see example in **Figure 6**).

2.1.2.3 Relative size

Figure 7 - Example of Relative Size - Monocular cue. Photo taken from Pexels [25].

The a priori knowledge related to normal object sizes infers knowledge about depth. For example, we know a priori that a building is bigger than a person, so when a picture shows a person having the same size of a building we are induced to logically think that realistically the person is closer to the viewer and the building is in the background (see **Figure 7**).

2.1.2.4 Shading and shadows

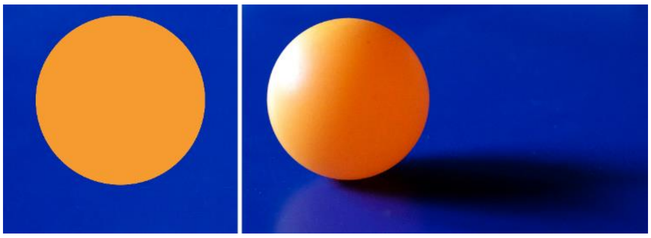


Figure 8 - Example of Shading and shadows - Monocular cue.

When objects are represented with shading and shadows casted on other objects, the viewer can perceive depth. For example, if we want to distinguish a circle from a sphere when drawing, we can add lighting and shading effects to provide the illusion of a three-dimensional object (see **Figure 8**).

2.1.2.5 Depth from motion

Figure 9 - Example of Depth from motion - Monocular cue. The arrow moves towards the viewer and gets bigger. Photo taken from Pexels [26].

Depth from motion is a monocular cue that provides distance estimations according to the change in size of objects in movement. When an object moves towards the viewer and gets closer, its relative size increases, and triggers the viewer's brain to perceive a distance change (see example in **Figure 9**).

2.1.2.6 Occlusions

Figure 10 *- Example of Occlusions - Monocular cue. Photo adapted from Pixabay [27].*

Occlusions are one of the strongest monocular cues, and work at any distance from the viewer. They rely on the fact that foreground objects occlude background objects (see example in **Figure 10**, where buildings in the foreground occlude the ones in the background, providing depth sensation).

2.1.2.7 Depth-of-field and Defocus Blur

Figure 11 *- Example of Defocus Blur - Monocular cue. Photo taken from Pexels [28].*

In photography, defocus blur can be used to separate a subject that is in focus from the background using small camera depth of field (distance between the nearest and the furthest objects giving a focused image). This gives the viewer the impression that the focused subject is located farther from the background and closer to the camera (see example in **Figure 11**, where the finger appears closer due to defocused background). This also applies to landscapes that present atmospheric haze.

2.1.2.8 Texture Gradients

Figure 12 *- Example of Texture gradients - Monocular cue. Photo taken from Pexels [29].*

This type of monocular cue is caused by the fact that textures get more and more blurred with distance due to perspective transformation. An example is the texture of the sand in a desert, that becomes more blurred when farther from the viewer (see **Figure 12**).

2.1.2.9 Motion parallax

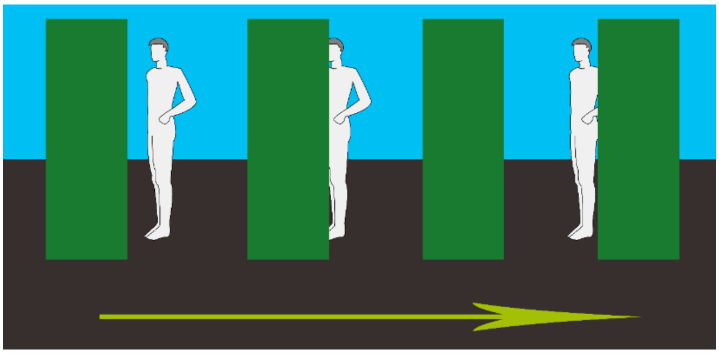


Figure 13 *- Example of Motion Parallax - Monocular cue. When the viewer moves from left to right, the person is hidden behind the green obstacle (which is closer to the viewer).*

When the viewer moves relatively to a planar ground, nearby objects move faster in the field of view than distant ones. This enables distance estimations and allows the possibility to see objects that are occluded behind foreground subjects (see **Figure 13**).

2.1.2.10 Accommodation

Accommodation refers to the variation of the lens shape and thickness (and thus its focal length), which allow the eye to focus on an object at a certain distance [23]. This happens in the human eyes when objects in real life are located to different distances from the viewer and need to be focused through eye lenses contraction or extension.

This monocular cue works in combination with binocular convergence (see paragraph 2.1.3.1).

2.1.3 Binocular cues

2.1.3.1 Binocular convergence

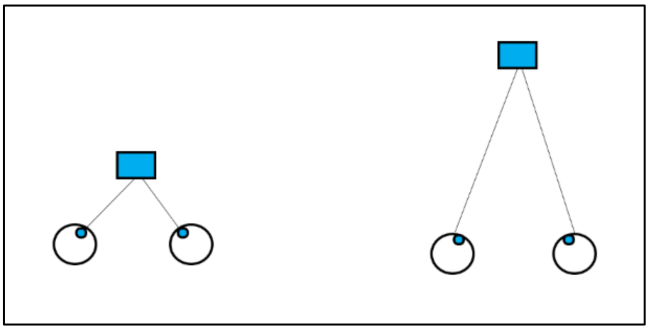


Figure 14 *- Scheme of binocular convergence - Binocular cue.*

When eyes focus on an object, they can change their angle to guarantee a clear visualization of the gazed point. This angle defines their convergence and is used by the visual system to judge distance (see **Figure 14**). The angle of convergence is smaller when the eye is fixating on far away objects.

In real life, convergence is combined with accommodation (see paragraph 2.1.2.10) to improve distance estimation and focus on the subject. On a display, these are in conflict because even if the eyes converge on the observed object accommodation does not change and is kept constant on the screen.

2.1.3.2 Binocular parallax

Figure 15 *- Example of Binocular Parallax, where each eye observes the scene from a different viewpoint and perceives 3D through stereopsis.*

Binocular parallax is defined as the relative displacement of an object viewed along the two lines of sight caused by viewpoint change (see example in **Figure 15**), which causes disparity between left and right views [23]. This induces stereopsis, which stimulates depth perception.

2.2 Direct and Indirect Observation Systems

When we compare direct and indirect observation of an existing place, the most important difference is that indirect observation requires a *medium* to be interposed between the human eye and the virtual environment. The medium involves among other things the use of a display (e.g. 2D/3D monitors, wall screens, HMD, etc.) and a camera (e.g. 2D/3D lens, and the involvement of IR-lights, laser devices, etc.) and a telecommunication network (e.g. the Internet or a dedicated one). The above needs introduce challenges affecting realistic perception of the remote shown environment.

Therefore, to achieve realistic results it is necessary to put our indirect observation system in the ideal situation where:

- **Cameras can reproduce as accurately as possible characteristics of the viewer's eyes**, by offering:
    - Similar field of views;
    - Dynamic converge and accommodation capabilities;
    - Variable pupil size according to the amount of light presented by the scene;
    - Extremely high-resolution sensors to record faithfully the rich number of colours and details of the environment;
    - Distance between the lenses comparable with the interpupillary distance (IPD) of the viewer's eyes.

  However, these features are still very hard to realize using currently available cameras.
- **The *medium* is not interfering with the natural observation of the virtual environment**, by providing high levels of both immersion and isolation from the real place where the viewer is standing. This can be achieved by using modern HMDs in virtual reality (see section 2.7.2).
- **The visual perception of the user is as accurate as possible**, to make the experience believable and convincing enough, so that the viewer has the illusion to physically be present inside the observed virtual environment.

We can generally refer to the indirect observation of a real place as a **remote visual observation** (see section 2.3) performed through a medium that shows its virtual representation.

2.3 Remote Visual Observation of Real Places

When a real place is not accessible due to its distance from observer (e.g. planets, nuclear dangerous areas for humans) or its time of existence (e.g. ancient cities, destroyed locations), a *remote visual observation* of its environment can still be performed by using a *medium* or a digital device to virtually reproduce its visual appearance.

Thanks to the enormous benefit provided by remote visual observations in performing hazardous tasks (e.g. bomb disarming, radioactive materials management) as well as in exploring real places without the need to physically visit them, researches into this area have gained a significant consideration during last five decades within a wide range of applications (e.g. planetary exploration [30] as shown in **Figure 16**, robotic control systems for tele-operations [31], [32], [33], e-tourism [34]).

Figure 16 *- The NASA Mars Exploration Rover, designed for remote observation of other planets. Picture taken from Pixabay [35].*

In order to capture visual information of the real place and show it indirectly to the observer, several approaches exist (e.g. paintings, photographs, 3D pictures (section 2.4), panoramic views (section 2.5), virtual reality applications (section 2.8)). However, not all of them are able to achieve the same accuracy and faithfulness of the represented environment. This is because they can provide the observer with different levels of *realism*, *sense of presence ("feeling of being there")*, *immersion*, and *comfort*. These characteristics can be used to estimate the overall quality of a remote visual observation and are further discussed in subsection 2.3.1.

2.3.1 Realism, Sense of presence, Immersion, Comfort

Realism is defined as the quality or fact of representing a person or thing in a way that is accurate and true to life. When we apply this definition to images and panoramas, it refers to the visual quality and accurate representation that these can deliver to the viewer, inducing the idea of observing something real even when it is not. In the context of remote observation, we can assume then that the representation of a remote environment to observe an existing place is very realistic when the visual information provided by the *medium* look very similar to the physical visualized existing place.

Sense of presence is defined as the subjective experience of being in one place or environment even when one is situated in another [36]. As Mel Slater argues, "presence is about form, the extent to which the unification of simulated sensory data and perceptual processing produces a coherent 'place' that you are 'in' and in which there may be the potential for you to act" [36]. We can generally conclude that the term *sense of presence* can be associated to the "feeling of being there".

According to Mel Slater [37] there is a considerable difference between the term *presence* and the term *immersion*.

Immersion refers to the objective level of sensory fidelity a VR system provides [38], and can be enhanced by adding audio, tactile and force feedback [39]. The distinction between immersion and presence is also motivated by the fact that sometimes different immersion systems may have indistinguishable perceptual impacts on people in terms of presence [37].

We use the definitions of *sense of presence* and *immersion* to estimate the quality of a virtual reality experience in relation to viewers' perception of the virtual environment.

Comfort refers to the presence or absence of *visual fatigue*, *motion sickness*, *headache*, and *eye strain*. In virtual reality it is essential to deliver a comfortable visualization of the virtual environment, and to make sure the viewer possible experience of any of the above-mentioned issues is limited.

2.3.2 Real-time vs Non-Real-time Remote Visual Observation

We distinguish two types of remote visual observations: *real-time* and *non-real-time*. The former provides an actual live representation of a real scene as it appears while being observed; the latter reproduces a place that was acquired sometime in the past, intended for future observations too.

An example of *real-time* remote visual observation is a telerobotic operation captured by digital devices, which allow the operator to understand the space surrounding the robot while it is being teleoperated.

In the context of both real-time and non-real-time remote visual observation of existing places, the use of **3D camera systems** (see section 2.4), **panoramic visualization** (see section 2.5), **3D visualization systems** (see section 2.7), and **virtual reality** (see section 2.8) may be very beneficial.

2.4 Cameras and 3D Acquisition

3D camera systems can imitate the behaviour of human eyes, to faithfully capture visual information of the real world and store them for remote observations. Within this purpose, *stereoscopic cameras* (see example in **Figure 17** and **Figure 18**) make use of two camera lenses, each representing left or right eye, to imitate the structure of the Human Visual System (HVS).

Figure 17 *- Example of stereoscopic webcam, presenting two lenses to imitate the HVS.*

Figure 18 *- Front view of the Fuji 3D stereoscopic camera.*

The term *stereoscopic* defines a device or technique that can provide the illusion of depth by means of *stereopsis* and binocular vision (see subsection 2.1.3). Together with *kineopsis*, which relates optical flow to depth and motion in space, *stereopsis* recovers depth from the disparity field of human vision [40].

2.4.1 Camera parameters

To achieve good performances and avoid alteration of the authentic appearance of reality, camera *intrinsic parameters* and *extrinsic parameters* need to be correctly analysed and configured.

2.4.1.1 Relevant internal parameters

Camera internal parameters represent camera intrinsic characteristics that are responsible for mapping into the camera image-plane the observed 3D world. Below a brief summary of main elements is reported, including those elements relevant when acquiring photographs and videos for visual observation of remote environment.

Focal length

Focal length is a measure of how strongly the optical system converges or diverges light. Specifically, it indicates the distance over which the initially parallel light rays are converged to a point of focus. The shorter is the focal length, the more powerful is the optical system as it can bend light rays more sharply, bringing them to a focus in a shorter distance.

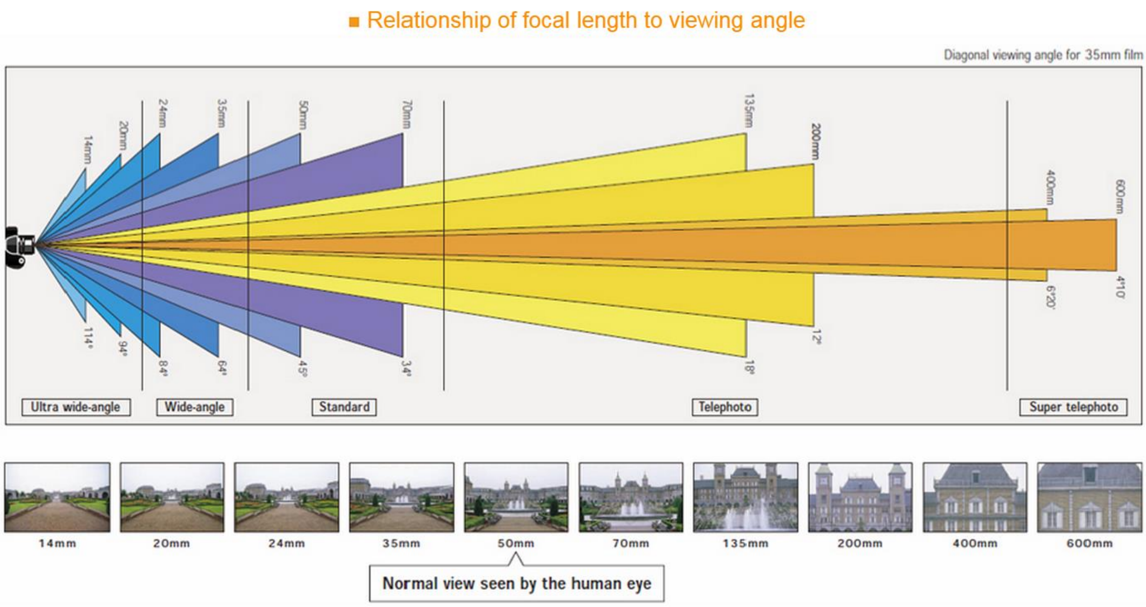


Figure 19 *- The effect of focal length over image perception and viewing angle. Illustration by Panasonic [41].*

In terms of image alteration, a longer focal length leads to higher magnification and a narrower field of view of a subject that is infinitely far away. By contrast, a shorter focal length provides lower magnification and a wide angle of view (see **Figure 19**).

Horizontal field of view (FOV)

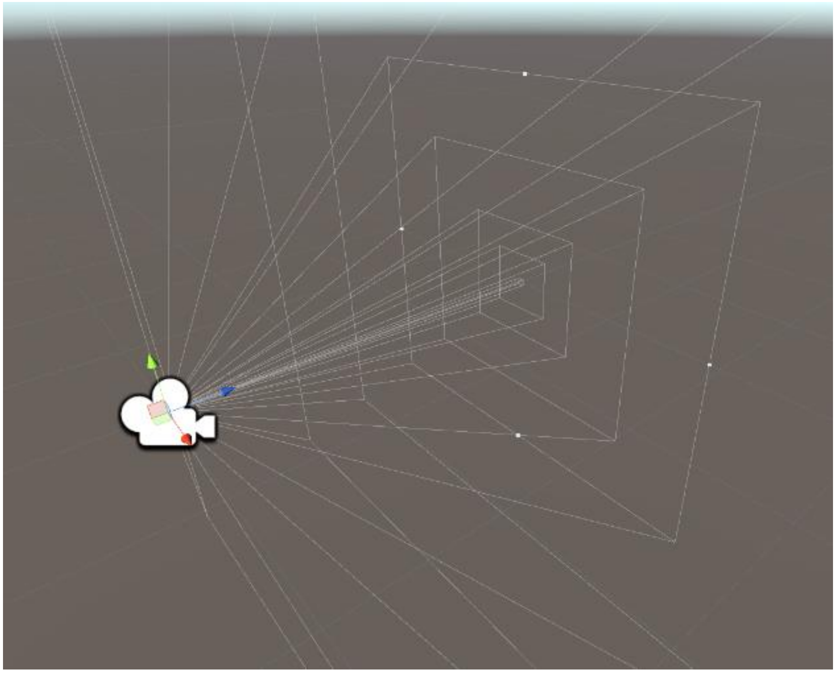


Figure 20 - Diagram in Unity [16] showing different field of views of the virtual camera.

Camera's horizontal field of view defines the maximum horizontal angle of vision that the camera can provide over the captured scene (see concept in **Figure 20**). It is related to camera focal length by **Equation 1** and **Equation 2**:

$$\text{Horizontal FOV} = 2 * \text{atan}\left(0.5 * \frac{\text{width}}{\text{focal length}}\right)$$

Equation 1 - Horizontal Field of View (FOV) as function of focal length

$$Vertical\ FOV = 2 * \text{atan}\left(0.5 * \frac{height}{focal\ length}\right)$$

Equation 2 - Vertical Field of View (FOV) as function of focal length

where *width* and *height* refer to the size of the sensor of the camera.

According to the field of view, we distinguish:

- **Fisheye lenses** (ultra-wide), with FOV over 180 ° and focal length 8-24mm;
- **Wide angle lenses**, with FOV 64°-84° and focal length 24-35mm;
- **Standard lenses**, with FOV 40°-62° and focal length 36-60mm;
- **Long focus lenses**, with FOV 31°-35° and focal length 61-84mm;
- **Medium telephoto lenses**, with FOV 10°-30° and focal length 85-135mm;
- **Super telephoto lenses**, with FOV 1°-8° and focal length over 300mm.

Image sensor format

The format of the image sensor of the camera is responsible for the amount of light details that are digitally recorded from the real environment. The larger the sensor size, the higher the amount of data reproducing faithfully the scene. Among different sizes, we distinguish two typologies:

- **CMOS sensors**, which usually consume less power but tend to offer lower sensitivity to light;
- **CCD sensors**, which usually consume more power but provide higher sensitivity to light.

Several camera lens types can be mounted on a camera, depending on the user's needs, which affect the focal length. We generally distinguish *fixed lenses* (which do not offer the possibility to adjust the focal length) from *varifocal lenses* (which provide variable focal length and focus point).

Exposure: Aperture, ISO, and Shutter Speed

When capturing a scene with a digital camera, light rays pass through camera lenses and reach the camera sensor to be transformed into digital inputs and stored in memory. Depending on the amount of light captured, we define **exposure** as the amount of light per unit area reaching the image sensor.

Camera exposure is controlled by three different parameters:

- **Aperture**, which controls the area over which the light can enter the camera: the larger the aperture, the higher the number of light rays impacting the camera sensor.
- **ISO speed**, which controls the sensitivity of the camera sensor to a given amount of light: the higher the value, the noisier the picture.
- **Shutter speed**, which determines how long the camera sensor will remain open to collect light rays from the scene: longer times cause motion blur as more light changes are recorded.

Camera exposure is essential in photography to guarantee all details of the captured environment to appear clear and free of over-exposed (see **Figure 21**) or under-exposed (see **Figure 22**) areas. To produce good quality photographs, HDR photography can be used (see section 2.9).

Figure 21 *- Example of panorama with over-exposed areas (bright areas).*

Figure 22 *- Example of panorama with under-exposed areas (dark areas).*

2.4.1.2 Relevant External parameters

We identify as relevant external camera parameters the position of the camera in the space (x, y, z), and its angular position (pan, tilt, verge).

In addition, stereoscopic cameras make use of two lenses, which can be placed at several distances (i.e. **baseline**) and different angles. Depending on the two lenses layouts, we distinguish:

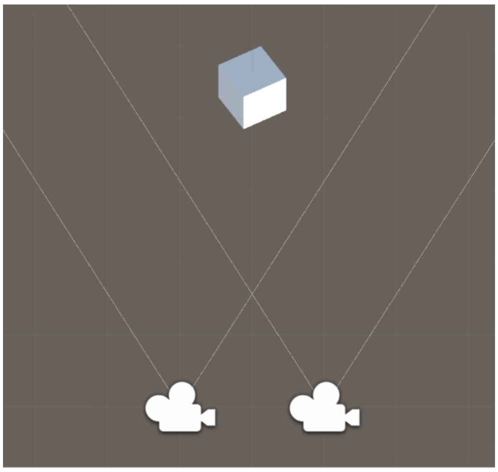


Figure 23 *- Stereoscopic camera with Parallel Configuration.*

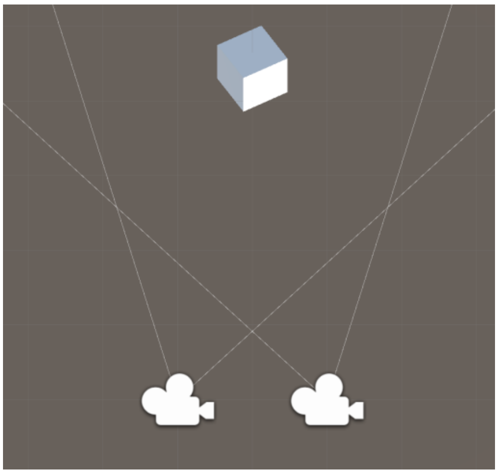


Figure 24 *- Stereoscopic camera with Toed-in Configuration.*

- **Parallel configuration** (see **Figure 23**), which implies a parallel alignment of the two camera lenses;
- **Toed-in (verged-in) configuration** (see **Figure 24**), which implies different directions of the two camera lenses.

Furthermore, it should be noted that viewers can only experience the 3D effect comfortably if they have the exact same binocular disparity as the stereoscopic 3D camera and watch the 3D media from the same shooting position [42]. Therefore, values of camera extrinsic parameters must be correctly configured within this purpose, depending on the visualization setup chosen for the viewer.

2.5 Panoramic Visualization

A visual illustration of real places that is significantly suitable for *non-real-time* remote visual observations derives from old *panoramic paintings* and modern *panoramic photography*, which offer a wide space representation of the observed environment enabling greater *sense of presence* for the viewer.

2.5.1 Panoramic paintings

Early illustrations enhancing visual observation of real places come from the *panoramic paintings* of Robert Barker [43], which was one of the first painters to portray a landscape in a full circle of 360 degrees as realistically as possible [44]. His famous "London from the Roof of Albion Mills" panoramic painting (**Figure 25**) is one of the first works of art surrounding the viewer completely like a real environment, giving to the observer the illusion to be physically present inside the panorama. The painting was shown to the public through an immersive studio stylised in **Figure 26**.

Figure 25 *- Full size panorama painting of the «London from the Roof of Albion Mills» by Robert Barker [45], completed in 1792.*

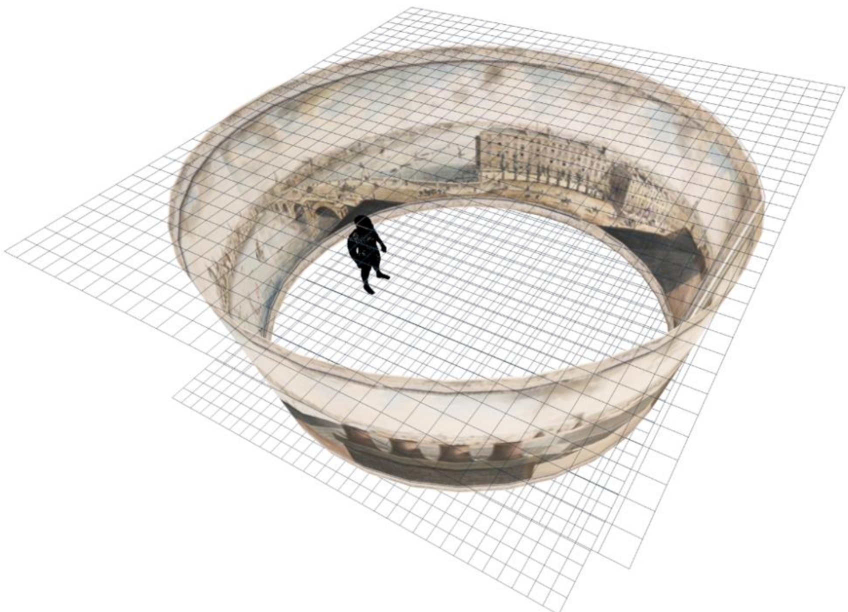


Figure 26 *- 3D concept representing the remote visual observation of the «London from the Roof of Albion Mills» through the panoramic painting of Robert Barker [45].*

The attempt to introduce a wider representation of reality through panoramic views in paintings expressed the will to immortalize real places in their entireness, to deliver a more realistic observation of their reproduction. This motivation was in line with some of the emerging artistic movements of the 19th century, which inspired artists to pursue a faithful reproduction of real life free of visual artefacts.

An example of refusal to reality's alteration in art comes from the *Realism* movement, which started in France with Gustave Courbet [46]. Differently from other artistic movements, Courbet painted scenes taken from the present, which was considered the only moment of truth representation of reality deprived of the influence of the painter's personal representation of past or future events. Later, this ideology inspired the development of new technologies able to capture on digital pictures instants of the present reality to accurately visualize them in the future.

2.5.2 Panoramic photography

The invention of photography has enabled the acquisition of detailed real scenes through the natural optical phenomenon of the *camera obscura*. This technological advancement has achieved cutting-edge levels of realism on account of higher resolution, wider color spectrum, and advanced image sensors, especially in the context of remote visual observation of real-life scenes through printed or digital photographs.

A natural evolution of this technology is *panoramic photography*, which recalls the same purposes that have been defined for panoramic paintings. A stylized example of a *cylindrical panoramic photography* of the Sicilian island *Lachea* (that I captured within the Visual 3D Fruition project [47]), is shown in **Figure 27**.

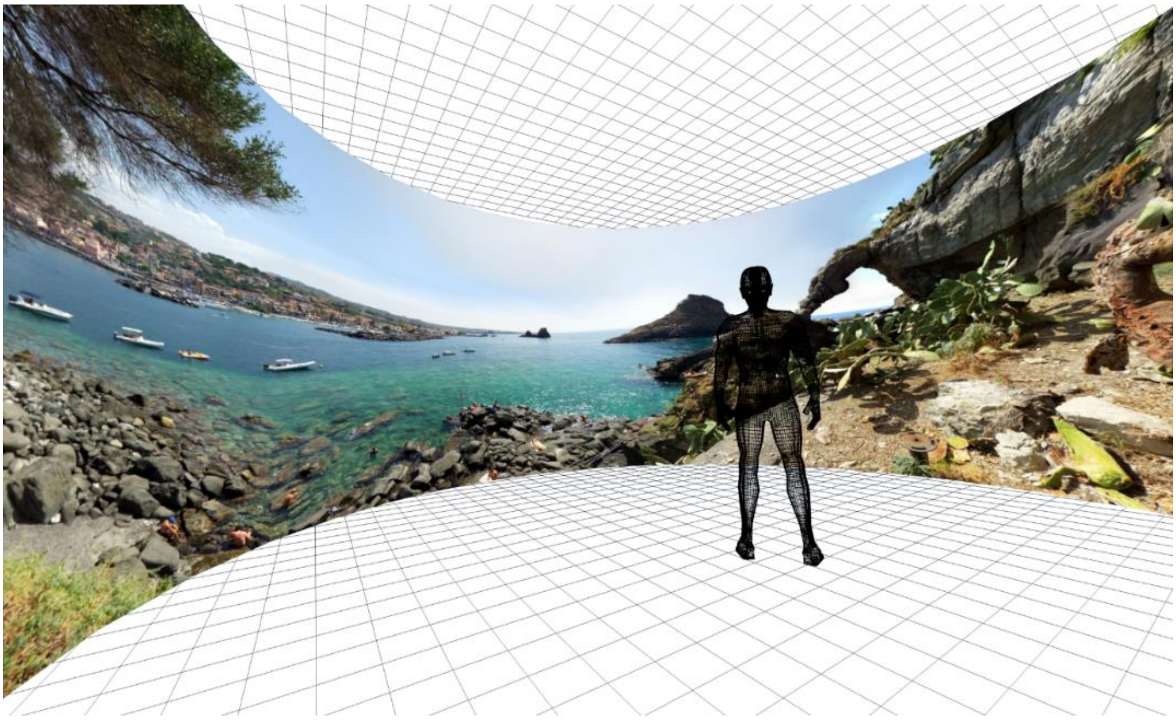


Figure 27 *- Section of Lachea Island panorama with cylindrical projection.*

Among *cylindrical panoramic layout* (*see* **Figure 27**), equirectangular (**Figure 28**) and *cubic panorama* layouts (**Figure 29**) exist.

Figure 28 *- Equirectangular panorama layout.*

Figure 29 *- Cubic panorama layout.*

2.5.3 The problem of *Stitching* and 2D panorama generation

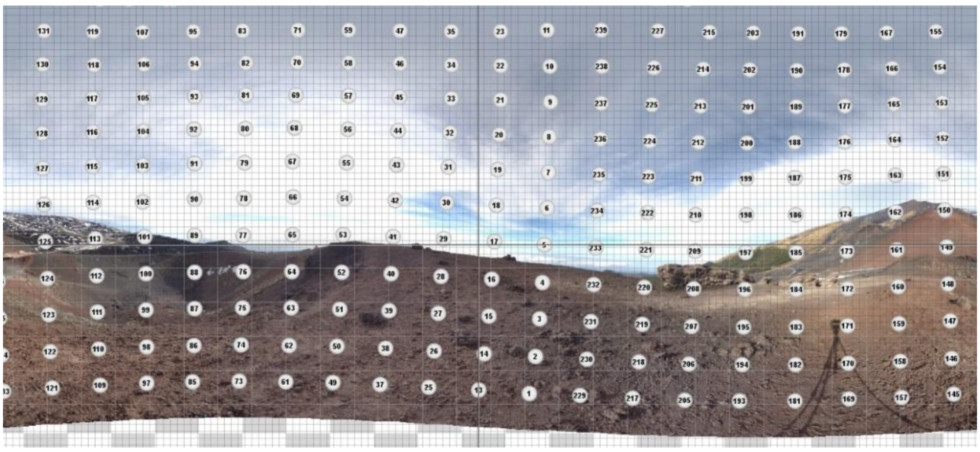


Figure 30 *- Example of a stitched panorama, where every number represents a single shot.*

Several cameras and techniques can be adopted to create panoramic photos of real environments. Since a single camera lens is not able to capture the entire spherical panoramic view of a real place in a single shot due to limits on its optics, photographers are forced to *stitch* together many photos that are taken from different angles, as shown in **Figure 30**. The process of merging single shots to produce a panoramic picture is called *stitching*.

Despite the number of photos that are used to compose the final panorama, several *stitching errors* may occur due to moving objects, illumination and color changes, lens distortions, or improper camera rotation while capturing each section of the real scene.

To solve occurring *stitching errors* researches devised several algorithms. Within this research investigation the algorithm used to generate panoramic photos was the ***Scale-invariant feature transform (SIFT)*** [48], patented [49] in 1999 by David Lowe and studied by several stitching researches [50] [51] [52]. The key stages of the SIFT algorithm are discussed below:

- <u>*Scale-invariant feature detection*</u>*.* In this stage each image is converted into a collection of *feature vectors*, which are lists of specific structures that describe each detected object of a scene through its characteristics (e.g. position, orientation, size, illumination, geometry). For this algorithm, selected features are invariant to image transition, scaling, rotation, and partially invariant to illumination changes and local geometric distortion. Through image processing algorithms that make use of the difference of Gaussians function, relevant information is stored into *SIFT keypoints*, which specify a 2D location, a scale, and an orientation. All *SIFT keypoints* are then stored into a database.
- <u>Feature matching and indexing</u>. *SIFT keypoints* are indexed and matched between images using optimized nearest neighbors algorithms. To do it, Euclidean distances and probability techniques are used, resulting in a collection of possible matching keypoints.
- <u>Cluster identification by Hough transform voting</u>. Using the *Hough transform* [53], which is a feature extraction technique to find imperfect instances of objects within a certain class of shapes, keypoints are compared to identify a possible pose of the 3D object of the scene they belong to, and keypoint clusters are generated.
- <u>Model verification by linear least squares</u>. By using mathematical calculations based on linear least squares to devise the model pose, each generated cluster is verified.
- <u>Outlier detection</u>. Outliers are removed by comparing the features of the images with the generated model pose. Using a probabilistic model and Bayesian probability the likelihood that the object is present in an image is calculated. Images where the object is present are then considered to be stitched together, taking into account the found SIFT keypoints.

Further information on other algorithms such as ***Speed up robust features (SURF)*** [54] [55] and ***ORB*** [56], [57], as well as additional material on stitching optimization techniques [58], [59], [60], [61] are suggested to the reader, even if not used in the context of this thesis.

2.6 Mono and Stereo Panorama Acquisition

The combination of 3D perception with panoramic visualization can offer higher levels of details and richer environment representations. However, panoramic visualization is still an open issue in terms of panorama acquisition, due to the difficulty of stitching several pictures avoiding distortions and visual errors, especially when dealing with 3D panoramas.

It's deemed useful to explain possible options and to understand the performance of different panorama acquisition methods and cameras, to refer to my published investigation [62]. In that study I analysed the following setups:

A. 3D Panoramas captured by **stereoscopic cameras** (see example in **Figure 31**);
B. 3D Panoramas captured by **smartphone app** (e.g. Google Cardboard Camera App);
C. 2D Panoramas captured by **single lens cameras**;
D. 2D Panoramas captured by **two wide lenses cameras** (see example in **Figure 32**);
E. 2D Panoramas captured by **single wide lens action cameras** (see example in **Figure 33**).

Figure 31 *- Fuji W3 Stereo Camera*

Figure 32 *- Ricoh Theta S Panoramic Full-spherical Camera*

Figure 33 *- Flylink Action Camera*

A summary of most relevant results is presented through **Table 1**, **Table 2**, **Table 3**, and **Table 4**.

The included advantages and disadvantages columns are considered useful to give readers an understanding of the achievable performance (according to systems' specifications).

3D Panoramas captured by stereoscopic camera	
Camera specifications	Observations
• Sensor Resolution: 10.0 Megapixel • Optical Sensor Type: CCD • Optical Sensor Size: 1/2.3" • System: TTL contrast detection • Focal Length Wide: 35mm • Max Aperture Wide: 3.70	• Advantages: ○ Higher resolution images ○ Automatic exposure for HDR panoramas • Disadvantages: ○ Expensive setup, long process ○ Editing needed to correct colours and distortion

***Table 1**. Summary of pros and cons of 3D panoramas captured by stereoscopic cameras.*

3D Panoramas captured by smartphone app	
Advantages	Disadvantages
• Easy and fast panorama acquisition • Multiplatform application • No need to buy expensive cameras or pan tilt units	• Non-spherical panorama (cylindrical) • Computed parallax from a single lens (possible errors) • Errors when moving objects in the scene

***Table 2**. Summary of pros and cons of 3D panoramas captured by smartphone app.*

2D Panoramas captured by two wide lenses cameras	
Camera specifications	Observations
• Number of Lenses: 2 • Video Resolution: 1920 x 1080 at • 30 fps/16 Mbps • Photo Resolution: 5376 x 2688 • (24 Megapixels) • Aperture: f/2.0	• Advantages: ○ Easy and fast acquisition ○ Cheap system, automatic stitching ○ Pan tilt unit not needed • Disadvantages: ○ Less resolution than panoramas created using pan tilt units ○ Large field of view of the lenses may introduce distortions

***Table 3**. Summary of pros and cons of 2D panoramas captured by two wide lenses cameras.*

2D Panoramas captured by single wide lens action cameras	
Camera specifications	Observations
• Lens: 170-degree wide-angle • Video Resolution: 1080P (1920 * 1080) 30FPS, 720P (1280 * 720) 30FPS, VGA (848 * 480) 30FPS, QVGA (640 * 480) 30FPS • Photo: 12M/8M/5M • Video format: MOV • The video coding: H.264	• Advantages: ○ Very cheap system ○ Less than 5 shots needed for an acceptable panorama • Disadvantages: ○ Low resolution ○ Distorted image to be rectified and stitched in post-production ○ Long post processing

***Table 4**. Summary of pros and cons of 2D panoramas captured by single wide lens action cameras.*

2.7 3D Visualization Systems

2.7.1 3D Display Systems to Visualize 3D images

One of the earliest 3D visualization system is the *stereoscope*, which was devised by Wheatstone in 1838 [63] [64]. This device uses mirrors (see **Figure 34**) to show on each eye a different perspective of the observed scene, and therefore induce stereopsis by binocular parallax providing depth perception.

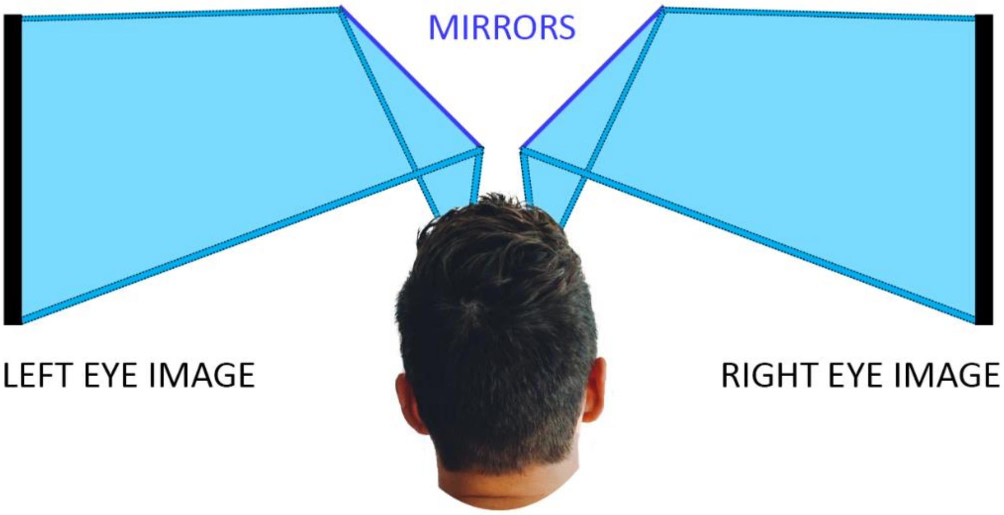


Figure 34 *- Concept representing Wheatstone's stereoscope.*

Modern 3D Visualization systems can be classified according to the stereoscopic-3D approach employed to display 3D images. According to Livatino et al. [33] the different approaches can be grouped as:

- **Passive stereo**, which makes use of space-multiplexed images that can be visualized using *anaglyph* glasses (with color filters), *polarized* glasses (with polarized filters, see **Figure 35**), or *separated* displays located very close to viewer's eyes (e.g. HMDs);
- **Active stereo**, which makes use of time-multiplexed images that can be visualized through *shutter glasses* (see **Figure 35**) offering LCD (liquid crystal display) shutter panels synchronized with the visualization display;
- **Autostereoscopic stereo**, which uses special reflecting layers together with *parallax barriers* and *lenticular sheets* [65] that provide the correct 3D perspective of the virtual environment, by showing only some portions of the screen in relation to viewer's eyes position, with no need for googles.

Figure 35 *- Polarized glasses (left) and Nvidia Shutter glasses (right).*

Among 3D approaches, different display systems can be used to visualize stereoscopic media. These are responsible for the degree of immersion, isolation from the surrounding environment, and image quality of the observed photographs. We can distinguish display systems into:

- **Handheld**, which are portable and are suitable for the passive stereo 3D approach (e.g. Smartphones, Tablets, LED screens);
- **Desktop / Laptops**, which are suitable for the active and passive stereo 3D approaches, and offer CRT or LCD displays;
- **Walls**, which offer large screens with different projection modalities and different display structures;
- **Rooms**, which can provide cubic or multi-part screens to enable high user involvement (e.g. CAVEs);
- **Tables**, which can offer different projection modalities and table-like displays;
- **Head Mounted Displays (HMDs)**, which offer very small displays located inside portable helmets at very short distance from the viewer, enabling higher levels of isolation from the surrounding space.

Recently a great interest has been paid towards HMDs because of the latest technology advances. This interest, motivated by the advantage of high isolation from surrounding space, has got new momentum in the last years because of the technology advances. HMDs (also referred as VR Headsets) well meet the needs that virtual reality observation has, to enhance quality of remote observations.

2.7.2 Virtual Reality Applications

Virtual reality is typically defined as a *medium* that allows users to visualize virtual environments, by using stimuli that induce human senses (i.e. sight, hearing, taste, smell, touch) and make humans believe that the visualized environment is real.

3D visualization represents the most important stimulus the virtual reality *medium* can rely on. This happens because the *Human Visual System (HVS)* is the most dominant human sense. Hence, the most effective for perceiving surrounding environments [7], due to the richness of information the visual world offers (e.g. vision allows the perception of distant, noxious, or fragile objects [8]), to the lack of accuracy of the other senses (e.g. audition has severe accuracy limitations in perceiving directions and distances), and to the level of reliability humans put over the sense of sight (studies prove that it can bias information coming from the other senses [9]).

Figure 36 *- HTC VIVE used for medical training. Photo taken from Unsplash [66].*

To provide 3D perception in virtual reality, all 3D display systems presented in section 2.7.1 can be adopted. However, Head Mounted Displays (HMDs) have recently gained a huge interest thanks to their portability and isolation performances. For this reason, we have recently assisted to the release of several HMDs on the market, such as the Oculus Rift [10] (see **Figure 1**), the HTC Vive [11] (see **Figure 36**), and the StarVR [12], which are today used for a wide range of applications, among them:

- **Training and education.** VR delivers learning experiences with unprecedented levels of interactivity;
- **Architecture and Planning.** VR makes it possible for engineers and manufacturers to experience their creations before they are built. This also reduces production costs and development times;
- **Entertainment and gaming.** Games in VR can offer real-life simulated environments and intuitive interfaces that allow players to participate more naturally with higher levels of realism;
- **Simulations.** VR make complex situations easier by offering a safe environment for training purposes with higher accuracy and interactivity (e.g. flight simulators, see **Figure 37**);
- **Conferencing.** VR brings digital workers together in digital meetings and conferences, removing the barrier of distance and enabling higher levels of telecommunication;
- **Help and healing.** VR can help people with phobias or pathological issues to better tolerate their fears and recover from traumas.
- **E-tourism.** VR offers the possibility to visualize places around the world realistically, regardless of the viewer's actual location. This has a great potential, especially for places not easily accessible to humans.

Figure 37 *- Example of Simulation, with child trying roller coaster in virtual reality. Photo taken from Unsplash [67].*

2.8 VR Headsets

HMDs can be classified into:

- **Desktop VR headsets**, which require high computations and a wired connection to an auxiliary computer or laptop that offers high-end graphic cards such as Nvidia GTX 1070 [68] or 1080 [69] (e.g. Oculus Rift [10] in **Figure 1**, HTC Vive [11] in **Figure 36**, StarVR [12]);
- **Mobile VR headsets**, which require a smartphone to be used as a display to show stereoscopic content, and are portable with no wires (see **Figure 38**);
- **Stand-Alone VR headsets**, which have no needs for further devices to work and can offer an integrated operative system to interact within virtual reality (e.g. Oculus GO, in **Figure 39**).

Figure 38 *- Example of Mobile VR headset with smartphone showing 3D panorama.*

Figure 39 *- Oculus GO, example of Stand-Alone VR headset.*

The very first *Head Mounted Display* (*HMD*) was developed in 1960 by Morton Heilig [5], which initially offered only stereoscopic 3D with wide vision and stereo sound. One year later, this innovative device inspired two Philco Corporation engineers to develop an upgraded version called *Headsight*, which included a magnetic motion tracking system. Thanks to the tracking system, the new version allowed viewers to change the orientation of a remote camera through the viewer's head rotations, resulting in the natural observation of a stereoscopic panoramic video stream [70]. However, the *Headsight* lacked the integration of computer and image generation, and only showed a panoramic view in relation to the actual movements of a real stereoscopic camera.

When virtual reality started to become more popular, new HMDs have been developed also to offer enhanced 3D visualization of computer-generated environments. Today many applications are available on the market (e.g. Oculus Store VR apps), offering virtual reality experiences based on remote observation of real environments (see example in **Figure 40**).

Figure 40 *- Virtual Reality used to visualize a remote environment.*

2.8.1 Mobile VR

Compared to Desktop VR, Mobile VR gives the possibility to use smartphones inside a headset to visualize virtual environments in stereoscopic panoramic mode. These devices can be purchased at considerably lower prices than other types of headsets as they do not need additional hardware or computer to work. We can classify Mobile VR headsets into:

- **Cardboards**, which are usually foldable and offer a very flexible but fragile structure (e.g. Google Cardboards, see **Figure 41**);
- **Rigid headsets**, which usually provide a plastic case with adjustable lenses and variable screen distance, and optionally buttons to control the smartphone via touch screen, Bluetooth, or direct connection (e.g. VR Shinecon in **Figure 2**, Samsung Gear VR in **Figure 42**).

Figure 41 *- Example of Mobile VR cardboard headset.*

Figure 42 *- Samsung Gear VR, example of Mobile VR Rigid headset.*

Despite the advantages of portability, low costs, and ease of use, the main disadvantage of Mobile VR headsets comes from the lack of guidelines relative to the smartphones to be used: advertisers simply suggest minimum and maximum display sizes compatible within the HMD, but usually without any indication on the characteristics of the lenses, on suggested display image resolution / pixel density, and on the general optimal configuration between smartphones and HMD. This cause viewers to be biased by HMD's incorrect configurations and believe that virtual reality is not suitable for high quality realistic observation of remote environments.

However, to reduce quality issues some companies like Samsung developed Mobile VR headset specifically designed for only a limited number of smartphones, which meet required specifications not only in terms of screen size but also of display technical characteristics.

2.9 High Dynamic Range Images and Virtual Reality

When shooting scenes that present very different light conditions (e.g. an indoor environment with the sunlight coming from a window and darkness within the areas not illuminated by the Sun), a single camera exposure value can be not enough to achieve a photograph free of over-exposed or under-exposed regions.

In such condition, the human eye continuously adapts the aperture of the pupil to the amount of light that is directed towards the retina, not only to achieve a clear view of the scene but also to prevent possible damages of the photoreceptor cells. This mechanism is called **eye adaptation**.

When capturing panoramic pictures, cameras can adapt as well to the intensity of light and change the exposure value automatically. However, when the photos need to be stitched each of them might present different exposure values, resulting in non-uniform colours and light representations.

To avoid this, photographs can decide to use **High-dynamic-range imaging (HDRI)** [71], which is an HDR technique to reproduce greater dynamic range of luminosity. The common procedure implies the acquisition of multiple photographs of the same scene taken with different exposure times, and then combined in post-production through different types of algorithms [72].

In addition to the problem of HDR capture of a real scene, the second problem is how to show the captured photograph on display, as current technology is still unable to reproduce accurately high-dynamic-range images on commonly used screens. To solve the problem, HDR images are converted into **Low-dynamic-range** (LDR) images, which are compatible with common display capabilities. This conversion is performed through **tone mapping operators** [73], which select the optimal appearance of the colours and approximate the appearance of high-dynamic-range images in a medium that has a more limited dynamic range (see **Figure 43**, which is the tone mapped panorama resulting by merging **Figure 21** and **Figure 22**).

Some editing software like Adobe Photoshop can store HDR images using a floating-point numeric representation that is 32 bits long (32-bits-per-channel); when storing images at 16 bits or 8 bits, luminance value are stored only from black to paper white, reflecting an extremely small segment of dynamic range in the real world. Adobe Photoshop offers a tool called *Merge to HDR Pro* that gives the possibility to create HDR images by combining multiple photographs captured at different exposures [74].

Figure 43 *- Example of tone-mapped HDR panorama, resulting from* ***Figure 21*** *and* ***Figure 22****.*

2.10 Eye Tracking in Virtual Reality

We can distinguish three main categories of eye tracking approaches:

- **Eye-attached tracking**, which measures the movement of a contact lens attached to the viewer's eye;
- **Optical tracking**, which does not require direct contact to the eye and usually relies on the use of infrared cameras and image processing algorithms to detect the position of the pupil and estimate the point of gaze or the motion of the eye relative to the head;
- **Electric potential measurement**, which makes use of electrodes around the eyes to estimate electric potentials.

Some virtual reality headset provides optical tracking functionalities to monitor viewer's eyes movements and understand the area of the display that is observed (e.g. FOVE VR headset [75] with optical tracking, see **Figure 44**). This represents a great advantage to improve the presentation of the virtual environment in relation to the portions of the scene that the viewer watches the most. Techniques such as **foveated rendering**, which reduces the rendering workload of the areas of the scene that are situated in the peripheral vision (outside the zone gazed by the fovea, which is the area of the eye responsible for sharp central vision), or **heat maps** to analyse viewer's attention (by showing areas of the virtual space that have been observed the most), are some examples.

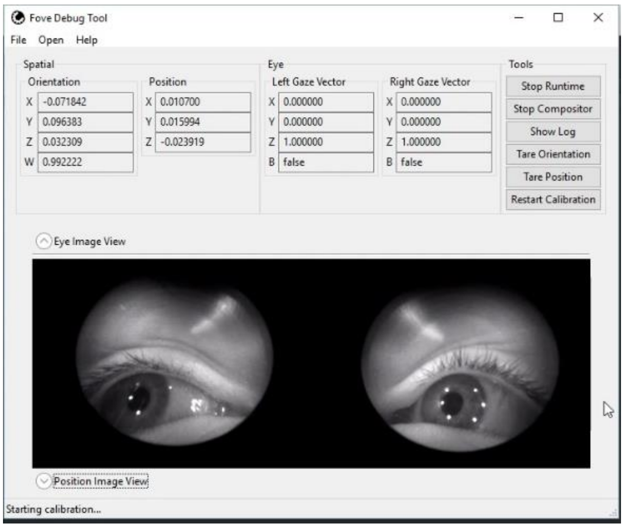


***Figure 44** - Example of Optical tracking, performed by the FOVE VR headset [75].*

2.11 Usability Evaluations in VR

The increasing amount of virtual reality applications and devices released on the market, and the different response of VR users to the quality of the experiences, highlights the need to test and scientifically evaluate achieved performances through usability evaluations.

This is also suggested by Livatino et al. in their work "Simple Guidelines for Testing VR Applications" [76]. The main reasons why the usability evaluations are needed, are to faithfully investigate human behaviour, visual perception, and task performances of users, when they are exposed to virtual reality observations. Guidelines and rules to perform usability evaluations in VR are deemed necessary: to avoid possible risks that could arise by the participation of test users, to properly use new devices, and most importantly to acquire data in an unbiased way.

To correctly perform a usability evaluation, researchers suggest to:

- Prepare in advance the specific design to be used for the user study, which includes preparation for testing environment and testing setup;
- Select the number of participants that should be involved based on their personal characteristics if necessary;
- Write forms to be compiled by test users (e.g. information sheet, consent form, questionnaires, task scenarios, data collection forms, ethical approval);
- Establish a schedule for the procedure to be followed to reduce the risk of biased results;
- Perform pilot studies to improve chosen design and analyse possible risks;
- Perform formal studies based on the outcomes of pilot studies;
- Analyse results through statistical tools (e.g. mean, median, frequency distribution, standard deviation, t-test).

2.12 Conclusion

This chapter offered an overview of the main topics discussed by this thesis, to help the reader achieve a better understanding of the subject and facilitate the reading of the following chapters.

An introduction to the Human Visual System and direct observation of existing places was presented, focusing on the role of stereoscopy, monocular cues, and binocular cues. Furthermore, the possibility to emulate the direct observation of real places through remote visualization was discussed, emphasising existing techniques for capturing panoramic and 3D pictures, and the use of 3D visualization systems.

Finally, the possibility to interact with a digital visualization system using VR headset was proposed, together with an overview of existing hardware and headsets for VR visualization systems (e.g. Desktop HMD, Mobile VR headsets, HMD with eye tracking capabilities). Furthermore, High Dynamic Range photography was presented as a possible improvement for the acquisition and visualization of observed environment in Virtual Reality, to faithfully reproduce its different lighting conditions.

CHAPTER 3. PROPOSED INVESTIGATION

3.1 Aims and motivations

This research investigation aims at understanding the role of technology setup and camera-display parameters in providing convincing visual realism, for reducing the perceptual gap between direct and indirect observation of real places when using Mobile VR headsets.

The research towards the pre-fixed aim is proposed to advance through the study and identification of most relevant features that are responsible for delivering a convincing perception of visual realism, while maintaining high sense of presence, suitable depth perception and comfortable viewing.

This study is motivated by the lack of setup and design guidelines on recently emerging VR headset systems, which potential is being increasingly confirmed towards accomplishing a comfortable and realistic visualization of remote environments. Therefore, outcomes of this investigation may represent meaningful knowledge to generate much improved indirect observations of existing environments, to the benefit of realism critical remote presentations. Applications include tele-intervention, tele-exploration and tele-observation of remote inaccessible areas, telerobotic, telemedicine, e-tourism, and entertainment.

The general concept for this investigation is to first learn from both established theories and recent literature contributions, to then proceed with sets of experiments aimed at providing or denying hypothesis related to VR headset viewing. The proposed methodologies combine what can be referred as ***traditional*** and ***action research*** [77] [78] (see **Figure 45**). This combination, which can be seen as a revolving traditional research approach, is believed very suitable for problems were theories about the mechanics of human visual systems and their interaction with displays and virtual reality systems, need to be combined with practical observations and experimental findings.

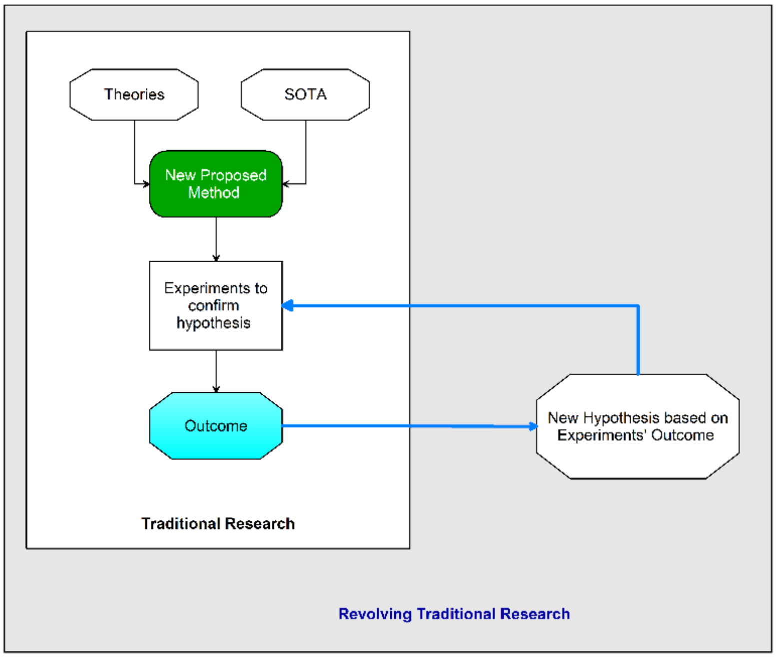


Figure 45 *- Revolving Traditional Research approach.*

3.1.1 Research Questions

This PhD work research questions can be summarized into the following fours interrogation statements:

1. **Display**. How display and image parameters influence visual perception of remote environments in virtual reality in terms of realism, presence, depth-perception and comfort?
2. **Environment**. How environment characteristics affect visual perception of remote environments in virtual reality in terms of realism, presence, depth-perception and comfort?
3. **Familiarity**. How place familiarity affects visual perception of remote environments in virtual reality in terms of realism, presence, depth-perception and comfort?
4. **Eye-Adapted HDR Viewing.** How effective is eye-adaptation driven HDR viewing in terms of realism, presence, depth-perception and comfort in VR?

3.1.2 Research Development Phases

The proposed research development plan includes three study and development phases:

1. **Systematic Review for Learning**. A deep study of most relevant camera-display parameters to understand the role they play in remote real-place observation. This phase is implemented through a systematic review of the state-of-the-art.
2. **Build Knowledge through Assessments**. A set of experiments aimed at assessing real-place observations after some relevant factors that have been identified as potentially relevant for VR headset observation, and of interest for VR system designers. These factors include the roles played by place familiarity, display technology, and visualized environment.
3. **Focused Evaluation**. A set of focused trials with the purpose of further investigate previously assessed elements within a specific VR headset technology. For this purpose, the use of eye-tracked headsets is further assessed.

Through the three phases this research aims at gathering enough information to provide VR systems' designers with targeted knowledge and guidelines for devising and developing the capture and visualization of real scenes to be used on VR headset displays.

Figure 46 shows timing for the proposed research development and the foreseen overlapping among the three phases.

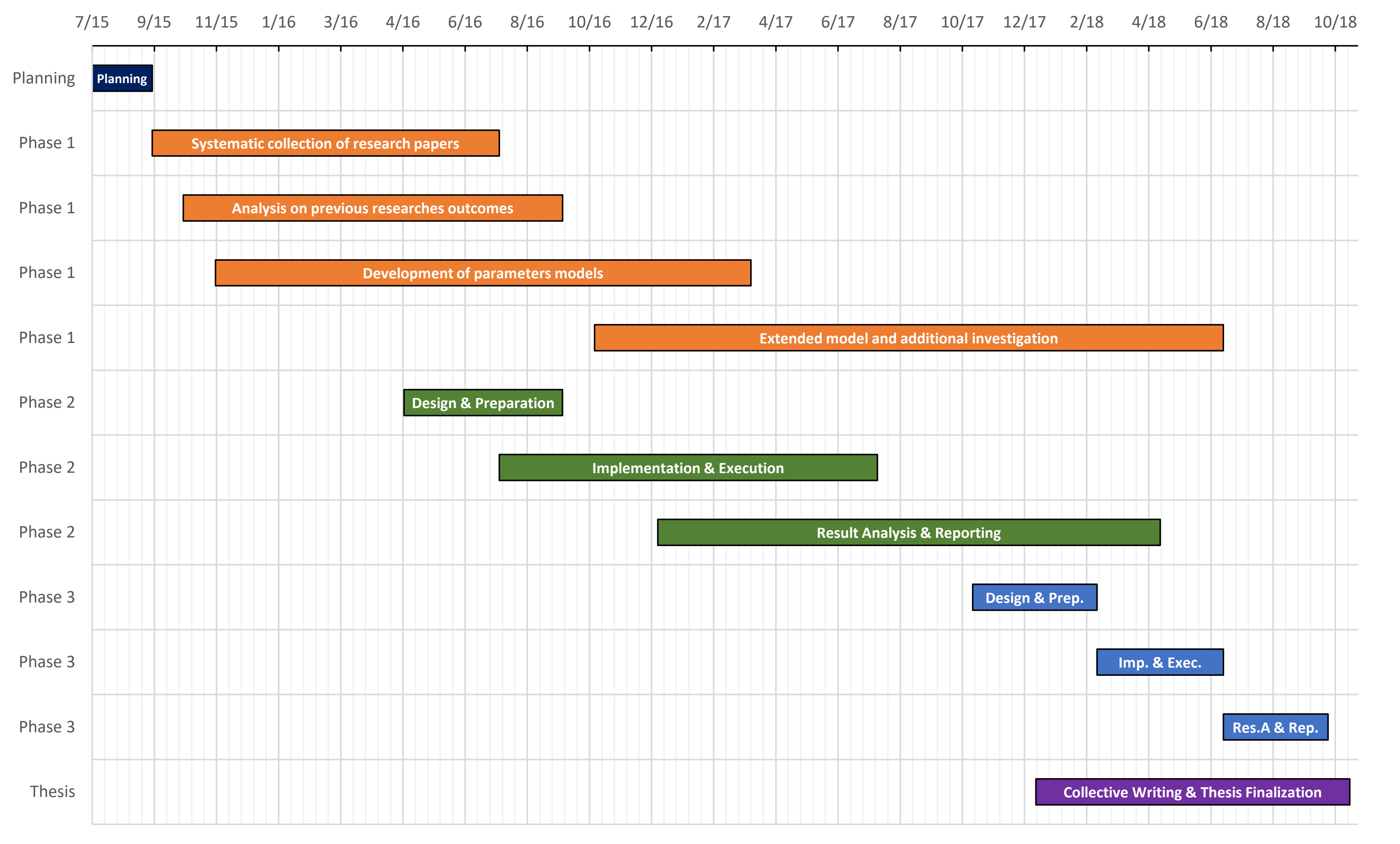


Figure 46 *- Gant Chart of the proposed Research Development phases.*

3.2 Phase One - Systematic Review for Learning: Role of camera-display parameters in remote real-place observation.

It is proposed a systematic literature review on the role of camera-display parameters within VR applications in relation to relevant human factors. The proposed review is articulated through four steps:

1. **Systematic collection of research papers** from the literature, discussing problems affecting indirect visual observation of remote places using 3D technologies (see section 4.1);
2. **Analysis on previous researches outcomes** considering collected papers from the previous step (see section 4.2);
3. **Modelling relevant parameters** to better analyse discovered interactions and setups for the remote visual observation system (see section 4.3);
4. **Extended model focused on presence and distance estimation** with further research papers collection and new entries (see section 4.4);

The outcome of this phase will provide a clear picture of findings and relevant relations between camera-display parameters and human factors solely based on previous works. This will represent for us a baseline to rely on when designing the experiments of phases two and three.

3.3 Phase Two - Build Knowledge through Assessments: Experiment the role of Familiarity, Environment, and Display

Several experiments will be planned and performed during phase two of this investigation. These will aim at gaining evidence on the role of camera-display parameters and human factors towards delivering a convincing perception of visual realism, while maintaining high sense of presence, effective depth perception, and comfortable viewing. The designs for these experiments will be based on outcome and findings of phase one. This will specifically affect the questionnaire design and result interpretation of the proposed user studies.

3.3.1 Experiment Objectives

The assessments that are proposed for phase two will group data and evaluate results after three research questions.

1. **Familiarity**. How place familiarity impacts the perception of the remote environment through virtual reality, in terms of realism, comfort, sense of presence, and depth?

The perceived realism is more carefully assessed by those who know the place compared to those unfamiliar with the location, who can nonetheless still judge based on appearance of well-known objects (e.g. trees) and phenomena. Furthermore, there is indication that perceived realism is also connected to the achieved sense of presence and other human factors such as comfort.

Different people are called to visit a place through a VR headset. For this purpose, the following two groups are defined:

- Experts. I.e. people who have already visited the observed place and know it well.
- Non-Experts. I.e. people who are viewing the place for the first time through a virtual reality display.

Then, a user study is conducted to compare the judgements in terms of realism and other involved human factors, provided by people who know the place well and those who do not.

2. **Environment**. How different environments such as indoors and outdoors affect the visual perception of the remote environment in terms of realism, comfort, sense of presence, and depth?

The appearance of the shown place may play a role on the perceived visual impression. Thus, two locations showing different atmosphere and illumination are assessed. Their main differences consist of:

- Outdoor vs Indoor
- Natural Illumination vs Artificial Illumination
- Wide-Color Spectrum vs Narrow-Color Spectrum
- Great to Medium Distances vs Small to Medium Distances

Then, a user study is performed to comprehend the role of the above listed (contrasting) elements and their consequence on perceived realism and other involved human factors.

3. **Display.** How display and image parameters influence the visual perception of the remote environment in virtual reality in terms of realism, comfort, sense of presence, and depth?

Display features such as size, resolution, brightness, color, etc., may affect the relevant human factors, thus the visual perception of virtual reality users. To test different values of such features and estimate their effects, two smartphone screens will be chosen having different specifications in terms of the following list of features:

- Image brightness, and display light intensity;
- Color intensity, hue intensity, saturation of display, vividness, and contrast;
- Display resolution, number of pixels and pixel density, and sharpness;
- Distortions due to used lenses;
- Display size;
- Display field of view.

Then, a user study is performed to compare the obtained results with those expected from the state-of-the-art, to understand and assess the role of the display.

3.3.2 Usability Evaluations

3.3.2.1 Design

The user studies will analyse usability of proposed technologies, including acquisition and display setups, in terms of the provided viewing experience and related relevant human factors. Tests are both within and in-between subject evaluations, depending on the specific experiment. The user studies are designed according to recommendations gathered from the literature [79] [80] [81] [82]. Participants have different levels of experience with VR devices and computer games and will operate under the same conditions.

The above design is applied to two sets of user studies:

A. ***Familiarity and Environment.*** Implementation and Results Analysis are described in Chapter 5.

B. ***Display and Environment***. Implementation and Results Analysis are described in Chapter 6.

3.3.2.2 Procedure

The participants are asked to observe several panoramas with different setups. Each person runs several trials on each VR facility and/or configuration. The task/facility sequence assigned to each participant is set according to a pre-determined schedule, to counterbalance sequence of tasks and avoid fatigue and learning effects. At the start of each trial each participant ensures to correctly and comfortably wear the headset. After observation, test users are asked for filling in a questionnaire. In doing so, each test conforms to the traditional approaches in terms of forms and questionnaires [76] with few additions, including the use of an information sheet, consent form and pre-test screening.

3.3.2.3 Extended Pilot test

During this investigation I was inspired by usability evaluations guidelines for pilot studies provided by Livatino and Koeffel [76], aimed at optimizing questionnaire design, evaluation procedures and data reliability. When designing the above described procedure concept to the proposed user studies, it is expected that the task sequences would play a relevant role to the outcome. This is particularly appropriate when testing VR applications. Therefore, I decided to setup the specific procedural steps according to an extended pilot test.

The outcome of this extended pilot test will result in an improved data collection and higher data reliability. Furthermore, by running the proposed extended pilot test it will be possible to gather useful feedback for improving the user study procedure. The design and outcome of the pilot test are presented in section 5.3.

3.3.2.4 Results Analysis

Once the trials are completed, answers will be stored on an Excel file and T-tests will be used from the Excel Data Analysis suite to analyse them.

Mean values and standard deviation of acquired data will be calculated, and the statistical significance of results will be measured by estimating the Student's *T* distribution for paired and unpaired comparison (depending on the specific group of test users) with repeated measures. When considering different sets, a *p*-value is going to be estimated. I decided to set p=0.05 as threshold, which is the conventional threshold used in literature according to Fisher [83]. Furthermore, the standard error of the Mean (SE) for each comparison will be estimated. Moreover, what follows is going to be used for each test:

- An ordinal scale of measurement based on the *5-Likert* and *7-Likert* scale.
- Plotted data in a bell-shaped normal distribution curve.
- A reasonable large sample size, which approaches a normal bell-shape curve.

Specifically, the sample size used for each of the VR setups (i.e. Lachea with Experts, Lachea with Non-Experts, Monello with Experts, Monello with Non-Experts) is uniform and equal to 20 participants.

In certain cases, a correlation matrix might be used to analyse the trend of each variable when other variables change. Furthermore, even if sometimes ANOVA could be used, I will prioritize combinations of several t-tests, to better understand which specific group of values presents significant differences.

Diagrams and perform graphical display will be plotted [76] in the proposed analysis while looking for evidence in the state of the art that backs or contradicts my findings. This way the field knowledge in this subject can become more solid for some aspects.

The diagrams will be presented together with text description of the displayed outcome. This will be followed by an "Analysis" section summarizing the most significant results and the most relevant presuppositions that have been confirmed by the relative usability evaluation. Eventually, a "Guidelines" section will summarize main experimentations outcome in terms of concise guidelines for system designers. Both the Analysis and the Guidelines are laid down looking at the evaluation outcome against the four elements that have been considered relevant contributors to remote visual observation: realism, comfort, presence and depth-perception.

3.4 Phase Three - Learning from Usability Results: Experiment the Role of Eye-Adaptation Driven HDR Viewing.

An experiment is planned aimed at further investigate previously assessed elements within a specific VR headset technology. I target the use of eye-tracking in VR headsets and its role towards delivering a convincing perception of visual realism, while maintaining high sense of presence, effective depth-perception and comfortable viewing. The experiment is designed to include lessons learnt from the previous two phases.

3.4.1 Experiment Objectives

The assessment that is proposed for phase three will group data and evaluate results after the following research question.

Eye-Adapted HDR Viewing. How effective is eye-adaptation driven HDR viewing in terms of realism, presence, depth-perception and comfort?

Human vision includes seeing high-range light-intensity images on the retina foveal area and the eye rapidly moves to observe difference portions of an area of interest. The pictured appearance of whatever we observe varies therefore according to this principle and mechanic.

While the above happens during direct observation of objects in real environment, one does not achieve the same experience during indirect observation of objects in real environment, and this is the case for observing through a VR headset too. In particular, the displayed image does not change its light-intensity according to eye movements, neither with large ranges of light-intensity changes.

Having today available both HDR image capture and eye-tracked VR headset, it is believed that there are to some extend the conditions for replicating natural direct viewing. For this reason, I decided to investigate the simulation of eye-adaptation in VR based on eye position and the appropriate light-intensity range of the targeted image area.

It is therefore presented a method capable of:

A. Generating different light-intensity ranges for different image-portions based on eye focused area. This will imply the use of HDR images.
B. Selecting the appropriate light-intensity range based on viewer's eye position.

With such system in place the purpose of this experiment is to assess its effectiveness towards a more natural viewing experience, therefore possibly with an impact on realism, presence, depth-perception and comfort.

Within this purpose, the following three different setups will be used for the user study:

- HDR static and static illumination;
- HDR dynamic with eye-adaptation based on the user's Head Rotations **only**;
- HDR dynamic with eye-adaptation based on the user's Head Rotations **and** Eyes Movements.

Therefore, I will run a user study to comprehend the role of the above listed setups to understand their effectiveness and consequence on perceived realism and other involved human factors. Further details are discussed on Chapter 7.

3.5 Conclusion

This chapter thoroughly presented the stages of the investigation proposed by this thesis. In particular, the aims and motivations to conduct this research were discussed, highlighting the need for design guidelines of VR headset systems, in the context of remote visual observation of existing places. This involved the introduction of this research's questions and the discussion of the development phases that were designed to answer them.

In phase 1 a systematic review for learning was introduced, with details on the procedure adopted for the papers collection. In phase 2 two usability evaluations were introduced to build new knowledge from what was learned with the previous systematic review. In phase 3 a third usability evaluation was introduced, as an additional investigation based on the outcomes of the previous two usability evaluations. This will focus on the use of eye-adapted HDR remote observation in VR.

In summary, the complete design of this research was described by this chapter, to help the reader follow the logic behind all the choices taken, and to better understand the structure and contents of the following chapters.

CHAPTER 4. SYSTEMATIC REVIEW FOR LEARNING: ROLE OF CAMERA-DISPLAY PARAMETERS IN REMOTE REAL-PLACE OBSERVATION

4.1 Systematic collection of research papers

Nowadays the fantastic potential of virtual reality makes it possible to observe real places remotely by making use of immersive displays, high quality stereoscopic panoramic cameras, 3D panoramic pictures, and 3D panoramic videos. This promising possibility has been empowered by the introduction of Mobile VR devices, which made this technology portable and accessible to the mass market.

However, in certain cases the visual representation of the captured environment does not look as faithful as real life. The responsibility of this lack of realism lies on technical characteristics of the chosen observation system, which governs the acquisition, reproduction, and visual perception of the portrayed location. This leads the viewer to still notice the difference between the physical appearance of the real place and the 3D panoramic photos depicting it through the Mobile VR system.

Furthermore, motion sickness and visual fatigue have been reported by many virtual reality users due to various system faults, including inaccurate system calibration, inadequate performances, and inconsistent sensory information. Despite the need to better understand how to improve the quality of the remote observation achieved by Mobile VR systems, it is still difficult to identify optimal setups, as well as having a clear picture of parameters significantly affecting the produced visual and spatial perception of 3D panoramas in VR.

To better understand previous researches in this field and identify significant parameters that affect realism, comfort, sense of presence, and depth perception in VR systems, this chapter presents a detailed analysis of the state-of-the-art that I performed for this thesis. This literature investigation is divided into two main stages:

1. **Preliminary study on 3D Acquisition and 3D Visualization Systems**, to understand the influence of 3D cameras and 3D displays over viewer's visual perception. This includes:
    a. A deep analysis on scene complexity, scene depth, and scene dynamics;
    b. Guidelines on stereoscopic camera setups, focusing on internal and external camera parameters;
    c. Display parameters, presenting guidelines to increase the level of realism, reduce discomfort, and enhance depth perception;
2. **Analysis focused on parameters influencing Presence and Distance Estimation in Virtual Reality**, to investigate human and artificial factors that have an influence over depth perception and sense of presence when using virtual reality devices.

All information retrieved from the state-of-the-art highlighted the absence of a global model that summarizes the influence of all identified parameters over realism, comfort, depth perception, and sense of presence in the context of 3D Visualization Systems.

To fill the gap identified in literature, and to conduct further investigations on a subset of detected significant parameters, I devised the global model of a 3D Visualization System Setup with the aim of producing a clear picture of elements affecting the final visual perception of the viewer in virtual reality and Mobile VR. Procedure used for literature data collection and design of the abovementioned 3D Visualization System Setup are presented in subsection 4.1.1.

4.1.1 Proposed Literature Review

For the purpose of this investigation, a literature review has been carried out to collect information on previous studies, and a comprehensive analysis of the state-of-the-art on stereoscopic 3D was performed in December 2015. To do it, a procedure inspired by PRISMA's flow diagram for systematic review papers [84] has been adopted. Within this purpose, the online research databases of *Science Direct* [85] and the *Thomson-Reuters Web of Science* (WOS) [86] have been chosen to find relevant scientific papers. The main stages of the above-mentioned procedure are summarized below:

- **Stage 1**: Selection of relevant keywords to be used for papers collection.
- **Stage 2**: Use of the chosen keywords to get a collection of papers from [85] and [86] for the current investigation.
- **Stage 3**: Papers scoring based on their titles and abstracts.
- **Stage 4**: Full-text papers reading, considering only papers having score higher than a threshold. Then, identification of parameters relevant to 3D realism, comfort and depth perception.
- **Stage 5**: Selection of the most important sentences from each paper, examination of the papers cited by each selected sentence, and reading of full text of cited papers to check if additional relevant parameters and topics were discussed.
- **Stage 6**: Grouping of all the collected sentences by topic, identifying the final complete list of relevant parameters. Identification of the main relevant areas affecting 3D realism (i.e. scene elements, 3D camera parameters, 3D display parameters) and distribution of the whole set of parameters into their relative belonging areas.

The complete workflow for the state-of-the-art analysis is shown in **Figure 47**.

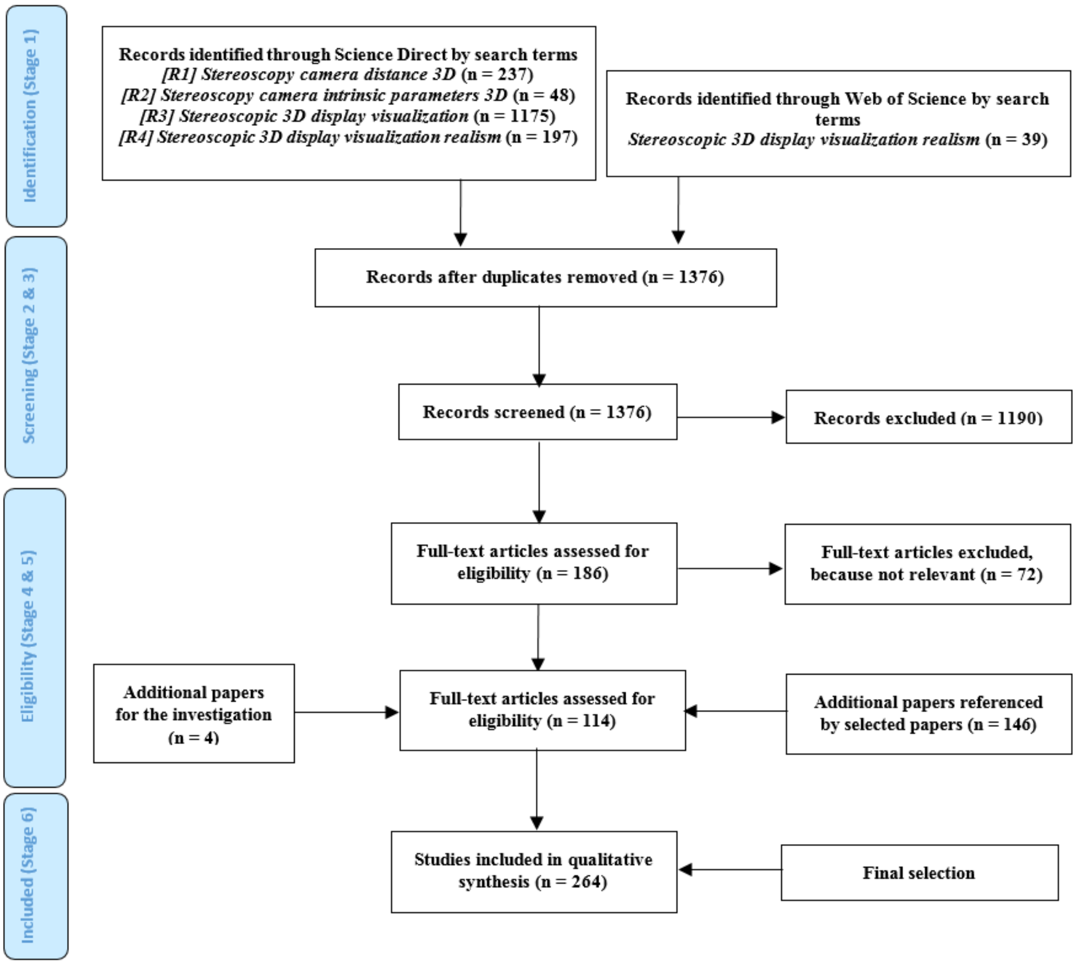


Figure 47 *- Workflow to collect investigated information from literature. This model is an extension of the PRISMA method, which is commonly used for medical researches.*

This plan was to capitalize on literature research, making collected data time-efficient and well-structured, so that this could be used to write a survey paper on the relative area of investigation.

4.2 Systematic Review Outcomes

Several researches discussed issues related to modern 3D visualization systems, in terms of realism, comfort, and depth perception. This section summarizes the most relevant discoveries and guidelines currently reported by the state-of-the-art.

To give the reader a clear overview of the topics discussed by this analysis, the following subsections have been defined:

1. **Scene elements**: discussions on the influence of the scene over the viewer's visual perception (see subsection 4.2.1). These will focus on:
    a. Scene Complexity
    b. Scene Depth
    c. Scene Dynamics

2. **3D Camera parameters**: discussions on guidelines and best practice for the setup of stereoscopic cameras, with the aim of creating higher quality 3D footage (see subsection 4.2.2).

3. **3D Display parameters:** discussions on guidelines and best practice for the setup of stereoscopic displays, with the aim of reducing discomfort, enhancing realism, and improving depth perception when viewing captured 3D footages (see subsection 4.2.3).

4.2.1 Scene elements

In this section features of the scenes that influence comfort, depth perception, and realism are discussed. For clarity, they are grouped and presented into the following paragraphs:

1. **Scene Complexity**, which refers to visual complexity of a scene and the subsequent impact on human factors (see 4.2.1.1);
2. **Scene Depth**, which refers to different cues of the scene that alter user's depth perception (see 4.2.1.2);
3. **Scene Dynamics**, which refers to moving objects and dynamic features of a visualized scene and their impact on visual perception (see 4.2.1.3).

4.2.1.1 Scene Complexity

Scene complexity refers to a portrayed scenario consisting of many different not aligned objects, which affect the amount of visual information presented to a viewer. In this paragraph, further details on this topic are presented. A summary of all guidelines for scene complexity is presented on **Table 5**.

4.2.1.1.1 Scene complexity and Visual discomfort: facts and issues

What contributes to a visually complex scene is the presence of:

(1) Many objects
(2) Complex geometry
(3) Complex textures and materials
(4) Cluttered appearance

The more complex a scene is, the more time a viewer needs to "digest" it [87] [88]. This may result in visual discomfort, especially on 3D image sequences where a viewer typically wishes longer observation time.

Wilson [89] reported that complex scenes can impact frame rate transmission, which varies inversely to complexity. This particularly affects computer generated scenes. If the result of such complexity is a lower frame rate, this can generate lags, causing possible dizziness [90], which will negatively affect viewer's observation, causing discomfort [89] [91] [92].

4.2.1.1.2 Scene complexity and Visual discomfort: guidelines and solutions

Studies prove that 3D content creators should consider that a scene with few visual details can benefit comfort, while this still delivers a higher level of realism [90]. A simple scene may also refer to simpler reflection models' appearance, which increase comfort. Furthermore, specular highlights should be avoided at all costs [93]. The same for *glossy reflections* as they cause a difference in depth impression called *highlight disparity*. This is a strong factor in the perception of *gloss* [94] [95] [96] [97] and material authenticity [98], and it causes discomfort depending on disparity values [94]. Templin et al. presented a *glossy-reflection* reduction technique that can preserve realism of the scene, enhancing comfort [94]. Experiments conducted by Blake and Bülthoff confirm that most realistic gloss impression occurs when highlights are located on a concave surface or behind it [94] [95].

Scene complexity		
Comfort	**Realism**	**Depth Perception**
Highlight disparity causes discomfort depending on disparity values	Highlight disparity helps material authenticity	Avoid glossy reflections to avoid highlight disparity and difference in depth impression
Use glossy-reflection reduction techniques to preserve realism of the scene, enhancing comfort		-
Avoid specular highlights	Prioritize highlights located on concave surfaces or behind them to achieve realistic gloss impression	-
Use scene with few visual details	-	-

Table 5. *Summary of the most important guidelines for comfort, realism, and depth perception in relation to scene complexity.*

4.2.1.2 Scene Depth

Scene depth refers to the different depth levels represented visually in an image. This affects comfort zone and depth perception of viewers.

Other than binocular vision, depth perception can be induced by monocular depth cues[1], with monocular and binocular cues being additive. Landy et al. [99] reviewed psychophysical studies of human depth cue combination, and classified depth cues based on scene content. They developed a model called *modified weak fusion* (MWF), which calculates depth cues' weighted linear combination and consistency. Their model also addresses several issues, such as *promotion*, *dynamic weighting*, and *robustness* of cues.

Reichelt et al. [100] analysed near-range depth cues by comparing visual performance and depth-range capabilities of 3D displays from a physiological point of view. Their study proves that consistency of vergence - accommodation cues is essential to guarantee a natural and comfortable 3D experience.

In the following, an overview of depth cues that have a relevant influence over comfort, realism and depth perception is presented. Within this investigation, facts and issues are discussed by providing guidelines and solutions. A summary of all guidelines for scene depth is presented on **Table 6**. This topic has been analysed as follows:

- Occlusion, disparity magnitude, and motion parallax (see subparagraph 4.2.1.2.1);
- Guidelines to prevent discomfort (see subparagraph 4.2.1.2.2);
- Depth intervals and the comfort zone: guidelines and solutions (see subparagraph 4.2.1.2.3);
- Depth cue consistency and depth perception (see subparagraph 4.2.1.2.4).

[1] These include motion parallax [151] [435], vertical size, focus, perspective [139] [434] [436], relative size, occlusion, shading, and spatial frequency of textures [110] [20].

4.2.1.2.1 Occlusion, disparity magnitude, and motion parallax

Occlusion is the strongest monocular depth-cue. It appears to be independent from the distance of objects [23]. The same can be said for less strong cues such as relative size and relative density.

After occlusion, ***motion parallax*** (motion perspective) is considered the strongest cue responsible for depth perception in S3D scenes.

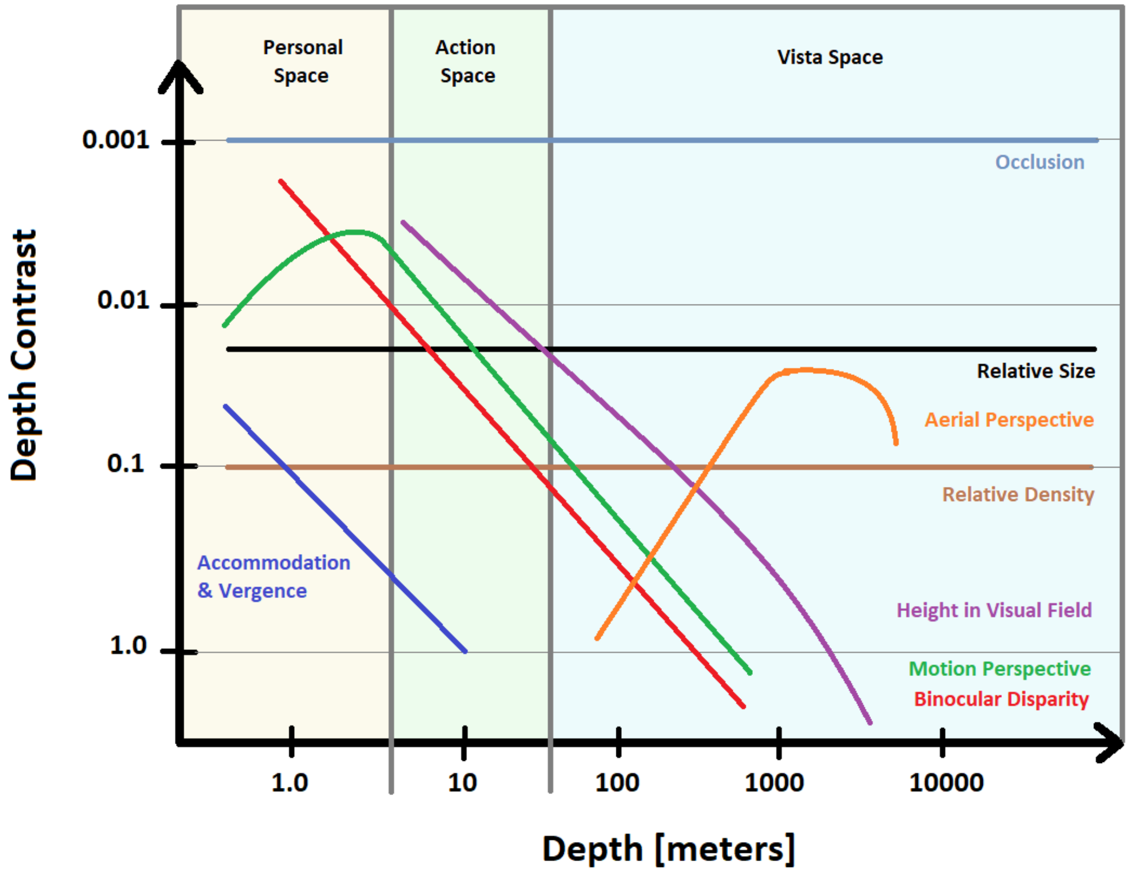


Figure 48 *- Sensitivity (just-discriminable depth thresholds) functions of common depth cues. The lower the threshold (depth contrast), the more sensitive the HVS is to that cue. This diagram was adapted from Masia et al. [101], which adapted studies made by Cutting et al. [20]*

Masia et al. [101], adapting the studies conducted by Cutting et al. [20], created a diagram (adapted in **Figure 48**) representing human sensitivity to common monocular and binocular depth cues. It is assumed there is a link between the sensitivity functions of disparity and of motion parallax [102] [103]. Bradshaw et al. studied such link and proved that the integration of disparity cues and motion parallax cues is nonlinear [104].

Disparity magnitude is one of the most accurate and efficient depth cues [23] [105], but it is also the most salient source of discomfort [106] [107] [108] [109]. The amount of disparity magnitude is regulated by:

(1) Distance between camera and objects in the scene (depending on the content of the scene);
(2) Camera separation (baseline);
(3) Cameras alignment;
(4) Viewer's eyes separation;
(5) Distance from screen;
(6) Size of display [110] [111] [112] [113] [114] [115] [116].

Furthermore, Didyk et al. considered the change in disparity magnitude (***disparity change***) between different images to produce a model on perceived depth [117]. Experiments prove that depth is perceived most effectively at curved or non-uniform surfaces, where second order differences of disparity are non-zero [117].

Li et al. [118] and Speranza et al. [106] examined the influence of binocular disparity change on discomfort [118], demonstrating that the amount of disparity change is more detrimental to visual comfort than absolute values of crossed and uncrossed disparities [106]. Additionally, experimental results prove that rapid changes in disparity magnitude over time might be a major source of discomfort [119] [120].

4.2.1.2.2 Guidelines to prevent discomfort

To prevent discomfort and adapt 3D images to different screen sizes, it is suggested to:

(1) Prioritize a restricted frequency of disparity changes [121].
(2) Pay attention to screen size [122], exposure duration, orientation, and spatiotemporal properties, which can cause discomfort [106] [123]. This is to guarantee a correct vision without diplopia.
(3) Use artificial blur to cover up imprecise disparity [23].
(4) <u>Choose disparity intervals according to screen size and viewing distance</u>. 3D movie makers adopt the "rule of thumb" to identify a comfortable amount of disparity and depth interval for their viewers, but this rule is incorrect because it does not consider the screen size and the viewing distance [93]. Therefore, rules based on screen size and viewing distance must be considered.
(5) <u>Don't exceed relative disparity between foreground and background</u>. Studies prove that relative disparity is more influent than accommodation-vergence conflict, because the vergence of the viewer's attention will switch between background plane and foreground plane: the higher the relative foreground-background distance, the higher the discomfort [124].

To do it, 3D producers can edit binocular disparity of images acquired by camera in post-production, using depth remapping and 3D warping techniques [42] [125] [126].

4.2.1.2.3 Depth intervals and the comfort zone: guidelines and solutions

D*epth interval* refers to the distance between the furthest and the closest object in the observed scene. In the study of Vlad et al. [90] a database of 3D images was used and classified according to the amount of depth intervals. Three categories were considered to evaluate the effect on realism, image quality and comfort of these intervals: small, average, and large depth interval. Large depth intervals provided more comfort to test users. Nonetheless, large depth interval's integrity is prevented by vergence-accommodation conflict [127].

To reduce discomfort and improve depth perception for viewers, 3D producers are suggested to follow these guidelines:

- <u>Scene depth consistency</u>. Ensure that objects of interest maintain the same depth by adjusting the depth brackets across scene changes [128]. This will facilitate the user to follow the visuals without stress.
- <u>Use a depth (or disparity) histogram storyboard</u> [42], which is a diagram representing the amount of depth in a 3D video sequence. This can help to control accurately amount of depth within scenes.
- <u>Stay within depth intervals</u>. Therefore, objects should not appear too close to the camera [129] [130] [131] [106] [132] [109]. The above should also be adapted to viewer's age [133] [134] [135] [136]. Objects close to the viewer within the comfort zone are perceived more

accurately. Studies prove that the accuracy of depth perception is higher for targets presented closer to the viewer, in accordance with the egocentric peripersonal space[2] perception [22].

- Use blending between scene cuts. Depth intervals could be exacerbated on scene cuts due to lags of fusion causing discomfort [128]. This is particularly relevant for horizontal image translation.
- Stay within the comfort zone. Depth intervals need to be well calibrated to ensure that objects in a S3D scene move within the comfort zone [23] [137] [119]. Such calibration mainly depends on screen size and viewing distance from the screen, in conjunction with spatial and temporal properties of the scene content [23]. In **Figure 49** (adapted from Winkler et al. [23]), values for accommodation and vergence distance are combined to show the size of the comfort zone.

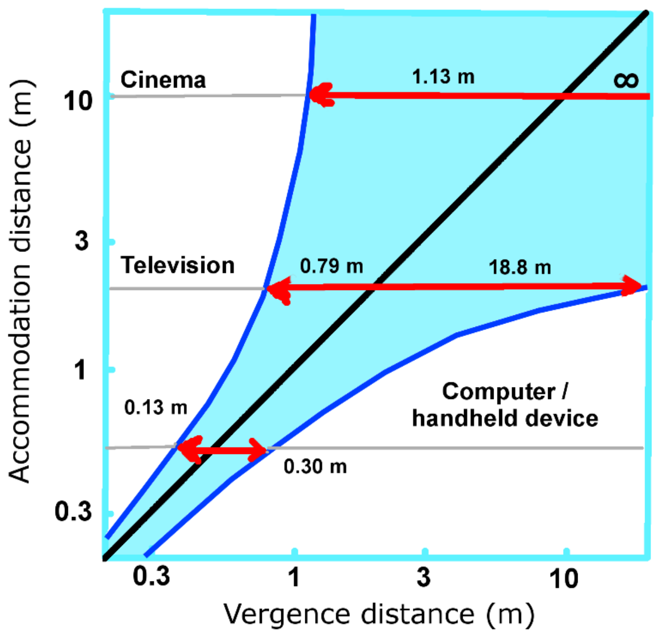


Figure 49 *- Estimated size of comfort zone, adapted from a diagram that was plotted by Winkler et al. [23] and estimated by Shibata et al. [138].*

The comfort zone size and shape vary depending on the viewer [139] [112]. Therefore, individual control over stereoscopic depth is desirable. Work by Christopher W. Tyler [140] showed that Panum's fusional area[3] depends on disparity magnitude and change.

To deal with variable-sized comfort zone, Shibata et al. provide a technique to define comfortable disparity ranges for average viewers [139].

Several comfort zone guidelines and values are suggested in literature:

- ±0.2D of depth of field (DOF) [118] [119] [137]
- ±1 arc degree of visual angle [118] [141] [106] on the basis of empirical measurements [142] [137] [141] [143] [106]
- ±3% of the screen width for 3D television [118] [144]

[2] PPS, which is the region of space immediately surrounding the viewer's body.

[3] Panum's fusional area is a space within which diplopia does not occur.

- 50mm in front and 60mm behind a desktop display viewed from 700mm (meaning that objects viewed in desktop displays will have compressed depth to remain within the comfort zone) [111]
- less than 70 arcmin (based on subjective assessment) [145] [109]
- 0.3 diopter (reciprocal value of distance) and 60 arcmin as limit of Depth of Focus and binocular disparity [145] [109].

In cinematography, crossed disparity (i.e. objects visualized off screen in front of it) should not exceed 2-3% of screen width, and uncrossed disparity (i.e. objects visualized behind the screen) should range within 1-2% of screen width. However, this is only a guideline and not a quantitative rule [138]. On a 30 foot cinema screen, practical disparity values are +30 pixels (behind the screen) and -100 pixels (in front of the screen) for a 2048 pixels video [146] [147].

4.2.1.2.4 Depth cues consistency and depth perception

Thresholds for disparity, occlusions, and cross talk are suggested to improve depth perception. Monocular occlusions (especially in cluttered environments) and monocular crosstalk significantly affect depth perception when values exceed 1%, whilst the threshold for binocular disparity values is 2-4% [148]. Other reasons leading to inaccurate depth estimation are focus cues and flash lighting. These could induce Cardboard effect on the viewer [149].

Compared to disparity and motion parallax, other depth cues are less influent on depth perception, but still need to be considered consistently to reduce erroneous depth estimations.

To improve the depth perception and the naturalness of the content, De Silva et al. discovered that a gradual change of object size can be used as an additional cue [150].

Hendrix et al. proved that the use of shadows and droplines in CGI (from the floor to the base of each object) can increase precision in 3D relative distance evaluation (but not altitude evaluation). Moreover, the addition of texture gradient on the horizontal ground of a scene seems to have the minor influence over depth perception compared to shadows and droplines [151].

Furthermore, depth sensation can be enhanced using a method proposed by Jung and Ko [152] [23] that relies on the Just Noticeable Depth Difference (JNDD) [150] [23].

Scene Depth	
Comfort improvement	**Depth perception improvement**
Preserve scene depth consistency across scene changes	Use a depth (or disparity) histogram storyboard
Choose disparity intervals according to screen size and viewing distance	Objects close to the viewer within the comfort zone are perceived more accurately
Don't exceed relative disparity between foreground and background	Large depth interval's integrity is prevented by vergence-accommodation conflict
Use large depth intervals	Depth sensation can be enhanced using Jung and Ko JNDD-based method
Use blending between scene cuts	Examine depth cues' sensitivity functions to enhance depth perception
Stay within depth intervals and the comfort zone	Depth is perceived most effectively at curved or non-uniform surfaces
Prioritize a restricted frequency of disparity changes	Artificial blur can be used to fill in the parts of visual space where disparity is imprecise

Table 6. *Summary of the most important guidelines for comfort and depth perception in relation to scene depth.*

4.2.1.3 Scene Dynamics

This refers to scene dynamics and how these affect the perception of an image or a sequence of images. This specifically applies to:

(1) Changes due to movements of camera and objects in a scene;
(2) Changes due to video editing adjustments, such as scene cuts, transitions, and visual effects.

Table 7 presents a summary of the most important guidelines for comfort, realism, and depth perception in relation to scene dynamics.

4.2.1.3.1 Movements and video editing: facts and issues

If the rate of these changes exceeds a certain threshold, it will negatively affect comfort, realism, and depth perception. This happens because retina's receptors have a limited perception of high temporal frequencies [153] [101] (within the *Window of Visibility*, the shape of which is discussed by [149] [154] [155] [156] [157]), and are unable to adapt to rapid illumination's changes [117].

Since 3D stereoscopic image sequences provide higher depth information than 2D movies, rapid depth variations may inhibit appropriate depth reconstruction [23] [158] [137], produce erroneous depth estimations [149], and cause visual discomfort [159] [143] [158] [137] [106] (especially in prolonged times [127] [133] [160]). Additionally, depth variations can cause blurred vision [161] [162], which negatively affects realism as well [163].

Furthermore, in 3D sequences two types of motion are usually presented:
(1) *Planar motion* [118], which involves movements that maintain the distance from the camera;
(2) *In-depth motion*, [118] [122] [159] which refers to motion along the viewing direction of the camera.

Experimental studies demonstrates that *in-depth motion* has a stronger influence over comfort than *planar motion* [118] [122] [124] [144] [164] [137] [106] [165], particularly when showing significant disparity offsets. This happens especially in the case of crossed eyes vision [144] and short viewing distance [163]. Furthermore, rapid depth variations can result more detrimental to comfort than a large *depth bracket* (distance between the object closest to the camera and the farthest) [23] [119] [124] [106] [137]. The consequence is uncomfortable viewing, even within the comfort zone [118] [119] [106] [124]. Additionally, *in-depth* movements that follow a *step pulse* function cause discomfort even within the depth of focus zone (in which normally visual fatigue is reduced) [119] [124].

Li et al. proved that the frequency of the oscillation of in-depth motion is not a significant factor in 3DTV, and that in some cases static objects far from background can induce a stronger visual discomfort than motion [118]. Nonetheless, supplementary investigations on this topic and on the role of the background's position in the scene are needed [118].

The issues above also apply to interactive 3D panoramic images viewed through immersive VR displays. In that context, VR sickness can also be caused by inconsistency between the degree of freedom (DoF) of VR motion platforms and the visual information presented on the display of the VR system [39].

4.2.1.3.2 Movements and video editing: guidelines and solutions

Possible solutions to the issues presented above are shown below and on **Table 7**:

1. Reduce the rate of changes in acquired images and use longer scene lengths. This may allow the viewer to catch whole scene details with less stress [166].
2. Avoid blurred vision caused by motion and unnatural amount of blur. Reducing unnatural blur [163] and blurred vision due to movements can improve both comfort and realism.
3. Prioritize long sequences and reduce scene cuts. 3D videos require longer scenes, giving more time to the viewer to properly analyse scenes' depth information.
4. Avoid depth discontinuities between scenes. When scene changes occur, a smooth transitions of depth distributions can mitigate the effect of temporal depth discontinuities [23] (which are major factors for visual fatigue [129]).
5. Reduce objects and camera speed. A reduced motion velocity guarantees more comfortable vision for the viewer. This is suggested by Du et al., who sustain that motion (planar and in-depth) together with disparity and spatial frequency of luminance contrast could be taken into account as a metric of comfort for 3D movies [101] [167].
6. Crosstalk in auto-stereoscopic displays may reduce dizziness. Interestingly, in the case of auto-stereoscopic displays, dizziness caused by rapid view switches can be reduced by crosstalk [23] [168]. However this has a cost on realism and depth perception: ghosting and crosstalk degrade 3D quality [23] [169] [148].
7. Introduce artificial blur to simulate depth beyond the depth range. Blur can also be beneficial in some cases to aid depth perception [170] [171] [172] [173] [174] [175] [176]. This happens because an artificial blurred object in a scene is perceived as a change of depth [150]. This allows to simulate depth levels beyond the depth range supported by the 3D display. However, blur degrades 3D shape content [149], and will be comfortable only if it is not introduced in a sharp region and if it is applied to objects behind the screen level [150]. Furthermore, artificial blur from *depth of focus* cues can increase image fusion range to improve comfort [163], and reduce the cardboard-cutout effect without negative impact on visual comfort.
8. Prefer *planar motion* to *in-depth motion* and interpolate *in-depth motion*. *In-depth motion* interpolation will avoid step pulse function motion and visual discomfort.

Scene Dynamics		
Comfort	**Realism**	**Depth Perception**
Reduce rate of changes due to movements and video editing		
Avoid blurred vision and unnatural amount of blur		Introduce artificial blur to simulate depth beyond the depth range
In-depth motion affects comfort more than *planar motion*	-	Prioritize long sequences and reduce scene cuts for better depth perception
Crosstalk in auto-stereoscopic displays may reduce dizziness	-	Ghosting and crosstalk degrade 3D quality
Step pulse function in-depth movements cause discomfort even within comfort zone	-	-
Inconsistency between degrees of freedom (DoF) may cause discomfort	-	-
Use longer scenes	-	-
Avoid depth discontinuities between scenes	-	-
Reduce objects and camera speed	-	-
Prefer *planar motion* to *in-depth* motion	-	-

Table 7. *Summary of the most important guidelines for comfort, realism, and depth perception in relation to scene dynamics.*

4.2.2 3D Camera relevant parameters

This subsection discusses the influence of stereoscopic camera's setup on comfort, realism and depth perception. For clarity, we classify 3D stereoscopic camera's parameters as:

(1) 3D camera internal parameters. These refer to focus, focal plane, depth of field (and depth of focus), field of view (FOV), focal length (see paragraph 4.2.2.1).
(2) 3D camera external parameters. These include camera layout and position, baseline (distance between left and right camera), horizontal disparity (related to the baseline), convergence distance, camera speed and motion, and vertical parallax (see paragraph 4.2.2.2).

These parameters need to be calibrated properly, to guarantee a comfortable and realistic vision. To do it, different calibration methods [149] [177] [178] [125] [111] [139] [179] [111] (i.e. geometric calibration [180] [181] [182] [183] [184], radiometric calibration [180] [184] [182] [185]), capture systems [186] [177] [178] [125] [111] [139] [179], and editing tools [177] [128] [187] exist.

The following paragraphs further analyse relevant parameter's impact on comfort, realism and depth perception.

4.2.2.1 Camera internal parameters

In this paragraph facts, issues and guidelines of some of the internal camera parameters that are most influent on comfort, realism and depth perception are presented. Specifically, it discusses focal length, field of view, depth of field and depth of focus. **Table 9** presents a summary of the most important guidelines for comfort, realism, and depth perception in relation to camera internal parameters.

4.2.2.1.1 Focal length and field of view (FOV): facts and issues

Depth perception is significantly affected by the value of a 3D camera's focal length. IJsselsteijn et al. [188], following Milgram and Krüger's studies [189], evaluated effects of changing the focal length on a camera, and demonstrated that a decrease in focal length will increase FOV and decrease disparity, perceived depth and size of objects. This is in accordance with **Equation 1** and **Equation 2** (see paragraph 2.4.1.1).

In addition, Koppal et al. proved that if camera focal length is higher than the eye focal length, cardboard effect occurs [128], producing depth distortions. Contrariwise, if camera focal length is lower than the eye focal length, pinching effect occurs [128].

Table 8 shows previously mentioned major stereoscopic effects. Furthermore, if a video includes subtitles, a variation of the focal length may cause unnatural depth perception [23].

Effect	*Heuristic or commonly held belief*	*Geometric explanation*
Cardboarding	Keep object "roundness" more than 20 percent.	Camera focal length (f_c) > eye focal length (f_e)
Pinching	Match the eye-camera field of view (FOV).	$f_c < f_e$
Gigantism	A narrow camera baseline causes this effect.	Camera baseline (B_c) < eye baseline (B_e)
Miniaturization	Avoid hyperstereoscopy.	$B_c > B_e$

***Table 8**. Effects of Cardboarding, Pinching, Gigantism, and Miniaturization, based on Koppal et al. [128] studies.*

Regarding the field of view, a variation may affect depth recovery of static or dynamic objects in a scene, causing over-estimation or under-estimation of depth [186] [190]. Studies prove that a narrow field of view decreases the feeling of presence in VR environments: it produces "tunnel vision" and monoscopic views, which alter depth range [39].

4.2.2.1.2 Focal length and field of view (FOV): guidelines and solutions

According to Koppal et al., human eyes' internal parameters values should be adopted for camera internal parameters (including focal length and FOV). This is to avoid distortions (i.e. cardboard effect, pinching effect) [128], improving realism.

For human eye's focal length, two standardized values are considered:

- 17 mm for an object distance of infinity [191] [192] [193] [194];
- 22-24 mm [191] [195].

Studies [196] [197] [198] [199] discovered the following approximated values for human eye's FOV:

- 60 degrees nasally (30 degrees considering the limitation due to brow)
- 60 degrees superiorly (45 degrees considering the limitation due to nose)
- 70 to 75 degrees inferiorly
- 100 to 110 degrees temporally (towards the temple)

Therefore, to increase level of realism in stereoscopic pictures, a stereo camera should use lenses with the suggested specifications, or a display should match these visual values.

4.2.2.1.3 Depth of field and depth of focus: definitions and guidelines

Depth of field (DOF) and depth of focus represent a range of distances for which an object is in focus for a given state of accommodation [163] [200]. The former refers to intervals in front of the eye, whilst the latter is its conjugate [142] within the retina, but both are generally considered equal [163] [201]. Furthermore, DOF's values vary inversely with pupil's diameter [142] (which in a camera corresponds to its aperture).

When stereoscopic pictures show objects outside certain ranges of depth of field and depth of focus, viewers experience visual fatigue [119]. As a result, depth of focus is used to calculate a comfortable viewing zone [124] [202] within which an image appears in sharp focus. For this reason, camera must be well calibrated, to avoid discomfort on the viewer.

According to the study of Chen et al. [145], the following values are suggested in literature for a comfortable vision of stereoscopic images:

- 0.3 diopter as limit of Depth of focus [145] [203], or a smaller value within ±0.2 diopters [124] [145] [119] [119];
- Between ±0.3 [145] [204] (which is the depth of field of a human eye and the area defined as comfort zone by Percival [129] [205]) and ±0.2 diopters as range of Depth of field [142] [137] [107] [206].

Furthermore, Yano et al [142] [137] [119] suggested that discomfort due to accommodation-vergence conflict is reduced within the limits of the human eye's depth of field. This happens because within that range gaze point (convergence) and focus point (accommodation) are coincident [119] [107]. In addition, they proved that stereoscopic HTDV images displayed within the depth of focus range induce the same level of comfort of images that are shown at the depth of the HDTV screen plane [119]. This means that, regardless of the position within the depth of focus range, stereoscopic HTDV images will be equally comfortable. Instead, discomfort occurs outside the depth of focus range.

For these reasons, it is suggested to set stereoscopic camera's parameters so that significant objects are located within the limits of depth of field and depth of focus, reducing discomfort.

Camera internal parameters		
Comfort	**Realism**	**Depth Perception**
Objects outside certain ranges of depth of field and depth of focus may cause visual fatigue	A narrow field of view decreases the feeling of presence in VR environments: it produces "tunnel vision" and monoscopic views, which alter depth range	
Depth of focus is used to calculate a comfortable viewing zone	Human eyes' internal parameters values should be adopted for camera internal parameters to avoid distortions and to improve realism	
Discomfort due to accommodation-vergence conflict is reduced within the limits of the human eye's depth of field	-	Depth perception is significantly affected by the value of a 3D camera's focal length
Regardless the position within the depth of focus range, stereoscopic HTDV images will be equally comfortable	-	If camera focal length is higher than the eye focal length, cardboard distortion occurs
Discomfort occurs outside the depth of focus range	-	If camera focal length is lower than the eye focal length, pinching distortion occurs
Significant objects should be located within the limits of depth of field and depth of focus, reducing discomfort	-	Subtitles + focal length variation = unnatural depth perception
-	-	A field of view variation may affect depth recovery of static or dynamic objects

***Table 9**. Summary of the most important guidelines for comfort, realism, and depth perception in relation to camera internal parameters.*

4.2.2.2 Camera external parameters

In this paragraph facts, issues and guidelines on relevant external camera parameters affecting comfort, realism and depth perception are presented. Camera lenses layout, baseline and horizontal disparity are discussed. **Table 10** presents a summary of the most important guidelines for comfort, realism, and depth perception in relation to camera external parameters.

4.2.2.2.1 Camera layout and position: facts and issues

Stereoscopic cameras have two possible lenses configurations [207]:

- Parallel configuration: the two lenses have their axes parallels reciprocally.
- Toed-in configuration: the two lenses converge their axes at one finite point.

Some stereographers argued that human eyes converge on objects in real life, and that similarly camera lenses should converge [93]. This belief is correct only if images are displayed directly to the retina and not on a flat screen, which would introduce keystone distortions [93] [149] [142] [112]: rectangular objects are distorted into trapezoids.

Besides keystone distortion, toed-in lenses introduce unnatural vertical and horizontal disparities [149], causing discomfort, incorrect depth perception and unrealistic mappings.

An additional problem due to toed-in lenses is the puppet-theater effect: foreground magnification is less than background magnification [149]. This difference, due to the distance from camera to object, was studied by MacAdams [208], who analysed geometric distortions.

Conversely, parallel lenses are not affected by the puppet-theater effect, because they keep foreground magnification and background magnification equal [149]. However, these lenses introduce excessive disparities with close objects [139].

Finally, vertical parallax introduced by a wrong alignment of camera lenses may cause discomfort to the viewer [209] [210], visual fatigue and eye strain [211] [188] [112].

4.2.2.2.2 Camera layout and position: guidelines and solutions

The state-of-the-art suggests replacing toed-in lenses with parallel lenses for the following reasons:

- <u>Parallel layout guarantees a correct representation of depth</u>. This is demonstrated by two experiments conducted by IJsselsteijn et al. [188], which obtained results in line with the ones presented by Yamanoue [212]. Furthermore, Daly et al. analysed perceptual issues in S3D and confirmed that the only way to produce geometrically correct perception is the use of parallel lenses [149]. Jones et al. also claimed that parallel lenses avoid vertical disparity compared to toed-in [111] [112] [213] [115] [214] [215] [216] [217] [218].
- <u>Parallel layout is the simplest and most practical configuration</u> [207].
- <u>Parallel provides a correct horizontal and vertical disparity of the retina regardless the viewer's convergence angle</u> [145].
- <u>With parallel lenses objects can be moved in depth on the screen plane without distortions</u> [142]. This is possible shifting left and right images on the screen.
- <u>Parallel preserves linearity during conversion from real space to stereo images</u> [145] [219]. This prevents the occurrence of puppet-theater effect [220] and vertical disparity [112].

Despite the advantages, parallel configuration is affected by high disparities for objects close to the camera. This may lead to discomfort and hinder image fusion. To reduce the risk of excessive disparities, a minimum distance between camera and objects should be considered depending on camera baseline.

To avoid puppet-theater effect with toed-in lenses, a predictive algorithm was developed [149] [221] using Fourier techniques.

Regarding vertical parallax, Woods et al. showed in their experiment that it should not exceed the limit of 7 mm, otherwise the image fusion of the stereoscopic image would be impossible [112]. Despite this theory, Allison [222] [142] speculated that larger displays can increase the vertical fusion range and reduce discomfort even when vertical parallax occurs.

4.2.2.2.3 Baseline and horizontal disparity: facts and issues

Horizontal disparity (*p*) is one of the strongest depth cues. Its relationship with perceived depth from the screen (d) and interpupillary distance (x_B) is regulated by **Equation 3** [150] [223]:

$$p \approx -\frac{x_{\mathrm{B}}}{v} d$$

Equation 3 *- Horizontal disparity (p) as a function of interpupillary distance (XB), perceived depth from the screen (d), and viewing distance (v).*

A known issue is that an increase of disparity increases depth perception but also causes significant discomfort. Experiments confirm that excessive horizontal disparity values are a major cause of discomfort for viewers [90]. These values depend on camera's baseline, which affects perceived depth, size of objects, and FOV [188] [189].

An increase of camera baseline causes:

- Increased disparity;
- Increased depth perception;
- Constant size of objects and FOV.

In addition, studies show that changing baseline has different effects with parallel and toed-in camera layouts:

- With parallel lenses (having a convergence point to infinite much farther than the object of interest), larger baseline shrinks the scene moving it closer to the viewer [149];
- With toed-in cameras (having a convergence object that is closer than the object of interest), it has the opposite effect [149].

It is not trivial to find a correct value for the baseline: using exact eye spacing does not guarantee the correct depth for the final viewer, unless objects size and depth matches the target display size and comfortable depth range (orthoscopic case) [111], which is a rare case.

Experiments evidence that Gigantism occurs [128] when camera baseline is lower than the Interpupillary distance (IPD) of the viewer (see **Table 8**). Furthermore, Miniaturization occurs [128] when camera is higher than the IPD (see **Table 8**). The common values for the IPD vary between adults (50-75 mm, with a mean value of 63mm) and children (around 40mm, down to five years old) [224].

4.2.2.2.4 Baseline and horizontal disparity: guidelines and solutions

To calculate camera's baseline, several approaches have been proposed:

- Kitrosser's Polaroid Interocular Calculator, using distance of furthest and closest objects to determine the baseline [215];
- Lipton's tables of calibration [178], assuming a maximum disparity value equal to the interocular separation of the viewer;

- Wartell system, which determines the value of the baseline so that the image scaled on display will have the scaled baseline equal to the viewer's interocular distance [111];
- Jones' et al. model [111], which consider projective transformation between camera/scene space and viewer/display space to reduce distortions and calculate baseline properly;
- Oskam et al. controller [139], which derive constraints for camera separation and convergence to guarantee depth within the comfort zone. Their innovative approach involves **dynamic baseline change** over time to increase viewer comfort. Besides, it is only a first attempt, and further research should be performed on adaptive stereoscopy (dependent on the scene content).
- Hyper stereo images approach [114], which adopt very large baseline (i.e. 5km) to understand the spatial shape of objects at very long distances (i.e. mountains distant 30km from the viewer);
- Kim et al. nonlinear mapping [225], which uses nonlinear functions to change the baseline enhancing the depth impression of the foreground while keeping the maximum disparities in the background bounded;
- The 1/30th rule of thumb of 3D [145] [125], which is the most common approach for stereographers. This method stipulates that the baseline should assume a value equal to 1/30 the distance between the camera and the first foreground object. However, it is only an empirical method for a rough estimation of the correct value.

Finally, Devernay et al. [226] proposed a complete geometrical analysis on the effect of baseline changes and viewport changes over the output image. Their work deeply analysed remapping methods to adapt perceived depth to different screen size and viewing distances.

Camera external parameters		
Comfort	**Realism**	**Depth Perception**
Do not exceed the limit of 7mm for vertical parallax, or image fusion would be impossible, leading to discomfort, unrealistic vision, and distorted depth perception.		
Calculate camera's baseline in advance, depending on the scene content (i.e. small baseline for very close objects, large baseline for hyper stereo images of landscapes). Furthermore, use remapping methods in post-production if baseline was not suitable for a correct vision.		
To avoid discomfort, choose a minimum distance between camera and objects depending on camera baseline	-	Parallel layout guarantees a correct representation of depth
-	-	With parallel lenses objects can be moved in depth on the screen plane without distortions
-	-	Use predictive algorithms to avoid puppet-theater effect with toed-in lenses

***Table 10**. Summary of the most important guidelines for comfort, realism, and depth perception in relation to camera external parameters.*

4.2.3 3D Display parameters

Recently, 3DTVs became very popular for domestic use. Nevertheless, Chen et al. claimed that none of the existing 3D systems is considered ideal yet [227] [202]. Furthermore, it is still difficult to preserve perceived depth of a 3D media when displayed on different screen sizes.

Several models have been proposed to adapt 3D content to different 3D displays (i.e. 3D warping [42], disparity remapping [177] [146] [228] [229], binocular parallax flexible manipulation [177] [225]). However, Azari et al. [230] assert that most of them does not consider factors like different viewing conditions, human perceptual factors [230] [231] [232] [233] [234], motion factors [230] [235], and erroneous measures due to cross-talk [230] [236] [237]: more advanced models should be proposed.

To facilitate the development of new optimized models, this subsection reports guidelines on known performant display specifications. For clarity, two groups of display parameters are defined:

(1) <u>Display's hardware specification parameters</u> (see paragraph 4.2.3.1), which include:
- Screen size and resolution;
- Field of View (FOV), framerate and latency;
- Color gamut, pixels, crosstalk;
- Optimal distance and position from the screen.

(2) <u>Display's hardware setup parameters</u> (see paragraph 4.2.3.2), which include:
- Color, luminance contrast and brightness;
- Parallax;
- Subtitles.

4.2.3.1 Display's hardware specification parameters

This paragraph discusses all those parameters that specifically address the hardware of stereoscopic 3D displays. For the convenience of the reader, a comprehensive summary of all presented guidelines is reported in **Table 11**.

4.2.3.1.1 Screen size and resolution: facts and issues

It is well known that the size of a display has a strong influence over perceived depth of stereoscopic 3D content [122]. However, according to the review by Winkler et al. [23] it also affects comfort. This is because a display size largely determines the optimal viewing distance, which in turn changes the size of the comfort zone for depth perception. Indeed, a larger comfort zone is observable in 3D cinemas, which offer a larger depth range compared to home 3D TVs and mobile phones.

In line with these results, Shibata et al. [238] [227] demonstrated that the small screen of a mobile device causes more discomfort and visual fatigue compared to larger displays. This is also confirmed by Cho et al. [122], who tested viewer's comfort using different in-depth motion values with different screen sizes.

However, when a S3D picture is displayed on a larger screen without content adaptation, it will show larger values of parallax. This causes higher disparity and alters depth perception [149].

Further experiments prove that screen size also affects depth of field of a scene and may introduce blur. This specifically applies to the experiment by Masia et al. [177], in which they compared three different automultiscopic displays: Holografika HoloVizio C80 movie screen, desktop and cell phone. Their results confirm that S3D images on smaller screens appear more blurred due to their limited depth of field.

Besides screen size, also screen aspect ratio has an impact on comfort and depth perception. Azari et al. analysed the state-of-the-art, reporting that screen ratio (width:height) between 5:3 and 6:3 are more pleasant to watch [230] [239]. Furthermore, they indicated that modern HDTV displays are designed wider compared to previous generation TVs [230] [240] because their ratio provide a better sensation of depth in both 3D and 2D images.

In terms of realism, screen resolution has a major effect. In particular, Masia et al. [101] stated that the mismatch between spatial resolution of captured images and display's resolution causes loss of detail, limiting the level of realism. This is because many displays and *HMDs* have been limited by costs and technical issues. Recently, higher resolution displays have been developed up to 8K resolution (HMD developed by *Pimax* and presented at CES 2017, with a 7680x4320 resolution). This promises higher levels of realism for S3D scenes thanks to 8K UHD displays.

4.2.3.1.2 Screen size and resolution: guidelines and solutions

To prevent depth alteration, camera should be configured taking into account the size of the screen that will be used. However, knowing the screen size during the shooting is not always possible, and the choice of the optimal configuration still remains a challenge [228].

To solve the problems related to different screen sizes and resolutions, several approaches have been proposed in the state-of-the-art.

To improve comfort maintaining a compelling 3D perception, Masia et al. [101] reviewed techniques to adapt the scene depth range to the size of the comfort zone provided by the display. These techniques relate to disparity retargeting methods, usually working on disparity maps to compress or expand disparity range.

To achieve higher comfort levels, Yan et al. [229] [228] proposed a linear mapping method to adjust depth range taking into account display size, pixel density and viewer distance. However, this approach causes issues with small objects in the scene. Another example of these techniques was presented by Lang et al. [146] [101].

Furthermore, retargeting methods can also help to reduce depth of field blur artifacts produced by smaller screen sizes. Masia et al. [177] proposed a function with dual optimizations:

1- Minimize difference between original and displayed scene in terms of perceived luminance and contrast;
2- Penalize loss in perceived depth.

Another method to adjust depth on different screen size is horizontal image shifting of left and right image pairs on S3D scenes. This is done to improve comfort reducing depth [149] [170]. However, with theatre's large screen this could produce exotropia (excessive uncrossed disparity) when watching very far objects in the scene and must be taken into consideration [149]. Furthermore, to reduce distortions due to retargeting stereoscopic content, Dekel et al. [241] proposed a method to adapt scenes to different screen size with minimum alterations.

Studies prove that a relation between screen size and parallax exists. In fact, common values for parallax to be shown on display are expressed in percentage of the horizontal screen size and reviewed by Tam et al. [142]:

- 1% for negative / crossed disparities and 2% for positive / uncrossed disparities of the horizontal screen size for cinema applications [142] [125];
- Larger values as high as 3% for smaller displays like TVs [142], considering that 3% of the horizontal TV size might be too small compared to the 3% of the one of a cinema displays.

In addition, Allison [142] [222] claims that larger screen can provide more comfortable vision for S3D videos that contain vertical offset, because the vertical fusion range is larger, and viewers can fuse stereoscopic images tolerating more vertical parallax problems.

About screen resolution, several solutions have been proposed to improve it (i.e. tilting projected images [101] [242] [243] [244] [245] [246], use more pixel per inch increasing pixel density [101]). Furthermore, Boev et al. [247] proposed a method to calculate the optimal resolution of a display for signals with a given apparent depth for multiview displays.

4.2.3.1.3 FOV, framerate and latency: facts and issues

As reported by Pfautz [248], in 1980 Hatada, Sakata and Kusaka proved that the subjective "sensation of reality" produced by a Visual Display System (VDS) is a function of the FOV. Furthermore, Pfautz reports that wider FOVs enhance the sense of presence [248]. However, such wide FOV values can be achieved at the cost of display resolution [89].

Important factors for realistic presence are also real-time responses and minimum latency (which is the time delay between an event occurring and its observation) [39]. One of the reasons is that a low frame rate may introduce a significant blur (defined as *hold-type blur*) in the perceived image [101]. This is caused by the interaction between the *Human Visual System (HVS)* and display [101] [249].

Piantanida et al. reported that a latency of 100ms causes motion sickness [91]. Furthermore, Moshell et al. reported through an experiment that higher latency values (such as 200ms) make it impossible to complete any task [92]. In VR and AR devices, latency is also due to tracking and rendering delays, causing users to feel a less-then-complete immersion in the watched VR scene [39].

4.2.3.1.4 FOV, framerate and latency: guidelines and solutions

FOV of 3D displays is today still under investigation, especially in VR and AR devices. Recently, new companies are working on new projects like *StarVR* [250] to use multiple display (dual 5.5" Quad HD Panels) on HMD to solve the problem of enhancing the FOV without losing resolution. However, this solution increases the price of these devices.

To reduce the hold-type blur effect, higher frame rates should be used. This is because objects in the scene remain in the same location for a reduced amount of time, and the smooth pursuit eye motion (SPEM) that naturally occurs in real life will not average object's position, perceiving it more clearly [101]. Furthermore, higher frame rates (HFR) in both capture and projection increases temporal resolution, improving perceived image resolution and reducing motion artifacts such as strobing and judder. HFR have been used in S3D filmmaking by James Cameron and Peter Jackson [161] [251].

According to experiments by Wilcox et al., 48fps and 60fps are more preferred than cinematic frame rates [161] (which are commonly 24fps [149] because they deliberately introduce blur on the movie to achieve "the film look" effect, causing objectionable flicker in strobed displays [161] [252]). To achieve higher frame rates (e.g. 100 or 200 Hz) without brightness or flickering alteration, interpolation techniques exist [253]. Kuroki et al. reported that for framerates of 250Hz with CRT displays jerkiness disappears completely [161]. In line with these results, Hoffman et al. proved that higher frame rates improve depth perception in S3D media, showing that 24fps causes judder for a variety of objects speeds [161] [254] [255].

To partially solve the problem caused by latency, Liang et al. designed Kalman filters to compensate delays in orientation data, reducing noise in positional data on VR devices [256].

4.2.3.1.5 Color gamut, pixels, crosstalk: facts and issues

Color gamut is the display parameter that controls screen color emission through hardware manipulation. This is one of the parameters that have been investigated by Masia et al. in their survey on computational displays. They reported that the range of colours that existing displays can reproduce is only a subset of the whole range perceivable by our visual system [101]. This limitation increases the gap between colours perceived watching a real place and colours perceived watching its digital reproduction, providing low realism on 3D displays.

About comfort, it is affected by the number of pixels (pixel density) [257], and crosstalk (which is a distortion that occurs when one view is also shown over the other [23] [258] [169]) of the 3D display. Crosstalk is commonly measured as the percentage of the intended signal. When its value is less than 5% it will not affect image perception, whilst for values greater than 25% it is unacceptable [247] [259]. These thresholds can vary depending on the local contrast of the content and on the white-to-black contrast ratio of the screen [247].

According to the stereoscopic mode used, crosstalk may be caused by:

- Color filters of anaglyph glasses not separating properly spectral components [23] [260];
- Tilted polarized glasses;
- Time synchronization not accurate in active shutter glasses;
- Imperfect multiplexing in auto-stereoscopic displays [23] [261].

However, even if crosstalk affects comfort and 3d quality negatively [23] [169] [260], it was also reported by Jain et al. [168] that it can mitigate dizziness in autostereoscopic displays when the viewer moves fast his position in the space, improving 3D depth perception.

An experiment on crosstalk was performed by Seuntiëns et al. [169], to investigate its effect on perceived image distortion, perceived depth and visual strain. Their results demonstrated that perceived image distortions increase with the amount of crosstalk, and that higher levels of crosstalk are more visible at larger camera baseline distances. Furthermore, crosstalk lower than 15% does not alter visual strain and perceived depth compared to baseline changes. These results are coherent with the experiment by Pastoor [169] [262], who found that crosstalk increases with increasing binocular parallax and also increasing contrast. Hanazato et al. [169] [263], in line with Lipton speculations [169] [264], reported that crosstalk up to 2% in natural images is unnoticed. This is because, compared to wireframe images or CGI that have hard edges and high contrast, natural scenes usually have soft edges and an amount of details that can mask crosstalk.

Tsirlin et al. [148] performed two experiments to analyse the effect of crosstalk on binocular disparity and on monocular cues. They found that crosstalk alters depth perception with both cues, but it has a stronger negative effect from monocular occlusions, degrading noticeably the quality of stereoscopic images.

A side effect caused by crosstalk is ghosting, which is a major factor influencing viewer satisfaction in 3D stereoscopic displays [148] [169] [259] [265]. As reported by Tsirlin et al., studies [148] [169] [259] [266] prove that ghosting increases with the increase of disparity. Furthermore, Pala et al. [267] conducted an experiment to understand the effect of crosstalk on observers, asking them to align rods in depth. They discovered that the presence of crosstalk causes an increase in perceived workload. Other studies prove that ghosting may introduce unnatural blur or accelerate the difficulty of accommodation, causing discomfort [124] [268].

4.2.3.1.6 Color gamut, pixels, crosstalk: guidelines and solutions

To reproduce a wider set of colours, ultra-wide color gamut displays using four [101] [269], five [101] [270] [271], and up to six color primaries [101] [272] [273] [274] have been designed, together with multi-primary displays based on projection [101] [275] [276] [277] [278]. An alternative strategy is color gamut mapping and manipulation, reducing or expanding the range according to original and target color gamut [101] [279]. Besides the increase of display color gamut, the color scheme model has been improved: sRGB color space has been extended by the scRGB in 2003 [101] [280] and later by the xvYCC (which supports nearly double the color gamut of sRGB) [101] [281], in order to work well with wider gamut displays.

About pixels, Mazikowski et al. [257] described the minimum number of pixels necessary to mask pixel edges visibility, which guarantees a comfortable view. They conducted their study to realize a VR CAVE system at Gdańsk University of Technology. Considering a distance equals to half the side of the screen, and considered that the normal human visual perception threshold (to distinguish object details) is 5' (angle minutes) [282], **Equation 4** and **Equation 5** apply:

$$tan(\propto) = \frac{\frac{x}{N}}{\frac{x}{2}}$$

Equation 4 *- Angular threshold of the human visual system in relation to pixel density of display.*

$$N = \frac{2}{\tan(\propto)}$$

Equation 5 *- Minimum number of pixels required to prohibit pixel distinction.*

where $\propto$ is the angular threshold of the human visual system (equal to 5'), N is the minimum number of pixels required to prohibit pixel distinction, and $\frac{x}{2}$ is the distance from the screen (where x is equal to the length of the side of the screen).

Besides comfort, display with high pixel density enhance realism thanks to the higher visual resolution they provide. Methods to achieve high pixel density were reported by Masia et al. [101], such as optical superposition and temporal superposition. To reduce costs, optical pixel sharing (OPS) can also be adopted [283] [284]. Another technique used is the sub-pixel rendering, which was adopted by Hara and Shiramatsu [285] to replace the standard RGB pattern with an RGGB pattern extending the apparent pass band of moving images and improving the perceived quality.

In terms of depth perception, the pixel shape on a 3D display has a significant role. Basu's study shows that non-squared horizontally finer pixels can potentially provide more accurate 3D estimation on 3D displays [286] [230]. Furthermore, Azari et al. [230] conducted subjective studies with parallel and convergent 3D configurations to verify this theory. They proved that, given a constant resolution, 3D visual experience can be improved by a finer horizontal resolution relative to the vertical display resolution. This suggests using pixel ratio with larger horizontal values than vertical values on 3D displays.

To reduce crosstalk in 3D display, Wang et al. [287] [23] proposed a method by correcting the luminance value of displayed images.

4.2.3.1.7 Optimal distance and position from the screen: facts and issues

According to Hayes's study [288], humans' depth perception is unequal at different viewing distances: over 6 feet distance depth starts to quickly vanish, disappearing completely after 100 feet distance. This is also shown graphically by Masia et al. (see **Figure 48**) [101], who analyse the effect of distance on monocular and binocular depth cues, introducing the Panum's fusional area (see **Figure 50**).

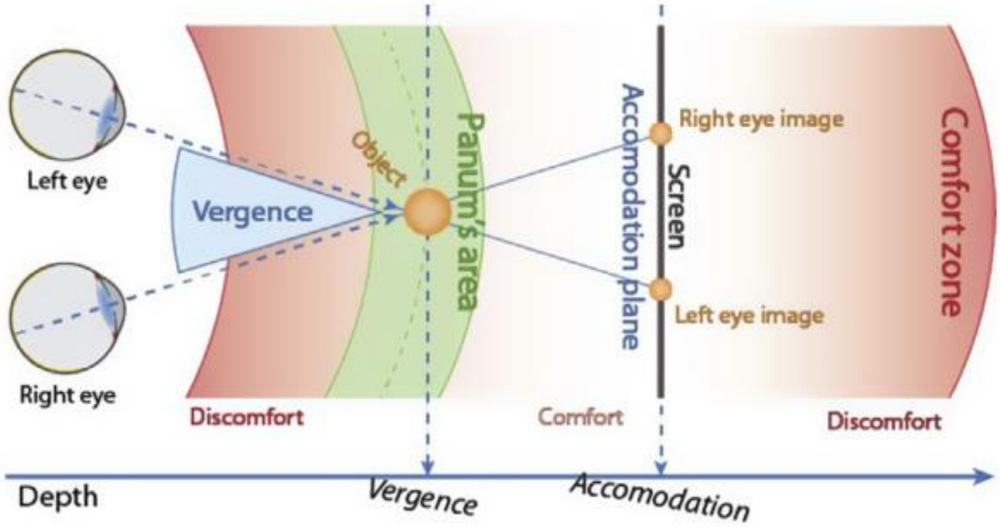


***Figure 50** – Panum's fusional area, accommodation - vergence planes, depth, distance from screen, and comfort zone. Figure from the paper by Masia et al. [101].*

Similarly, Obrist et al. proved that observer's position and the distance between observer and 3D display can affect the perception of a stereoscopic scene [227]. In their experiment, they discovered that people watching a 3D movie on a 3D TV from an armchair felt more immersed than those standing. Furthermore, they highlighted the need to investigate people 3D visual perception by changing several contextual influences (such as viewing angle, light conditions, sound level, and social context).

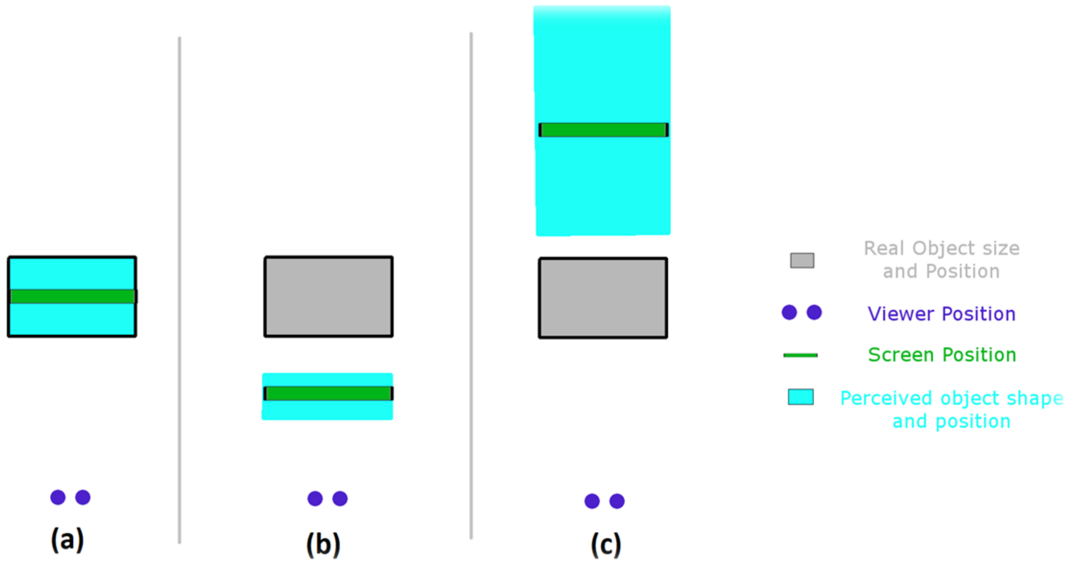


__Figure 51__ - Normal perceived object (a), Compressed perceived depth due to screen too close (b), Expansion of perceived object due to screen too far (c).

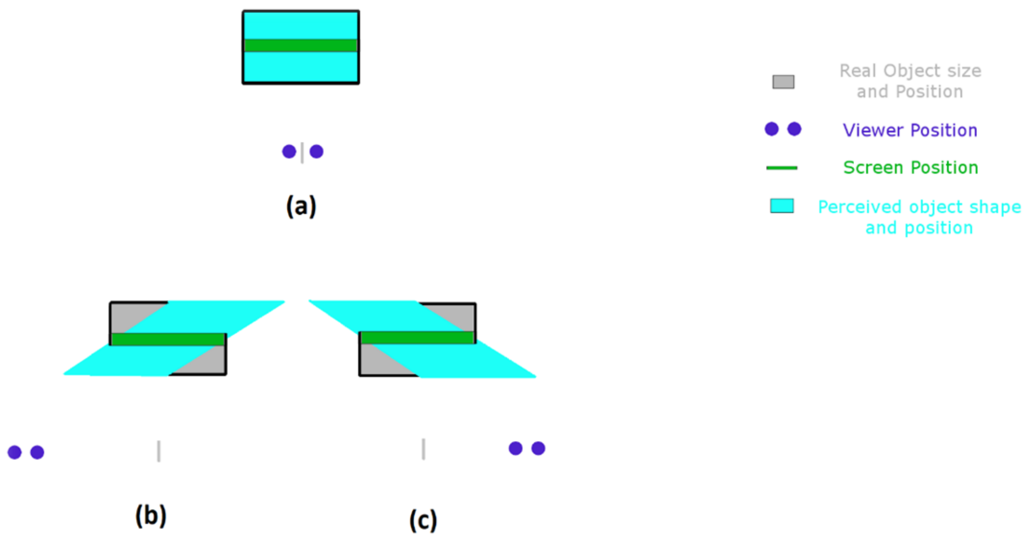


__Figure 52__ - Normal perception of object shape (a), Distortion due to offset left position of viewer from the screen (b), Distortion for offset right position of viewer from the screen (c).

Daly et al. discussed perceptual issues in stereoscopic signal processing [149], reporting that viewing distance, shifted viewer position and rotations have an influence over observer's stereoscopic vision. When a viewer is too close to a 3D display, the scene undergoes a compression in depth (see **Figure 51b**). Vice versa, the scene undergoes an expansion in depth (see **Figure 51c**). When a viewer is not aligned with the centre of the screen, objects in front of the screen and beyond of the screen move in opposite directions shearing the 3D scene (see **Figure 52**). Furthermore, pitch has no influence, whilst yaw and roll produce vertical disparities.

De Silva et al. [150] analysed the effect of viewing distance's changes on the so called "just noticeable difference in depth" (JNDD) (which is the minimum depth variation the human visual system can perceive). Other studies analysed the effect of the viewing distance, proving that the JNDD varies linearly with observer's viewing distance [150] [20], and that binocular disparity is inversely proportional to the square of observer's viewing distance. In addition, perceptual researches demonstrated that the same relation exists for stereoacuity [146]. This again shows how depth perception is affected by the distance between display and observer. Furthermore, Woods et al. [112] reported in their experiment that the range of viewable parallax increases with increased viewing distance. This is in line with Lipton's theory [248] [289], which demonstrates that the distance allowed between near and far points on a stereoscopic scene increases with the viewing distance. Furthermore, a study by Yano et al. discuss the relation between viewing distance and depth of focus [137].

On *Head Mounted Displays* (*HMDs*), Wilson [89] wrote a review discussing eyes and LCDs positions. Reviewing Howarth studies [89] [290], he reported that if the distance between the LCDs does not match the observer interpupillary distance (IPD) discomfort gradually rises. Furthermore, prismatic distortions may occur when the HMD is not perfectly aligned to observer's eyes. Other studies have been conducted by Lambooij et al. [163], which reported that at short viewing distance fast in-depth motion can induce visual discomfort on the observer.

Besides observer's position, viewing angle is another factor affecting depth perception and comfort. When a 3D display is not viewed on its "sweet spots" (which offer the optimal perception of the 3D image), the human visual system is unable to perceive the scene correctly, and distortions occur ("lopsided keystone" distortions) [23] [93]. For example, when an observer assumes an oblique position less 3D immersion will be perceived, but also less motion sickness [23] [291].

With autostereoscopic displays it is more challenging to find the optimal viewing position. This is because only in a few positions the observer is able to suppress properly the view of each eye from the other eye [105], perceiving 3D correctly. With 3D glasses instead, more positions can achieve this easily.

Other distortions (such as shape distortions) caused by incorrect observer positions have been discussed by Didyk et al. [117] .

4.2.3.1.8 Optimal distance and position from the screen: guidelines and solutions

Recently, several investigations have been performed on observer-3D display possible layouts, with the aim of improving 3D perception by setting a viewing optimal distance and position. This is showed by Nee et al. analysis [292], which reported that position tracking and calibration have been central topic in 2008 researches.

Kim et al. discussed the problem of comfort when watching a 3D footage and pointed out that a comfortable view can be achieved only when observers match same position and binocular disparity of the 3D camera [42]. With their investigation they speculated that a real-time stereoscopic rendering system is needed to compensate different viewing conditions and could potentially reduce perceived visual fatigue. Within this purpose they proposed a real-time algorithm model, which adopts eye-pupil detection to measure the viewer pupillary distance and map eyes location in 3D space. Their results prove that the real-time match between viewer eyes position and camera lenses position should be implemented via hardware acceleration, since real-time stereoscopic rendering is impossible with only a software implementation due to the complexity of the SIFT and optical flow calculations.

Winkler et al. considered the effect of viewing distance on depth perception. They reported that viewing distance is largely dependent on the screen size, and proposed depth grading and depth alterations to adapt 3D content to individual viewing conditions [23]. In their paper, they also show graphically (see adaptation in **Figure 49**) the impact of viewing distance on the comfort zone.

Furthermore, several solutions have been developed considering observer's head position. Mazikowski et al. [257] designed a sphere that contains the observer in the centre of a CAVE. This is to maintain his body on a fixed position. Their system allowed the user to move in the virtual environment perceiving 3D correctly, without the need to track head position. This is possible because no variation on viewer-display distance occurs in their CAVE.

To deal with changes in observer's head position, Juang et al. [293] developed a simulator (called SimCrane 3D+ and used for medical training) using *kinesthetic vision*, which continuously adapts the 3D virtual environment to the observer's position in the space. This allows a realistic and more accurate 3D perception of the scene. A similar approach was proposed by Solari et al. [209], who designed an adaptable asymmetric camera frustum using motion tracking, to consider the different positions of observer's eyes.

To adapt a 3D image to different viewing conditions (distance and position, screen size), left eye and right eye images can be shifted adjusting depth range. However, Daly et al. reported that pushing images backward too much may lead to discomfort due to exotropia (excessive uncrossed disparity) [149]. Editing tools also exist to manipulate stereoscopic images for these purposes, such as the viewer-centric editor for stereoscopic cinema by Koppal et al. [128] [294].

Yano et al. [137] conducted an investigation on visual fatigue, reporting that higher viewing distances and higher brightness can reduce visual fatigue [137] [107] for 3D HDTV/HDTV images.

For further readings the review written by Bando et al. [121] is recommended. This discusses visual fatigue caused by stereoscopic images and possible solutions.

Display hardware specification parameters		
Comfort	**Realism**	**Depth Perception**
Adapt scene depth range to size of display comfort zone	Use wider FOV in HMDs to improve sense of presence	Choose screen size in accordance with S3D camera setup
Larger screens can provide more comfort in S3D having vertical parallax	Use higher resolution displays	Use retargeting methods to reduce depth of field blur artifacts
Use horizontal image shifting in S3D to reduce depth on different screens and improve comfort	Use ultra-wide color gamut displays	Adapt scenes to different screen size to reduce distortions and depth alterations
Reduce delays to improve comfort by using delay compensation techniques like Kalman filters	Reduce hold-type blur effect by using higher frame rates and increasing temporal resolution	Higher frame rates improve depth perception in S3D media
Use suggested minimum number of pixels necessary to mask pixel edges visibility	Use display with high pixel density to enhance realism	Non-squared horizontally finer pixels can potentially provide more accurate 3D estimation. Use pixel ratio with larger horizontal values than vertical values
Reduce crosstalk and increase pixel density. Note that crosstalk is more visible at larger camera baseline distances and increases with increased contrast and increased binocular parallax	-	Crosstalk alters depth perception negatively. However, it can sometimes mitigate dizziness in autostereoscopic displays when the viewer moves fast
Comfortable view can be achieved only when observers match same position and binocular disparity of the 3D camera. Higher viewing distances and higher brightness can reduce visual fatigue	-	Use proper viewer's position from the screen to avoid distortions depending on screen size
On HMDs the distance between LCDs should match the viewer's IPD	-	Align the HMD with user's eyes to avoid distortions
Oblique position of viewer reduces motion sickness	-	Oblique position of viewer introduces distortions

***Table 11**. Summary of the most important guidelines for comfort, realism, and depth perception in relation to displays hardware specification parameters.*

4.2.3.2 *Display's hardware setup parameters*

This paragraph discusses parameters that can be changed by the viewer regardless of the hardware of the display used. A summary of the most relevant guidelines is shown on **Table 12**.

4.2.3.2.1 Color, luminance contrast, brightness: facts and issues

Stereoscopic cameras, which use two lenses, can be affected by color mismatch between left and right images. When this occurs, poor and uncomfortable 3D perception is experienced [23].

Besides, image's luminance has an impact on 3D perception: Wang et al. [287] proved that crosstalk can be reduced by correcting luminance of images displayed on a slanted lenticular 3D screen and proposed a method to regulate it.

Didyk et al. [295] conducted experiments on luminance and depth perception, demonstrating that a certain magnitude of luminance contrast is required to make disparity visible. Furthermore, they show that with low contrast and blurry patterns stereopsis is weak, and that the *Cornsweet illusion* can be used to enhance perceived depth [177] [296].

In addition, Masia et al. discuss the contrast sensitivity function [101], showing similarities with the disparity sensitivity function, which though has a peak at a different spatial frequency according to the research of Bradshaw et al. [297] [101]. Furthermore, studies prove that visual perception is limited by physiological factors, which are modelled as contrast sensitivity function [247] [298] [299].

In terms of realism, Masia et al. [101] reported in their survey that natural scenes have luminance values ranging within 12-14 orders of magnitude, whilst simultaneous luminance values vary from 4 to 6 orders of magnitude (for further information they suggest to read the paper by Xiao et al. [300]). However, the *HVS* can perceive up to four orders of magnitude, taking advantage of the *dynamic adaptation,* which allows a shift between magnitude orders to adapt to the light of the scene. Furthermore, the *HVS* is more sensible to near-threshold variations in contrast and less sensitive at high contrast levels [177] [301].

In terms of comfort, Du et al. [167] [101] showed that spatial frequency of luminance contrast (together with disparity, motion in depth, and motion on the screen plane) has a relevant influence over observer's visual perception.

Besides, brightness has an influence over depth perception. This is reported by studies discussing illusions such as the Craik–O'Brien–Cornsweet illusion, suggesting similarities in the mechanisms of brightness and depth perception [101] [117] [295].

An important consideration on brightness is proposed by William Brown [302], who clarifies that the darkness of the room is fundamental to guarantee a good experience to observers in cinemas. He analysed 3D effect achieved by the movie Avatar, supporting his speculations.

4.2.3.2.2 Color, luminance contrast, brightness: guidelines and solutions

To enhance realism, displays that support wider ranges of luminance contrast should be used. In their survey Masia et al. [101] explained that CRT and LCD display can only show about 2 orders of magnitude, whilst HDR displays can enrich the visual experience. Furthermore, HDR displays use local dimming and a dual modulation with different resolution [101] [303]. This shows that in terms of realism HDR displays perform better than normal displays.

Furthermore, Masia et al. discussed the Bloch's law [101] [304], which states that the detectability of a stimulus depends on the product of luminance and exposure time. However, this only applies to short time duration of around 40ms [101] [304]. Besides, analysing observer's preference of higher frame rate movies, Wilcox et al. also reported that an increased luminance benefits dynamic visual acuity for high-speed motion more than for slow/stationary targets [161].

To enhance depth perception, Didyk et al. [295] suggest using high luminance contrast value for significant amount of disparity.

To reduce visual fatigue on the observer, studies suggest high values of brightness in conjunction with a large viewing distance [137] [107], reducing the conflict between convergence eye movement and accommodation.

4.2.3.2.3 Parallax: facts and issues

There is evidence that excessive value of parallax on a 3D display may cause visual discomfort [142] [109] [305], even when motion in the scene is small [158]. This happens because observers are unable to fuse correctly left and right images to perceive depth with very large parallax.

Further studies have been conducted by Nojiri et al. [206] [142], demonstrating that visual comfort is highly correlated to the overall range of distribution of the screen parallax. In support of this speculation, they proved in another study [158] [142] that when the bottom of the screen appears closer to the viewer than the upper screen, observers perceive the 3D image more comfortably. In addition, it was proved that uncrossed disparities might be more comfortable than crossed disparities.

4.2.3.2.4 Parallax: guidelines and solutions

Tam et al. [142] reported that parallax values should remain within the comfort zone depending on the screen size. Analysing the state-of-the-art, expressing parallax as the percentage of the horizontal screen size, they conveyed the following:

- For cinema applications, 1% negative/crossed disparities and 2% for positive/uncrossed disparities are suggested [125] [142];
- For common 3D television, values as high as 3% are suggested [142].

According to Nojiri et al. investigation, to achieve a comfortable vision for observers the scene should always appear with closer objects located on the bottom, whilst farther objects positioned on the top [158].

4.2.3.2.5 Subtitles: facts and issues

Even if adding subtitles to 3D movies seems a simple process, it is surprisingly complex [23] [159]. A few studies have been performed to investigate the impact of subtitles in 3D movies, showing problems in terms of depth mismatches [23] [306]. Especially when object occlusion is not handled properly, subtitles can induce visual discomfort on observers. Furthermore, when multiple subtitles are introduced on the scene at different depths, more visual discomfort may occur.

Winkler et al. [23] reported that depth perception can be altered by subtitles when internal parameters of the 3D camera change while maintaining subtitles fixed.

Depth discontinuities around subtitles and due to window violation can induce discomfort too [159]. Lambooij et al. investigated the impact of video characteristics and subtitles on comfort in 3D TVs and proved that the insertion of subtitles is subjected to possible depth discontinuities, which cause discomfort (i.e. objects in front of the screen that are occluded by subtitles). However, their results proved that other perceptions such as 3D experience are not affected by subtitles.

4.2.3.2.6 Subtitles: guidelines and solutions

Winkler et al. [23] reported that the best way to introduce subtitles on a 3D movie without affecting observer comfort is to maintain their depth similar to the overall depth bracket of the scene. They claim that it is important to check depth range of the scene before subtitles insertion, to prevent possible visual discomfort.

Display hardware setup parameters		
Comfort	**Realism**	**Depth Perception**
High values of brightness in conjunction with a large viewing distance reduces the conflict between convergence eye movement and accommodation, improving comfort	Increased luminance benefits dynamic visual acuity for high-speed motion	High luminance contrast value for significant amount of disparity enhances depth perception
Closer objects should always appear located on the bottom of the screen rather than on the top	Use of HDR images on HDR displays is suggested	-
Keep screen parallax within comfort zone	-	-
Maintain subtitle's depth similar to the overall depth bracket of the scene	-	-

***Table 12**. Summary of the most important guidelines for comfort, realism, and depth perception in relation to displays hardware setup parameters.*

4.3 Modelling Relevant Parameters

4.3.1 Investigated parameters and groupings

In the context of 3D capture and 3D visualization systems a huge number of parameters influencing depth, comfort, realism, and sense of presence was found and discussed in literature. Therefore, in this section I propose parameters groupings and classifications to achieve a clearer understanding of the major interactions existing between them. Within this purpose, I devised several model versions during the literature analysis, together with a final version that was used as a point of reference for the proposed set of investigated parameters.

Model – Version 1

For this model three outputs were considered to measure the visual quality of the remote observation: **depth perception**, **comfort**, **level of realism** (including sense of presence). Furthermore, a **3D System Setup** has been proposed as the middleware between the outputs and the parameters affecting them. Then, input parameters and groupings have been identified. As a result, we distinguish:

- **Camera**: parameters related to 3D camera configuration (e.g. camera baseline, camera intrinsic and extrinsic parameters);
- **Display**: parameters related to 3D display configuration (e.g. screen size, resolution, pixel density, display brightness, display color gamut).

After further revision, two more groups of parameters influencing the three outputs were identified:

- **Viewer**: parameters related to viewer characteristics (e.g. distance of the eyes, age, viewing distance from the screen).
- **Scene**: parameters related to different scene layouts (e.g. cluttered, presenting different depths) and frequency of changes (e.g. moving objects, viewer's movements).

The abovementioned model is shown in **Figure 53**.

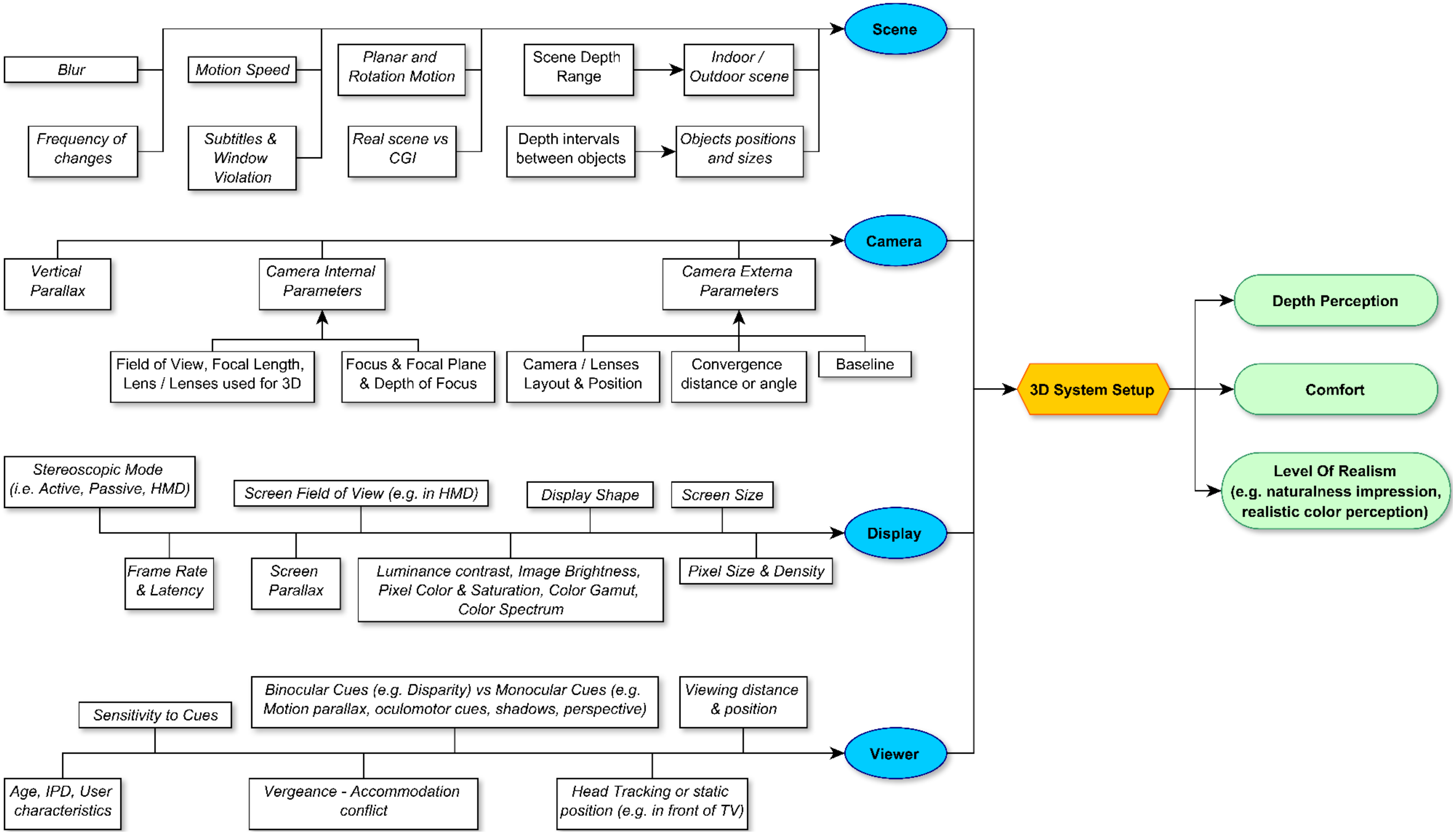


Figure 53 - Model – Version 1.

Model – Version 2

After further revisions, I decided to update the system layout by reducing the number of parameters and by excluding the ones that were considered less relevant (display shape, motion tracking of the HMD, pixel density and pixel size, real scene vs CGI).

Furthermore, the "Scene" group has been merged with the "Camera" group and renamed as "Camera & Acquisition". Furthermore, I included display viewport (defined as graphical field of view, and HMD FOV for VR headsets) as it was reported to have a significant influence over distance estimation performances. Therefore, an updated version of the model is shown in **Figure 54**.

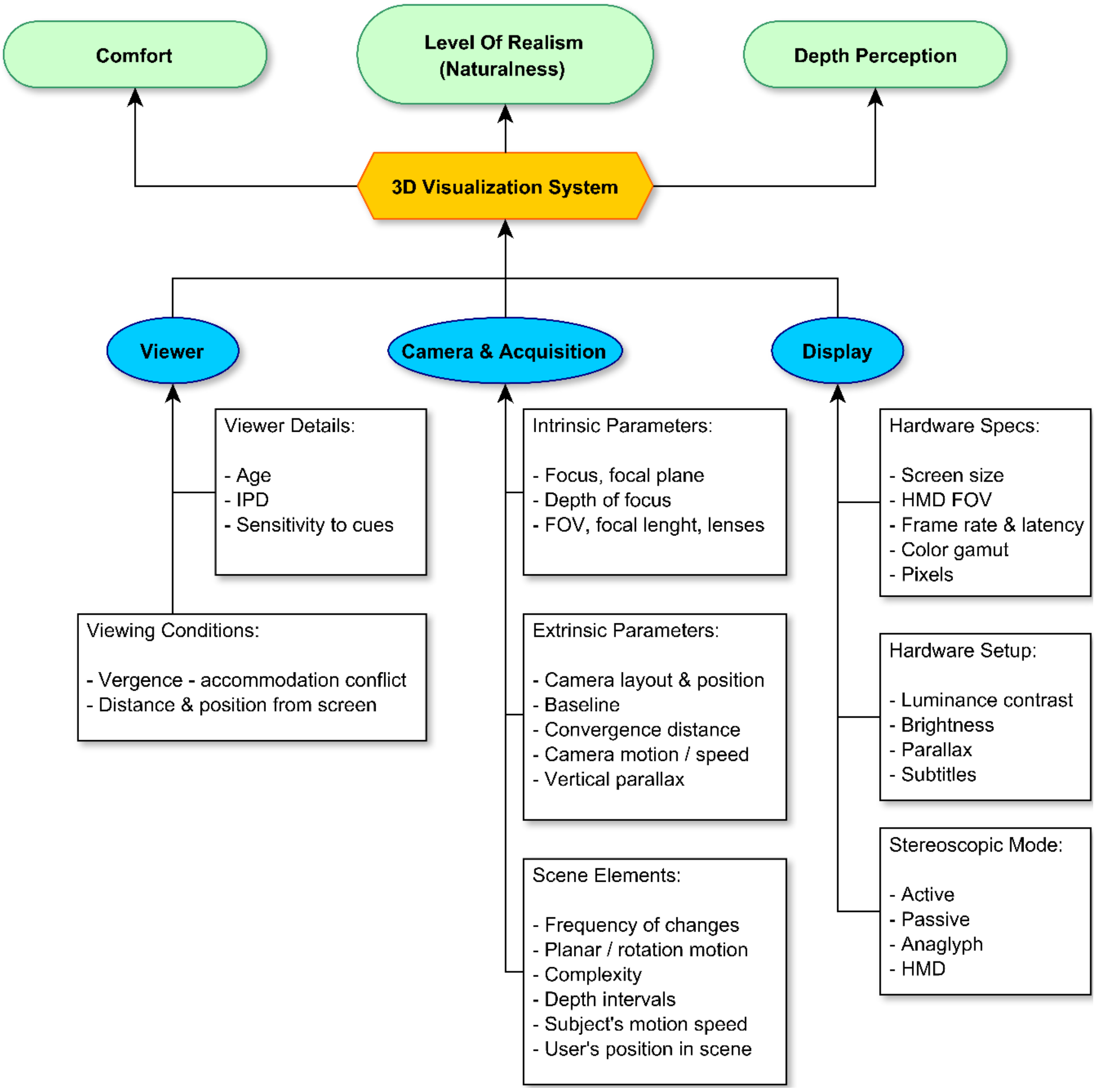


***Figure 54** - Model – Version 2.*

Model – Version 3

Thanks to further readings and revisions analysing both known relations and hypothetic interactions between parameters, I decided to focus on **3D camera**, **3D display**, and **Human eye**. Within this context, I designed an interaction diagram of these parameters (which is shown in

Figure 55) to help me further investigate the topic.

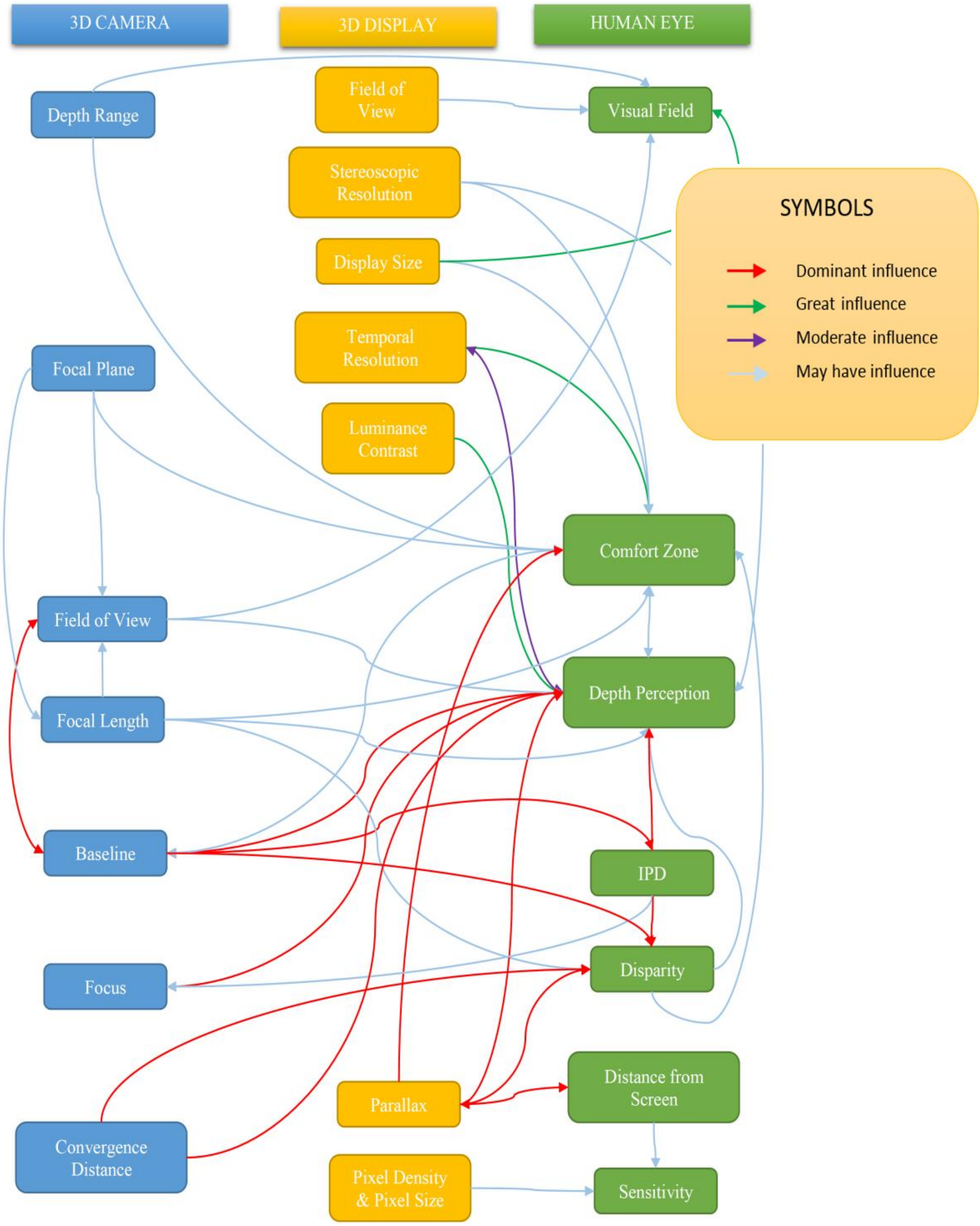


Figure 55 - Model – Version 3. Interactions between parameters are shown by the coloured arrows (which identify dominant, great, moderate, and possible influences).

Model – Final Version

To proceed the investigation, I refined what was shown in the previous models by considering *scene elements* as an independent group of parameters. This was a reasonable choice as the content of the scene is independent from camera and display parameters. Furthermore, "Human eye" was included into display parameters as it mainly refers to optimal distance and position of the viewer from the screen. These considerations led to the development of a final model, described in **Figure 56**.

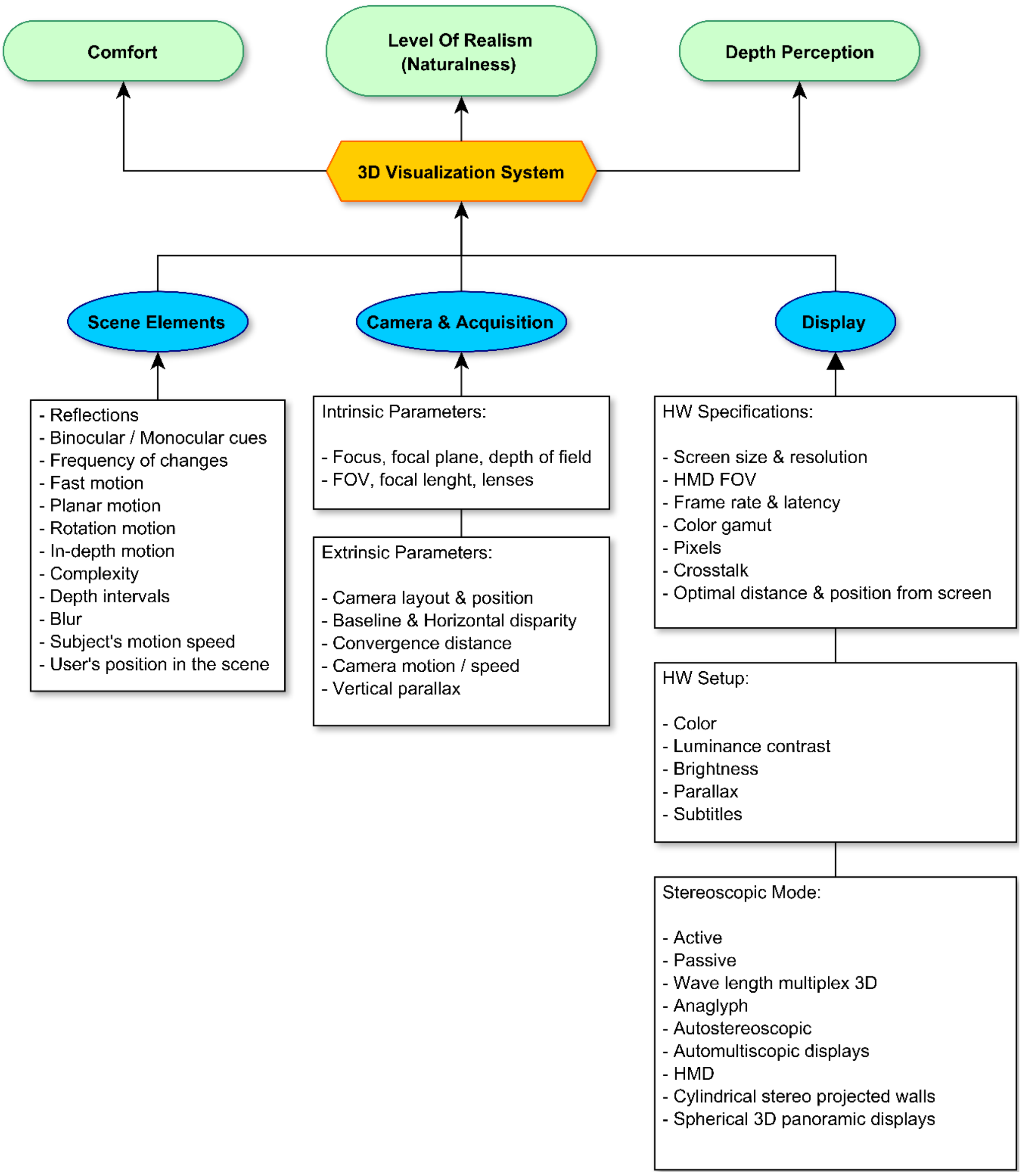


__Figure 56__ - Model – Final Version. This was used as a base to better organize the presented investigated parameters.

4.4 Extended Model: Focus on Presence and Distance Estimation

This section describes an extended version of the previously built model, focused on *presence* and *distance estimation*, which according to further readings resulted as relevant outputs to be examined per se. This new focus also required new literature entries to be considered, in the context of Virtual Reality remote observation.

The diagram shown in **Figure 57** describes the links between all relevant parameters having an impact on presence and distance estimation. The type of impact for each of them is classified as: *enhance*, *reduce*, *influence*, *no Influence*. To give more validity to the whole structure of the proposed diagram, each row in **Table *25*** (in Appendix E) associates the citations in **Figure 57** with the related references that were found during this second analysis of the state-of-the-art.

Furthermore, **Table 13** and **Table 14** summarize the most important parameters that reported an influence over sense of presence and distance estimation respectively.

4.4.1 Analysis on parameters influencing Remote Visual Observation in Virtual Reality

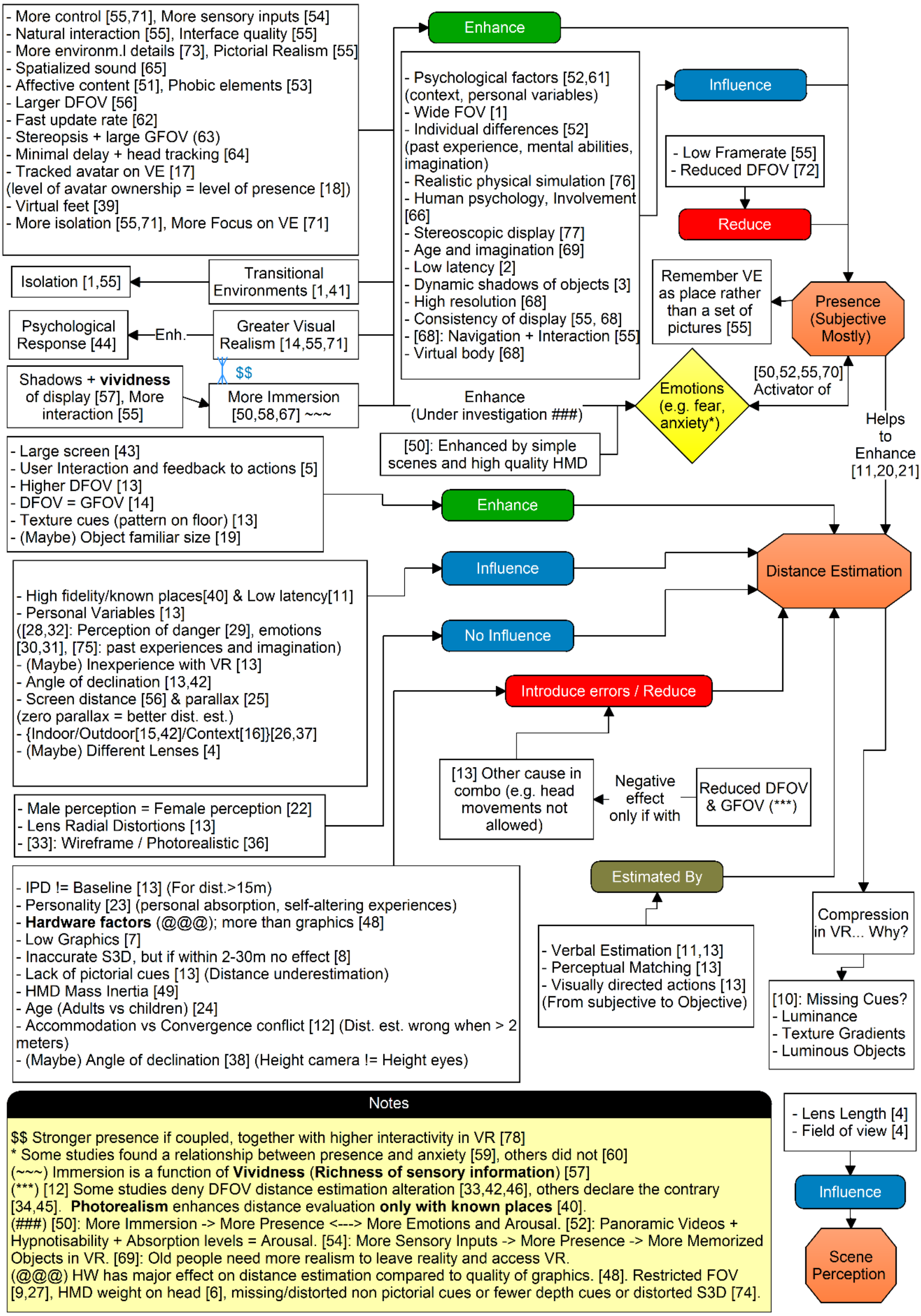


__Figure 57__ - Diagram resulting from an extended analysis of parameters influencing sense of presence and distance perception in virtual reality. Note that the reference numbers on the diagram refer to references in **Table 25**, *in Appendix E.*

Parameters having an influence over sense of presence	**Psychological factors (context, personal variables) [51] [52],**
	Wide FOV [53],
	Individual differences (past experience, mental abilities, imagination) [51],
	Realistic physical simulation [54],
	Human psychology and involvement [55],
	Stereoscopic display [56],
	Age and imagination [57],
	Low latency [58],
	Dynamic shadows of objects [59],
	High resolution [60],
	Consistency of display [60] [61],
	Navigation and interaction [60] [61],
	Virtual body [60];

***Table 13**. Parameters having an influence over sense of presence. Note: reference numbers refer to the references of this thesis in accordance with what reported by* **Table 25**.

Parameters having an influence over distance estimation	**High fidelity / known places [62],**	
	Low latency [63],	
	Personal variables [64],	
	Perception of danger [65] [66] [67] [68] [69] [70]	Threatening objects [71],
		Rival group [72],
		Pointy objects that evoke aversion [73] [74],
		Violation of personal space [75],
		Intrusive object for distances shorter than 100 cm [76],
		Intrusive cones violating observer's personal space [77],
	Emotions [65] [66] [67] [68] [69]	Hills are perceived steeper after 1 hour of run [78] [79],
		Water is perceived larger when thirsty [80],
		Space perception is altered by mental and bodily states [79],
		Desired objects are felt as nearer or closer [81] [82],
	Past experiences and imagination [83] [84] including potentially inexperience with VR [64],	
	Angle of declination [64] [85],	
	Distance from screen [86] and Screen parallax [87] (in their experiment they reported that zero and negative parallax provided less distance overestimation, whilst positive parallax provided increased distance underestimation),	
	Indoor or Outdoor scenes [64] [88] and environmental context [89] [90] [91] [92]	
	Different lenses [93]	

***Table 14**. Parameters having an influence over distance estimation. Note: reference numbers refer to the references of this thesis in accordance with what reported by* **Table 25**.

4.5 Conclusion

In this chapter a thorough systematic review of existing researches on scene elements, 3D cameras and 3D visualization systems was presented. The outcomes of this study were guidelines to design systems for improved acquisition and visualization of three-dimensional scenes.

All investigated features were closely analysed to understand their influence on user's comfort, depth perception, and perceived level of realism. Among them, the complexity of the scene, the depth of the environment and its dynamics were discussed in detail, providing recommendations based on what was found in literature. Furthermore, internal and external camera parameters were analysed and discussed, extending the proposed guidelines. The same procedure was carried for 3D visualization systems, through a focused investigation on hardware specifications (e.g. screen size and resolution, FOV, color gamut, and optimal distance from the screen) and hardware setups (e.g. color, luminance contrast, brightness, parallax, and subtitles).

At the end of the chapter, several models have been presented and used to better organize the content of the systematic review. Within this purpose, diagrams have been proposed to facilitate the reader comprehension of the logic behind the design of the systematic review, and to deliver a clearer understanding of each investigated feature and its relative guidelines.

Among the diagrams that have been presented, the one in **Figure 57** showed a detailed analysis of the parameters influencing sense of presence and distance estimation, in the context of VR remote observation of existing places. In the next chapters, what has been reported by this diagram and by the whole systematic review will be used to discuss and compare results of the performed usability evaluations of this thesis.

CHAPTER 5. USER STUDY - FAMILIARITY AND ENVIRONMENT

This chapter describes the first experimentation related to the proposed Phase Two (Build Knowledge through Assessments), focused on *familiarity* and *environment*. The chapter starts by introducing the proposed evaluation's design (section 5.1), and it continues by presenting implementation (section 5.2), procedure and extended pilot test (section 5.3), user study's results (section 5.4), results analysis (section 5.5), and guidelines (section 5.6).

5.1 Evaluation Design

Following the aim of the investigation (see subsection 3.3.1), I decided to examine the role of **familiarity** and **environment** in Mobile VR observation. The goal is to understand their influence on realism, comfort, presence, depth perception. To do it, a user study with a specific design is proposed.

5.1.1 Familiarity

Users taking part to this user study need to have dissimilar knowledges of the place they will visualize in Mobile VR, and will be clustered into two groups. Within this purpose, we define:

- ***Experts***: test users who have previously visited in real life the places they will observe through Mobile VR. These will be considered as ground truth for this evaluation.
- **Non-Experts**: test users who have never visited in real life the places they will observe through Mobile VR.

5.1.2 Environment

To comply with the requirements of this experiment objectives (see subsection 3.3.1), I chose the **Lachea island** and the **Monello cave** as environments for the Mobile VR observation. These sites have dissimilar characteristics, which match well with the proposed objectives. A summary of the main characteristics is shown in **Table 15**.

Lachea Island	Monello Cave
Outdoor environment	Indoor environment
Large open-ended distance range for objects and landscape, with several objects presented afar from viewer (generating low disparity values)	Small close-range bounded spaces, with many objects presented close to viewer (generating high disparity values)
Dynamic environment representing moving objects	Static environment with no moving objects
Natural sun-light illuminations and reflections due to sea waves	Several shadows on objects from point light-source artificial illumination
Several shadows on objects due to the presence of direct Sun light onto the scene	Several artificial shadows caused by artificial illumination
Wider range of bright and soft colours with wide levels of contrast due to natural elements (e.g. trees, rocks, grass, sea, sky, clouds).	Primarily monochromatic and dark appearance due to the predominance of rocks all over the scene.
Expected more stitching errors on the sky and sea due to smooth textures and moving objects while capturing the scene	Expected minimum stitching errors due to texture variation in most of the scene elements and absence of movement

***Table 15**. Comparison between environmental features of Lachea island and Monello cave.*

The two chosen sites were captured within the Visual 3D Fruition project [47]. Several three-dimensional views of those sites were acquired to be then turned to 3D panoramas observable through a VR headset. The below **Figure 58** and **Figure 59** show the two sites.

Figure 58 - Lachea Island 3D 360 panorama in Sicily. The Mobile VR tour is accessible at http://www.cutgana.unict.it/VirtualTours3D/IsolaLachea/mobile.html

Figure 59 - Monello's Cave 3D 360 Panorama in Sicily. The Mobile VR tour is accessible at http://www.cutgana.unict.it/VirtualTours3D/GrottaMonello/mobile.html

5.1.3 Usability variables

For the aim of the investigation, the following set of independent and dependent variables is considered.

Independent variables:

- Observed Environment. This is an outdoor panorama illuminated by natural sunlight (Lachea island) or an indoor panorama illuminated by artificial light (Monello cave).
- Group of test user. This distinguishes test users belonging to one of the two groups defined above (Expert, Non-Expert).

Dependent variables:

- Realism. It measures the level of perceived realism achieved by visualizing the observed environment on the Mobile VR display. It relies on the definition of *perceptual realism*, which addresses the perception of stereoscopic images as being a truthful representation of reality [36].
- Comfort. It measures the occurrence of *visual fatigue*, *eye strain*, and *discomfort* while watching the 3D panoramas on the Mobile VR display, by using users' subjective evaluations.
- Presence. It is generally defined as users' subjective sensation of *'being there'* [307] in a scene depicted by a medium [308]. Following a more rigorous definition given by Slater and Usoh [309], presence is "the extent to which human participants in a virtual environment allow themselves to be convinced while experiencing the effects of a computer-synthesized virtual environment that they are somewhere other than where they physically are".
- Depth perception. It addresses the perception of space in VR, and the achieved 3D effect performances when viewing the panorama on the VR display.

For each dependent variable (*factor*) the following set of features are defined and investigated:

- **Realism**
  - Level of realism compared to real life
  - Level of realism compared to photos and videos
  - Level of realism in terms of objects deformations / natural elements on the scene
  - Level of realism in terms of emotions
  - Influence of horizontal size of display over Realism
  - Realism of the 3D effect perceived in VR
- **Comfort**
  - Discomfort
  - Eye strain or visual fatigue while watching panorama
  - Max duration of time during which watching the panorama remains comfortable
- **Presence**
  - Perceived sense of "presence"
  - Isolation
- **Depth Perception**
  - 3D Depth impression
  - Distance perception
  - Color contribution to 3D perception
  - Lights and shadows contribution to 3D perception
  - Distorted 3D perception in far or close objects
  - Initial time needed before perceiving the 3D of the panorama

5.1.4 Participants

Test users are selected taking into consideration their background experiences (i.e. experience with videogames, experience with 3D displays, experience with *HMDs* and virtual reality), gender, age, vision issues, use of glasses, interpupillary distance (IPD).

The complete set of users consists of 40 people, of which 12 females and 28 males.

Twenty participants have never visited the real places, which they viewed through the VR headset. Therefore, they belong to the "**group of Non-Experts**". The remaining 20 participants belong to the "**group of Experts**" since they visited the places in real life before viewing them in VR.

All test users paid attention consistently to the panoramas they watched.

5.1.5 Statistical tools

When comparing indoor and outdoor panoramas within the same group of people to look for significant judgemental differences, **within subject** t-test is used. When results between Experts and non-Experts are compared, **in-between subject** t-test analysis is performed.

In detail, the following assumptions are made to perform several *paired and unpaired t-test* [310], in line with literature guidelines [311] [312]:

- Data are collected from a simple random sample of 20 test users for each of the 2 defined groups, which are selected from a representative, randomly selected portion of the total population.
- Homogeneous, or equal, variance exists when the standard deviations of samples are approximately equal.
- The sample size used for each of the VR setups (i.e. Lachea with Experts, Lachea with Non-Experts, Monello with Experts, Monello with Non-Experts) is uniform and equal to 20 participants.

Collected data satisfy all above-mentioned paired t-test requirements. Furthermore, each test user visualizes the two panoramas in VR and evaluates them using the same questionnaire. This implies the collection of repeated measures for each test user.

Average, *standard deviation*, and *variance* are calculated for each setup on each question. Furthermore, the following comparisons are considered:

- Comparison between Experts and Non-Experts viewing Lachea;
- Comparison between Experts and Non-Experts viewing Monello;

Finally, the comparison between Lachea and Monello is performed considering all test users omitting their belonging group.

5.2 Implementation

The 3D VR panoramas of Lachea island and Monello cave have been developed within the Visual 3D Fruition project [47], by using the following setup:

- **Fuji 3D W3 camera**, to capture the real environment;
- **Tripod**, to keep the camera still and avoid blurred pictures while shooting;
- **Web Viewer**, to visualize the 3D panoramas online;
- **Mobile VR headset**, to use a smartphone and visualize the panorama in 3D;
- **iPhone 6**, smartphone to show on display the web viewer with the 3D panoramas.

Fuji 3D W3 camera

Figure 60 *- Fuji 3D W3 camera.*

The **Fuji 3D W3 camera** (see **Figure 60**) is a 3D camera with a baseline comparable to the average IPD of the human eyes. This enables less 3D distortions when capturing the real environment, and more accurate depth. For this reason, it was chosen to capture the two sites of Lachea island (**Figure 58**) and Monello's cave (**Figure 59**). Within this purpose, I started practicing at the University of Hertfordshire by acquiring indoor photographs of the UH halls (**Figure 61**). The resulting captured 3D photos of Lachea and Monello were used to build 3D 360 panoramas for the proposed user studies.

Figure 61 - Hutton Hub 3D 360 panorama, University of Hertfordshire. The Mobile VR tour is accessible at https://alessioregalbuto.com/3DPanorama/UH/

Web Viewer

Figure 62 *- Web Viewer to display the 3D panorama in side by side mode.*

Within the Visual 3D Fruition project [47], a web viewer (see **Figure 62**) was developed and published online [313], enabling users to access the panoramas through their Mobile VR headsets and perceive them in 3D thanks to the side by side stereoscopic mode.

Mobile VR headset

Figure 63 *- Mobile VR cardboard headset, with fixed IPD and fixed distance from screen.*

To guarantee a comfortable visualization of the panorama, a robust Mobile VR headset is chosen to provide adjustable IPD and variable distance from screen. For this reason, the idea of a Mobile VR cardboard headset is discarded due to its fixed IPD and distances, and to its fragile structure (see **Figure 63**). Instead, a **VR Shinecon Mobile VR headset** (see **Figure 64**) is chosen for the user study.

Figure 64 - *Mobile VR plastic headset, with adjustable IPD and distance from screen.*

iPhone 6 smartphone

Figure 65 - *iPhone 6 used for the user study.*

The display of an **iPhone 6** (see **Figure 65**) is used to show test users a panorama of the **Lachea island** (outdoor environment) and a panorama of the **Monello cave** (indoor environment) on a Mobile VR headset. The chosen smartphone display is deemed suitable because of its characteristics and image quality, capable of delivering sufficient representation of the different panorama elements indicated in **Table 15**.

5.2.1 Making 3D Panoramas for Virtual Reality

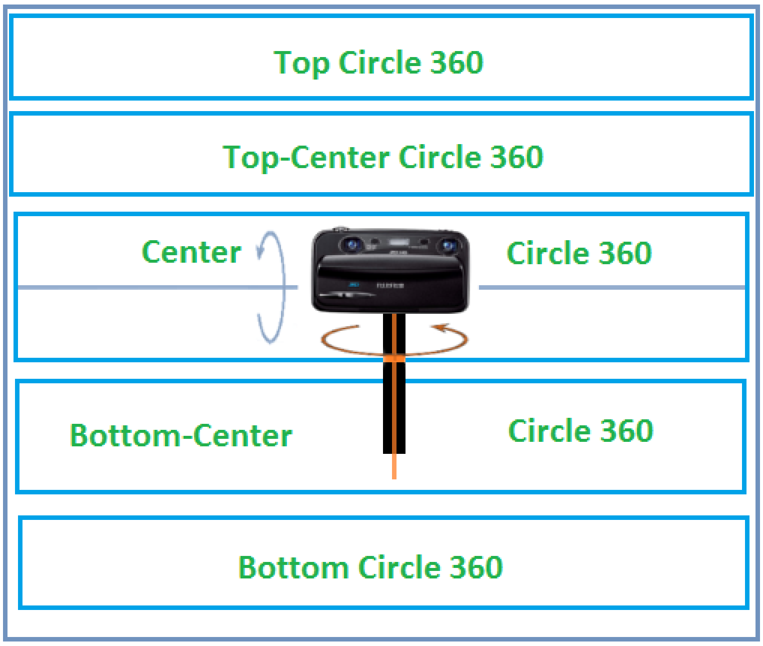


Figure 66 - Diagram showing the 3D camera rotations to capture the spherical panoramas.

The photo acquisition process uses a tripod and the Fuji FinePix REAL 3D W3 stereoscopic camera. In **Figure 66** the 3D camera stands over the tripod and makes two different rotations: the first one refers to the orange axis (Z axis) and the second one refers to the dark cyan axis (X axis). For each rotation, a collection of photos is taken. This process can be also automated using a **pan tilt unit** (see **Figure 67**).

Figure 67 - GigaPan Epic 100 pan tilt unit.

On average, the number of photos per 45 grades of horizontal or vertical rotation ranges from 5 to 8. Considering that the minimum number of photos for a single horizontal rotation is 5, 5*8 = 40 photos are collected per horizontal rotation. Assuming this calculation, for each of the 5 zones of the sphere each panorama takes at least 40*5 = 200 photos. Despite the extremely long acquisition times, I patiently collected 457 photos for the Monello's cave and 499 photos for the Lachea island.

Each acquired 3D photo is stored using the MPO format, which is then converted into left eye and right eye pictures using the software *StereoPhotoMaker* for batch conversion.

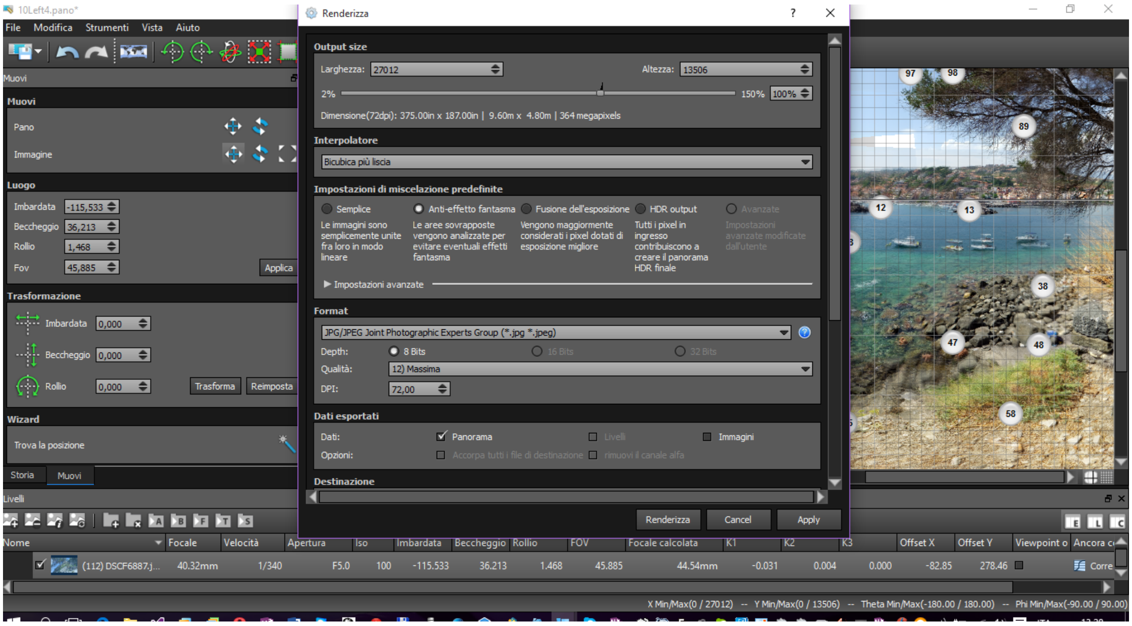


Figure 68 *- Autopano Giga 4, used to make Lachea, Monello, and UH 3D panoramas.*

Left and right eye pictures are then loaded separately in *Autopano Giga 4* (see **Figure 68**), which is used to create a single wide range image from hundreds of photos per panorama. The stitching algorithm used by Autopano Giga 4 is the **SIFT** (**Scale-invariant feature transform**), which detects and describes local features of each picture. The reason that led to the use of *Autopano Giga* instead of *Microsoft ICE* is that the latter does not provide any anti-ghosting mechanism. For further information on the comparison between the two software, check Appendix F.

Once ready, these images are loaded into *Panotour* to create the virtual tours as a web viewer. **Figure 69** shows the folder structure that is generated and used for the virtual reality online tour.

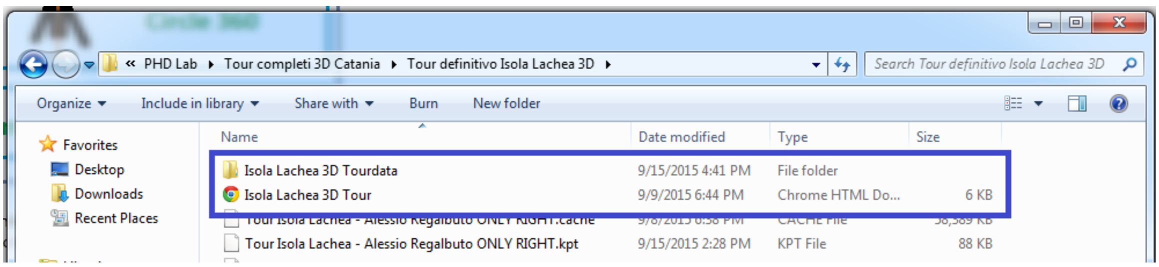
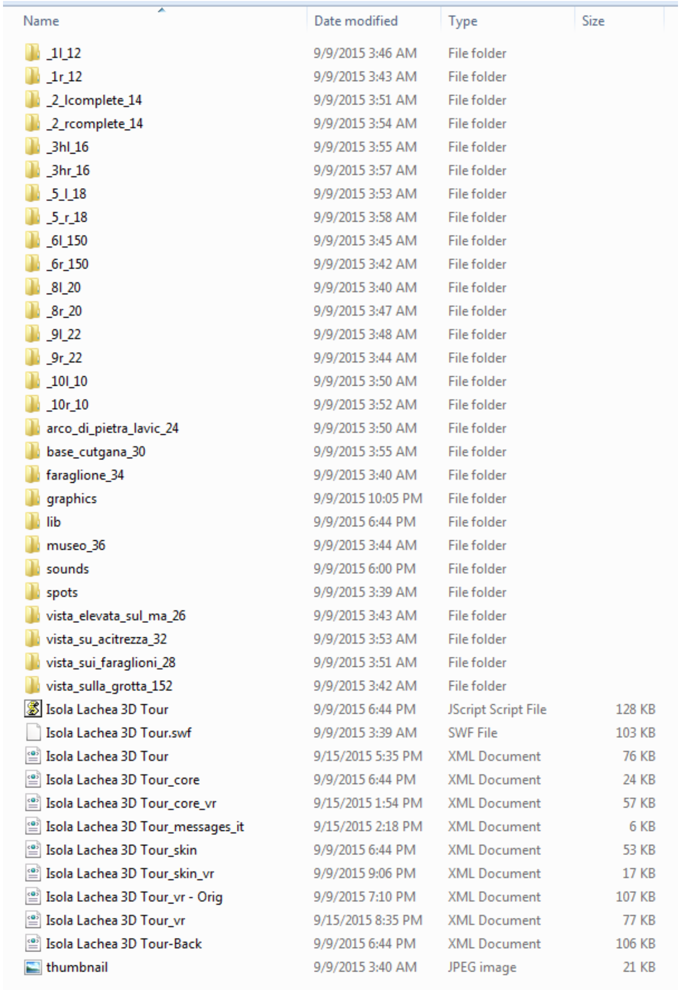


Figure 69 *- Folder structure of the generated Web viewer for the Lachea Island.*



Name	Date modified	Type	Size
_1l_12	9/9/2015 3:46 AM	File folder	
_1r_12	9/9/2015 3:43 AM	File folder	
_2_lcomplete_14	9/9/2015 3:51 AM	File folder	
_2_rcomplete_14	9/9/2015 3:54 AM	File folder	
_3hl_16	9/9/2015 3:55 AM	File folder	
_3hr_16	9/9/2015 3:57 AM	File folder	
_5_l_18	9/9/2015 3:53 AM	File folder	
_5_r_18	9/9/2015 3:58 AM	File folder	
_6l_150	9/9/2015 3:45 AM	File folder	
_6r_150	9/9/2015 3:42 AM	File folder	
_8l_20	9/9/2015 3:40 AM	File folder	
_8r_20	9/9/2015 3:47 AM	File folder	
_9l_22	9/9/2015 3:48 AM	File folder	
_9r_22	9/9/2015 3:44 AM	File folder	
_10l_10	9/9/2015 3:50 AM	File folder	
_10r_10	9/9/2015 3:52 AM	File folder	
arco_di_pietra_lavic_24	9/9/2015 3:50 AM	File folder	
base_cutgana_30	9/9/2015 3:55 AM	File folder	
faraglione_34	9/9/2015 3:40 AM	File folder	
graphics	9/9/2015 10:05 PM	File folder	
lib	9/9/2015 6:44 PM	File folder	
museo_36	9/9/2015 3:44 AM	File folder	
sounds	9/9/2015 6:00 PM	File folder	
spots	9/9/2015 3:39 AM	File folder	
vista_elevata_sul_ma_26	9/9/2015 3:43 AM	File folder	
vista_su_acitrezza_32	9/9/2015 3:53 AM	File folder	
vista_sui_faraglioni_28	9/9/2015 3:51 AM	File folder	
vista_sulla_grotta_152	9/9/2015 3:42 AM	File folder	
Isola Lachea 3D Tour	9/9/2015 6:44 PM	JScript Script File	128 KB
Isola Lachea 3D Tour.swf	9/9/2015 3:39 AM	SWF File	103 KB
Isola Lachea 3D Tour	9/15/2015 5:35 PM	XML Document	76 KB
Isola Lachea 3D Tour_core	9/9/2015 6:44 PM	XML Document	24 KB
Isola Lachea 3D Tour_core_vr	9/15/2015 1:54 PM	XML Document	57 KB
Isola Lachea 3D Tour_messages_it	9/15/2015 2:18 PM	XML Document	6 KB
Isola Lachea 3D Tour_skin	9/9/2015 6:44 PM	XML Document	53 KB
Isola Lachea 3D Tour_skin_vr	9/9/2015 9:06 PM	XML Document	17 KB
Isola Lachea 3D Tour_vr - Orig	9/9/2015 7:10 PM	XML Document	107 KB
Isola Lachea 3D Tour_vr	9/15/2015 8:35 PM	XML Document	77 KB
Isola Lachea 3D Tour-Back	9/9/2015 6:44 PM	XML Document	106 KB
thumbnail	9/9/2015 3:40 AM	JPEG image	21 KB

Figure 70 *- List of files inside the generated folder "Isola Lachea 3D Tourdata".*

The data of the tour is made up of a list of folders containing the tiles of the panorama divided per resolution (because for web optimization there are tiles of the panorama with very low resolution to work fast in mobile devices) and a list of xml files (see **Figure 70**). In details, the "*vr.xml" represents the description of the hotspots of the tour for the virtual reality mode.

However, the generated tour was initially in 2D mode due to a limitation of *Panotour*. To solve the problem, I manually needed to edit the "*vr.xml" file and set the stereoscopic visualization for all the hotspots of the tour. For more information on the edited code check Appendix D. Furthermore, it was necessary to remove the hotspots of the right panoramas from the generated code as the right eye was already mapped within the left hotspots through the stereoscopic mode.

Finally, the website is uploaded online and made accessible to mobile phones, to enable the virtual reality mode in Mobile VR through portable headsets.

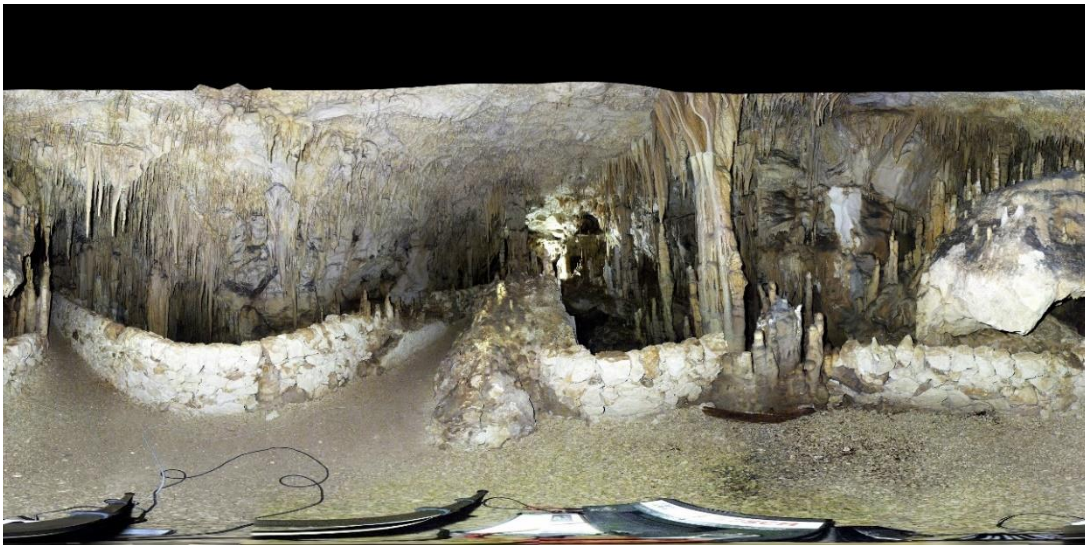


Figure 71 *- Example of incomplete stitching of panorama in AutoPano*

Despite the use of third-party software, sometimes the generated panoramas of AutoPano have been unable to correctly stitch the roof or the sky of the scenes due to the lack of control points or to the amount of distortions of objects in 3D (see **Figure 71**).

To solve the problem, I needed to manually add more control points in AutoPano, with the supplementary help of the software *InPaint*, which was used to rebuild the surface of remaining black holes in the scenes (see **Figure 72**).

Figure 72 *- Example of manually corrected panorama through Control Points manipulation and InPaint.*

Finally, *GIMP* is used to edit the contrast of some panoramas and to add further details to the scenes. This is one of the few editing software able to manage images with more than 30K pixel per side. Considering that a small panoramic picture produced with *AutoPano* starts from a resolution of 27074x13532 pixels (the aspect ratio of the spherical panorama is always 2:1), this software is perfect to match our needs. For further information on all the other issues and troubleshooting during the 3D panorama implementations check Appendix F.

5.3 Procedure and Extended Pilot Test

5.3.1 Pilot Test - Objectives

To improve the design of both the first and the second usability evaluations (see Chapter 6) before their execution, several pilot tests are proposed, with the aim of achieving the following objectives:

1. **Devise the best usability procedure to obtain a more reliable data acquisition**. The following modalities are considered to check which one is more suitable:
    - Modality 1 - Watch panorama and then answer questions;
    - Modality 2 - Watch panorama while answering questions;
    - Modality 3 - Watch panoramas when necessary to answer questions.

2. **Devise the best way to group questions according to investigated parameters**.

3. **Devise the best way to improve test-user's comprehension and panoramic viewing optimization.** Collect answers from test users and learn from results. This is useful to:
    - Evaluate how clearly questions are presented to test users and if they are useful to comprehensively answer what the research is investigating on;
    - Identify potentially ineffective questions and decide if additional questions are needed;
    - Check if the configuration of the panoramas is good enough to conduct trials appropriately.

5.3.2 Pilot Test - Strategy

To achieve the proposed objectives several groups of trials are performed, each of them with different questionnaire layouts and slightly different sets of questions.

Group 1 testing modality 1

Test users are asked to try each VR panorama. For each panorama, they use the HMD to view the scene, then they remove it and answer to a long list of questions related to what they have seen. Answers are written on paper, using a multiple-choice question form.

Group 2 testing modality 2

Questions are asked orally by a test monitor and reported on paper by the test monitor while the test user wears and uses the HMD. This is believed to introduce several advantages:

- More control over the answers given by test users;
- The chance to explain better a question if a test user does not understand it, orally;
- Record answers while the user watches the panorama, so that it is possible to focus more on the details requested by the test monitor. Therefore, questions relay on actual test user's vision rather than on test user's visual memory.

Group 3 testing modality 3

For this modality the list of questions is reduced. Furthermore, it is better structured and divided into 3 steps:

1. **Registration**: the test monitor explains to the test user the purpose of the evaluation;
2. **Training**: the test user wears the HMD, having the possibility to calibrate the IPD and distance from the eyes. This allows the test user to feel comfortable with the device, and to have a first experience with a virtual reality panorama, without stress.
3. **Evaluation**: once that the test user is comfortable with the VR device, a question paper form is provided, containing multiple choice questions. Test user is asked to read one question, wear the two HMDs to think about the answer, and then answer on paper. Then, this is repeated for each question.

5.3.3 Pilot Test - Procedure & Implementation

The total number of test users involved for each representative group of trials ranges between 2 and 5, which is in line with previous design adopted by usability experts [79] [80] [81] [82].

To meet this investigation's requirements, including the ones of the second usability evaluation discussed by Chapter 6, the following implementation choices are taken:

- **Familiarity**: people having different familiarity with the visualized environment are required. To meet this requirement students of University of Hertfordshire have been selected as test users. This is because they have high familiarity with the panoramas of University of Hertfordshire and low familiarity with the 3D panoramas of Lachea's island.
- **Environment**: an indoor environment and an outdoor environment are required. For this reason, 3D panoramas of University of Hertfordshire 3D are used to represent an indoor environment, whilst 3D panoramas of Lachea's island represents an outdoor environment.
- **Display**: two displays having different hardware specifications are required. For this reason, one iPhone 6 and one LG G3 are chosen and used inside a Mobile VR headset to visualize the panoramas.

5.3.4 Pilot Test - Outcome

Outcome for Modality 1: Watch panorama and then answer questions

This modality appears to be suitable for a possible usability evaluation design, but it presents several disadvantages on results:

- The long list of questions might have stressed too much test users, who might have answered to some of them with low attention;
- The fact that these answers were given without control by a test monitor could have introduced some errors for people answering questions even if not understanding completely their meaning;
- Since test users watched the panorama before answering with no chance to watch it again, answers referred to visual memory only, which might have led to improper responses.

Outcome for Modality 2: Watch panorama while answering questions

Despite the advantages, some other problems are introduced by this new approach:

- Test users operated with the HMD for a long time, and some of them could have experienced more discomfort for this reason, altering results in the pilot test.
- Some test users might have not paid enough attention to the last questions, due to their need to put the HMD off with the purpose of reducing their discomfort.

Outcome for Modality 3: Watch panoramas when necessary to answer questions

This modality presents several advantages:

- Test user can take more time to answer without stress, and can watch again the panorama to focus on the details related to the questions;
- Test user can experience less discomfort, not having to wear the HMD for the entire length of the test;
- The test monitor can check each answer and be available to clarify any doubt on questions.

The only disadvantage is that the test with this modality lasts more, but this is not a problem since the given answers for the evaluation are more accurate and reliable.

Outcome from collected answers and statistical analysis

First results address the following possible hypothesis:

- Smaller displays seem to provide less discomfort in terms of how long the panorama could be watched without visual fatigue.
- Larger screens, higher resolution, and higher pixel density seem to offer higher level of "*presence*" (defined as the "feeling of being there"), especially with indoor environments.
- Smaller screens appeared to provide higher sense of presence. However, this hypothesis should be rechecked due to the low number of test users in the context of the pilot test.
- Larger screens and higher resolution and higher pixel density appear to facilitate correct distances estimation with the used FOV on the experiment.

- Smaller screens seem to be more affected by distortions of the lenses of the VR headset, because the distances of some objects looking the centre and the border of the screen have significant differences.

Regarding the scenes, the following considerations are devised:

- Indoor scenes seem to provide more realism, having higher and better depth perception due to more objects close to the camera and a background consequently less flat. Furthermore, considering that the tested panorama was a cave, these effects could also be explained by the large amount of depth planes within the scene, and the highly-cluttered environment.
- Lights and shadows seem to have greater contribution in indoor scenes than in outdoor scenes.
- Indoor environments seem to provide higher sense of "*presence*" than outdoor environments. This happens probably because the environment is closed, and all the objects within it are well perceived in 3D by the viewer, who can feel himself more immersed inside the virtual scene.

Furthermore, when comparing LG G3 with iPhone 6 specifications, LG performed better than iPhone in terms of:

- Scene realism;
- Realism provided by better colours;
- Realism provided by lights and shadows displayed from the scene;
- Transmitted emotions;
- Natural appearance of colours and contrast;
- Presence ("feeling of being there");
- Comfort (iPhone can be watched without discomfort for less than 10 minutes, whilst LG between 10 and 30 minutes);
- 3D depth impression.

These results were used to confirm Modality 3 as the most effective design for the user studies. Therefore, Modality 3 is used for all the proposed usability evaluations. **Figure 73**, **Figure 74**, **Figure 75**, and **Figure 76** show some of the trials of the test users during the performed pilot tests.

Figure 73 *- Explaining the purpose of the usability evaluations to the test user.*

Figure 74 *- Letting the test user try Mobile VR for the first time and adjust lenses distance to achieve a comfortable view.*

Figure 75 *- Taking notes of test user's answers while watching the panorama.*

Figure 76 *- Asking test users to evaluate in real life the distance of objects they previously saw on the 3D virtual panoramas. This was useful to establish the best field of view and parameters for the usability evaluations.*

5.3.5 Chosen Procedure

To perform this first usability evaluation, the following steps are followed:

1. Each test user watches a VR panorama using an iPhone 6 inserted into a plastic Mobile VR HMD.
2. After the viewing of the virtual panorama, each test user is requested to answer a list of questions related to realism, comfort, depth perception, and presence. While answering, test user can view multiple times the panorama, to make more precise and accurate estimations.
3. After the viewing of both panoramas, a comparative questionnaire is proposed.
4. At the end of the usability evaluation, each test user is asked to leave a feedback on the experience.

5.3.6 VR Viewing Scheduling

It is decided to predefine a scheduling to regulate which panorama each test user must visualize first. Half of the test users visualizes the Lachea panorama first, whilst the other half observes Monello panorama first. The purpose is to reduce the risk of overestimation and bias for people who tried this technology for the first time, as well as the underestimation due to tiredness during the last part of the evaluation.

5.4 Results

Answers to the questionnaire have been collected using both questionnaires on paper and on a digital online document created via *Google Forms* [314]. Data has then been digitalized and processed as follows. From now on, the following charts will show in orange p-values < 0.05 resulting from the paired t-test comparison of each of the of VR setups. Significant differences will be highlighted in green in the following tables. Questions of this user study used a 7-Likert scale and multiple-choices.

5.4.1 Realism

5.4.1.1 Realism – Experts vs Non-Experts viewing Lachea island

The diagrams below show overall that significant differences were found for Lachea. Realism achieved in VR compared to Photos and Videos was evaluated considerably better by Experts. Non-Experts strongly believed that a larger display size can increase the level of realism.

5.4.1.1.1 Level of realism compared to real life

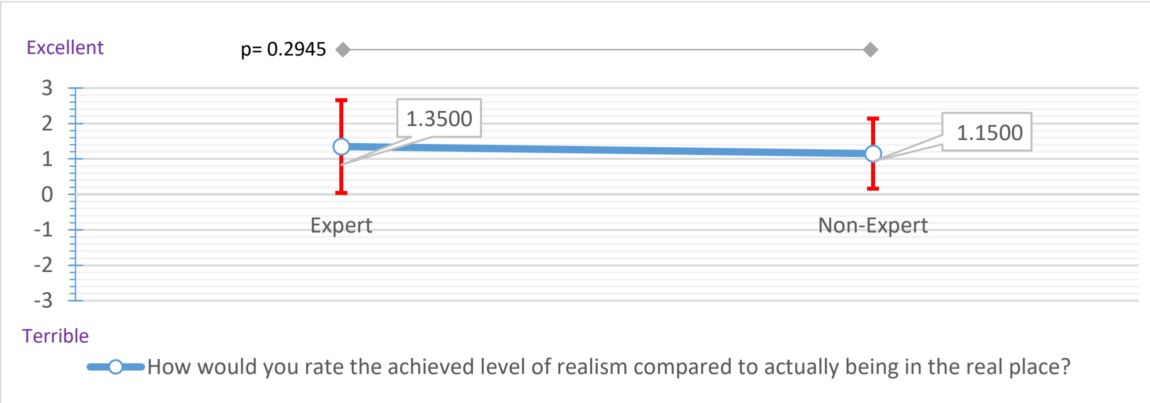


Figure 77 *- Experts and Non-Experts showed no significant difference and evaluated good the achieved VR realism compared to real life.*

5.4.1.1.2 Level of realism compared to photos and videos

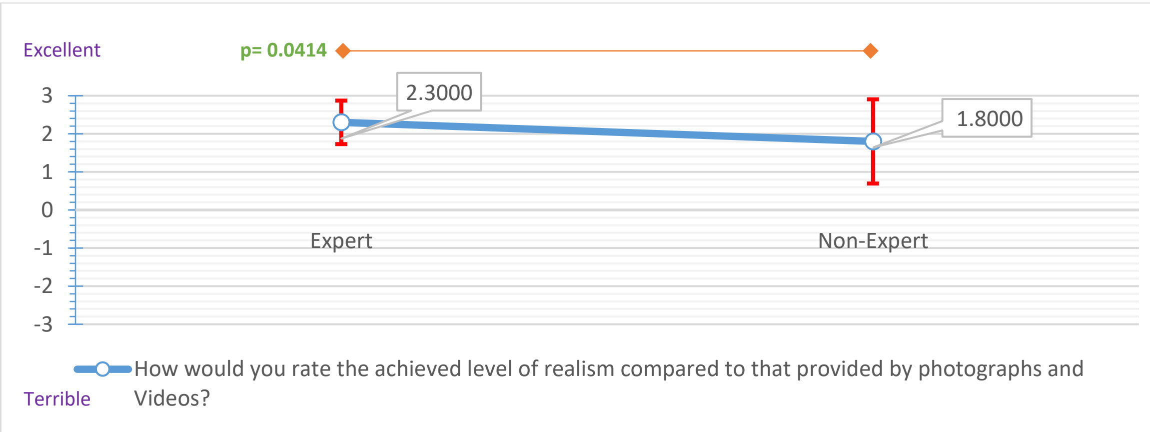


Figure 78 *- Experts evaluated very good the VR realism compared to photos and videos, whilst Non-Experts significantly underestimated this score considering it just good.*

5.4.1.1.3 Level of realism in terms of objects deformations / natural elements on the scene

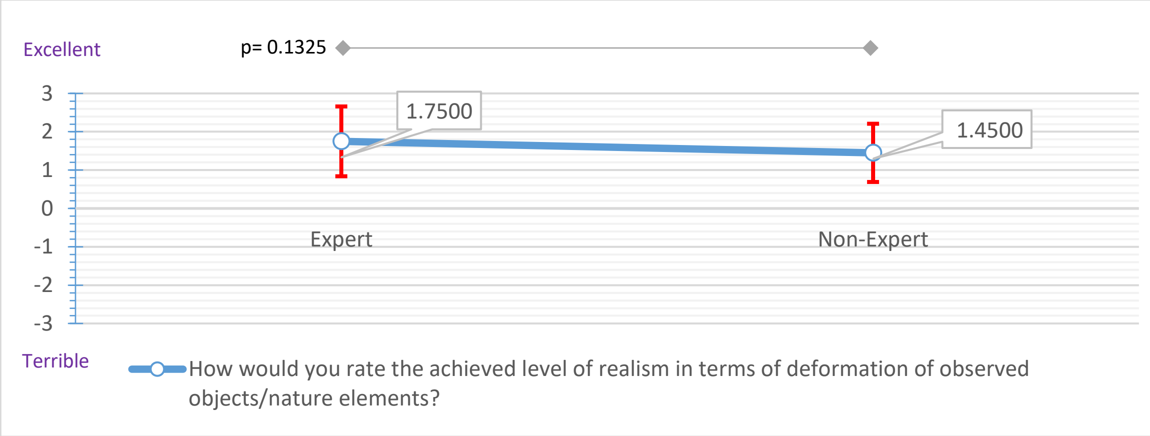


Figure 79 *- No significant difference was found between Experts and Non-Experts for realism in terms of object deformations / nature.*

5.4.1.1.4 Level of realism in terms of emotions

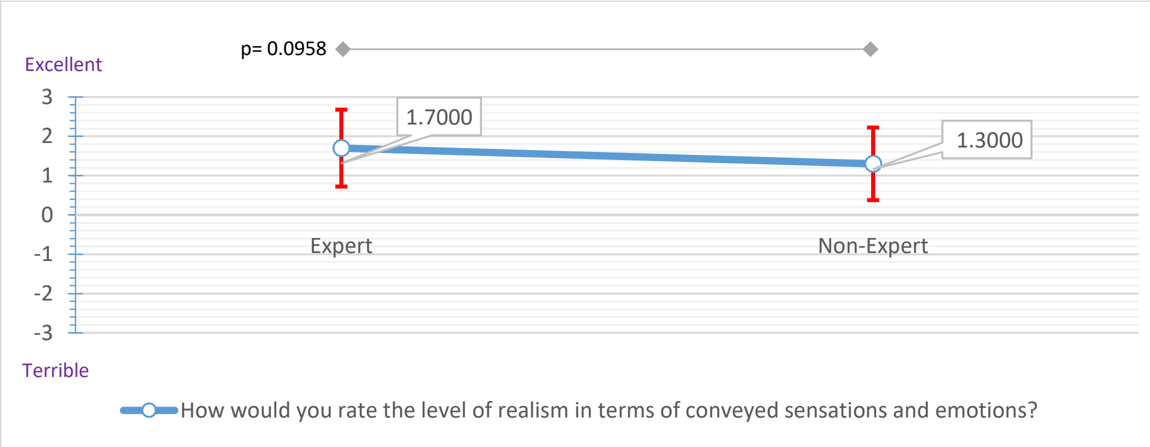


Figure 80 *- No significant difference was found between Experts and Non-Experts for realism in terms of emotions.*

5.4.1.1.5 Influence of horizontal size of display over Realism

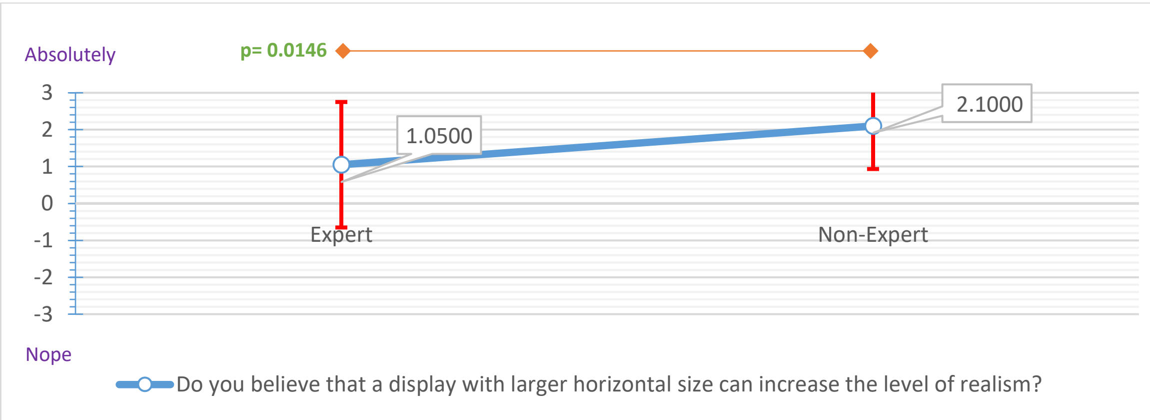


Figure 81 *- Non-Experts answered with significantly higher scores compared to Experts about the influence of horizontal screen size on realism. In detail, Non-Experts considered this influence very high, whilst Experts only good.*

5.4.1.1.6 Realism of the 3D effect perceived in VR

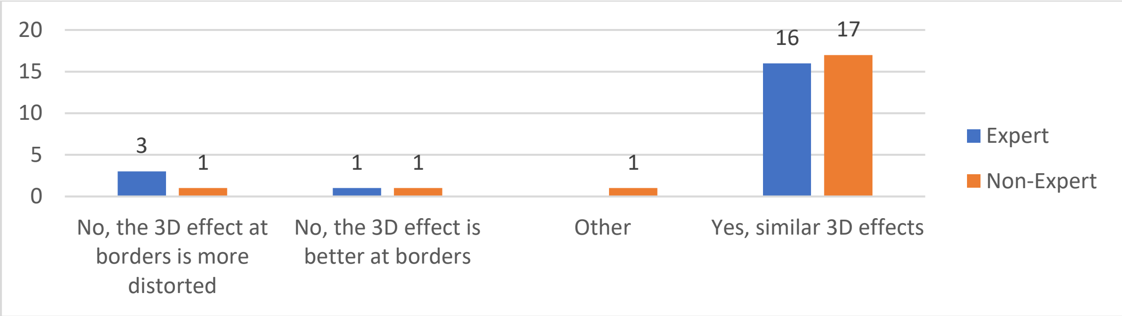


Figure 82 *- Number of test users grouped by their answer to the question <Is the realism of the 3D effect perceived at the screen borders similar to that perceived at the screen centre?>.*

Figure 82 shows that there was no significant difference between Experts and Non-Experts. Both groups evaluated equally the realistic perception of 3D in the centre and in the borders of the VR display.

5.4.1.2 *Realism – Experts vs Non-Experts viewing Monello cave*

Significant differences were found for Monello cave. Non-Experts strongly believed that a larger horizontal display size can increase the achieved level of realism.

5.4.1.2.1 Level of realism compared to real life

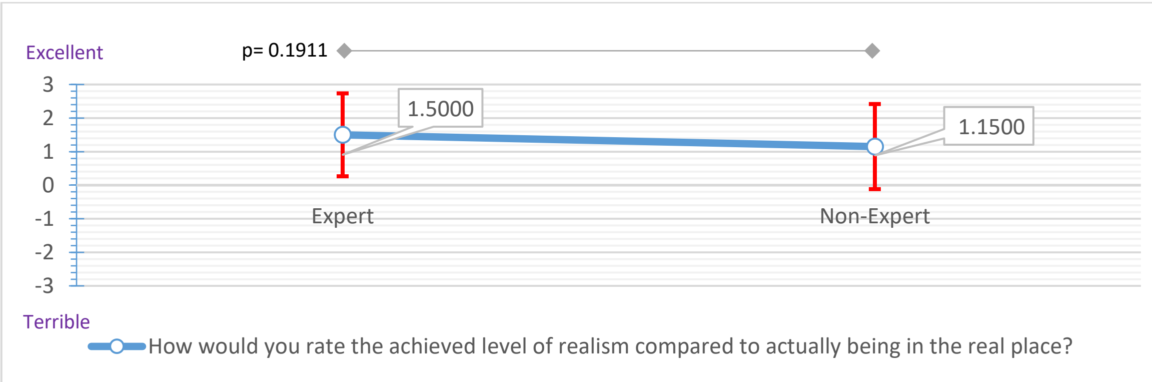


Figure 83 *- No significant difference was found between Experts and Non-Experts. Both groups evaluated the realism good compared to real life.*

5.4.1.2.2 Level of realism compared to photos and videos

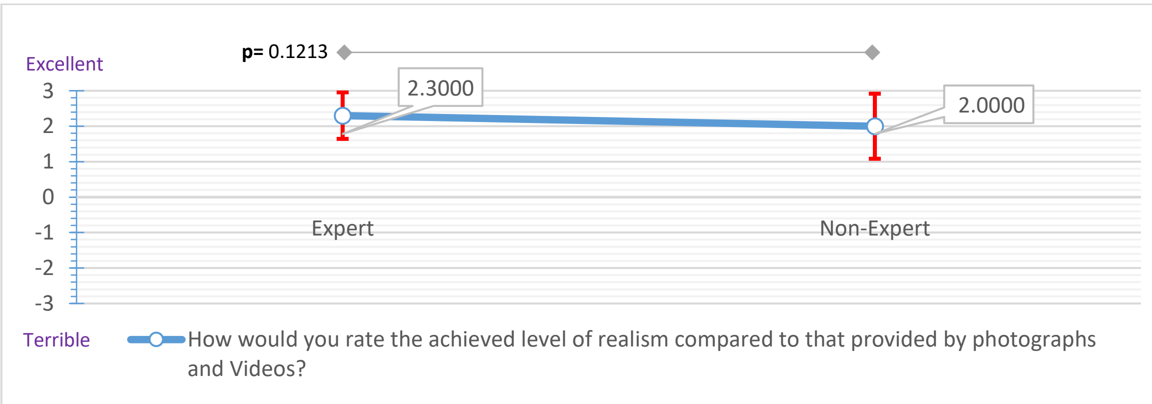


Figure 84 *- No significant difference was found between Experts and Non-Experts. Both groups evaluated realism very good compared to the one of photos and videos.*

5.4.1.2.3 Level of realism in terms of objects deformations / natural elements on the scene

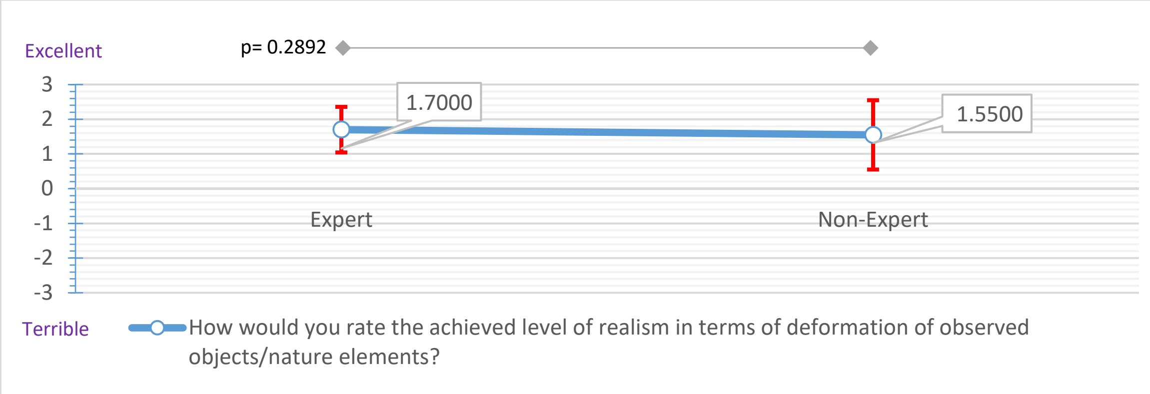


Figure 85 *- No significant difference between Experts and Non-Experts. Both groups evaluated positively the achieved level of realism in terms of deformations / nature.*

5.4.1.2.4 Level of realism in terms of emotions

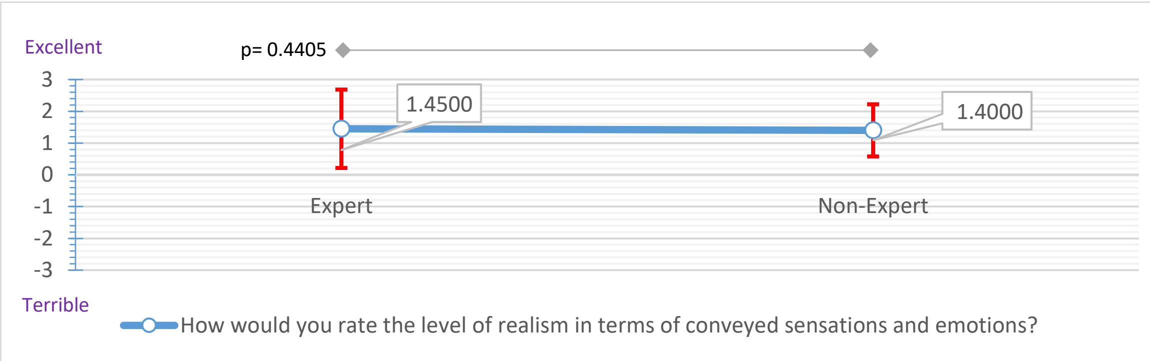


Figure 86 *- No significant difference was found between Experts and Non-Experts, which evaluated realism in terms of emotions good.*

5.4.1.2.5 Influence of horizontal size of display over Realism

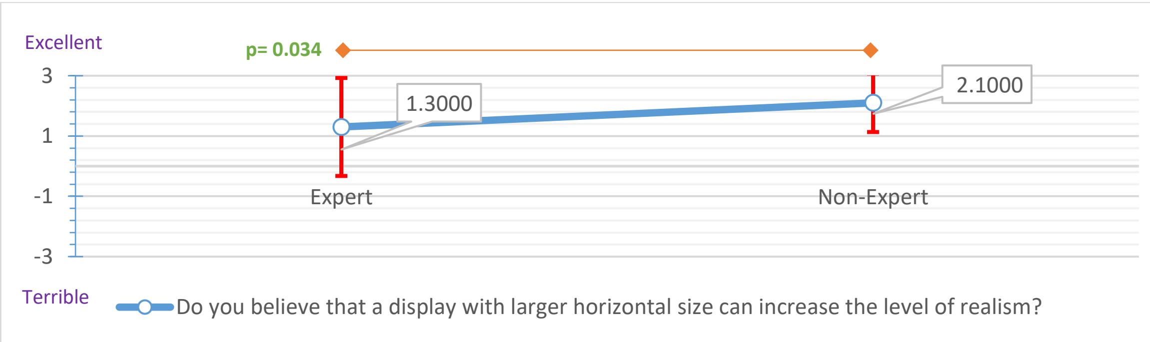


Figure 87 *- Non-Experts gave significantly higher score than Non-Experts about the influence of horizontal screen size on Realism.*

5.4.1.2.6 Realism of the 3D effect perceived in VR

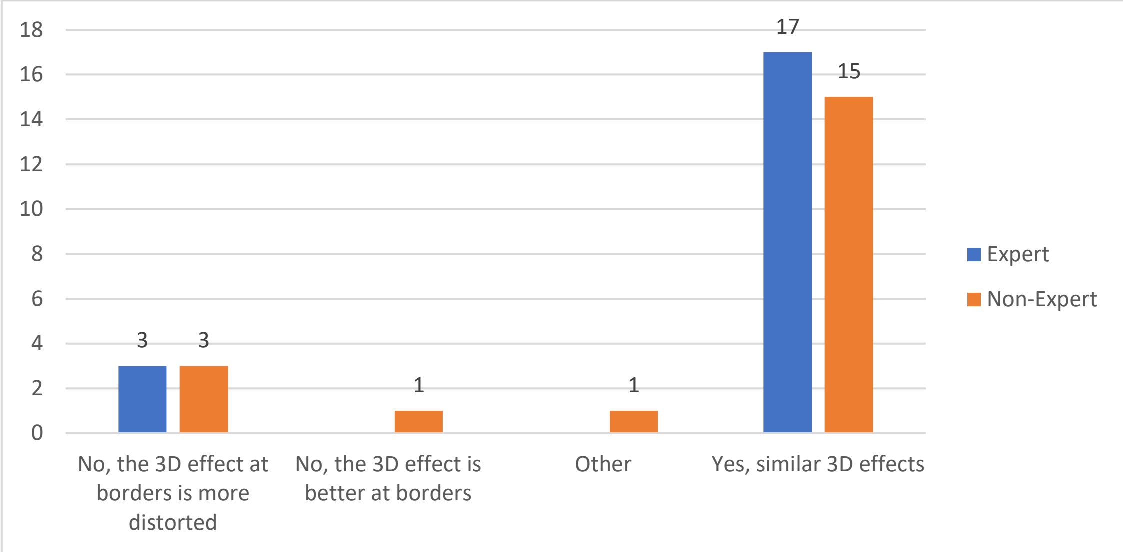


Figure 88 *- Number of test users grouped by their answer to the question <Is the realism of the 3D effect perceived at the screen borders similar to that perceived at the screen centre?>.*

Figure 88 shows no significant difference. Both groups evaluated equally the realism of the 3D effect achieved in the centre and in the borders of the display.

5.4.1.3 Realism – Comparative questionnaire for Lachea vs Monello

From a first analysis that considered a comparative questionnaire on Lachea island and Monello cave, Experts and Non-experts rated the achieved level of realism as shown in **Figure 89**.

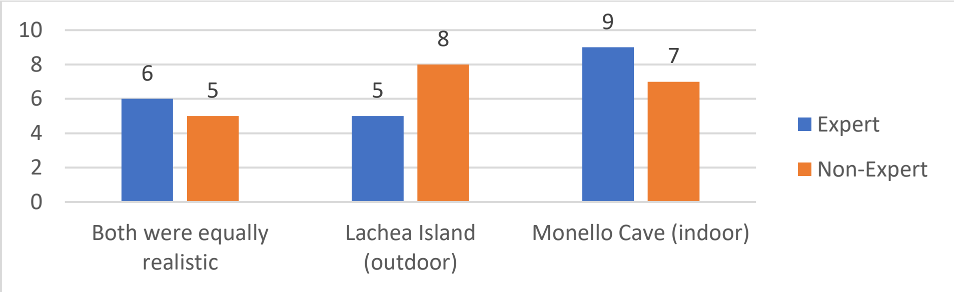


Figure 89 *- Number of test users grouped by their answer to the question <Which one of the two tours looked more realistic to you? (feeling of being there)>.*

In **Figure 90** the overall realism of the 3D perception achieved by both panoramas is shown.

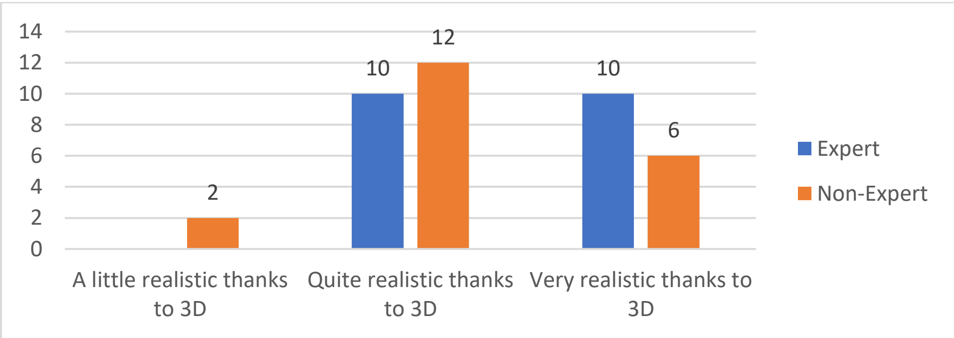


Figure 90 *- Number of test users grouped by their answer to the question <Overall, how much realistic do you rate the 3D panoramas?>.*

Overall, no significant results were found considering each group individually while comparing Lachea and Monello.

5.4.2 Comfort

5.4.2.1 Comfort – Experts vs Non-Experts viewing Lachea island

5.4.2.1.1 Discomfort

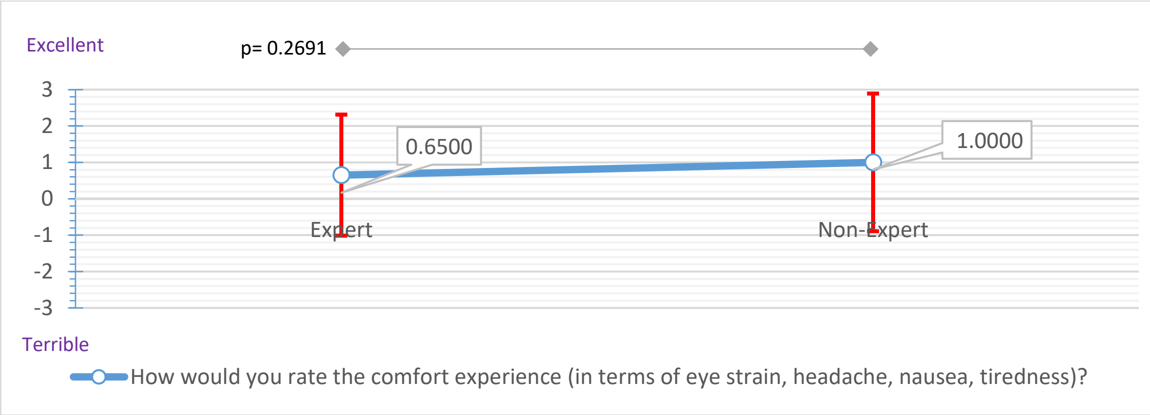


Figure 91 *- No significant difference was found between Experts and Non-Experts, which evaluated the overall comfort positively.*

5.4.2.1.2 Eye strain or visual fatigue while watching panorama

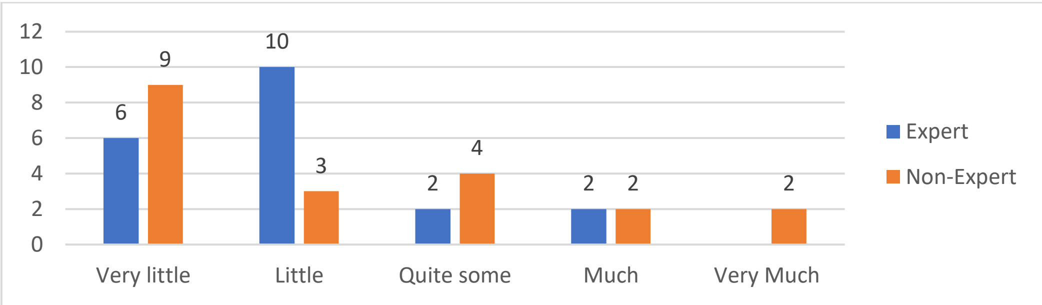


Figure 92 *- Number of test users grouped by their answer to the question <Did you experience discomfort during 3D viewing of the panorama?>.*

Figure 92 shows that Experts experienced more discomfort than Non-Experts with the Lachea panorama. This might be due to more stitching errors of the panorama combined with the richer background experience of Experts.

5.4.2.1.3 Max duration of time during which watching the panorama remains comfortable

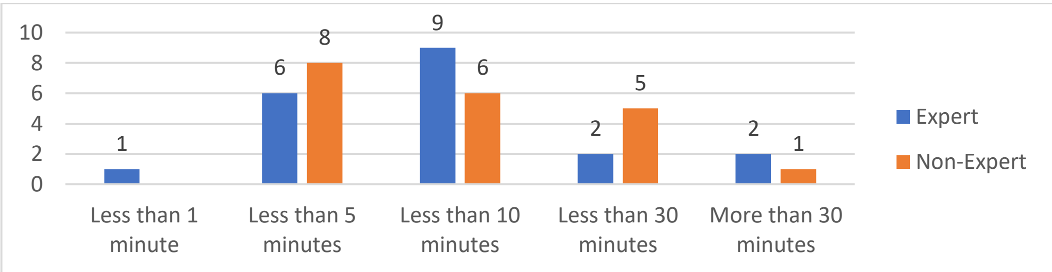


Figure 93 *- Number of test users grouped by their answer to the question <How long do you think you can watch the 3D panorama before getting tired?>.*

Figure 93 shows that Non-Experts can enjoy a comfortable view before getting tired for a slightly longer time than Experts.

5.4.2.2 *Comfort – Experts vs Non-Experts viewing Monello cave*

5.4.2.2.1 Discomfort

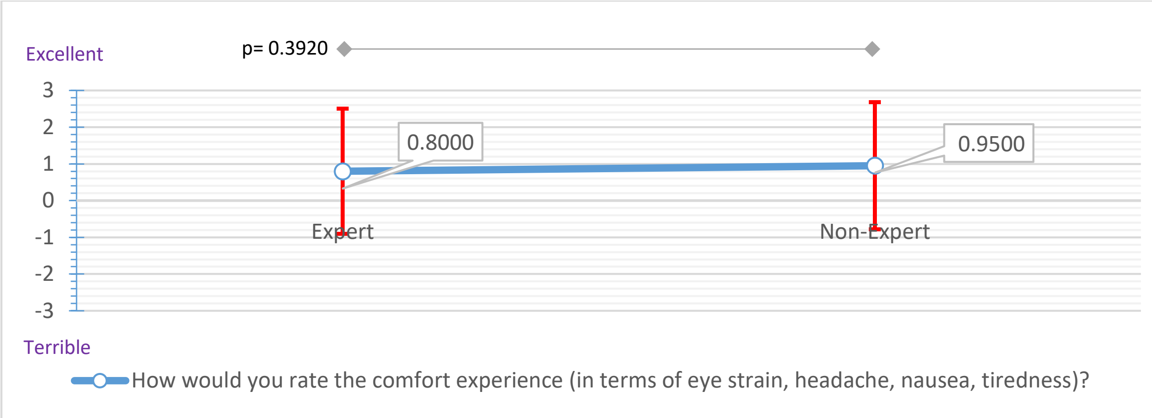


***Figure 94** - No significant difference was found between Experts and Non-Experts, which evaluated the overall comfort positively.*

5.4.2.2.2 Eye strain or visual fatigue while watching panorama

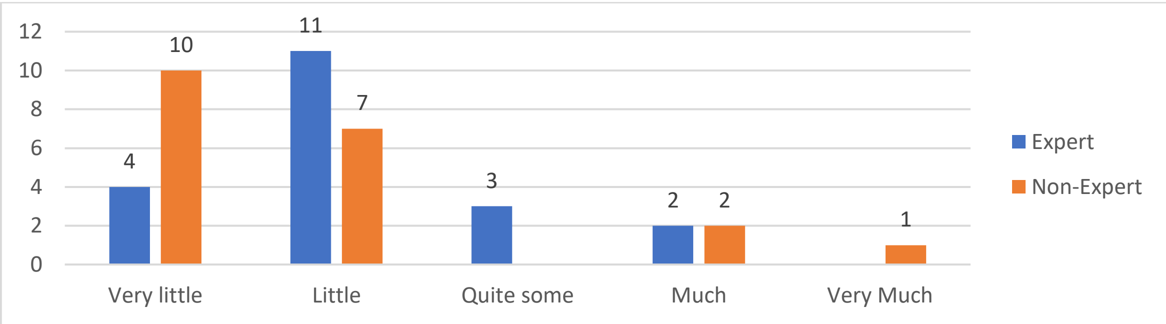


***Figure 95** - Number of test users grouped by their answer to the question <Did you experience discomfort during 3D viewing of the panorama?>.*

Figure 95 shows that Experts experienced slightly more discomfort than Non-Experts with the Monello panorama.

5.4.2.2.3 Max duration of time during which watching the panorama remains comfortable

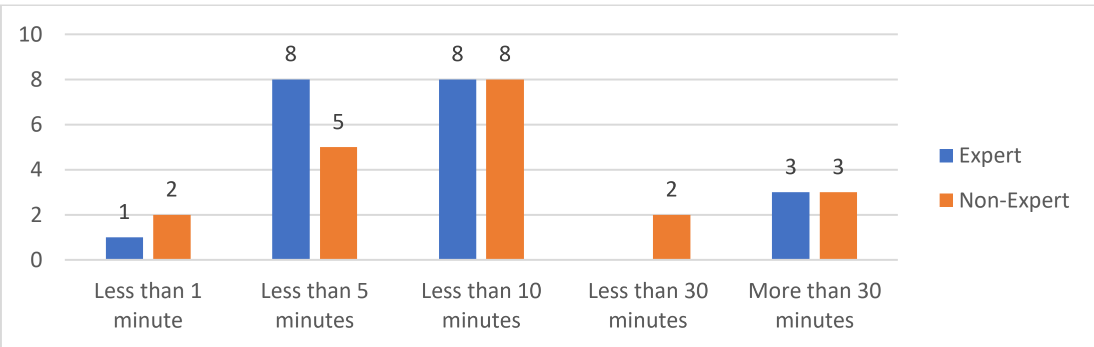


***Figure 96** - Number of test users grouped by their answer to the question <How long do you think you can watch the 3D panorama before getting tired?>.*

Figure 96 shows that Experts and Non-Experts can enjoy a comfortable view before getting tired, for a comparable duration of time.

5.4.3 Presence

5.4.3.1 Presence – Experts vs Non-Experts viewing Lachea island

5.4.3.1.1 Perceived sense of "presence"

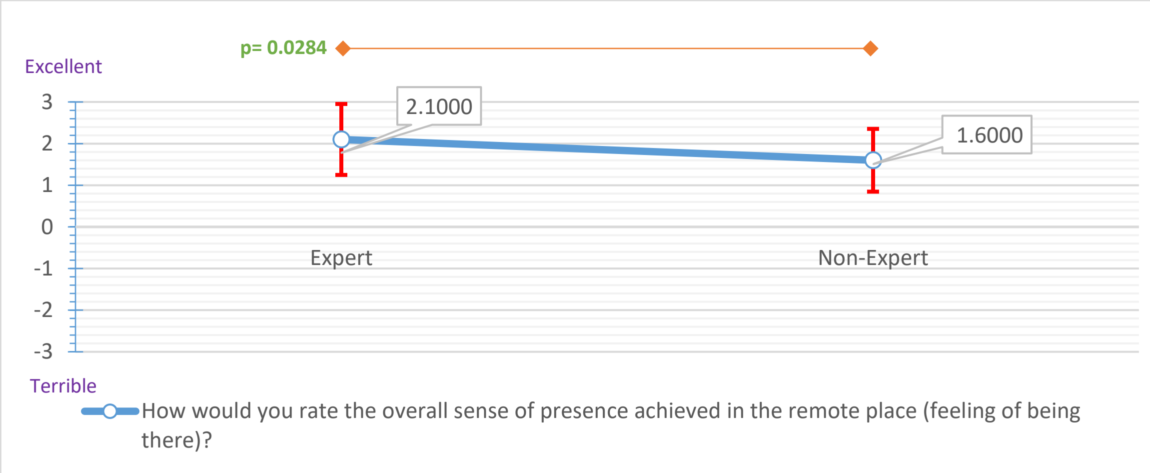


Figure 97 *- Experts rated the sense of presence significantly higher than Non-Experts. This is also suggested by **Figure 98**.*

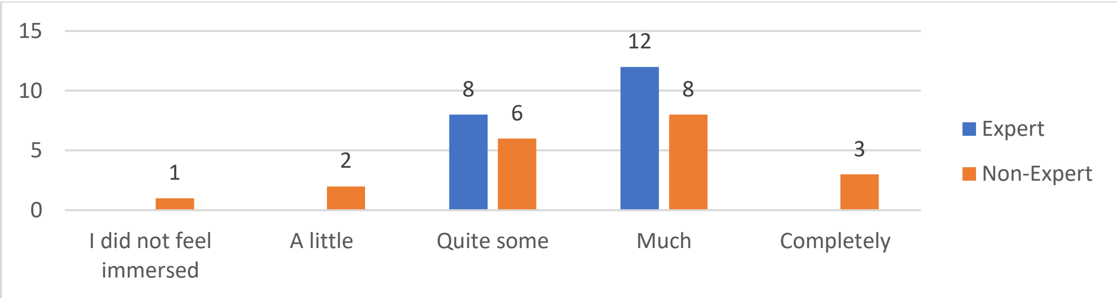


Figure 98 *- Number of test users grouped by their answer to the question <How much did you feel "immersed" into the panorama?>.*

5.4.3.1.2 Isolation

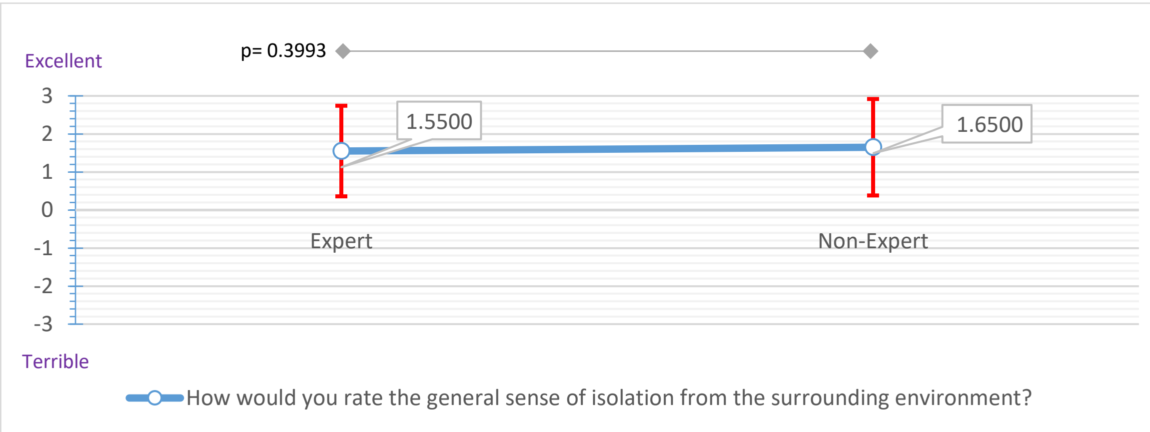


Figure 99 *- No significant difference was found. The general sense of isolation was equally rated as good by both Experts and Non-Experts.*

5.4.3.2 Presence – Experts vs Non-Experts viewing Monello cave

5.4.3.2.1 Perceived sense of "presence"

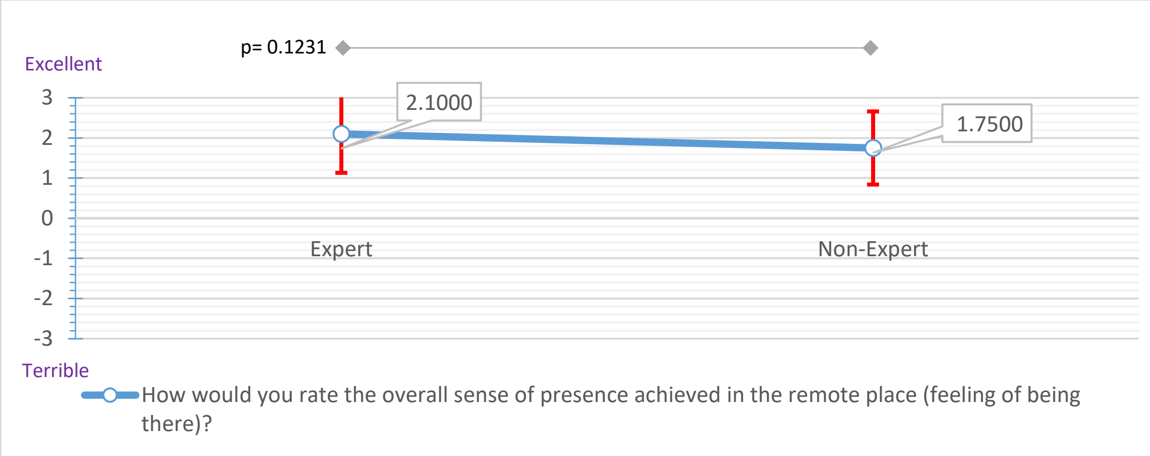


Figure 100 *- The general sense of presence was almost equally rated as very good by both Experts and Non-Experts.*

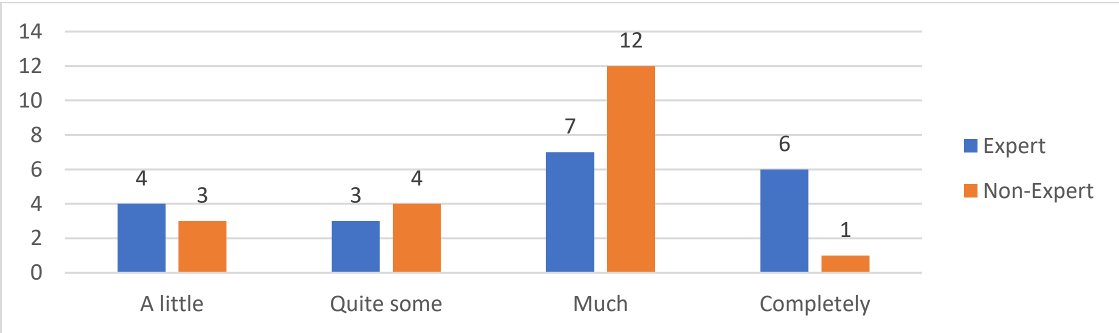


Figure 101 *- Number of test users grouped by their answer to the question <How much did you feel "immersed" into the panorama?>.*

5.4.3.2.2 Isolation

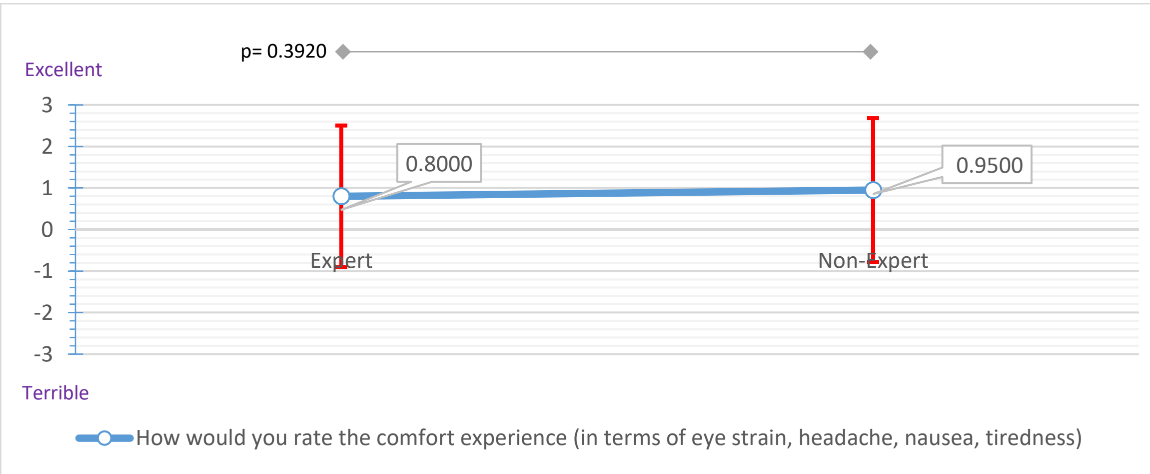


Figure 102 *- No significant difference was found. The general sense of isolation was equally rated as good by both Experts and Non-Experts.*

5.4.4 Depth perception

5.4.4.1 Depth perception – Experts vs Non-Experts viewing Lachea island

5.4.4.1.1 3D Depth impression

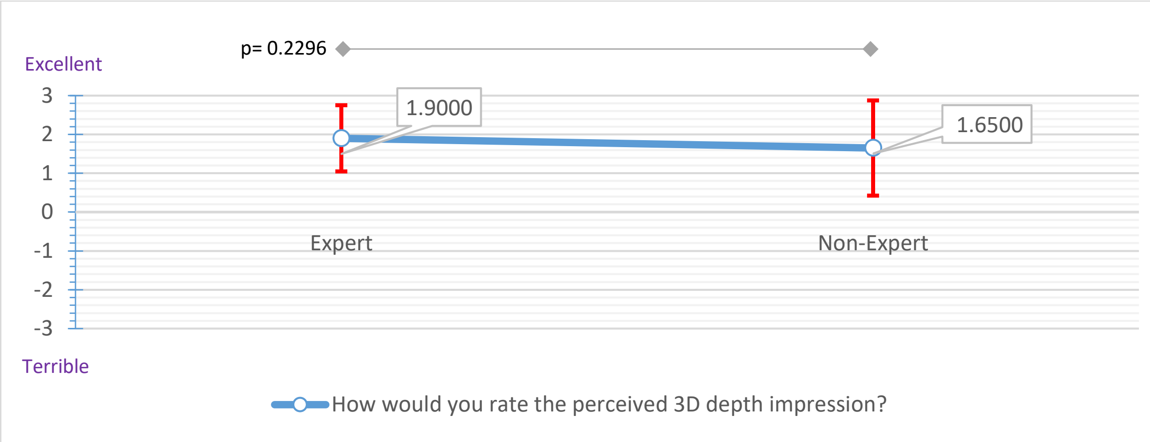


Figure 103 *- Depth impression was rated with no significant difference between Experts and Non-Experts, which considered good the overall achieved 3D perception.*

5.4.4.1.2 Distance perception

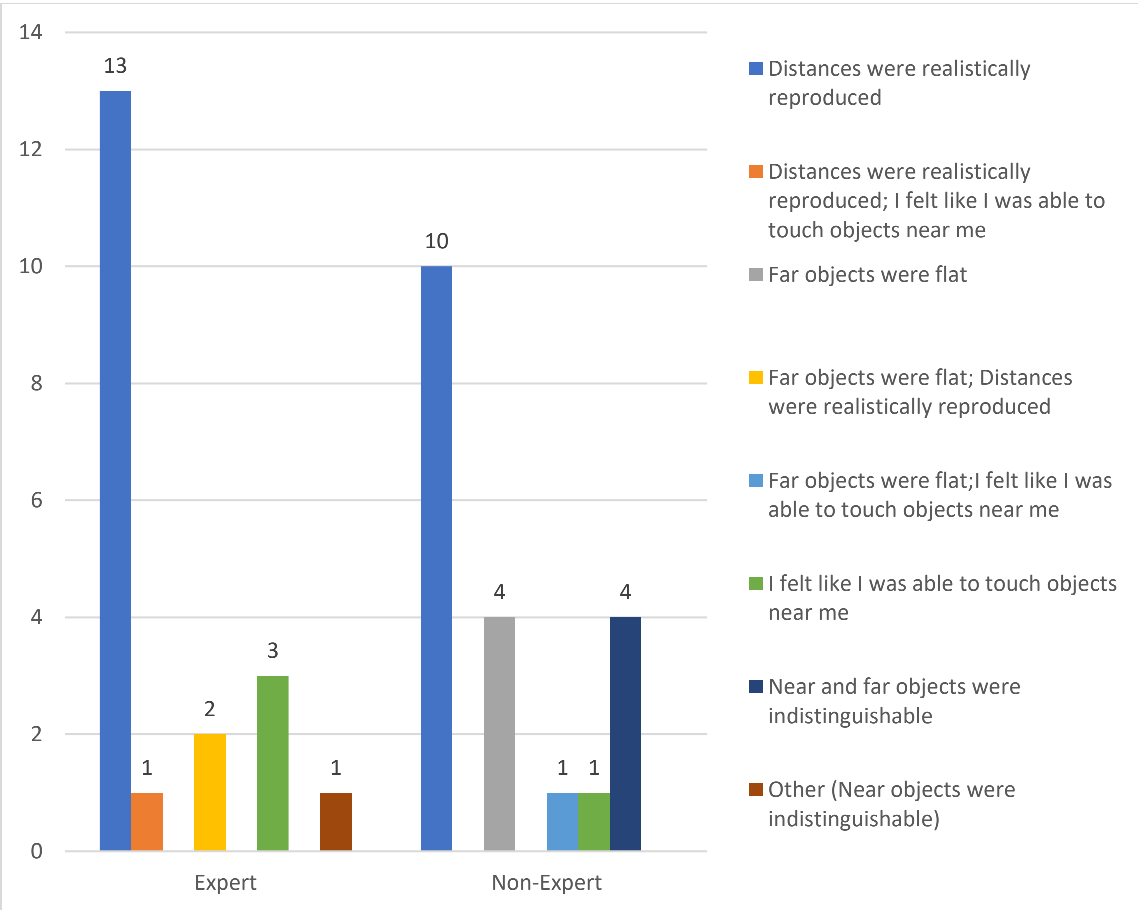


Figure 104 *- Number of test users grouped by their answer to the question <How did you perceive distances in 3D? >.*

Figure 104 shows that Non-Experts experienced issues in distinguishing close and far objects in the scene. In addition, some of them perceived far objects flat. By contrast, Experts had almost none of these issues. Furthermore, some Experts felt like able to touch object close to them and evaluated reproduced distances realistically accurate.

5.4.4.1.3 Color contribution to 3D perception

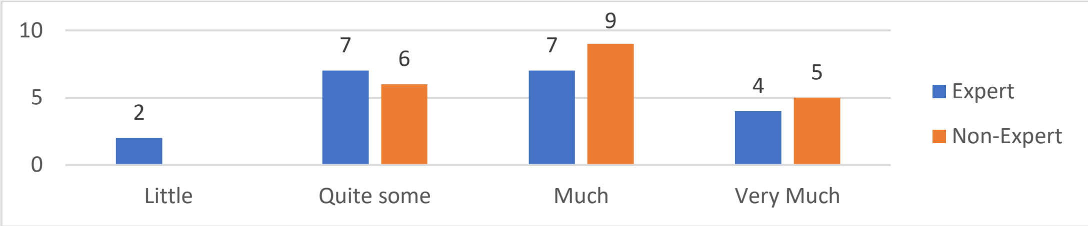


Figure 105 *- Number of test users grouped by their answer to the question <How much do you think that colours have contributed to the perception of your 3D?>.*

Figure 105 shows similar results for both **Experts and Non-Experts**, which **rated color influent on the 3D perception** of the scene.

5.4.4.1.4 Lights and shadows contribution to 3D perception

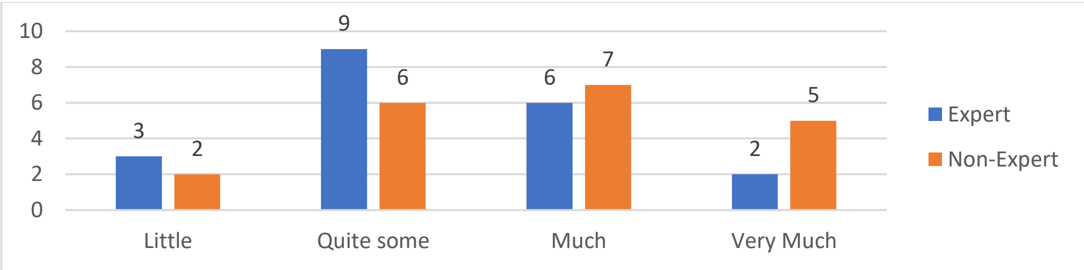


Figure 106 *- Number of test users grouped by their answer to the question <How much do you think that lights and shadows have contributed to the perception of your 3D?>.*

Figure 106 shows that **Non-Experts considered lights and shadows very influent on the 3D perception**. Experts by contrast rated that influence less significant to the final 3D perception of the scene.

5.4.4.1.5 Distorted 3D perception in far or close objects

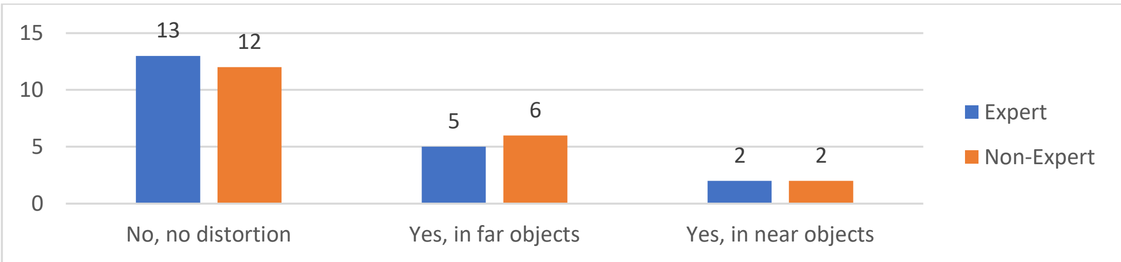


Figure 107 *- Number of test users grouped by their answer to the question <Did you perceive a distorted 3D?>.*

Figure 107 shows **similar results for Experts and Non-Experts**. Only a **few test users noted some 3D distortions in far objects or near objects**. The remaining test users experienced an **overall accurate 3D perception of the scene**.

5.4.4.1.6 Initial time needed before perceiving the 3D of the panorama

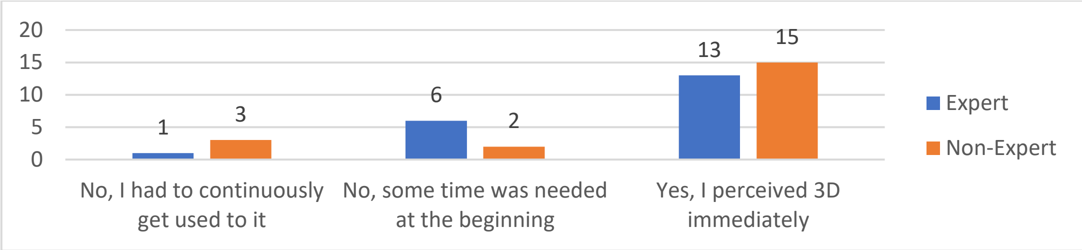


***Figure 108** - Number of test users grouped by their answer to the question <Did you perceive the 3D immediately or your eyes needed some time to get used to it?>.*

Figure 108 shows that **some of the Experts required some time at the beginning** before perceiving the 3D of the scene properly.

5.4.4.2 Depth perception – Experts vs Non-Experts viewing Monello cave

5.4.4.2.1 3D Depth impression

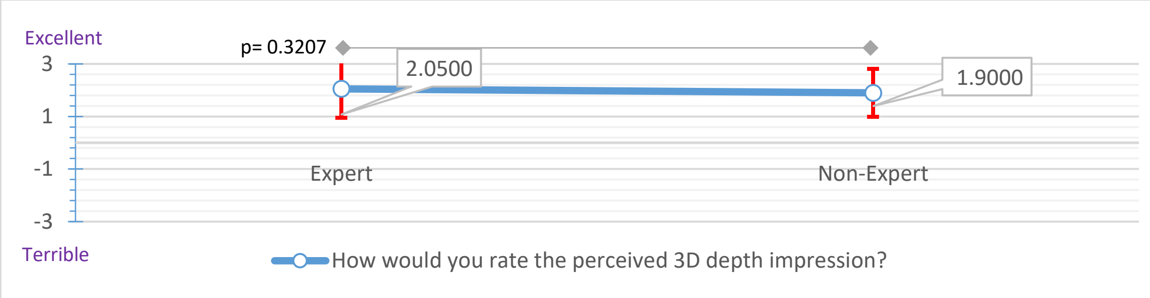


Figure 109 *- Depth impression was rated with no significant difference between Experts and Non-Experts, which considered good the overall achieved 3D perception.*

5.4.4.2.2 Distance perception

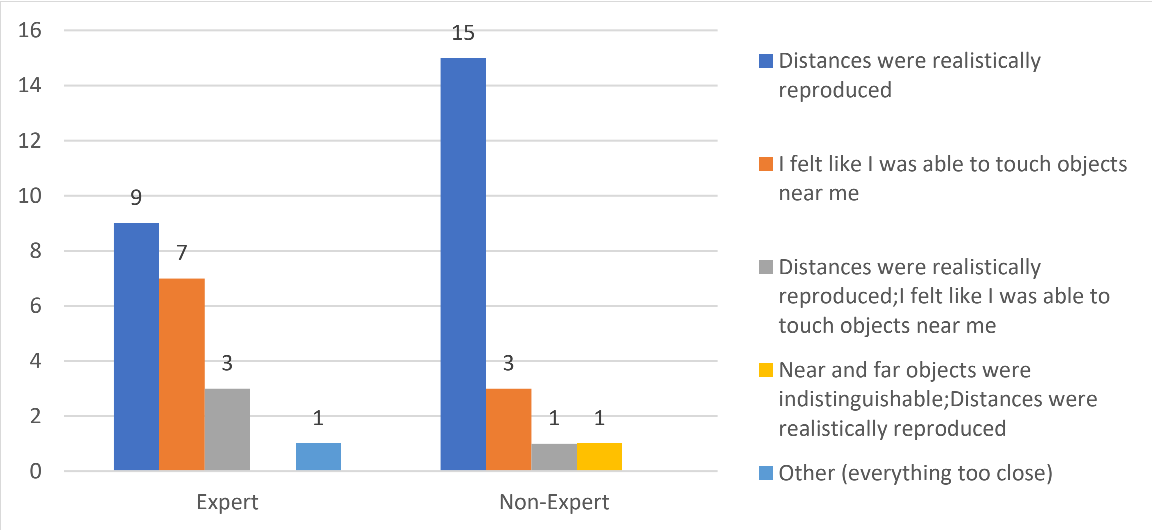


Figure 110 *- Number of test users grouped by their answer to the question <How did you perceive distances in 3D? >.*

Figure 110 shows that Non-Experts had almost no issue at all with distances, which were considered realistically reproduced. Experts felt more like able to touch objects near them, and a few of them did not considered distance realistically reproduced. This suggests that with indoor scenes and background experience the VR visualization can dramatically enhance the feel of presence of test users.

5.4.4.2.3 Color contribution to 3D perception

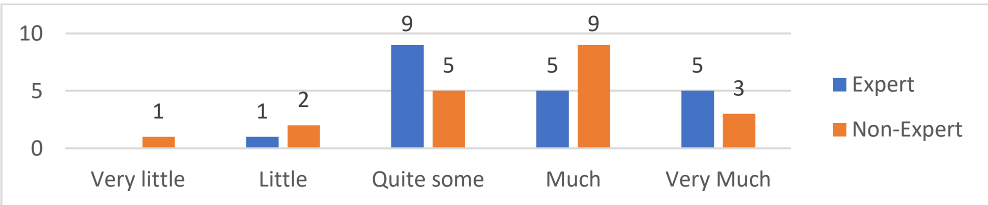


Figure 111 *- Number of test users grouped by their answer to the question <How much do you think that colours have contributed to the perception of your 3D?>.*

Figure 111 shows that **Experts considered colours slightly less influent on 3D perception** compared to Non-Experts. However, both groups rated the influence high.

5.4.4.2.4 Lights and shadows contribution to 3D perception

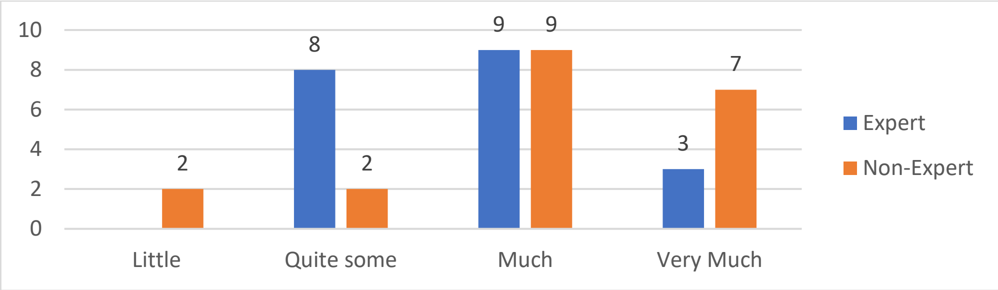


Figure 112 *- Number of test users grouped by their answer to the question <How much do you think that lights and shadows have contributed to the perception of your 3D?>.*

Figure 112 shows that **Non-Experts compared to Experts considered lights and shadows more influent on the 3D perception** of the scene.

5.4.4.2.5 Distorted 3D perception in far or close objects

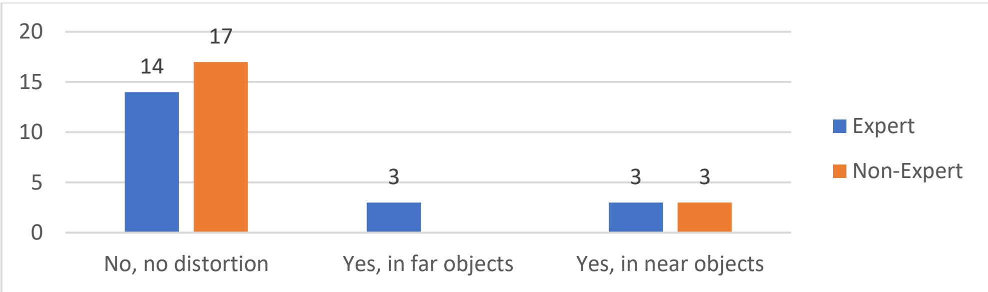


Figure 113 *- Number of test users grouped by their answer to the question <Did you perceive a distorted 3D?>.*

Figure 113 shows similar ratings in terms of **3D distortions** perceived, which were **almost absent in the scene**.

5.4.4.2.6 Initial time needed before perceiving the 3D of the panorama

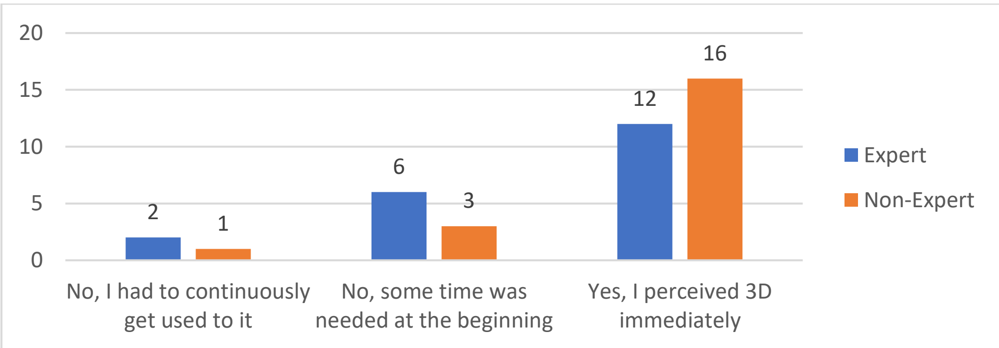


Figure 114 *- Number of test users grouped by their answer to the question <Did you perceive the 3D immediately or your eyes needed some time to get used to it?>.*

Figure 114 shows that **some of the Experts required some time at the beginning before perceiving the 3D of the scene properly**.

5.5 Analysis

Table 16 shows a summary of the most significant results and of the most relevant presuppositions that have been confirmed by the performed user study.

5.5.1 Realism

5.5.1.1 Outdoor Environments

Realism of VR compared to Realism of photos and videos. For outdoor environments, people with high familiarity gave a higher score to the realism of its virtual reality remote observation compared to photos and videos (see **Figure 78**). This generally means that Experts found a lot of similarities between how the place is reproduced in their memory and how it is reproduced in virtual reality, which leads them to consider the VR remote observation more realistic than the one of 2D photos and 2D videos. This result is also confirmed by literature [315] [316], which reports that presence pushes people to remember a virtual environment more as a place than as a set of pictures.

Realism in terms of horizontal size of display. Another interesting result is that people that have seen the place for their first time only in virtual reality thought that the horizontal size of the display used in the Mobile VR headset has a strong influence on the perceived realism of the environment. This is curious because people with high familiarity instead considered it less influent (see **Figure 81**). A possible reason is that Experts already know the place and associate already the missing portions of the panorama shown on screen with the memory they have of the place, which somehow completes their virtual vision. In literature I found other researches stating that larger horizontal displays can enhance realism and presence [317] [318] , however they did not specifically distinguish people with high familiarity from people with low familiarity.

5.5.1.2 Indoor Environments

Realism in terms of horizontal size of display. Like for outdoor environments, results proved that Non-Experts consider very strong the influence of the horizontal size of the display on realism, whilst Experts gave lower scores to it. This again can be explained by the luck of familiarity with the place of Non-Experts and their desire to explore it more than Experts, which already know it in real life.

5.5.1.3 Indoor vs Outdoor

Realism of indoor vs Realism of outdoor. From the comparative results, Non-Expert judged the overall achieved level of realism of the outdoor environment slightly more than Experts (see **Figure 89**). I believe this is due to the emotions they felt while observing the sunlight and the natural elements of the island (e.g. trees, sea, rocks) for the first time, and to the fact that they had no reference in real life to compare the virtual environment with the real one. Some of the Experts instead might have considered the panorama less impressive than the view they saw in real life, which caused them to consider it less realistic.

5.5.2 Comfort

5.5.2.1 Outdoor Environments

Eye strain or visual fatigue. Results show that Experts experienced a little more eye strain or visual fatigue (see **Figure 92**) than Non-Expert. I believe this was caused by the amount of small stitching errors that were present in the Lachea panorama (see subsection 5.1.2). Therefore, this might suggest that Non-Expert pay less attention to visual errors as they are more focused in exploring the remote environment for the first time, differently from Experts. Furthermore, Experts have more means to compare the real place with the virtual one and might experience more discomfort as they better notice any existing visual difference. This is also confirmed by the fact that Experts were generally unable to watch the panorama without experiencing discomfort for as long as Non-Experts (see **Figure 93**).

No other significant difference was reported between Experts and Non-Experts in terms of general discomfort, which was relatively low.

5.5.2.2 Indoor Environments

Eye strain or visual fatigue. Like for outdoor environments Experts experienced a little more eye strain (see **Figure 95**) than Non-Expert. Therefore, I make the same assumptions that were made for outdoors for the indoor environments.

However, this time both Experts and Non-Experts were able to watch for a similar amount of time the panorama before experiencing discomfort (see **Figure 96**). This result might have several explanations:

- The panorama of Monello Cave has fewer stitching errors than the Lachea, therefore it might have caused less discomfort in long terms to Experts, due to fewer visual mismatches.
- The amount of light in the indoor environment is lower than the one outside and might have caused less discomfort in long terms also to Experts, as the eye has been less constricted and more comfortably dilated, allowing longer comfortable vision of the scene.
- The scene was mainly monochromatic, and the lower presence of colours might have led Experts to watch the scene for longer times even if they knew the place before.

No other significant difference was reported between Experts and Non-Experts in terms of general discomfort, which was relatively low at a similar level of the one of outdoors.

5.5.3 Presence

5.5.3.1 Outdoor Environments

Perceived sense of presence. Experts felt significantly more present into the remote outdoor environment than Non-Experts (see **Figure 97** and **Figure 98**). A possible reason is that the emotions of being in an island with a sunny day and the good memories of that environment in real life pushed test users to feel more "teleported" into that remote environment. I was not surprised by this result as the same trend was obtained also for realism, which can be considered related to the sense of presence as well in this case. There are some studies that proved that emotions and sense of presence are strongly coupled [319] [320] [316] [321] [322]. I plan to further investigate this aspect through specific questions on the "feel of being on holiday" in the next user study, to further analyse the coupling emotions – sense of presence (see Chapter 6).

No significant difference between Experts and Non-Experts was reported in terms of isolation: both groups felt quite isolated from the room (see **Figure 99**).

5.5.3.2 Indoor Environments

No significant difference between Experts and Non-Experts was reported in terms of:

- Perceived sense of presence. However, slightly more Experts felt more immersed in the indoor panorama than the outdoor panorama (compare **Figure 98** and **Figure 101**).
- Isolation. However, both Experts and Non-Experts seemed to feel slightly more isolated with the indoor panorama than the outdoor panorama (compare **Figure 99** and **Figure 102**).

5.5.4 Depth Perception

5.5.4.1 Outdoor Environments

Distance perception. According to **Figure 104**, some of the Non-Experts had issues distinguishing close and far objects, others perceived far objects flat. By contrast, Experts had almost no issue, some of them felt like able to touch close objects and evaluated distances realistically reproduced.

Since Experts were practically extremely satisfied with the realism of the distances of objects in the scene, and since they can be considered a ground truth as they can compare real life with virtual scene, it can be assumed that the used virtual reality system was very accurate for depth reproduction of outdoor environments. However, it is curious that Non-Experts did not perceive the virtual environment in the same way.

I believe this is because Non-Experts have not as much familiarity with the place as Experts, and in virtual reality they are not entirely able to evaluate distances without the actual presence of their body. This was proved by some researches, which highlighted the need to represent the viewer's body using avatars in virtual reality to improve distance estimations [323] [324] [325].

Lights and shadows contribution to 3D perception. A small group of Non-Experts considered lights and shadows very influent on the 3D perception, compared to Experts (which reported lower influence, see **Figure 106**). This might be a mechanism of Non-Experts to compensate their low familiarity with the place and the absence of an Avatar representing their real body in virtual reality. I believe this pushed them to relay more on monocular cues to better estimate depth and space around them. There are some researches that prove how texture cues can enhance distance and depth estimation [323]: it is believed these also apply to shadows of objects in the scene, which in the outdoor environment are very strong due to the presence of the sunlight (see subsection 5.1.2). Some studies were performed in literature on the effect of shadows on depth perception, however some of them found no significant results to prove it [326]. This makes this investigation valuable for the results obtained and suggests that further investigation might be carried out.

Initial time needed before perceiving 3D. A few Experts (see **Figure 108**) needed some time to perceive the scene in 3D, and some Non-Experts needed to continuously get used to it. I believe this is due to the low background experience of those test users towards 3D and virtual reality devices, or to an inappropriate configuration of the IPD distance of the Mobile VR headset.

No other significant difference between Experts and Non-Experts were found in terms of:

1. Depth impression (both groups expressed very good scores);
2. Color influence on depth perception (much influence for both groups);
3. Distorted 3D perception (both groups perceived almost no distortions or some distortions in far objects only). These distortions might have been caused by some of the stitching errors present in the Lachea panorama (see subsection 5.1.2).

5.5.4.2 Indoor Environments

Distance perception. In **Figure 110** Experts felt more the sensation to be able to touch close objects than Non-Experts. I believe this is due to their higher familiarity with the place. Non-Experts were able to perceive distances more realistically than Experts this time. This might be due to the stronger depth perception caused by the closed environment of Monello cave, which presents higher parallax and more objects close to the viewer (see subsection 5.1.2).

Color contribution to 3D perception. Color contribution to depth perception was considered slightly less important by Experts than Non-Experts. Considering that the Monello cave is almost monochromatic (see subsection 5.1.2), I assume that Non-Experts were aided to focus more on depth without being distracted by wide ranges of colours like the ones of Lachea. This brings us to consider the possibility that a simpler S3D scene (including one with less colour changes) can help the viewer to better estimate depth. This is in line with some of the studies that this investigation found in the state-of-the-art analysis (see paragraph 4.2.1.1). Furthermore, it is possible that the difference with Experts exists because some of them might have not focused too much on the virtual perceived depth, as they relied more to their memory of the real place.

Lights and shadows contribution to 3D perception. Lights and shadows were considered more influent on depth perception for Non-Expert than for Experts. The assumptions that I made on lights and shadows contribution for the outdoor scene apply to the indoor scene as well, but in addition they are enforced by the artificial light that enlightened the space surrounding the viewer causing a lot of shadows to appear between the rocks.

Initial time needed before perceiving 3D. A few Experts (see **Figure 114**) needed some time to perceive 3D. I think that this result is again due to the low background experience of those test users towards 3D and virtual reality devices, or to an inappropriate configuration of the IPD distance of the Mobile VR headset.

No other significant difference between Experts and Non-Experts were found in terms of:

1. Depth impression (both groups expressed very good scores, slightly higher than outdoors). The higher score than the one of outdoors is justified because the viewer has closer objects and a closed environment with Monello cave, which means more disparity and stronger depth perception.
2. Distorted 3D perception (both groups reported almost no distortion).

	Realism	Comfort	Presence	Depth Perception
Significant differences (Lachea)	- Experts judged VR realism better than photos and videos - Non-experts believed a larger horizontal display enhances realism	- Experts experienced more discomfort and more quickly	- Experts rated sense of presence higher	- Experts had almost no issues distinguishing close and far objects - Non-Experts believed lights and shadows are very influent on 3D - Few Experts needed time to see in 3D
Similarities (Lachea)	- Good realism similar to real life - Good realism in terms of objects deformations / natural elements - Good realism for emotions - Equally realistic perception of 3D in centre and borders of display	- Overall comfort was positively evaluated - Similar time before getting tired of watching the VR panorama	- Good sense of isolation	- Good overall 3D perception - Colours are influent on the 3D perception - Overall accurate 3D perception of scene, with few exceptions
Significant differences (Monello)	- Non-experts strongly believed that a larger horizontal display would have enhanced realism	- Experts experienced more discomfort	---	- Experts felt like touching close objects (experience + indoor scene enhance presence), but quite less realistically reproduced distances - For Experts colours are slightly less influent on 3D perception - For Non-Experts lights and shadows are more influent on 3D - Few Experts needed time to see in 3D from the beginning
Similarities (Monello)	- Good realism similar to real life - Excellent realism compared to photos and videos - Good realism in terms of objects deformations / nature - Good realism for emotions - Equally realistic perception of 3D in centre and borders	- Overall good comfort - No difference for the time before perceiving discomfort watching	- Slightly lower sense of presence for Experts but not significantly - Good sense of isolation	- Good overall 3D perception - 3D distortions almost absent in the scene

***Table 16**. Summary of the outcome of the usability evaluation on familiarity and environment.*

5.6 Guidelines

Based on the previous sections' results and analysis, a Guidelines section is provided to summarize main outcomes in terms of concise guidelines for system designers. The guidelines are laid down looking at the outcome of the proposed evaluation against the four elements that have been considered as relevant contributors to the VR remote visual observation: realism, comfort, presence and depth-perception.

- **How to improve realism?**
  - Use Mobile VR rather than photos and videos to observe remote environments, especially when viewers already know the place.
  - Use very large horizontal size of display inside the Mobile VR HMD, especially when people do not know the place and observe it remotely for their first time.
  - Give preference to natural features (e.g. trees, sea, rocks, sunlight), especially when viewers have never visited the place in real life before.

- **How to improve comfort?**
  - Avoid stitching errors when generating the 3D panoramas, especially when viewers know the place. However, when viewers do not know the place, stitching errors might be less noticed as they will be more focused in exploring the place rather than carefully analyse its details and errors.
  - Avoid excessive brightness so that test users experience less visual fatigue.
  - Keep the scene simple, so that the user avoids uncomfortable vision in 3D.

- **How to improve presence?**
  - Enhance emotions for the viewer, by prioritizing nice looking scenes (e.g. wide range of colours, holiday like settings, images to elicit emotions). This specifically applies to people that have real life memories related to the visualized place.
  - Enhance immersion, by providing visual and audio feedback to make the experience more interactive and emotionally involving. This outcome first resulted by the literature analysis that suggests high immersion to enhance emotion, which in turns activates presence. This was also confirmed through this first usability evaluation.
  - Be sure to guarantee a good level of isolation from the place where the viewer stands in real life, to avoid distractions from the virtual experience.

- **How to improve depth perception?**
  - For a better distance estimation, use visual aids (e.g. arrows, grids, or virtual avatars) to allow the viewer to have points of references that can help better evaluations of surrounding spaces.
  - Provide monocular cues (e.g. lights and shadows, texture cues, relative sizes), so that viewers can solve possible binocular cues conflicts and better perceive depths of the scene. This covers an important role especially for people that have no familiarity with the place and try Mobile VR for the first time.
  - Consider an appropriate time to allow viewers to adjust IPD and distance from the screen accordingly, so that they can enjoy an optimal HMD setup and better 3D perception.
  - Prioritize objects that are close to the viewers but avoiding excessive disparity.

- Prioritize scene designs having several and large depth intervals. This will guarantee more virtual spaces between background and foreground objects, showing more distinctly their position within the 3D space, improving relative distance estimations.
- Use simple scenes to allow the user to perceive shapes and objects in 3D with low occurrence of visual conflicts.

5.7 Conclusion

In this chapter the procedure and outcomes of the first usability evaluation proposed by this thesis were presented.

To perform this first user study, an evaluation design was proposed focused on the *familiarity* of users with the observed environment in VR and on the characteristics of the observed *environment*. In detail, the visual perception of test users that visited the observed places in real life before visualizing them in VR (Experts) and test users that visualized these places in VR only (Non-Experts) were compared, in combination with indoor and outdoor observed panoramas.

Variables used for this evaluation, test users' data, and information on the statistical tools adopted to calculate results were also discussed. Furthermore, a detailed overview of the Virtual Reality panoramas' implementation and of the pilot-test-driven approach that has been used to improve the usability evaluation design were discussed.

All results were presented towards each usability variable's influence over realism, comfort, presence, and depth perception. These were then thoroughly analysed and compared with what was previously found in the systematic review. Finally, guidelines to improve the setup of the VR system were recommended, with the aim of achieving a more realistic, comfortable and faithful 3D remote observation of the visualized environments.

CHAPTER 6. USER STUDY - DISPLAY AND ENVIRONMENT

This chapter describes the second experimentation related to the proposed Phase Two ("Build Knowledge through Assessments). The experimentation focuses on *display* and *environment*. The chapter starts by introducing proposed evaluation's design (section 6.1), and it continues by presenting implementation (section 6.2), procedure and extended pilot test (see section 6.3), user study's results (section 6.4), results analysis (section 6.5), and guidelines (section 6.6).

6.1 Evaluation Design

Following the aim of the proposed investigation (see subsection 3.3.1), I decided to examine the role of **display** and **environment** in Mobile VR. My goal is to understand their influence on realism, comfort, presence, and depth perception. To achieve this, I propose this user study and devise specific designs to better analyse display characteristics and visualized environment.

6.1.1 Display

The design that is chosen to investigate the role of display involves the selection of two smartphone displays compatible with a Mobile VR headset and presenting dissimilar specifications. Within this purpose, I decide to use an **iPhone 6** and a **LG G3** smartphones, as they provide different resolution, different pixel density, and different screen size among other differences, while both being able to portrait the relevant characteristics of the presented environments. Such selection was performed by carefully looking at both display specifications. A comparative table is shown in **Table 17**, with an indication of what is considered the best expected performance for each feature (bold in the table).

Parameter	iPhone 6	LG G3	Best
Max brightness in Nits (Higher is better)	**559.2962**	379.322	iPhone 6
Black levels in Nits (Lower is better)	**0.3647**	0.4337	iPhone 6
Contrast Ratio at 100% Brightness (Higher is better)	**1,534.0**	875.0	iPhone 6
Average White point in K – Closer to 6504K is better	**6,515**	7,244	iPhone 6
Grayscale Accuracy (Lower is better)	**1.9683**	3.6935	iPhone 6
Saturation Accuracy (Lower is better)	**1.1929**	4.7599	iPhone 6
Gretag MacBeth ColorChecker Accuracy (Lower is better)	**1.7645**	3.9702	iPhone 6
Display type	Retina HD: LED-backlit IPS LCD	IPS LCD	-
Display Resolution	1334×750 pixels	**1440 x 2560 pixels**	LG G3
Pixel Density	326 ppi	**538 ppi**	LG G3
Pixel size	0.078 mm	**0.047 mm**	LG G3
Screen Size	4.7 inches	**5.5 inches**	LG G3

Table 17. Technical specifications and display comparison of iPhone 6 [327] and LG G3 [328].

6.1.2 Environment

To comply with requirements of the proposed experiment objectives (see subsection 3.3.1) I decide to select **Lachea island** and **Monello cave** panoramas, as they present many environmental differences (see 5.1.2) and performed very well in the previous user study (see Chapter 5).

6.1.3 Usability variables

According to the aim of the investigation, the following set of independent and dependent variables is analysed.

Independent variables:

- Type of Mobile VR display. This refers to the mobile device (iPhone 6 or LG G3) that was put inside the Mobile VR headset to view the 3D panorama.
- Observed Environment. This is an outdoor panorama illuminated by natural sunlight (Lachea island) or an indoor panorama illuminated by artificial light (Monello cave).

Dependent variables:

- Realism. It measures the level of perceived realism achieved by visualizing the Observed Environment on the Mobile VR display. It relies on the definition of *perceptual realism*, which addresses the perception of stereoscopic images as being a truthful representation of reality [36].
- Comfort. It measures the occurrence of *visual fatigue*, *eye strain*, and *discomfort* while watching the 3D panoramas on the Mobile VR display, by using users' subjective evaluations.
- Presence. It is generally defined as users' subjective sensation of *'being there'* [329] in a scene depicted by a medium [308]. Following a more rigorous definition given by Slater and Usoh [309], presence is "the extent to which human participants in a virtual environment allow themselves to be convinced while experiencing the effects of a computer-synthesized virtual environment that they are somewhere other than where they physically are".
- Depth perception. It addresses the perception of space in VR, and the achieved 3D effect performances when viewing the panorama on the VR display.

For each dependent variable (*factor*) the following set of features were defined and investigated:

- **Realism**
  - Level of realism in terms of image resolution
  - Level of realism in terms of image sharpness
  - Level of realism in terms of image definition and not pixeled
  - Level of realism in terms of image vividness
  - Level of realism in terms of image brightness
  - Level of realism in terms of intensity of colours
  - Level of realism in terms of contrast
  - Level of realism compared to real life
  - Level of realism compared to photos and videos
  - Level of realism in terms of objects deformations / natural elements on the scene
  - Realism of the 3D effect perceived in VR
  - Level of realism in terms of emotions
  - Influence of colours over perceived emotions
  - Influence of horizontal size of display over Realism

- **Comfort**
    - Discomfort
    - Eye strain or visual fatigue while watching panorama
    - Max duration of time during which watching the panorama remains comfortable
- **Presence**
    - Perceived sense of "presence"
    - Presence and Tunnel vision
    - Isolation
    - Influence of isolation over emotions
    - Induced emotions from the panorama visualization
    - Influence of emotions over the sense of presence
    - Emotions caused by coming back to real life after the VR experience
- **Depth Perception**
    - 3D Depth impression
    - Distance perception
    - Color contribution to 3D perception
    - Lights and shadows contribution to 3D perception
    - Distorted 3D perception in far or close objects
    - Influence of 3D perception over emotions
    - Initial time needed before perceiving the 3D of the panorama

6.1.4 Participants

The complete set of test users consists of 20 people, of which 6 females and 14 males.

All participants have never visited in real life the places they viewed through the VR headset before. Therefore, this sample was defined as a "**group of Non-Experts**".

The following additional information on test users are collected to investigate possible effects of personal differences on dependent variables:

- **Personal information**
    - Age of test user
    - Gender
    - Use of glasses in real life
    - Eye problems such as astigmatism, myopia, presbyopia, hypermetropia
    - Previous experience with computer games
    - Previous experience with 3D displays or 3D digital images
    - Previous experience with virtual reality
    - Interpupillary distance

6.1.5 Statistical tools

A within subject design is adopted. The following assumptions are made to perform several ***within-subject paired t-test*** [310], in line with literature guidelines [311] [312]:

- Data are collected from a simple random sample of 20 test users, which are selected from a representative, randomly selected portion of the total population.
- Homogeneous, or equal, variance exists when the standard deviations of samples are approximately equal.
- The sample size used for each of the 4 VR setups (i.e. Lachea iPhone 6, Monello iPhone 6, Lachea LG G3, Monello LG G3) is uniform and equal to 20 participants.

Collected data satisfy all above-mentioned paired t-test requirements. Furthermore, each test user visualizes the 4 VR setups and evaluates them using the same questionnaire. This implies the collection of repeated measures for each test user, using the 4 different VR setups.

Instead of a MANOVA, multiple t-tests in Excel are preferred to obtain further details on possible significant differences resulting from the 4 VR setups. To do it, all possible coupling combinations of values for the independent variables are considered as follows:

1. Lachea iPhone 6 vs Monello iPhone 6;
2. Lachea iPhone 6 vs Lachea LG G3;
3. Lachea iPhone 6 vs Monello LG G3;
4. Monello iPhone 6 vs Lachea LG G3;
5. Monello iPhone 6 vs Monello LG G3;
6. Lachea LG G3 vs Monello LG G3.

For each of the six combinations, the p-value resulting from each *paired t-test* performed on each question is calculated.

6.2 Implementation

The same implementation of the previous user study (see Chapter 5, section 5.1.5) is used during this test. Furthermore, additional questions are asked specifically addressing the effects of display and image characteristics on the viewer visual perception of the scene.

6.3 Procedure and Extended Pilot Test

Guidelines from the previous extended pilot test (see section 5.1.2) are followed, to improve design and questionnaire of this user study.

6.3.1 Chosen Procedure

After careful revision, the following steps are chosen for the evaluation procedure:

1. Each test user watches a VR panorama using an iPhone 6 or an LG G3 inserted into a plastic Mobile VR HMD. A total of 4 VR setups (phone-panorama combinations) is defined. Phone and panorama are established by the VR setups scheduling in **Table 18**. Each panorama is shown for less than 5 minutes.
2. After the viewing of the virtual panorama, each test user is requested to answer a list of questions related to features addressing *Realism*, *Comfort*, *Depth Perception*, and *Presence*. To make more precise and accurate estimations test user can view multiple times the panorama while answering the questions.
3. After completing all tasks, a comparative questionnaire is proposed to compare the 4 VR setups.
4. At the end of this usability evaluation, each test user is asked to leave a feedback on the experience.

6.3.2 VR Viewing Scheduling

Test user ID	Lachea island + iPhone 6 (SET A1)	Monello cave + iPhone 6 (SET A2)	Lachea island + LG G3 (SET B1)	Monello cave + LG G3 (SET B2)
#1	First setup	Second setup	Third setup	Fourth setup
#2	Fourth setup	Third setup	First setup	Second setup
#3	First setup	Second setup	Fourth setup	Third setup
#4	Third setup	Fourth setup	First setup	Second setup
#5	First setup	Third setup	Second setup	Fourth setup
#6	Second setup	Fourth setup	First setup	Third setup
#7	First setup	Third setup	Fourth setup	Second setup
#8	Fourth setup	Second setup	First setup	Third setup
#9	First setup	Fourth setup	Second setup	Third setup
#10	Second setup	Third setup	First setup	Fourth setup
#11	Second setup	First setup	Third setup	Fourth setup
#12	Third setup	Fourth setup	Second setup	First setup
#13	Second setup	First setup	Fourth setup	Third setup
#14	Fourth setup	Third setup	Second setup	First setup
#15	Second setup	Third setup	Fourth setup	First setup
#16	Fourth setup	First setup	Second setup	Third setup
#17	Second setup	Fourth setup	Third setup	First setup
#18	Third setup	First setup	Second setup	Fourth setup
#20	Fourth setup	Second setup	Third setup	First setup
#21	Third setup	Second setup	Fourth setup	First setup

***Table 18**. VR setups scheduled for each test user. Each row defines the order in which the 4 VR setups are shown to each test user. To reduce possible bias due to the order of visualization of the panoramas, as many different combinations as possible were used.*

Table 18 presents the scheduling that is proposed for each of the selected test users.

6.4 Results

Answers to the questionnaire have been collected using a digital online document created via *Google Forms* [314]. Data has then been processed as follows. From now on, the following charts will show in orange p-values < 0.05 resulting from the paired t-test comparison of each of the 6 combinations of VR setups.

6.4.1 Realism

6.4.1.1 Level of realism in terms of image resolution

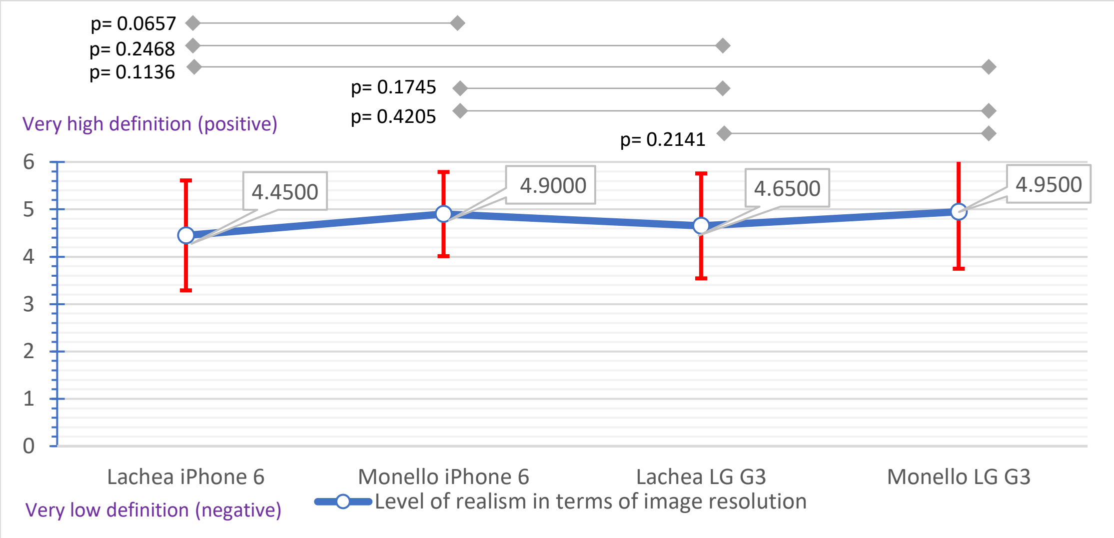


***Figure 115** - No significant difference resulted for panorama quality in terms of image definition (See Appendix B – question R2.6) [min = 0, Max = 6].*

Figure 115 shows that there is almost a significant difference between Lachea and Monello using iPhone 6 (p-value = 0.06), with the latter having higher quality image definition. However, this difference is not visible using the LG G3, and it is the only difference reported by results.

6.4.1.2 Level of realism in terms of image sharpness

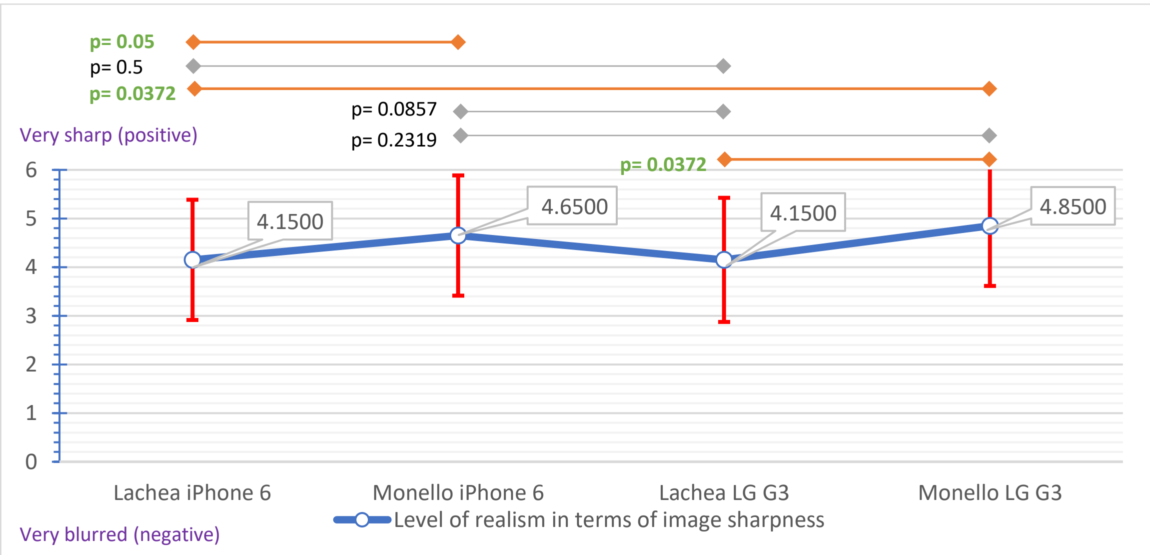


***Figure 116** - Significant differences for level of realism in terms of image sharpness (See Appendix B– question R2.7) [min = 0, Max = 6].*

Figure 116 reports three p-value < 0.05. This confirms the existence of a significant difference on perceived realism in terms of image sharpness.

6.4.1.3 Level of realism in terms of image definition and not pixeled

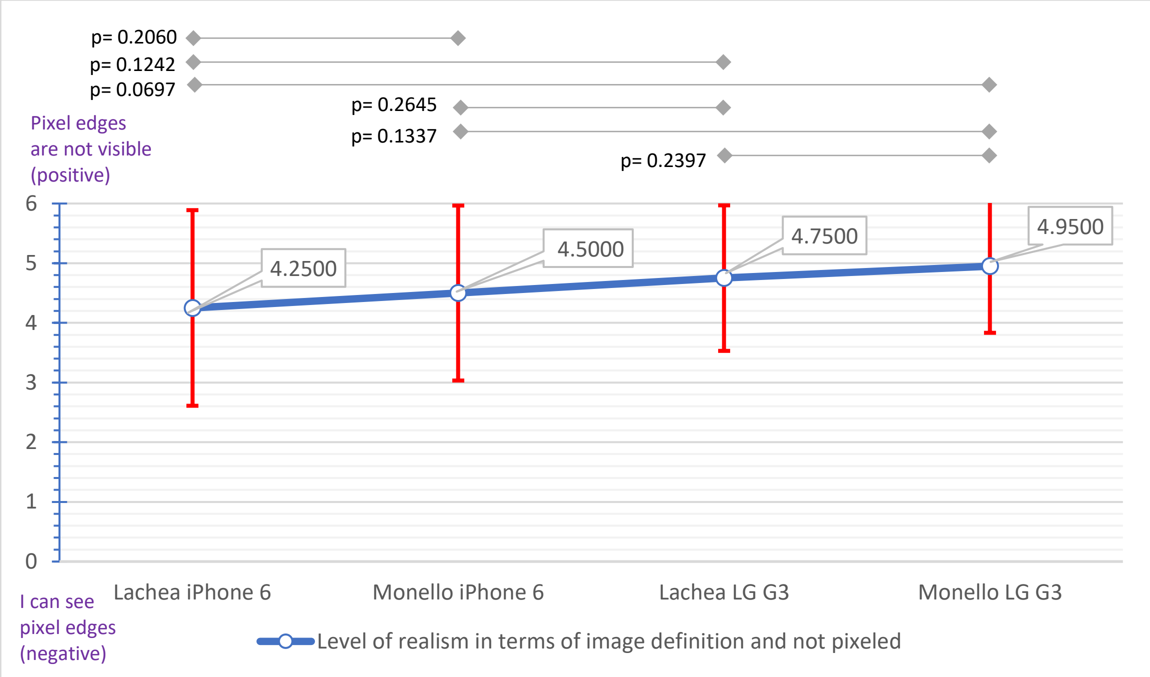


Figure 117 *- No significant difference resulted for level of realism in terms of image definition and not pixeled (See Appendix B – question R2.8) [min = 0, Max = 6].*

Figure 117 shows no significant difference in terms of image definition and not pixeled. Furthermore, the influence of pixel edges over the achieved level of realism was questioned. Results are shown in **Figure 118**, with no significant difference found.

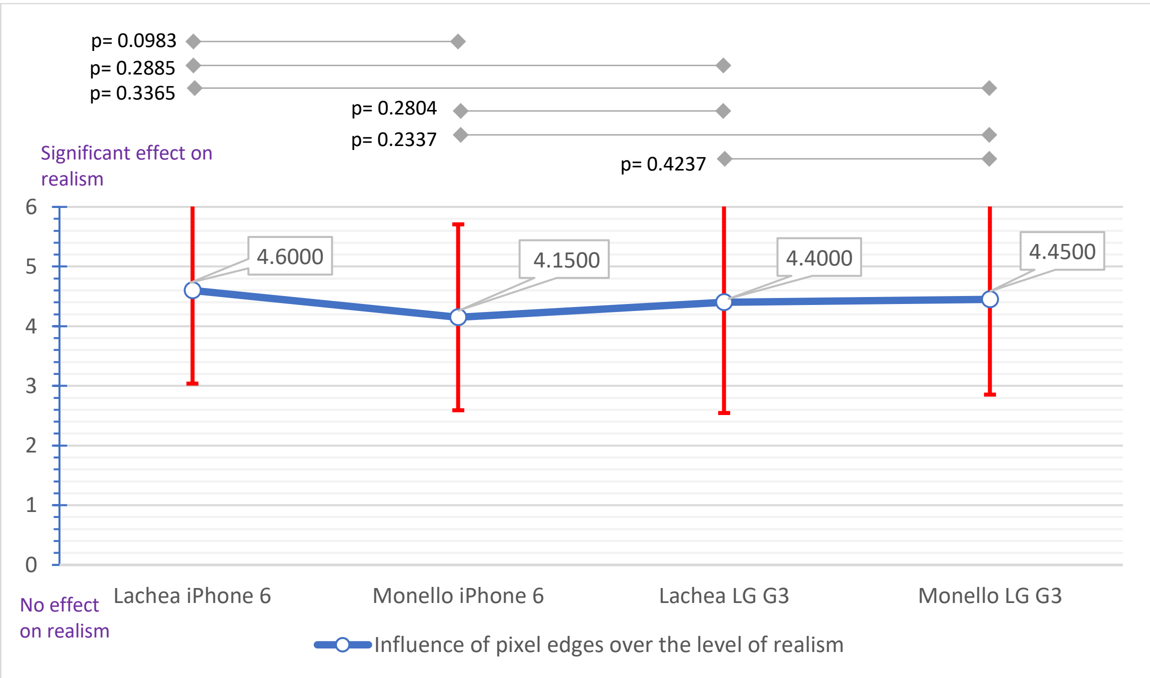


Figure 118 *- No significant difference resulted for the influence of pixels edges over the level of realism (See Appendix B – question R2.9) [min = 0, Max = 6].*

6.4.1.4 Level of realism in terms of image vividness

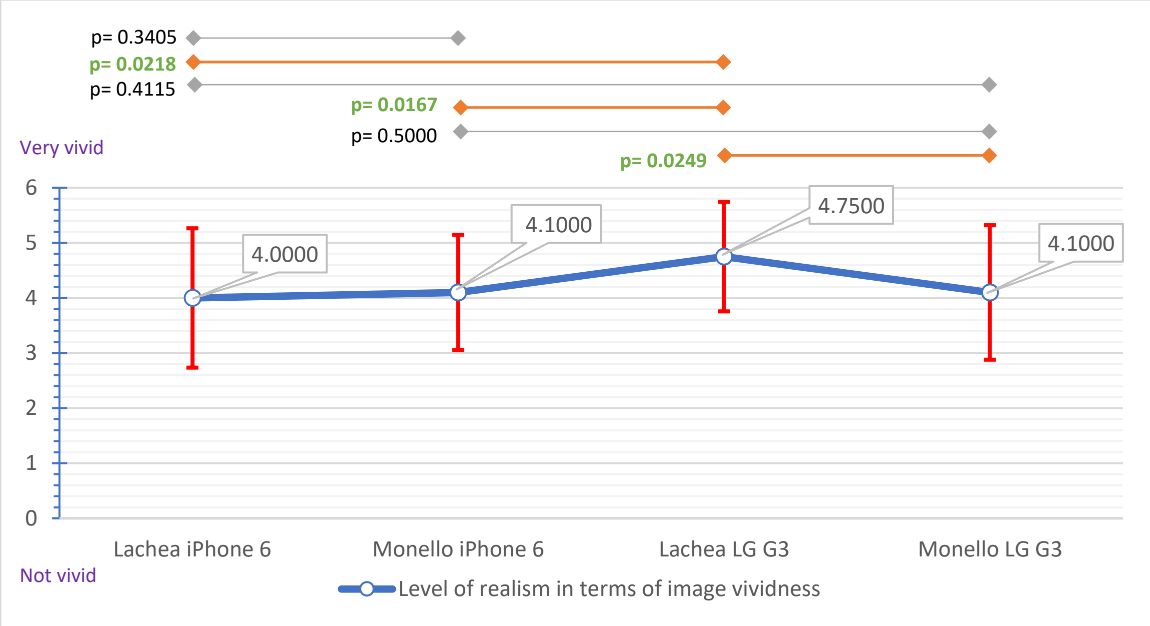


Figure 119 *- Three significant differences were found for realism in terms of image vividness (See Appendix B – question R2.10) [min = 0, Max = 6].*

Figure 119 reports several significant differences for image vividness.

Furthermore, the influence of display vividness over the achieved level of realism was questioned. Results are shown in **Figure 120**, which reports a significant difference between Monello iPhone 6 and Lachea LG G3.

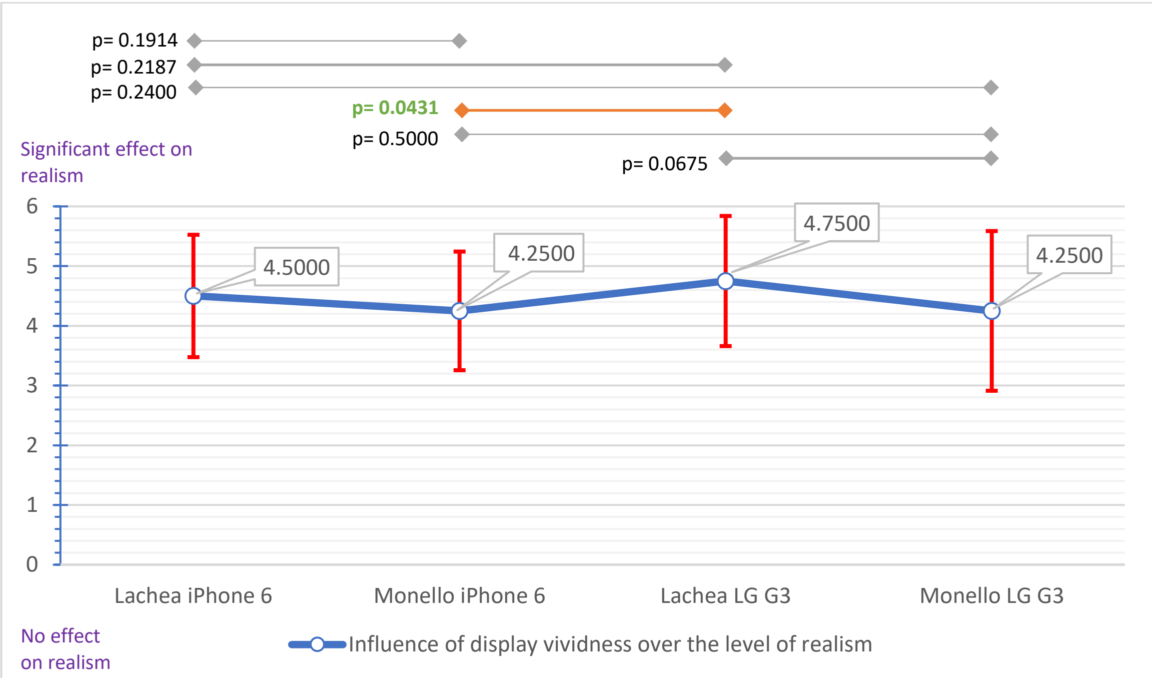


Figure 120 *- One significant difference was found for the influence of display vividness over the level of realism (See Appendix B – question R2.11) [min = 0, Max = 6].*

6.4.1.5 Level of realism in terms of image brightness

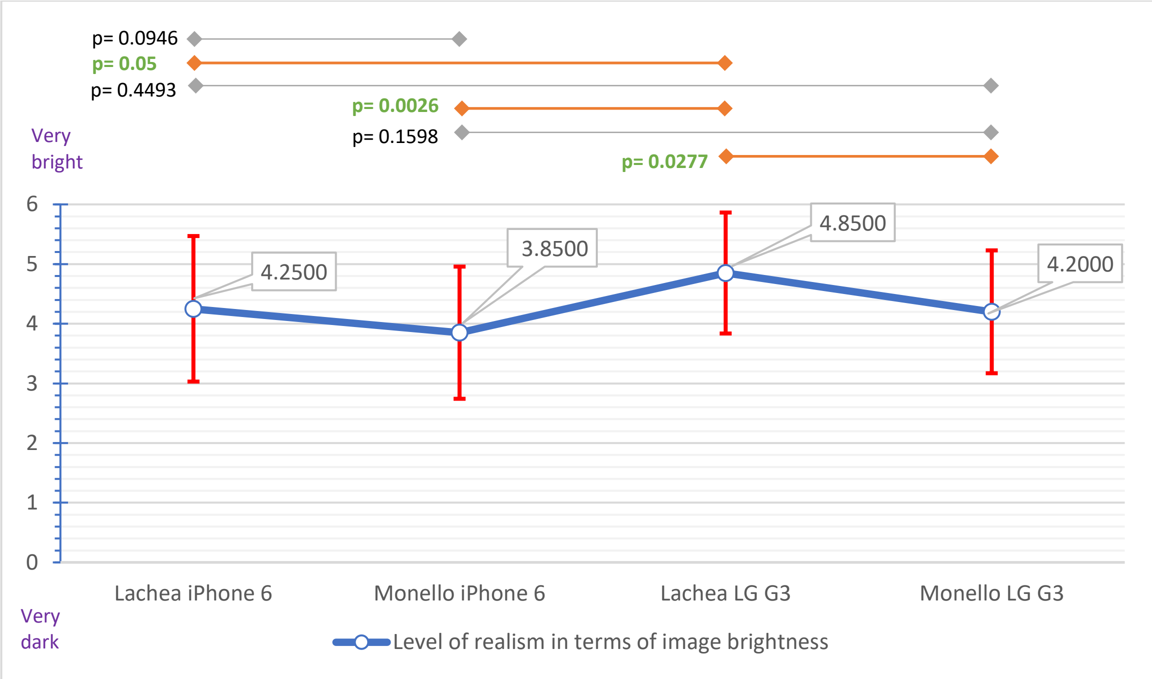


Figure 121 *- Three significant differences were found for realism in terms of image brightness (See Appendix B – question R2.12). [min = 0, Max = 6].*

Figure 121 presents several significant differences for image brightness. Furthermore, the influence of image brightness over the level of realism was investigated. Results are shown in **Figure 122**, which reports that Lachea LG G3 performed significantly better than Monello iPhone 6.

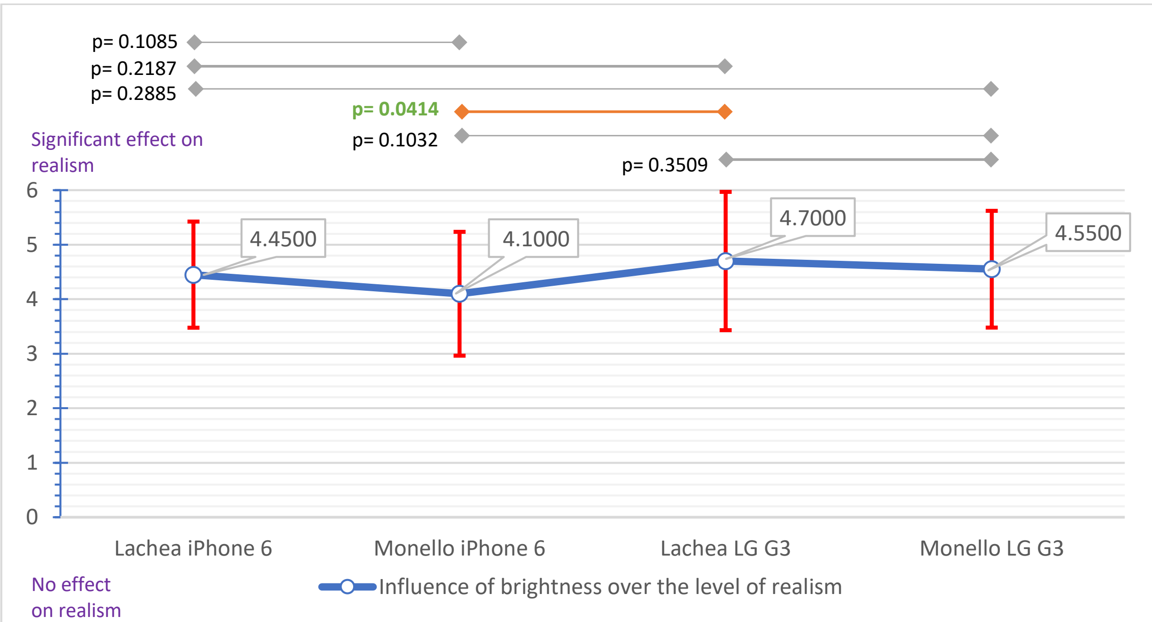


Figure 122 *- A significant difference was found for the influence of brightness over the level of realism (See Appendix B – question R2.13) [min = 0, Max = 6].*

6.4.1.6 Level of realism in terms of intensity of colours

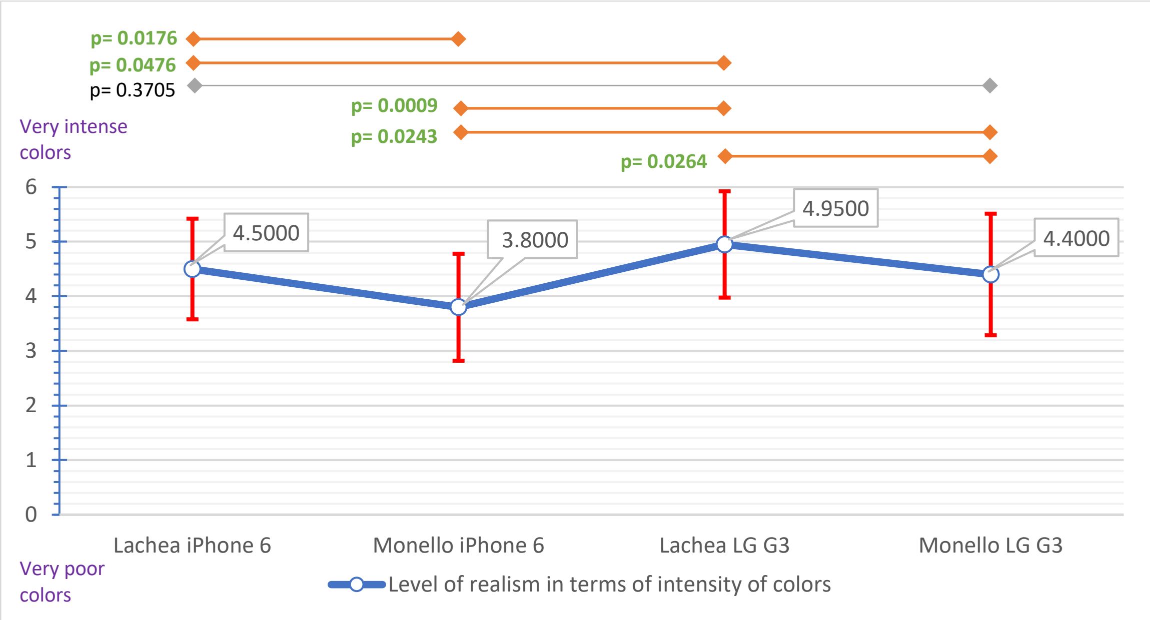


Figure 123 *- Five significant differences were found for level of realism in terms of color intensity (See Appendix B – question R2.14) [min = 0, Max = 6].*

Figure 123 reports that almost every coupling of phones and panoramas presents several significant differences for realism in terms of color intensity.

Furthermore, the influence of the intensity of colours over the level of realism was investigated. Results are shown in **Figure 124**. Only a significant difference was found between Monello iPhone 6 and Lachea LG G3, which reported higher scores.

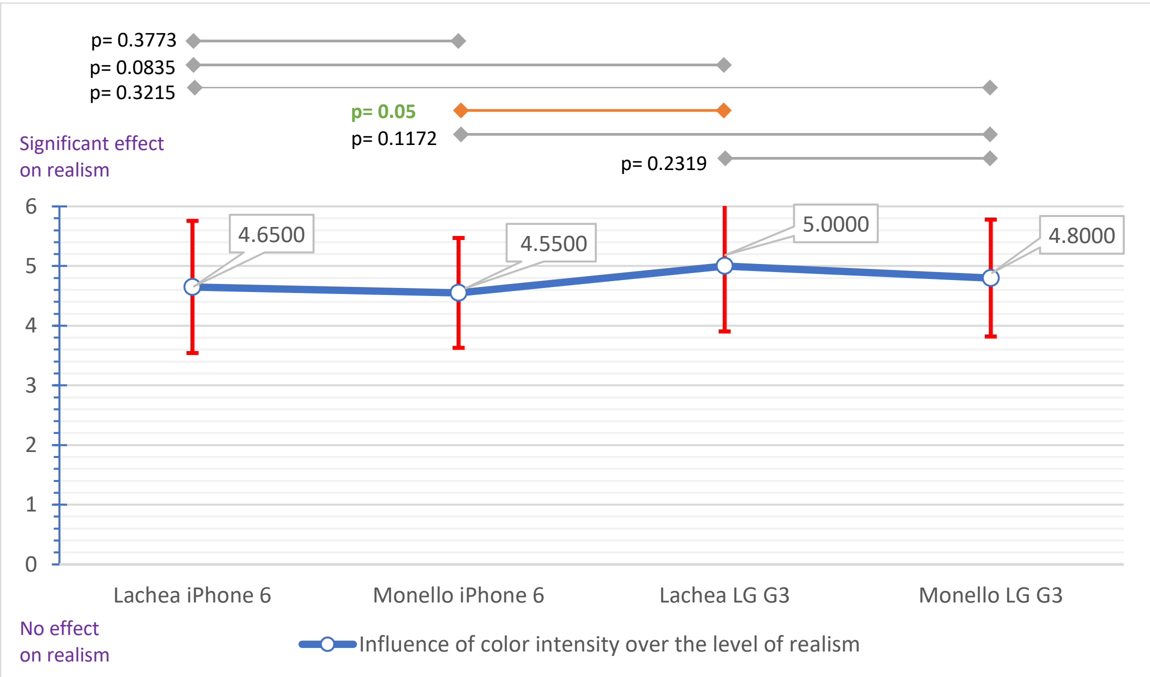


Figure 124 *- A significant difference was found for the influence of color intensity over the level of realism (See Appendix B – question R2.15) [min = 0, Max = 6].*

6.4.1.7 Level of realism in terms of contrast

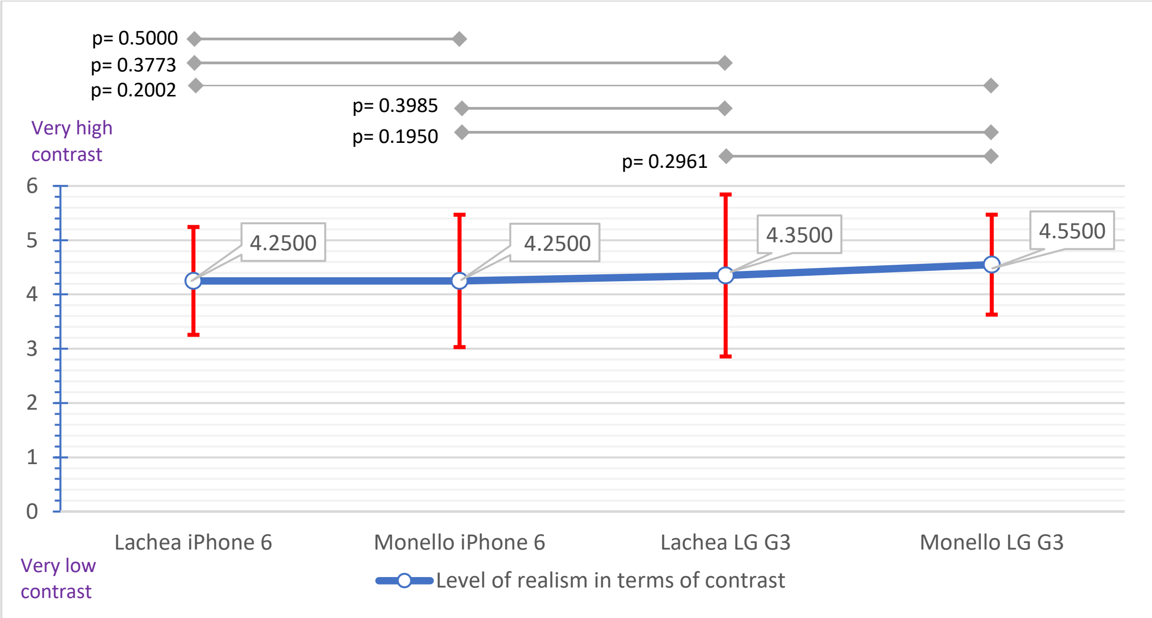


***Figure 125** - No significant difference was found for realism in terms of contrast (See Appendix B – question R2.16) [min = 0, Max = 6].*

Figure 125 reports no significant difference for image contrast. Furthermore, the influence of contrast over the level of realism was investigated and results shown in **Figure 126**, which reports several significant differences.

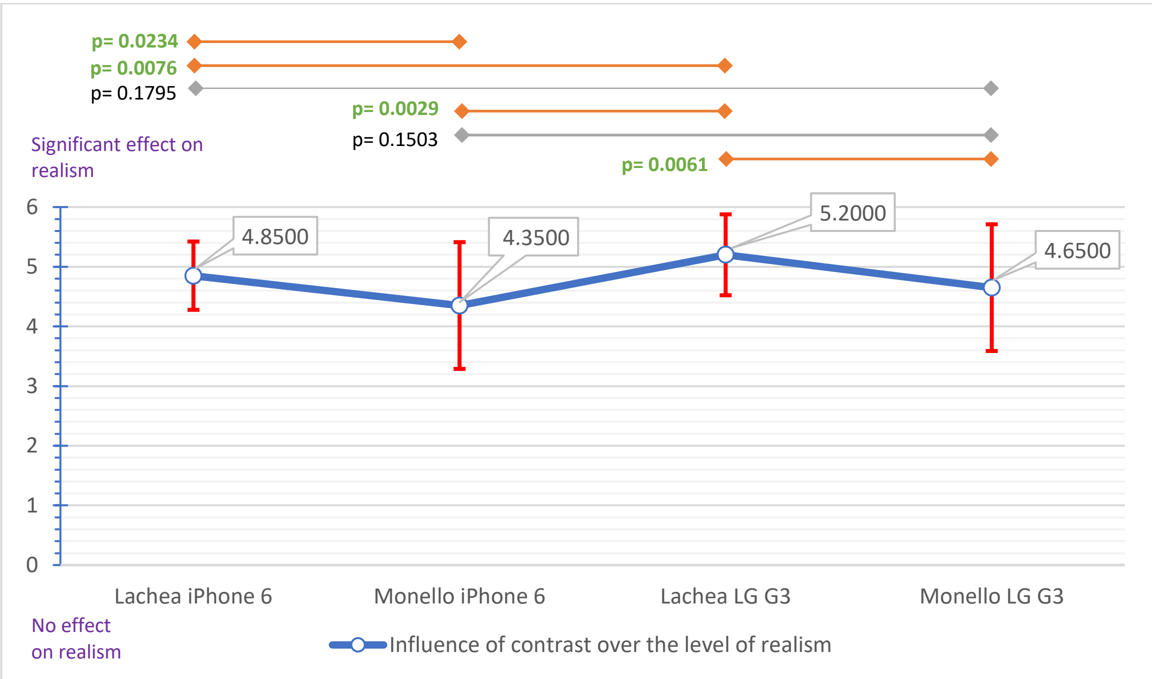


***Figure 126** - Four significant differences were found for the influence of contrast over the level of realism See Appendix B – question R2.17) [min = 0, Max = 6].*

6.4.1.8 Level of realism compared to real life

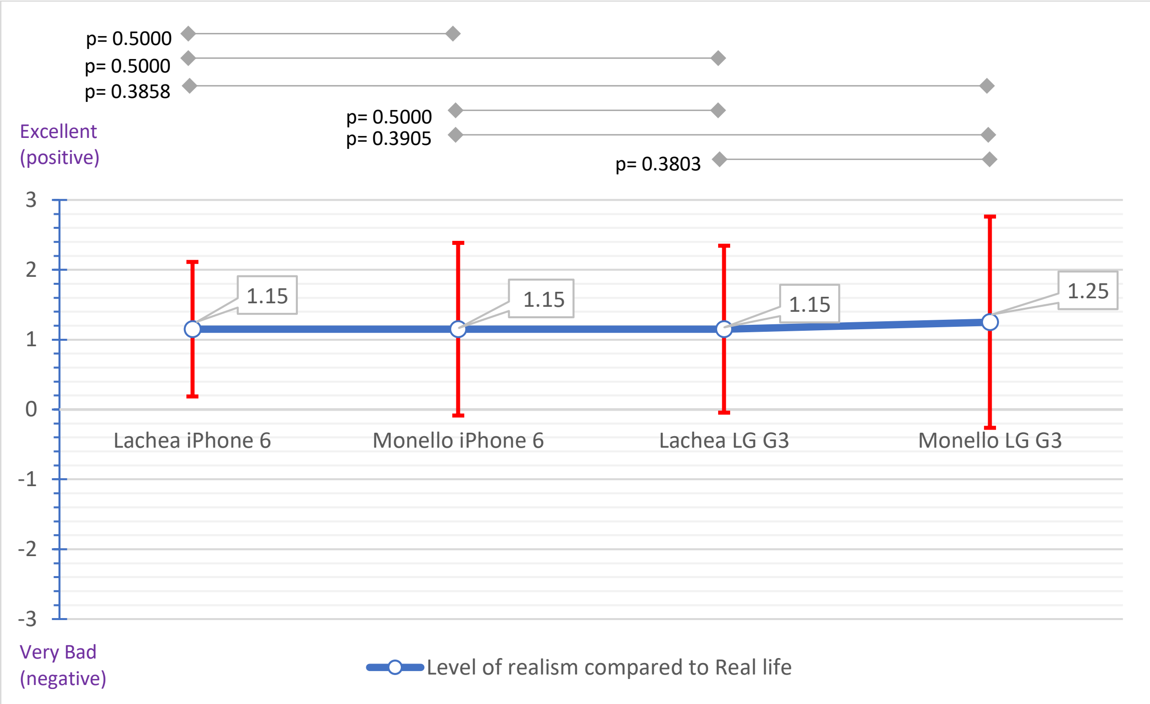


Figure 127 *- No significant difference resulted for realism compared to Real life (See Appendix B – question R2.1) [min = -3, Max = +3].*

Figure 127 reports no significant difference in terms of achieved realism compared to real life. All displays and panoramas achieved good scores.

6.4.1.9 Level of realism compared to photos and videos

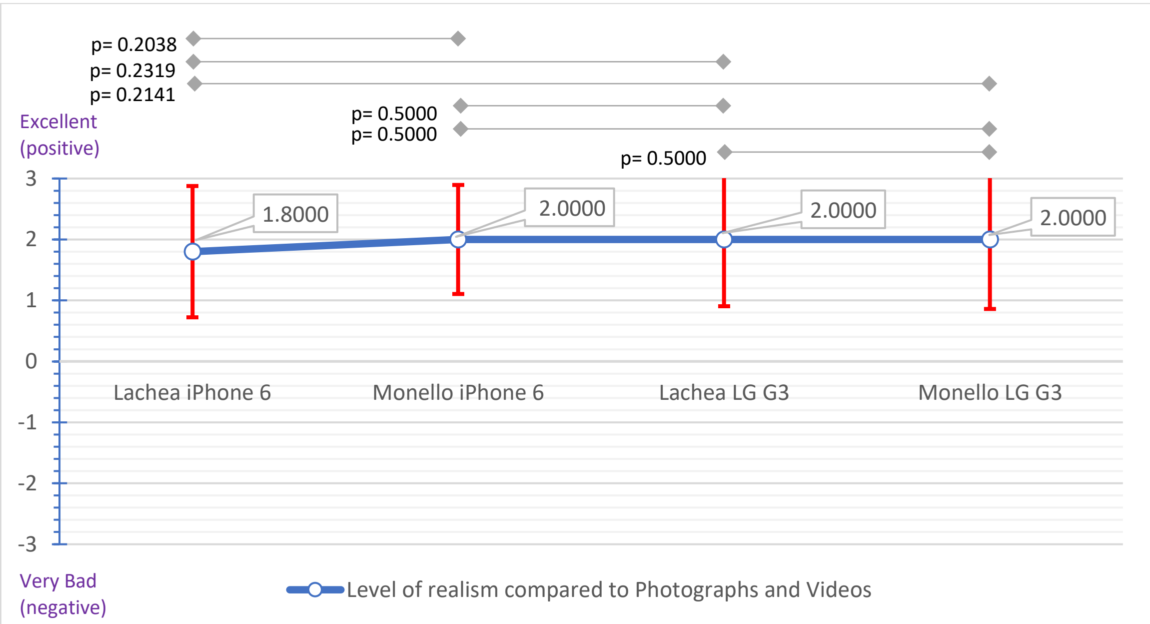


Figure 128 *- No significant difference in terms of realism compared to Photos and Videos (See Appendix B – question R2.2) [min = -3, Max = +3].*

Figure 128 reports no significant difference in terms of achieved realism compared to photos and videos. All panoramas and displays achieved high scores.

6.4.1.10 Level of realism in terms of objects deformations / natural elements on the scene

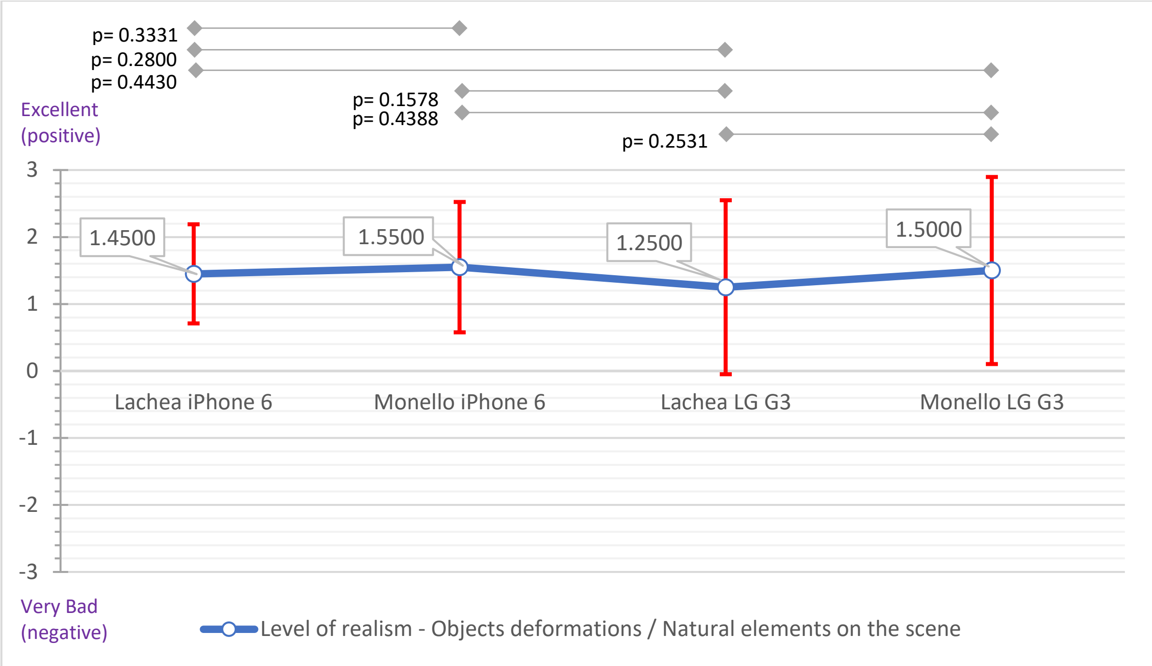


Figure 129 *- No significant difference in terms of objects deformation and natural elements of the scene (See Appendix B – question R2.3) [min = -3, Max = +3].*

Figure 129 reports no significant difference for realism in terms of objects deformation and natural elements on the scene. All displays and panoramas achieved good scores.

6.4.1.11 Realism of the 3D effect perceived in VR

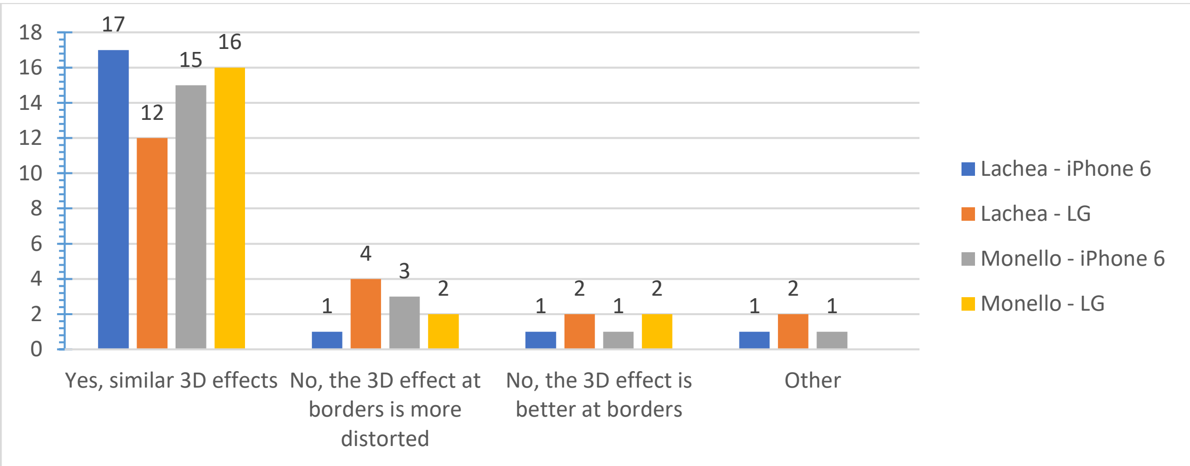


Figure 130 *- Number of test users choosing each answer to the question <Is the realism of the 3D effect perceived at the screen borders similar to that perceived at the screen centre?> (See Appendix B – question R1.1).*

According to **Figure 130** test users perceived similar 3D effects in the centre and borders of the screen with almost all setups excluding Lachea LG (40% of the 20 people disagreed), which induced slightly more distortions on the borders or on the centre of the screen. It can be assumed that when using a larger screen more distortions are introduced if the lenses are not adapting to the new size of the screen, reducing the level of realism achieved by the provided 3D effect.

6.4.1.12 Level of realism in terms of emotions

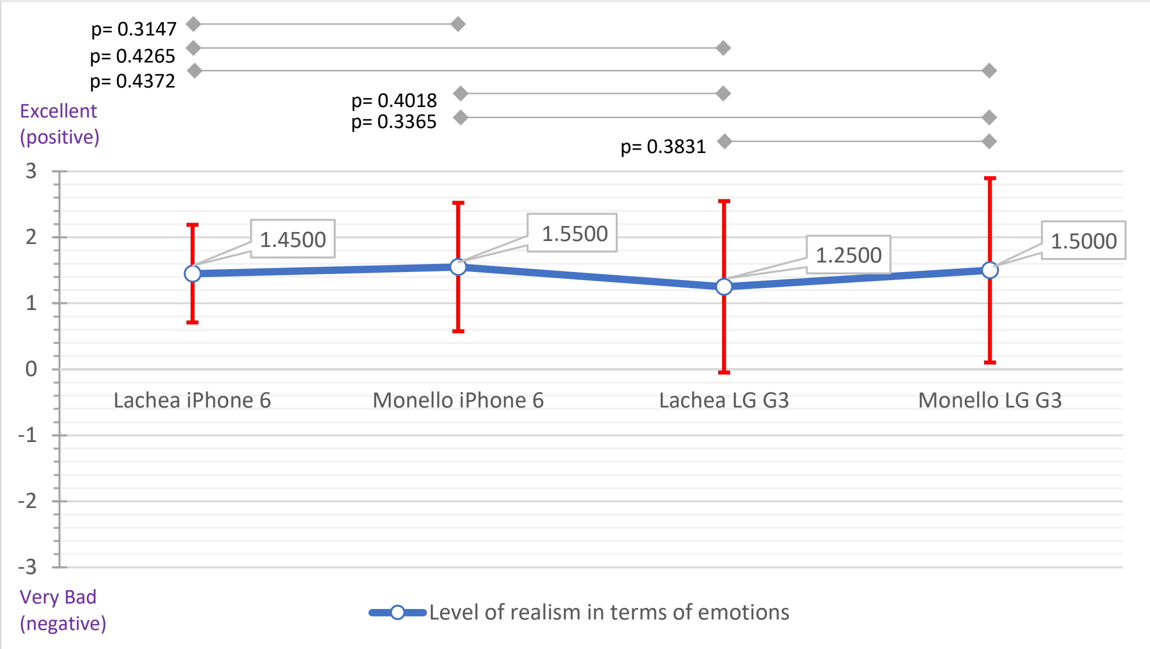


Figure 131 *- No significant difference resulted for realism in terms of emotions (See Appendix B – question R2.5) [min = -3, Max = +3].*

Figure 131 reports no significant differences for realism in terms of emotions, with all phones and panoramas achieving good scores.

6.4.1.13 Influence of colours over perceived emotions

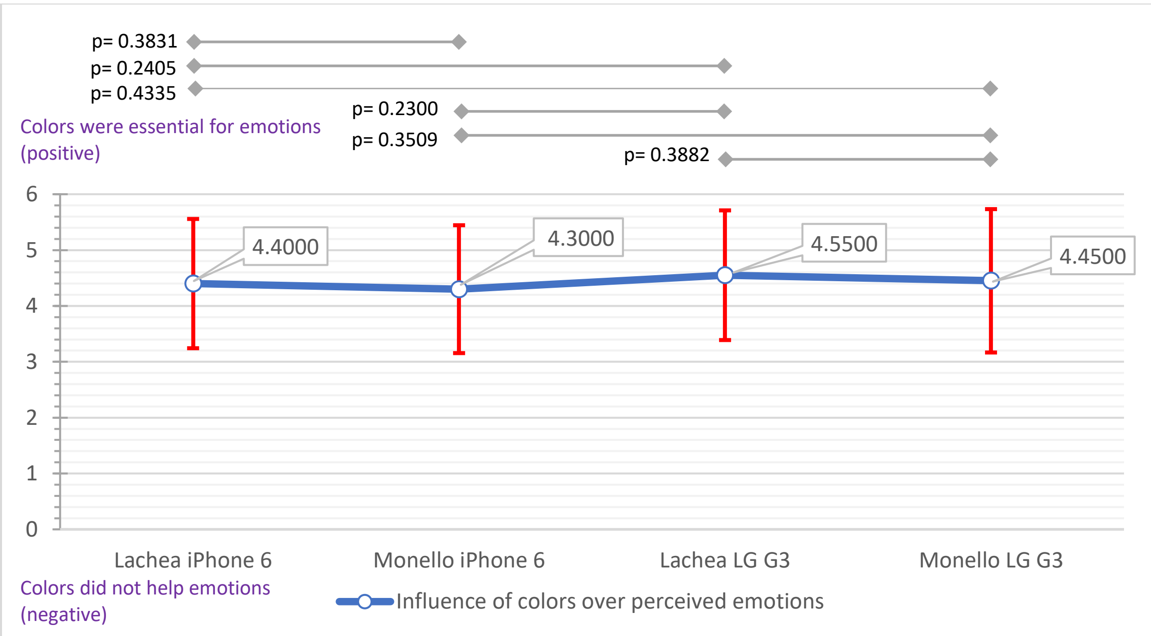


Figure 132 *- No significant difference was found for the influence of colours over perceived emotions (See Appendix B – question R2.18) [min = 0, Max = 6].*

Figure 132 reports no significant difference for the influence of colours over perceived emotions. All panoramas and phones achieved quite high scores.

6.4.1.14 Influence of horizontal size of display over Realism

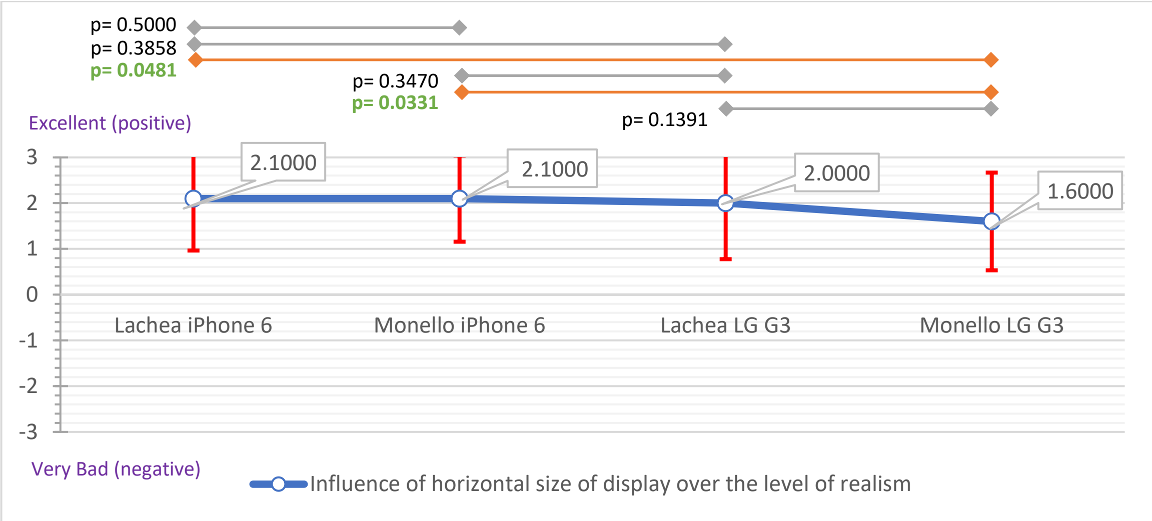


Figure 133 *- Two significant differences were found for the influence of horizontal size of display over realism (See Appendix B – question R2.4) [min = -3, Max = +3].*

Figure 133 shows several significant differences for the effect of horizontal display size on perceived level of realism.

6.4.2 Comfort

6.4.2.1 Discomfort

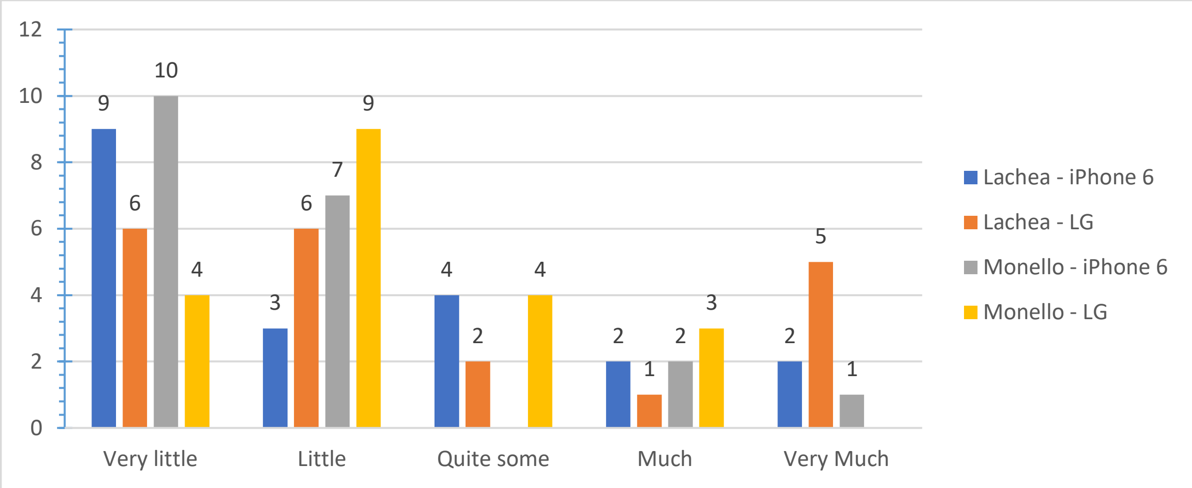


Figure 134 *- Number of test users choosing each answer to the question <Did you experience discomfort during 3D viewing of the panorama?> (See Appendix B – question C2.2).*

In **Figure 134** both Lachea (45%) and Monello (50%) with iPhone 6 caused “very little” discomfort compared to the other setups. Furthermore, Monello LG was judged with “little” discomfort scores (45%). The most uncomfortable setup was Lachea LG (25% of test users answered “very much” discomfort).

6.4.2.2 Eye strain or visual fatigue while watching panorama

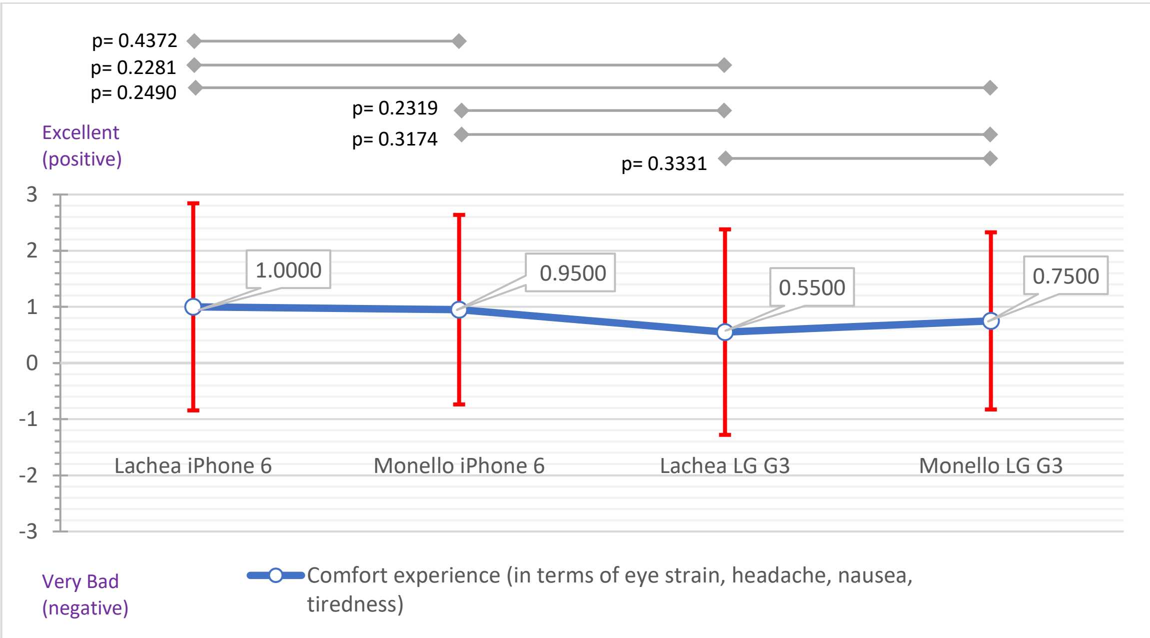


Figure 135 *- No significant difference was found for comfort experience, which was considered quite good (See Appendix B – question C2.2) [min = -3, Max = +3].*

In **Figure 135** all phones and panoramas achieved a similar comfort level in terms of eye strain, headache, nausea, and tiredness.

6.4.2.3 Max duration of time during which watching the panorama remains comfortable

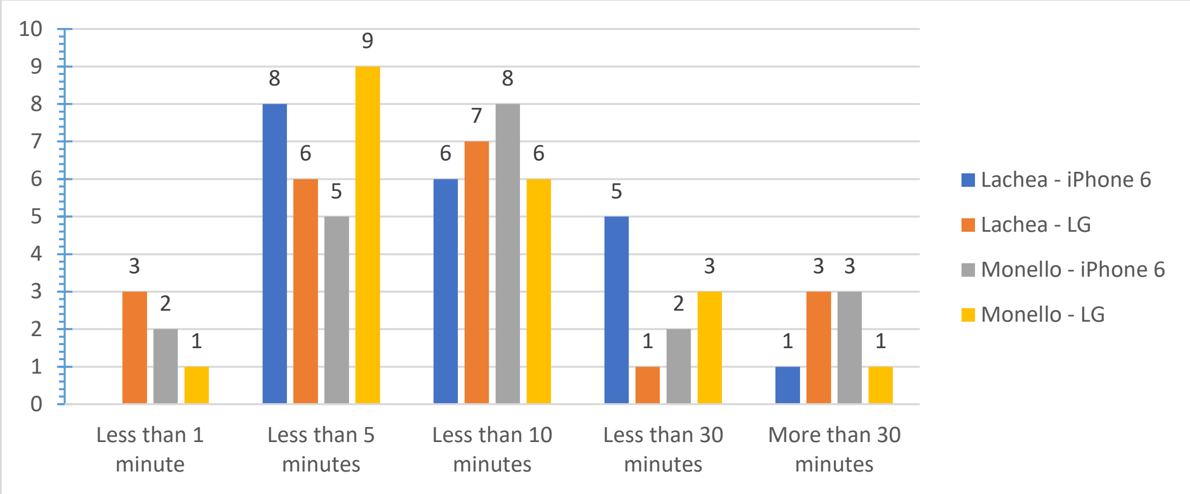


Figure 136 *- Number of test users choosing each answer to the question <How long do you think you can watch the 3D panorama before getting tired?> (See Appendix B – question C1.2).*

Figure 136 reports that test users think that Lachea iPhone 6 can be watched for more than 1 minute without discomfort (0% answered "less than 1 minute"). This shows that Lachea with iPhone 6 provided one of the best results in terms of how long people can watch the panorama before experiencing discomfort. Furthermore, the average answer for all VR setups ranged between 1 and 10 minutes.

6.4.3 Presence

6.4.3.1 Perceived sense of "presence"

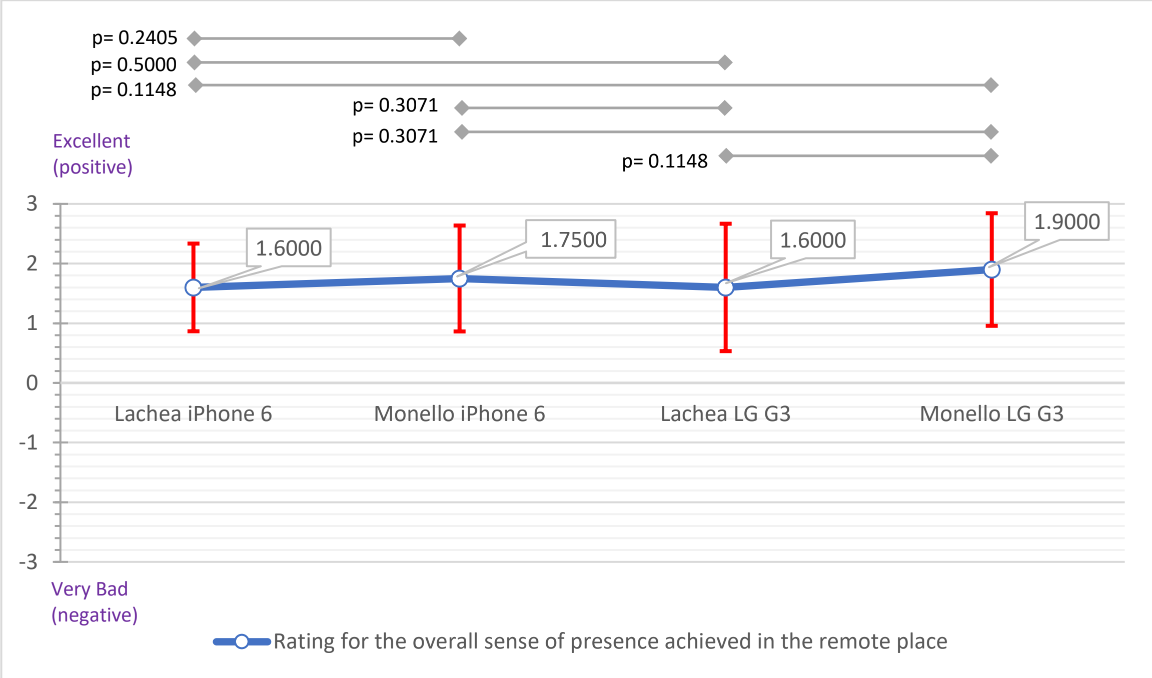


Figure 137 *- No significant difference was found on the overall sense of presence, which was considered very good (See Appendix B – question P2.2) [min = -3, Max = +3].*

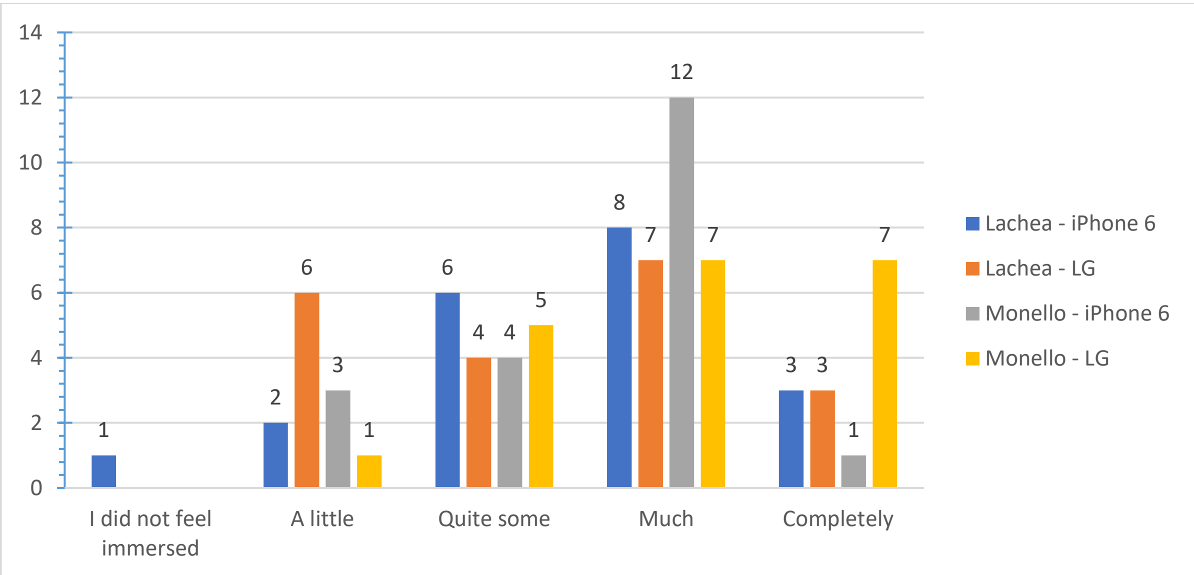


Figure 138 *- Number of test users grouped by their answer to the question <How much did you feel "immersed" into the panorama?> (See Appendix B – question P1.1).*

In literature some papers refer to "immersion" as the objective quantifiable measure related to the characteristics of the VR system [330]. However, for this specific question the term "immersion" was used to refer to the subjective "feeling of presence" in the Virtual Environment and was explained to test users before answering.

6.4.3.2 Presence and Tunnel vision

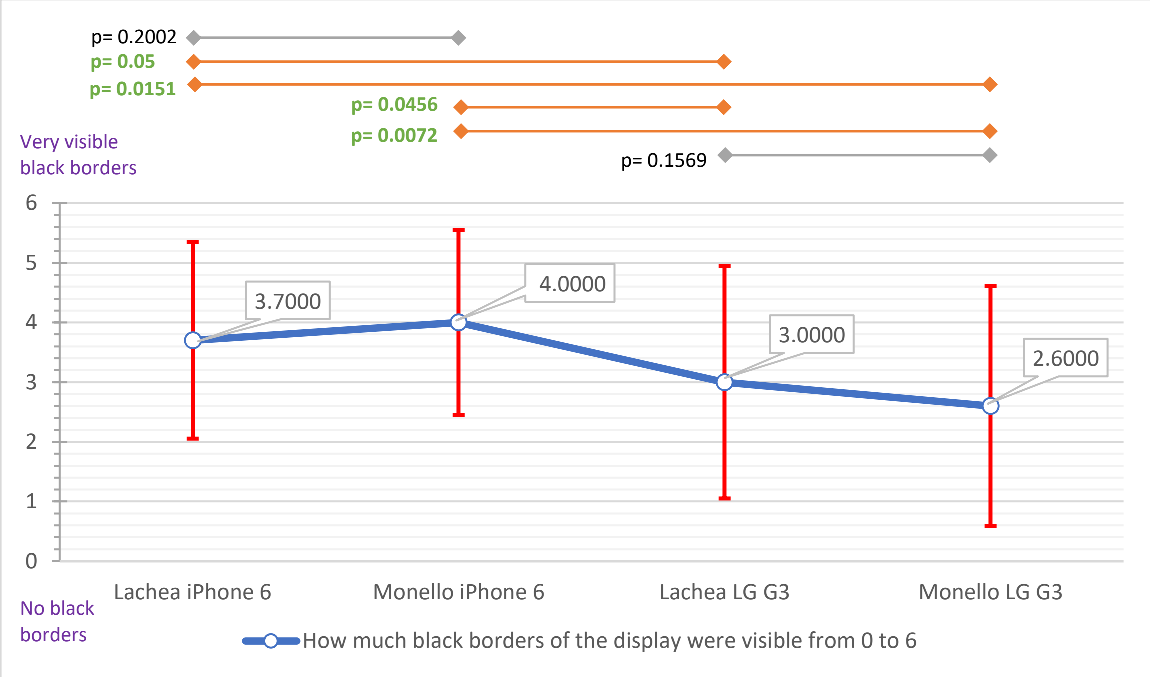


Figure 139 - Several significant differences were found for tunnel vision (See Appendix B – question P2.3) [min = 0, Max = +6].

Figure 139 reports significant differences on how much the black borders of the screen were visible while viewing the panorama.

Furthermore, test users were asked to evaluate the influence of the tunnel vision over the achieved sense of presence. Results are presented in **Figure 140**.

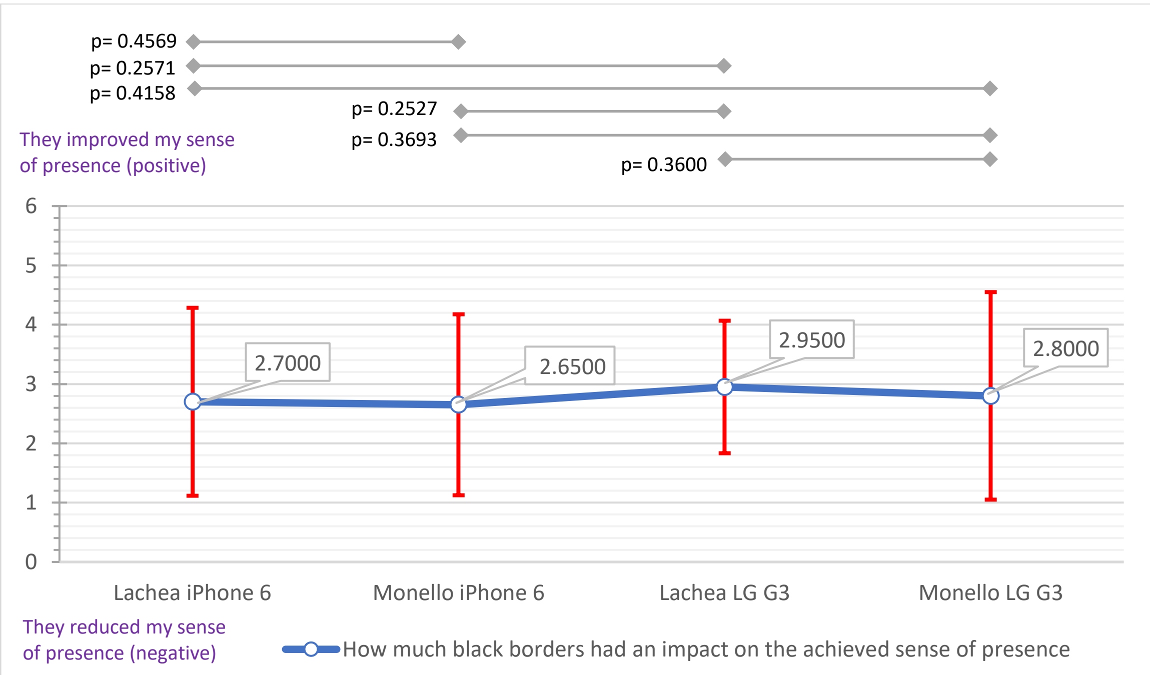


Figure 140 - No significant difference was found on the influence of tunnel vision over sense of presence (See Appendix B – question P2.4) [min = 0, Max = +6].

In **Figure 140** all the averages of the VR setups shown a neutral influence of tunnel vision over the achieved sense of presence. It can be assumed that **despite the difference in the tunnel vision between the two mobile phones the sense of presence was not altered.**

6.4.3.3 Isolation

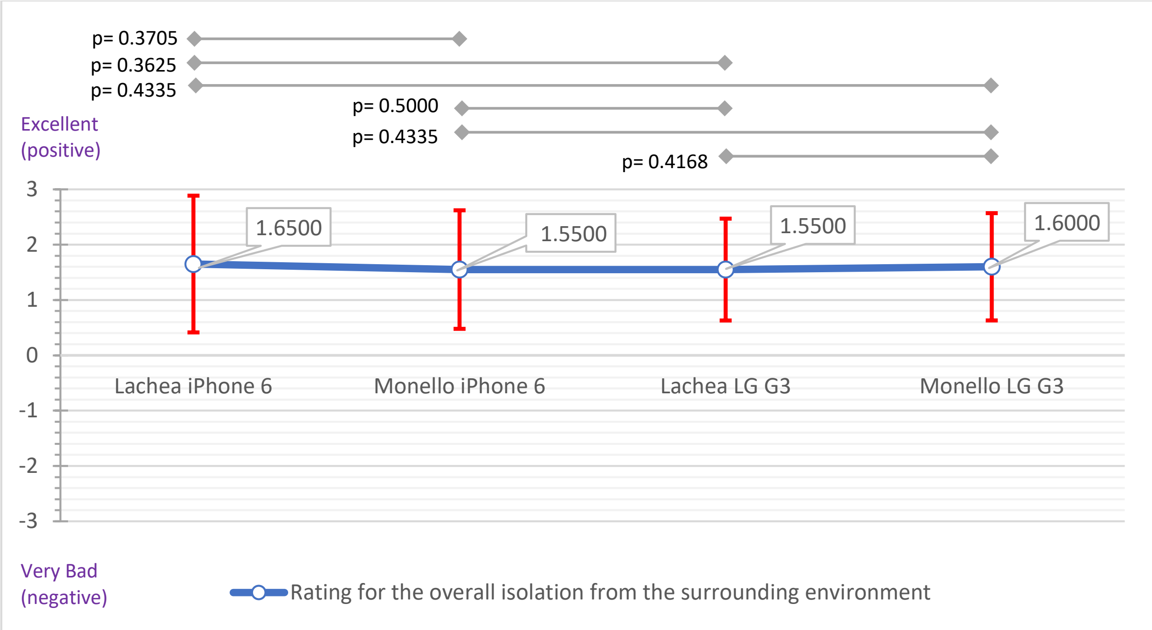


Figure 141 *- No significant difference was found in terms of isolation, which was considered very good with any phone and panorama (See Appendix B – question P2.2) [min = -3, Max = +3].*

6.4.3.4 Influence of isolation over emotions

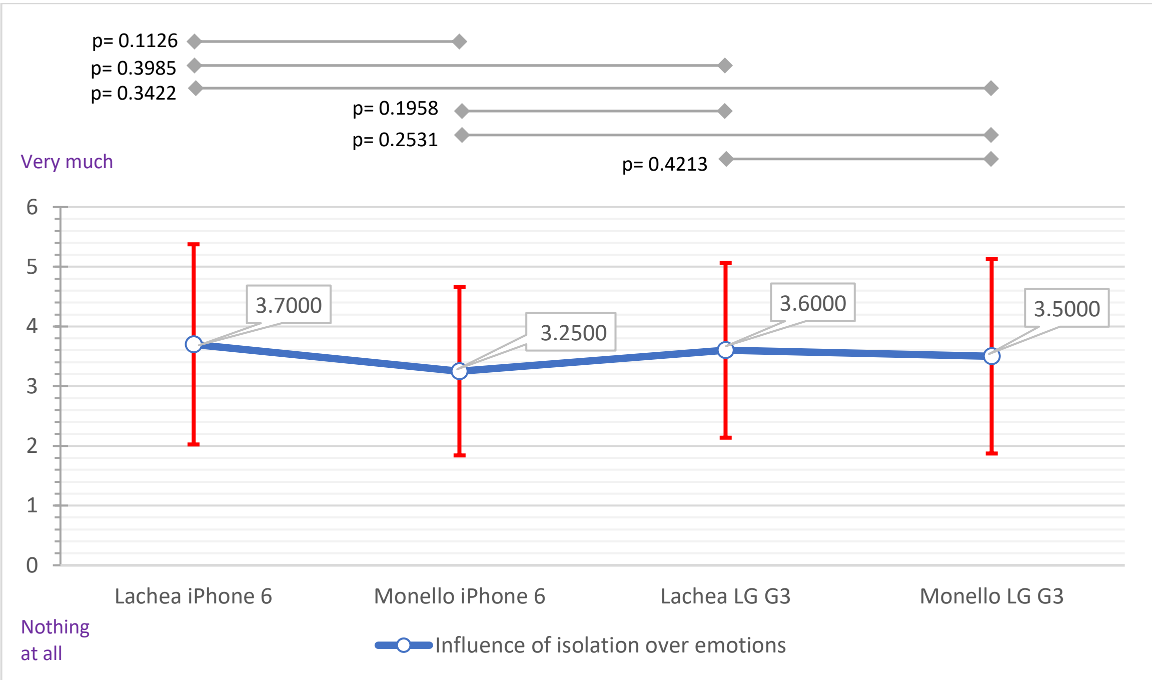


Figure 142 *- No significant difference was found for the influence of isolation over emotions, which was considered quite high (See Appendix B – question P2.12) [min = 0, Max = +6].*

6.4.3.5 Induced emotions from the panorama visualization

In the following, emotions elicited by viewing the VR panoramas have been analysed.

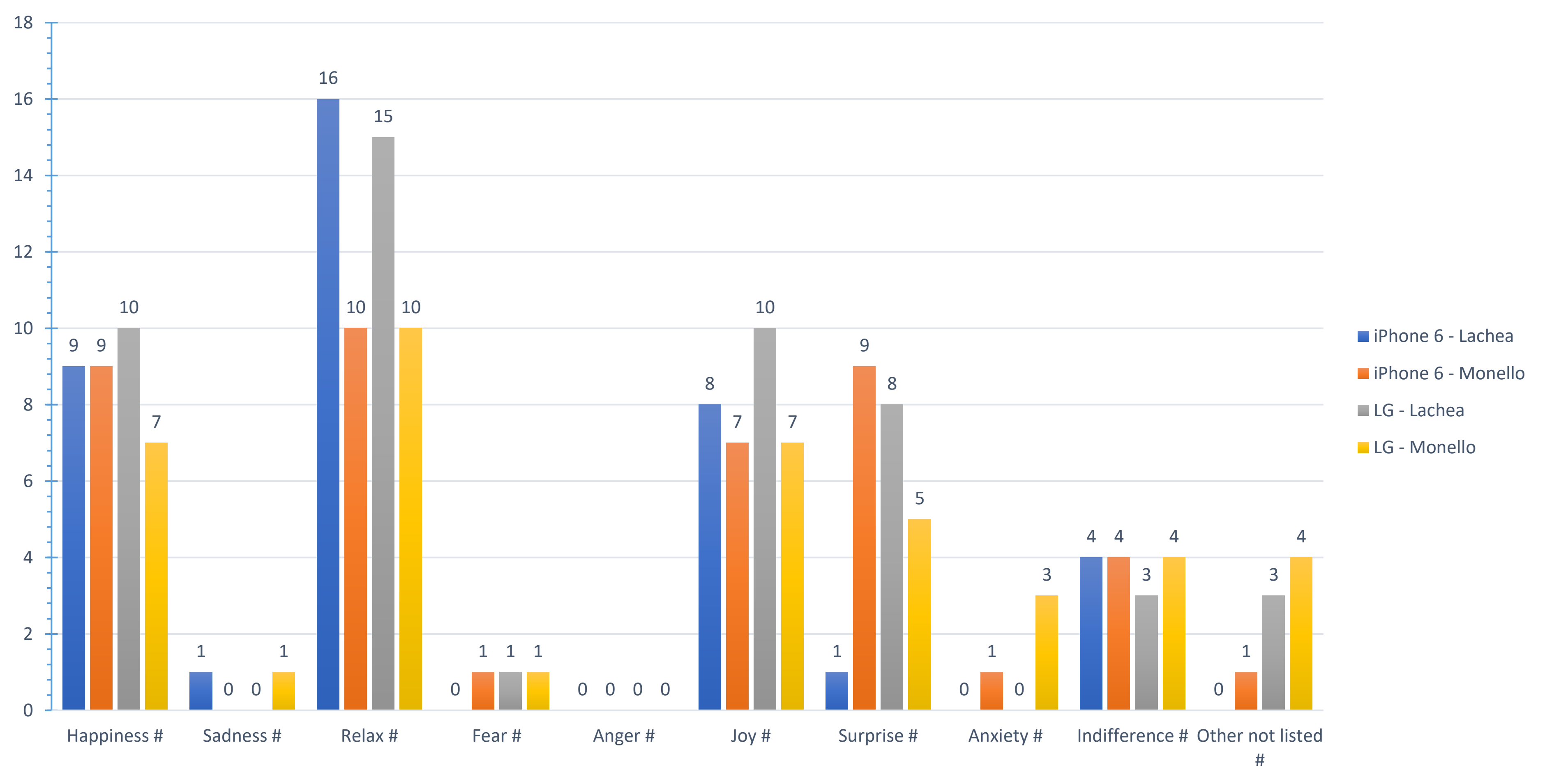


Figure 143 *- Emotions provided by each panorama according to phone used. For each emotion, the number shown identifies how many test users encountered that emotional state when watching the panorama with each specified setup (See Appendix B – question P1.2).*

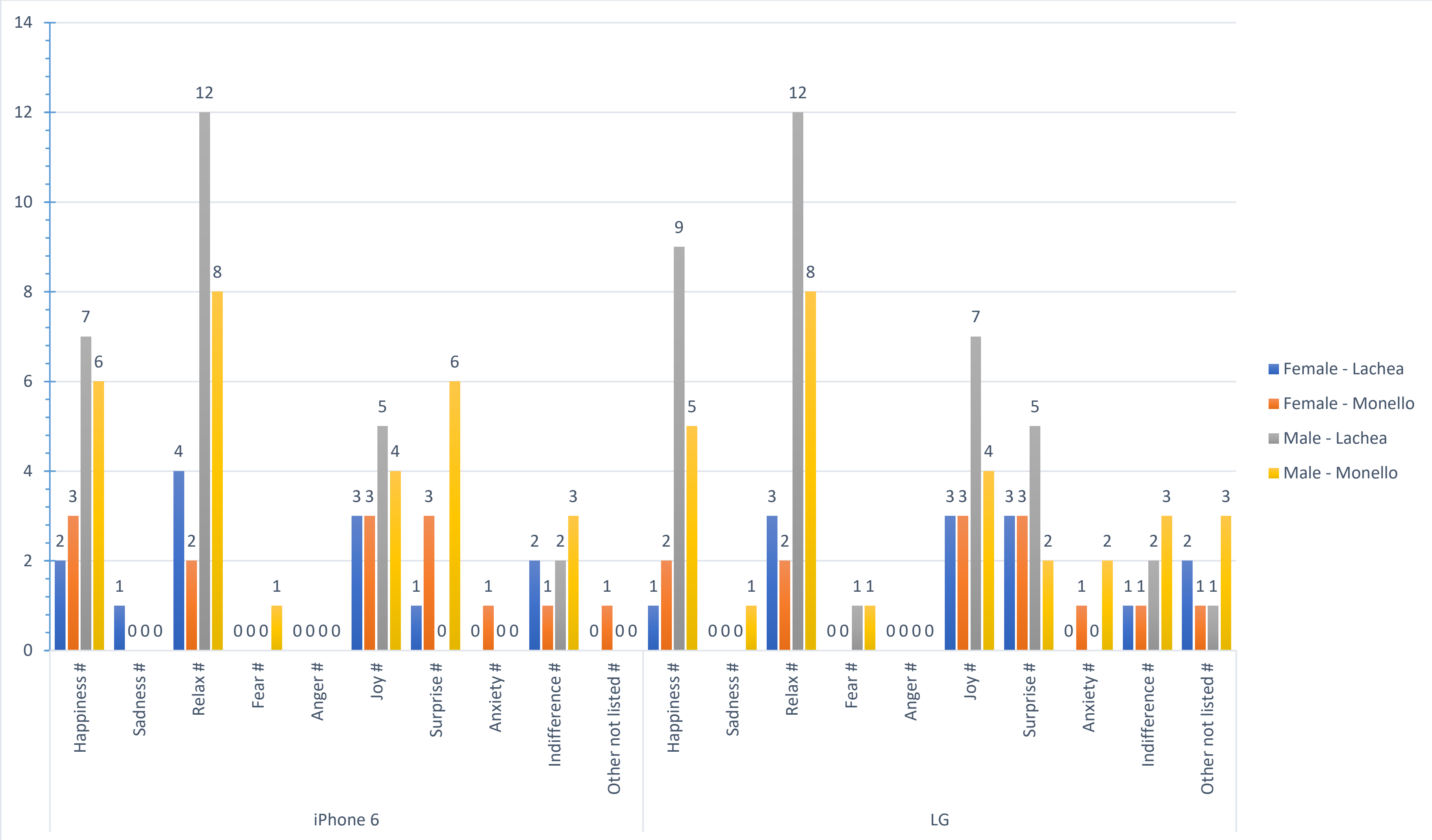


Figure 144 *- Emotions provided by each panorama according to phone used and gender of test user. For each emotion, the number shown identifies how many test users encountered that emotional state when watching the panorama with each specified setup (See Appendix B – question P1.2).*

Among results, **Figure 143** reports that no panorama induced “anger”, while an extremely low number of test users experienced “fear” or “sadness”. According to **Figure 144** the male sample (70% of test users) was more emotional compared to the female sample (30% of test users).

Furthermore, I asked to test users several questions on how much each of them felt a specific emotion, using the 6-Likart scale. The following are the results.

6.4.3.5.1 Positive / Negative overall emotions

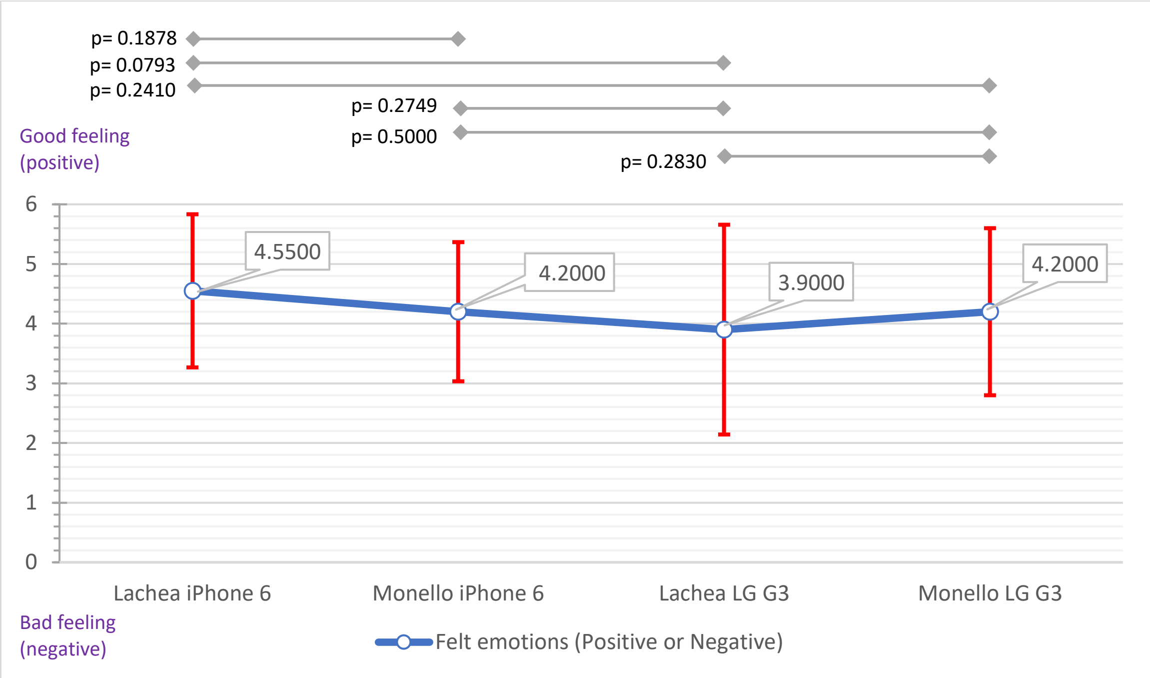


Figure 145 *- No significant difference was found in terms of felt emotions, which were perceived as very good feelings (See Appendix B – question P2.5) [min = 0, Max = +6].*

6.4.3.5.2 Enjoyment

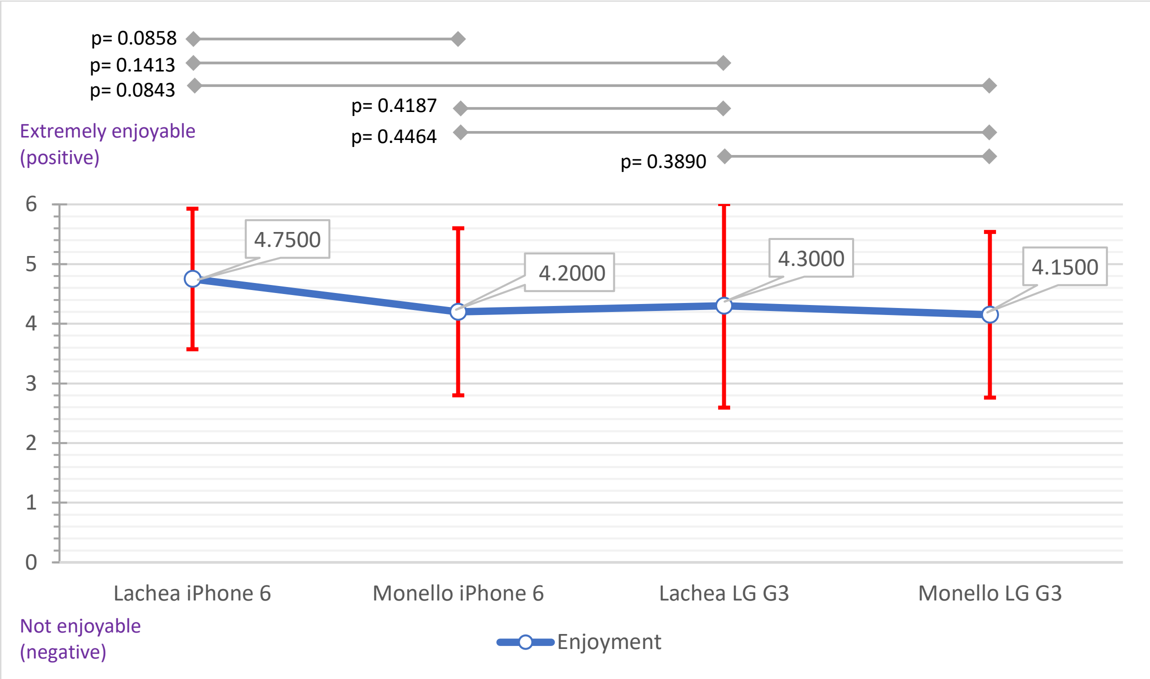


Figure 146 *- No significant difference was found in terms of enjoyment, which was considered very high (See Appendix B – question P2.6) [min = 0, Max = +6].*

6.4.3.5.3 Happiness

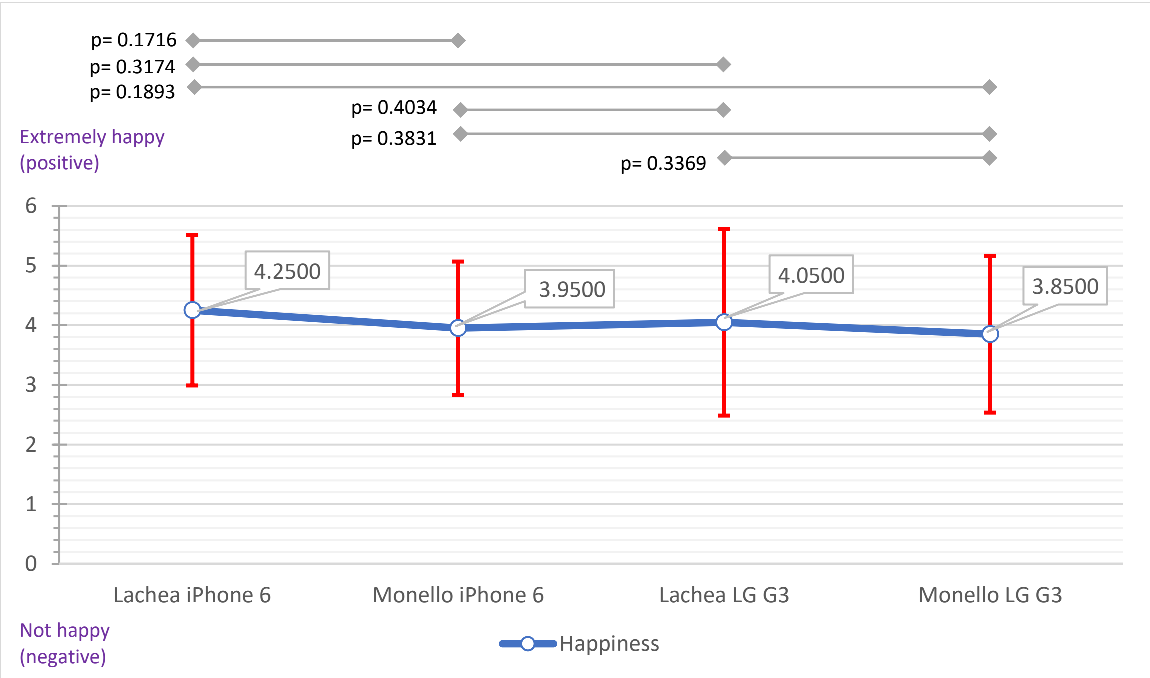


Figure 147 *- No significant difference was found in terms of happiness, which was considered very high (See Appendix B – question P2.7) [min = 0, Max = +6].*

6.4.3.5.4 Sadness

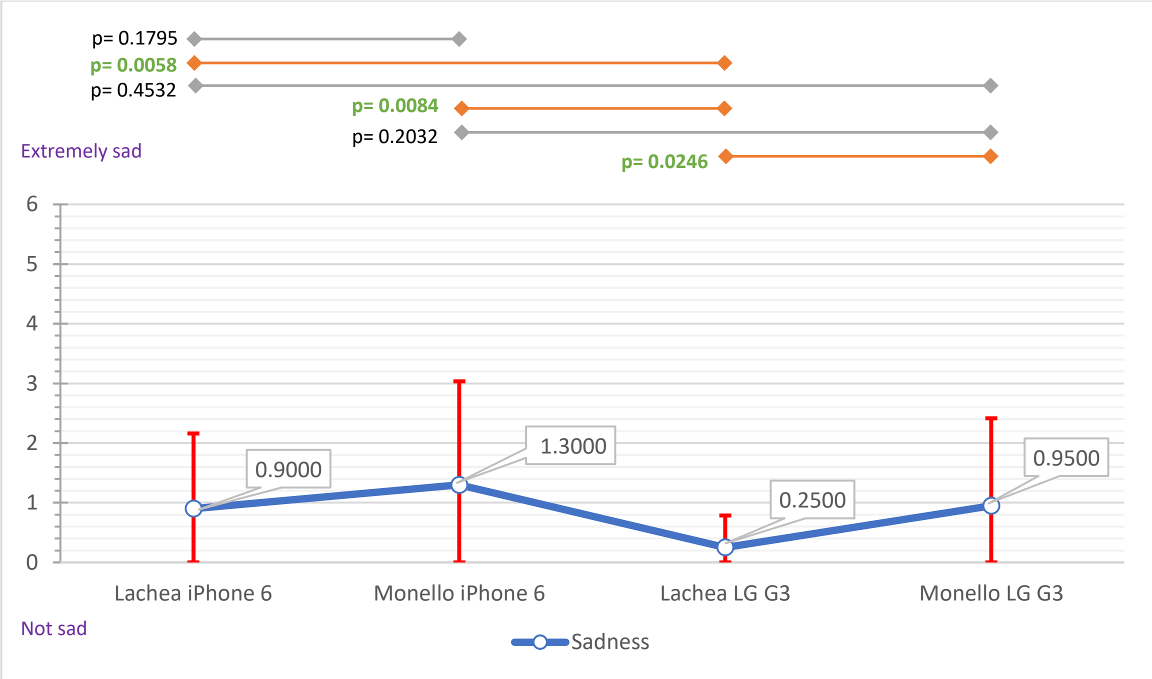


Figure 148 *- Three significant differences were found in terms of perceived sadness, which was considered very low (See Appendix B – question P2.8) [min = 0, Max = +6].*

6.4.3.5.5 Scariness

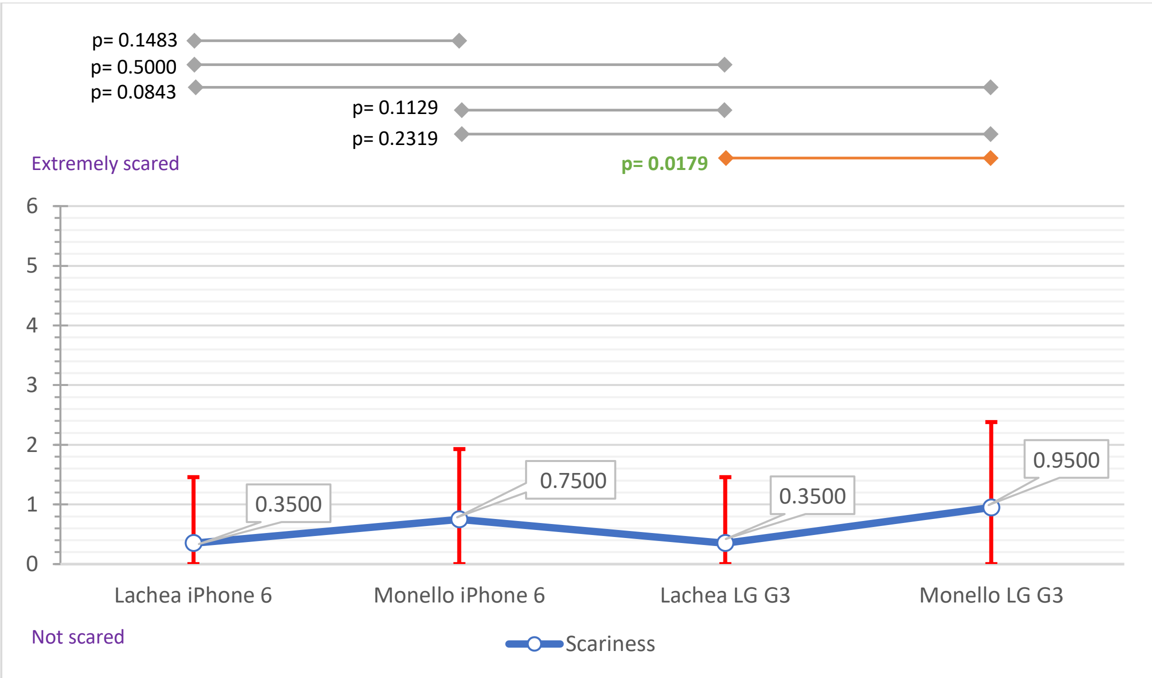


Figure 149 *- One significant difference was found in terms of perceived scariness, which was considered very low (See Appendix B – question P2.9) [min = 0, Max = +6].*

6.4.3.5.6 Relax

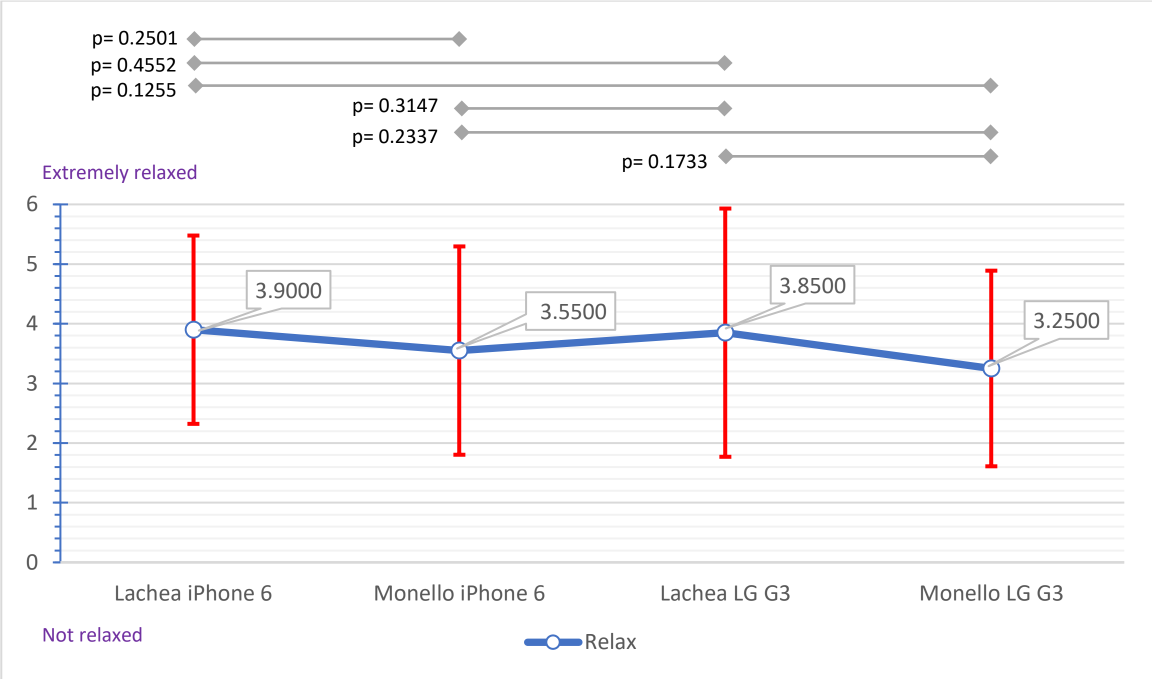


Figure 150 *- No significant difference was found in terms of perceived relax, which was considered high (See Appendix B – question P2.10) [min = 0, Max = +6].*

6.4.3.5.7 Feeling of like going on holiday

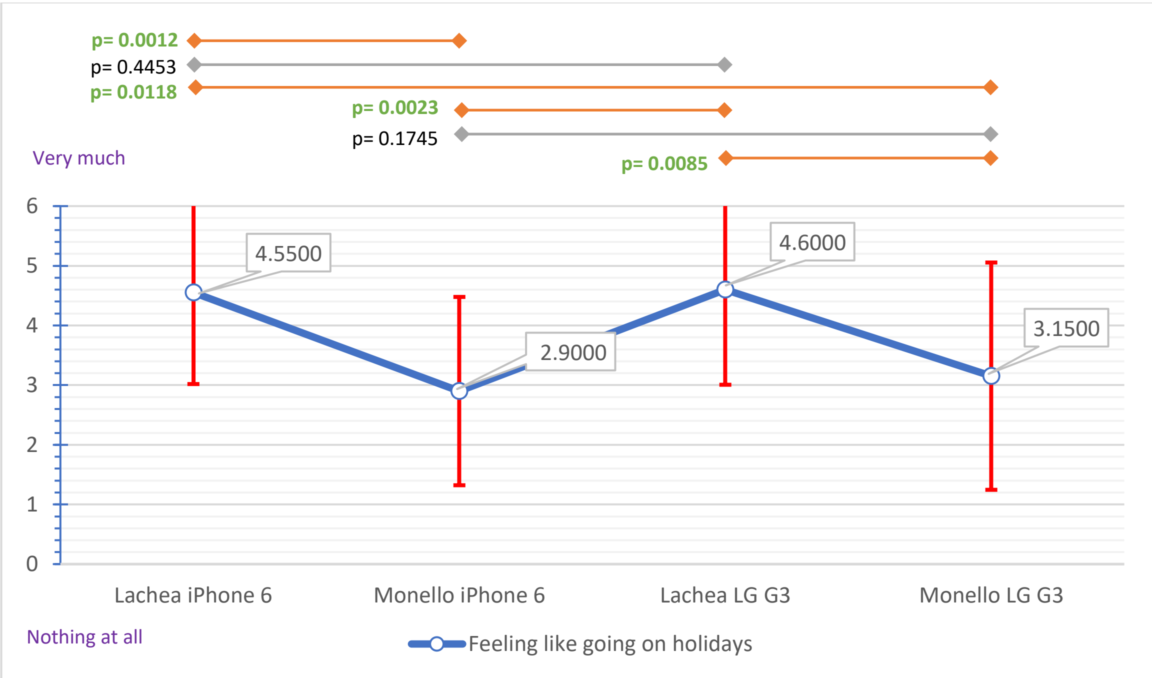


Figure 151 *- Four significant differences were found in terms of feeling of like going on holiday (See Appendix B – question P2.11) [min = 0, Max = +6].*

6.4.3.6 Influence of emotions over the sense of presence

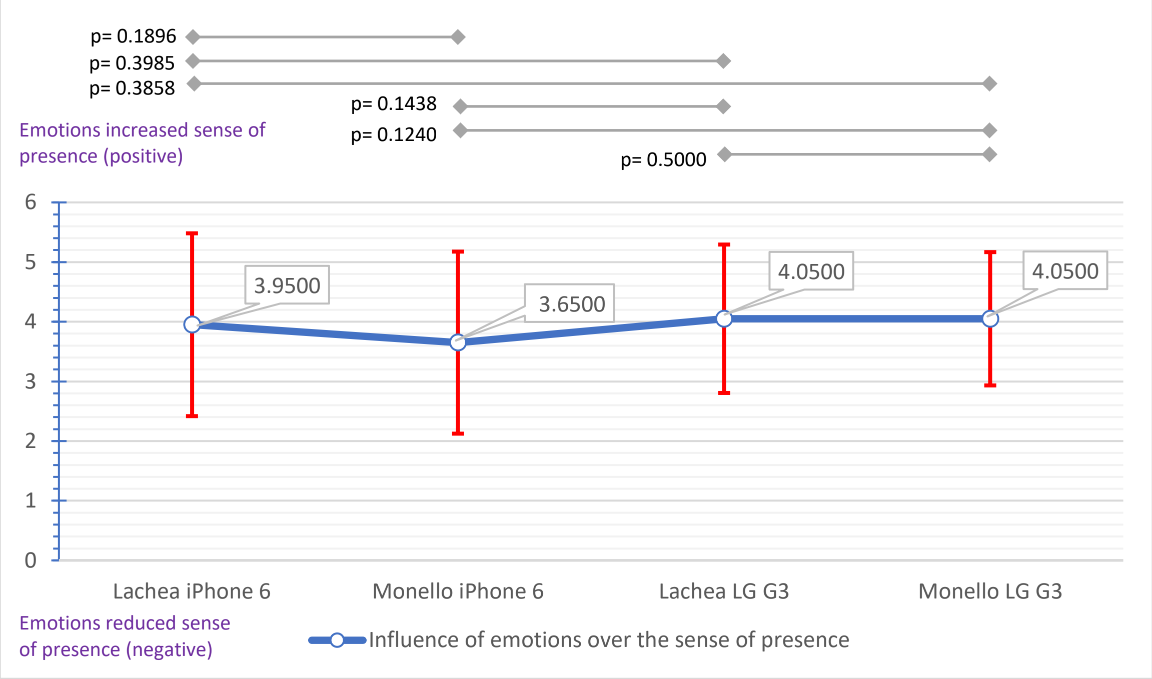


Figure 152 *- No significant difference was found in terms of influence of emotions (which resulted high) over the achieved sense of presence (See Appendix B – question P2.13) [min = 0, Max = +6].*

6.4.3.7 Emotions caused by coming back to real life after the VR experience

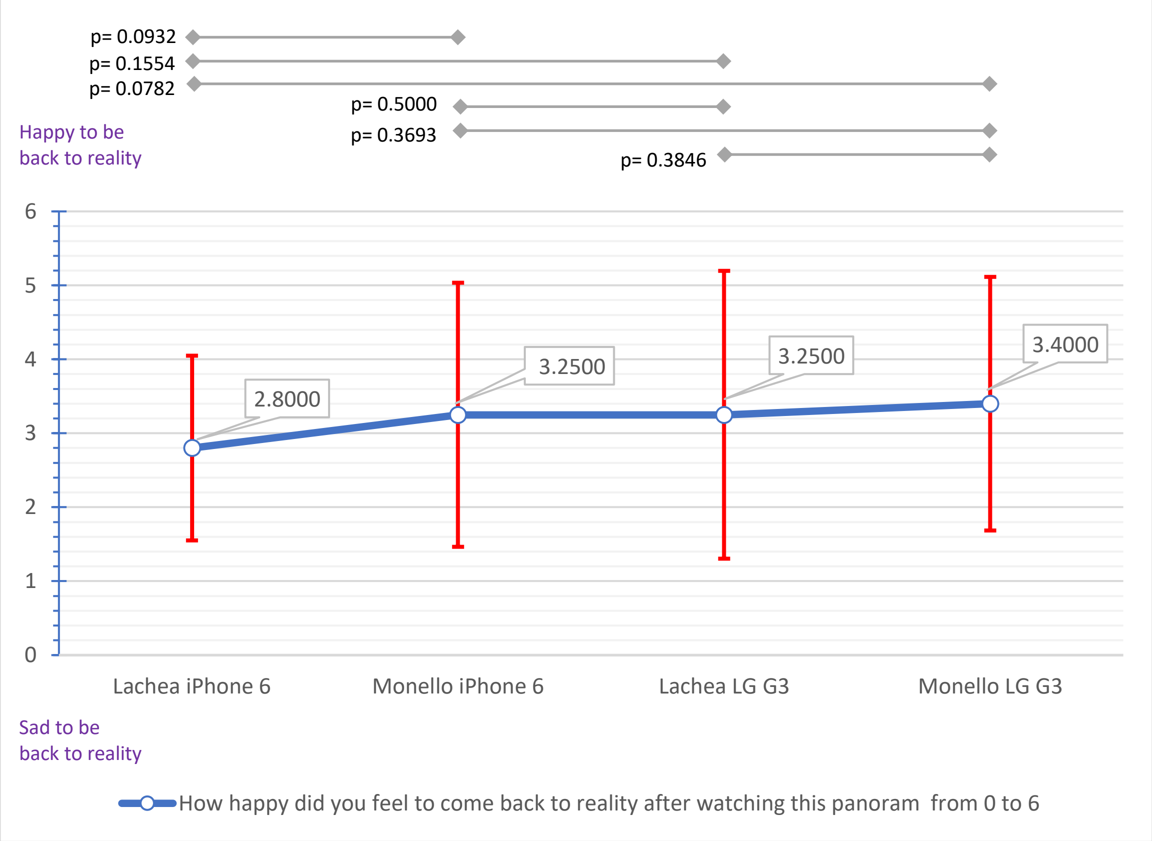


Figure 153 *- No significant difference was found in terms of emotions caused by coming back to real life after the VR experience (See Appendix B – question P2.14) [min = 0, Max = +6].*

6.4.4 Depth perception

6.4.4.1 3D Depth impression

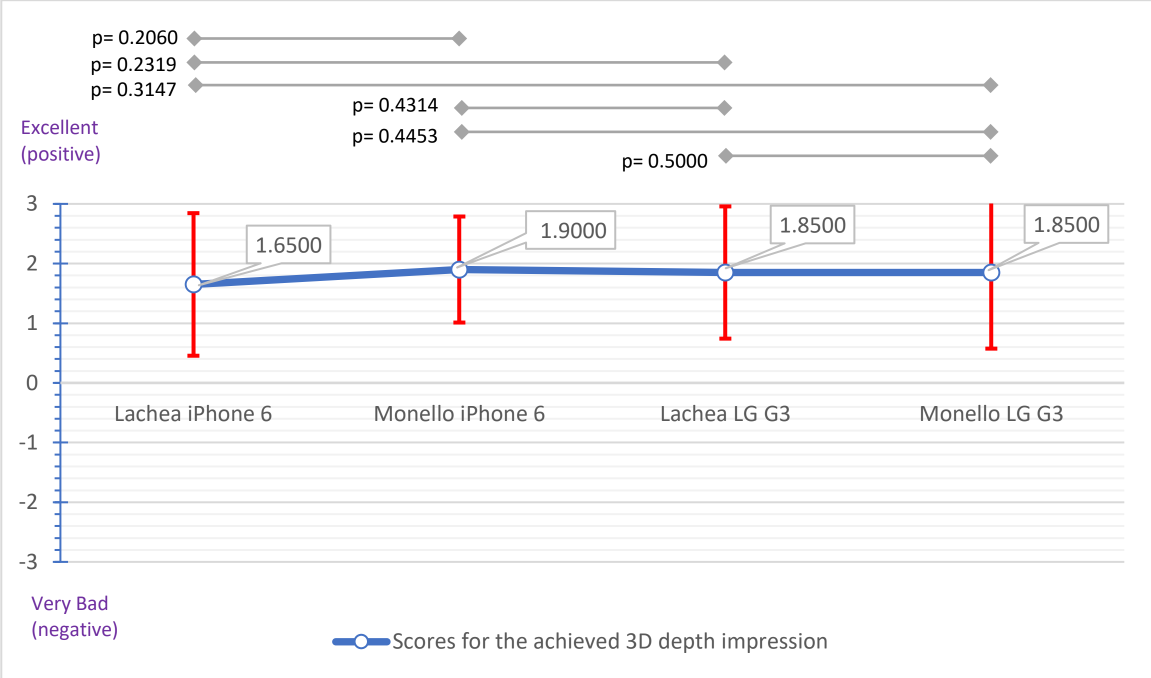


Figure 154 *- No significant difference was found in terms of achieved 3D depth impression, which was considered very good (See Appendix B – question D2.1) [min = -3, Max = +3].*

6.4.4.2 Distance perception

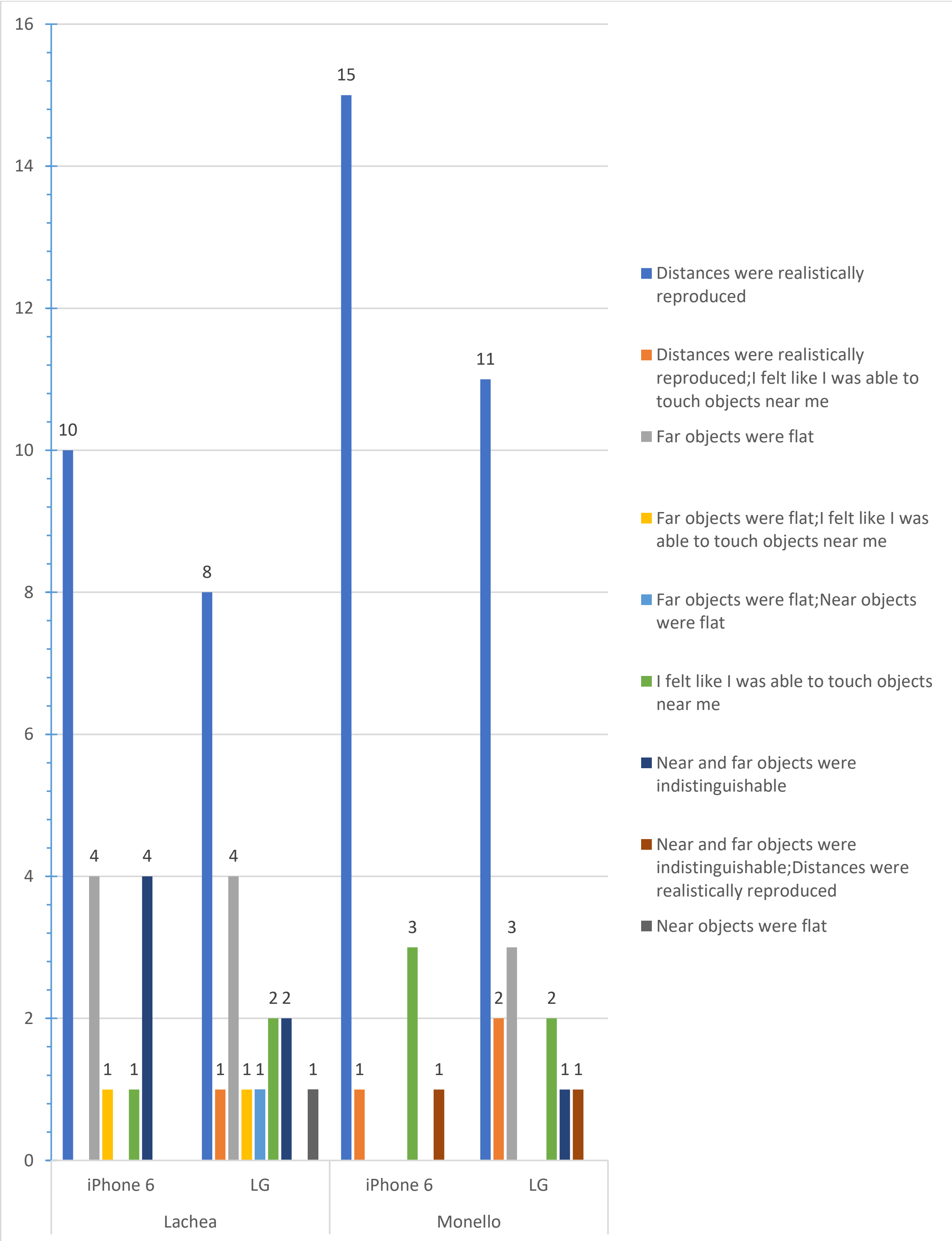


Figure 155 *- Number of test users grouped by their answer to the question <How did you perceive distances in 3D? > (See Appendix B – question D1.5).*

6.4.4.3 Color contribution to 3D perception

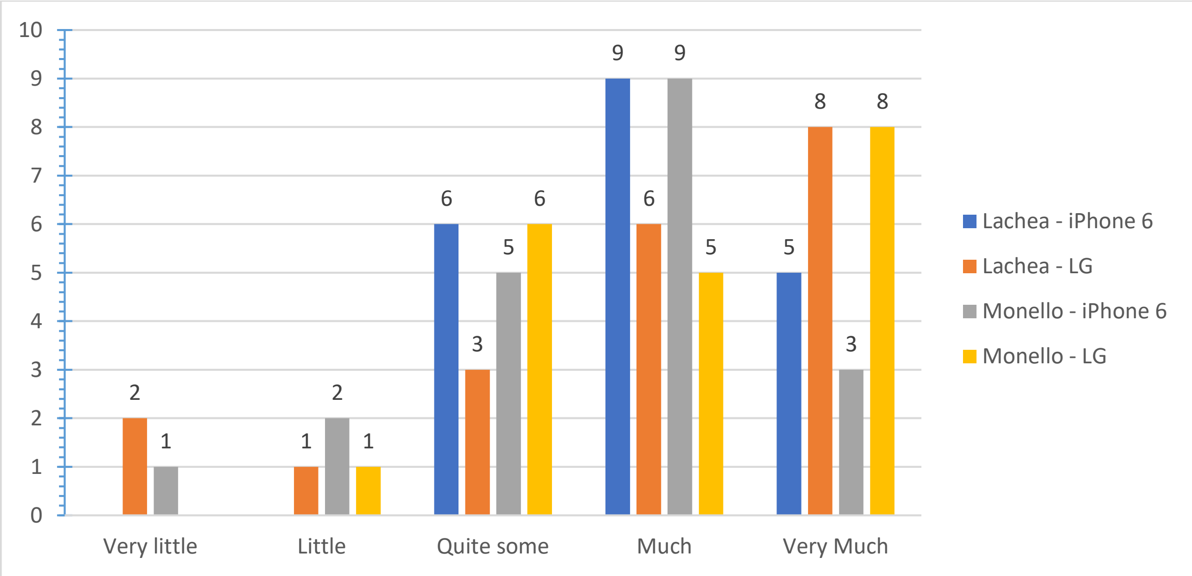


Figure 156 *- Number of test users grouped by their answer to the question <How much do you think that colours have contributed to the perception of your 3D?> (See Appendix B – question D1.3).*

6.4.4.4 Lights and shadows contribution to 3D perception

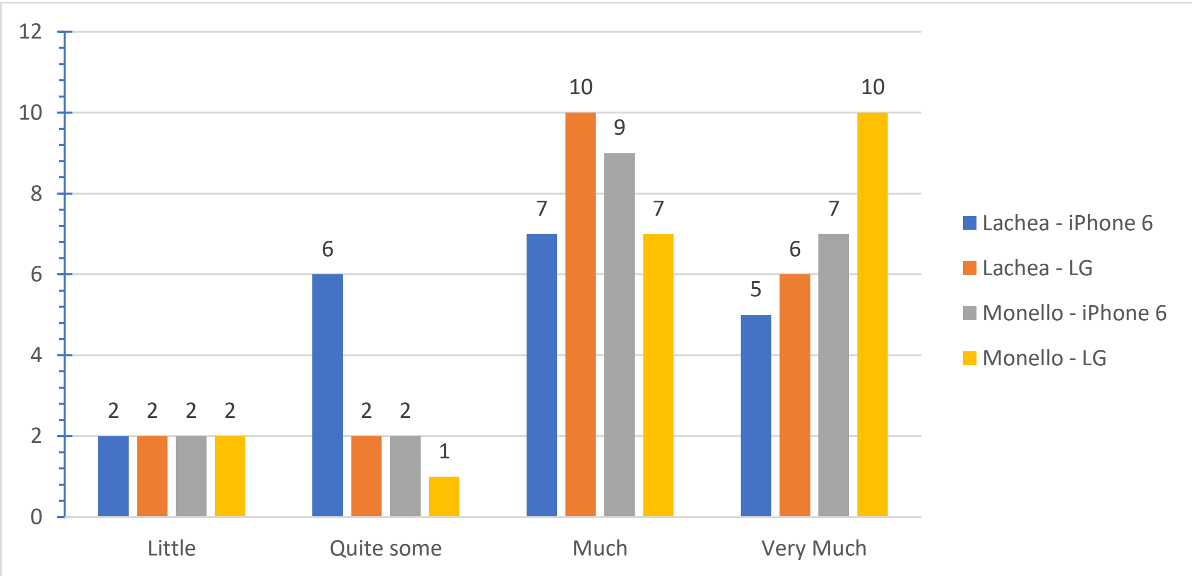


Figure 157 *- Number of test users grouped by their answer to the question <How much do you think that lights and shadows have contributed to the perception of your 3D?> (See Appendix B – question D1.4).*

6.4.4.5 Distorted 3D perception in far or close objects

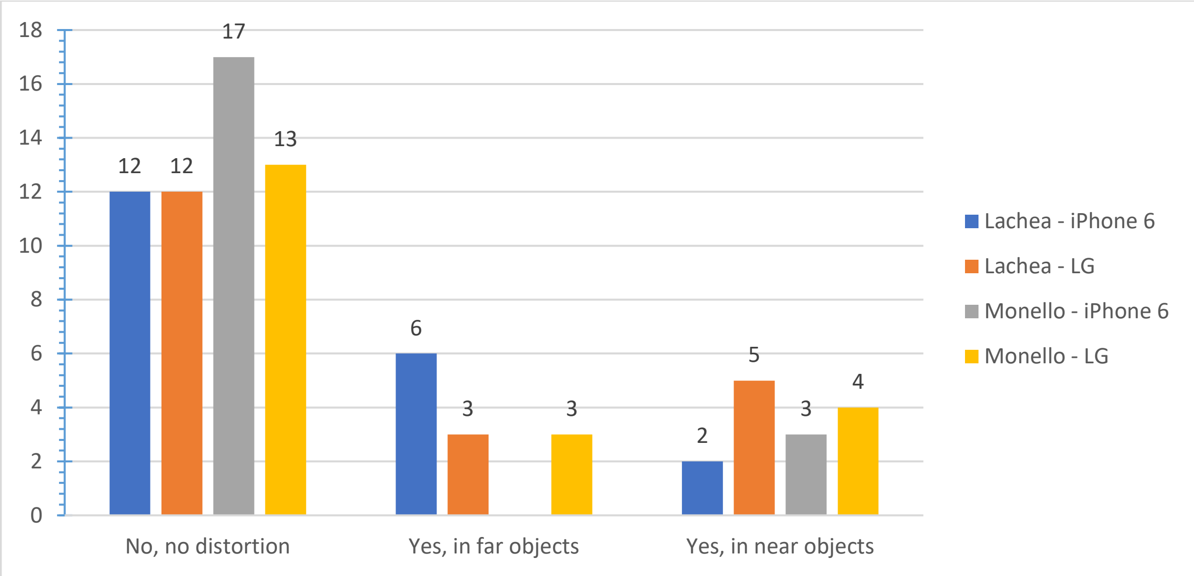


Figure 158 *- Number of test users grouped by their answer to the question <Did you perceive a distorted 3D?> (See Appendix B – question D1.1).*

6.4.4.6 Influence of 3D perception over emotions

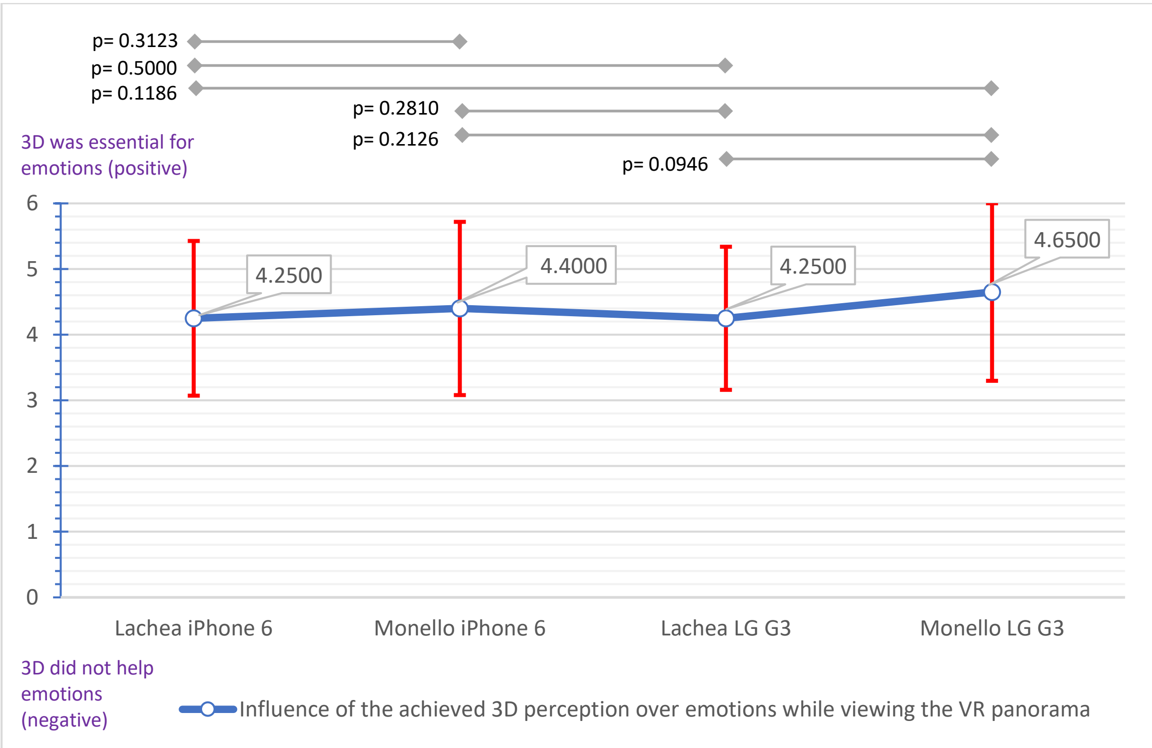


Figure 159 *- No significant difference was found for the influence of 3D perception over emotions, which was considered very high (See Appendix B – question D2.2) [min = 0, Max = +6].*

6.4.4.7 Initial time needed before perceiving the 3D of the panorama

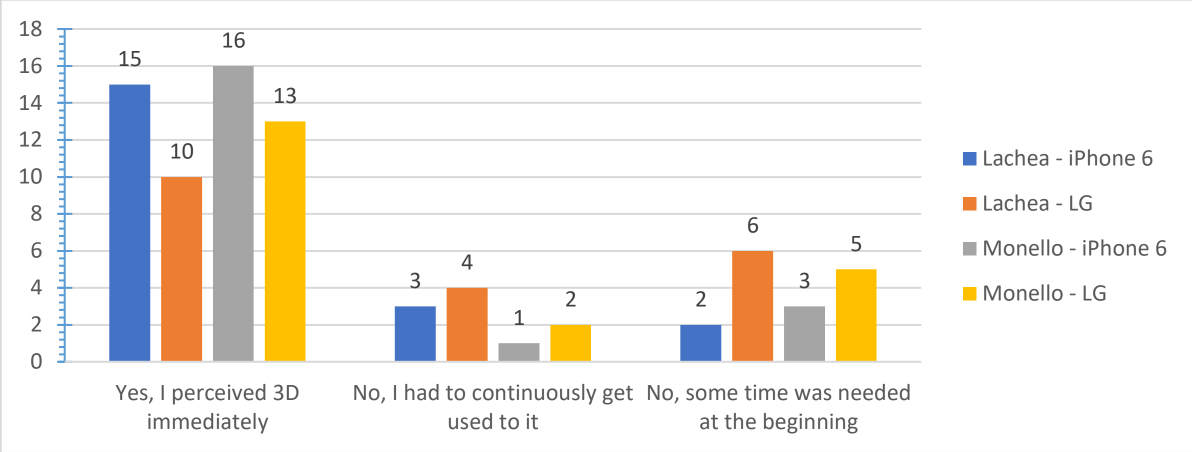


Figure 160 *- Number of test users grouped by their answer to the question <Did you perceive the 3D immediately or your eyes needed some time to get used to it?> (See Appendix B – question D1.2).*

6.5 Analysis

What follows is a summary of the most significant results and of the most relevant presuppositions that have been confirmed by this second usability evaluation.

6.5.1 Realism

6.5.1.1 Outdoor vs Indoor

Image resolution. Almost a significant difference was reported between outdoor and indoor when using iPhone 6 (see **Figure 115**, p-value = 0.06), with indoor achieving higher level of realism. However, this difference was not reported when using the LG G3. I believe this is because Monello panorama had a lot of dark areas and less colours than the Lachea (see subsection 5.1.2), and its luck of details probably made differences in resolution almost unnoticeable to the viewers. Therefore, Lachea was rich in details and the iPhone 6 lower resolution screen deteriorated them.

These results suggest that when a lot of object details are presented inside the panorama (e.g. Lachea), the use of a low-resolution display reduces the level of realism achieved. However, the level of realism is not altered when higher resolution displays are adopted regardless of the amount of image details.

Image sharpness. Monello achieved significantly higher scores than Lachea when using both phones (see **Figure 116**). The reason may reside on the better stitching of the Monello panorama compared to the Lachea island panorama, which reduced blurred vision in certain areas of the pictures.

Image vividness. Lachea performed significantly better than Monello only when using the LG G3 (see **Figure 119**). This combination of outdoor environment and LG G3 scored the best level of realism in terms of vividness. A possible reason is that the Lachea island presents a wider color spectrum compared to Monello cave (see subsection 5.1.2), and the LG G3 was able to provide the best color vivid representation compared to the iPhone 6 thanks to the higher pixel density and higher resolution, which allowed more colour shades displayed on the panorama.

Literature suggests that the coupling shadows - vividness of display might provide more immersion [326], which in turn enhances the sense of presence [319] [331] [332]. Therefore, I expect to find also a high sense of presence when Lachea is viewed on LG G3. Furthermore, when test users were asked whether they considered vividness influent on the level of realism (see **Figure 120**), there was no significant difference between Lachea and Monello when using the iPhone 6, and almost a significant difference when using the LG G3 (p-value = 0.067). This suggests that the wider spectrum of the Lachea panorama and its richness in details (see subsection 5.1.2) helped the realistic perception of the remote environment when using a phone with comparable specification to the LG G3.

Image brightness. **Figure 121** reported a significant difference between Lachea and Monello when using the LG G3, and almost significant when using the iPhone 6 (p-value = 0.09). This can be due to the elements of the outdoor scene (e.g. sunlight, sea reflections, natural illumination), which offered a more natural bright light reproduction than the artificial illumination of Monello panorama. The best score was achieved by the coupling Lachea and LG G3. The same result was shown for image vividness. This might prove a possible correlation between level of realism in terms of image vividness and in terms of image brightness. The performed literature analysis did not find evidence of such correlation in previous studies.

Furthermore, when I asked test users how much they thought image brightness was important for realism (see **Figure 122**), they reported that for panoramas like Lachea viewed with LG G3 it was significantly more influent than panoramas like Monello viewed on iPhone 6. This again suggests that displays with specifications like the ones of the LG G3 should be used when viewing remote environments presenting rich details and natural illumination.

Intensity of colours. Regardless of the phone used, the Lachea panorama achieved significantly higher scores than Monello panorama (see **Figure 123**). This is justified by the wider color range of the outdoor panorama, which presents natural elements (e.g. sea, rocks, trees, sand) and natural illumination (e.g. sunlight, sea reflections) more than the Monello cave (see subsection 5.1.2). This suggests that realism in terms of color intensity is higher when reproducing natural outdoor landscape with wide color spectrum and many natural features.

Furthermore, **Figure 124** shows that test users considered the intensity of the color very influent on the achieved level of realism for almost any panorama and display, and that expectations on color intensity for any type of scene (especially the ones with very rich colour details) is very high and influent on perceived level of realism.

Image contrast. No difference was found between Lachea and Monello: both performed very well in terms of contrast regardless of the phone used (see **Figure 125**). However, when I asked test users how much image contrast is influent on the level of realism (see **Figure 126**), very significant differences were reported between Lachea and Monello regardless of the phone used. This might be justified by the fact that viewers perceive the scene more realistically when wider color spectrum of Lachea panorama and wider levels of contrast are shown. Results suggest that when a scene contains lots of colours, light reflections, and natural illumination (e.g. landscape, forest, beach) image contrast covers an important role for the realistic perception of the remote environment.

No significant difference was reported in terms of:

- Image not pixeled and pixel edges. Both panoramas achieved very high scores (see **Figure 117** and **Figure 118**). This means that scene features of indoors or outdoors do not influence significantly this level of perceived realism.
- Realism compared to real life. Both Lachea and Monello performed equally, with good realism compared to real life (see **Figure 127**). I believe this is due to the good impression Mobile VR provided to "Non-Expert" test users for remote observation, regardless of the environment viewed.
- Realism compared to Photos and Videos. Both panoramas achieved high scores for realism compared to photos and videos. This suggests that regardless of the environment, Mobile VR can provide higher levels of realism than photos or videos.
- Realism of objects deformation and natural elements. Both panoramas achieved good scores (see **Figure 129**). This suggests that objects and natural elements depicted in the Lachea panorama had the same level of realism of the rocks of the cave in Monello panorama without relevant differences.
- Realism of the 3D effect. Both panoramas performed equally, with almost all test users considering the 3D impression good in both centre and borders of screen (see **Figure 130**).
- Realism in terms of emotions. Both panoramas achieved good scores (see **Figure 131**).
- Influence of colours over emotions. Both panoramas achieved high scores (see **Figure 132**).
- Influence of horizontal screen size over Realism. When using the same phone, no significant difference between Lachea and Monello was reported (see **Figure 133**). This suggests that the scene features alone do not significantly affect the amount of tunnel vision and the consequences on realism.

6.5.1.2 iPhone 6 vs LG G3

Image vividness. The LG G3 was able to provide the best color vivid representation of Lachea panorama compared to the iPhone 6 (see **Figure 119**). This is curious because iPhone 6 is supposed to have a better color representation than the LG G3 (see **Table 17**). This suggests that the color reproduction skills of the LG G3 were not the only specifications that allowed a better realism for image vividness. This means that in Mobile VR headsets the combination of other factors owned by the LG G3 (e.g. larger screen, higher resolution, higher pixel density) performed better for Lachea than the iPhone 6. Furthermore, when I asked test users whether they considered vividness influent on the level of realism (see **Figure 120**), there was a significant difference between Lachea using the LG G3 and Monello using the iPhone 6, with the former performing better. This suggests that coupling wide color spectrum with a display like LG G3 can enhance the realism of the panorama.

Image brightness. **Figure 121** shows that LG G3 performed significantly better in terms of image brightness than iPhone 6 when viewing the Lachea island only. This suggests that when the remote environment presents a lot of details and natural illumination, a display like the LG G3 can enhance the level of realism. However, **Table 17** shows that iPhone 6 seems to reproduce brightness better compared to LG G3. This could suggest that inside a Mobile VR headset the size of the screen or the pixel density are more important in terms of realistic image brightness reproduction. This is also confirmed by **Figure 122**, which reported that test users considered more important image brightness for realism with the Lachea panorama viewed on LG G3. The performed literature analysis did not find similar results in previous researches.

Intensity of colours. **Figure 123** shows that a significant difference exists between iPhone 6 and LG G3 in terms of achieved color intensity, with the LG G3 standing out for both panoramas. This is an unexpected result considered **Table 17** because technical specifications proved that the iPhone 6 is better in terms of color fidelity compared to the LG G3. Therefore, it can be assumed that inside the Mobile VR headset the resolution, pixel density and size of the screen have a major effect on the perception of colours of the remote environment, regardless of the colour's representation capabilities of the display. This is better visible by comparing **Figure 161** with **Figure 162**.

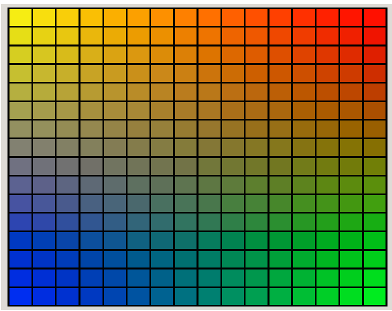


Figure 161 *- Example of high-resolution image with many colour shades.*

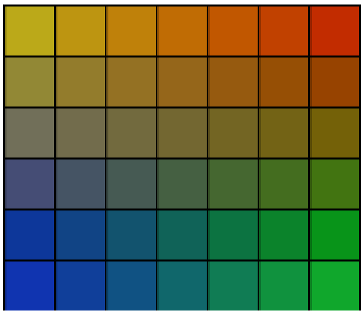


Figure 162 *- Example of low-resolution image of the same picture in* ***Figure 161****, showing less colour shades.*

Furthermore, **Figure 124** shows that test users have very high expectations for color intensity for any type of panorama visualized, and that the display of LG G3 had a major influence on level of realism when coupled with the Lachea compared to Monello on iPhone 6. This suggest choosing a display with specifications comparable to the one of the LG G3 to obtain higher levels of realism, especially with scenes rich in details and colours.

Image contrast. Both phones achieved high scores for image contrast. The result shown in **Figure 125** was unexpected, because according to **Table 17** the iPhone 6 was supposed to perform better compared to LG G3 in terms of contrast ratio. We can assume that display contrast ratio is not the only specification having an influence on contrast perception when the display is used inside a Mobile VR headset.

Despite the similar results between iPhone 6 and LG G3 for image contrast, a significant difference was reported when asking test users how much image contrast influenced realism (see **Figure 126**). Indeed, they reported that the image contrast of Lachea on LG G3 had a significantly higher influence on realism than on iPhone 6. This might confirm that in Mobile VR other specifications (e.g. pixel density, screen size, screen resolution) are responsible for the influence of image contrast on the level of realism, regardless of the contrast representation capabilities of the display. I did not find in my literature analysis similar results.

Realism of the 3D effect. **Figure 130** shows that Lachea viewed on LG G3 provided a different 3D perception in the centre and at the borders of the screen to 40% of the test users. Therefore, Lachea on LG G3 induced slightly more distortions on the borders or on the centre of the screen. I believe that when using a larger screen more distortions are introduced on the panorama if the lenses are not adapting to the new size of the screen. This reduces the level of realism achieved by the provided 3D effect. To solve the problem, retargeting methods from literature are suggested to accurately configure 3D perception on different displays (see subparagraphs 4.2.3.1.1 and 4.2.3.1.2), even when using panoramic stereoscopic pictures.

Influence of horizontal screen size over Realism. **Figure 133** showed that the iPhone 6 had slightly higher scores than LG G3, especially when comparing Monello viewed on LG G3 with both panoramas viewed on iPhone 6. I believe this is due to the smaller horizontal size of the iPhone 6, which reduces the device field of view (DFOV) of the Mobile VR headset. Therefore, viewers experienced more tunnel vision with iPhone 6, and noticed its influence on realism more than with the LG G3. This suggests that regardless of the panorama a phone with larger horizontal size can avoid tunnel vision, benefitting realism.

No significant differences were found for:

- Image resolution (see **Figure 115**), which was considered high with both phones (values between 4 and 5 on a scale of 6). The absence of significant differences between the phones was an unexpected result as image resolution is supposed to have a major effect on the amount of details of the image [101], which in turn should reduce the level of realism. However, even if this generally applies to 2D and 3D displays, I think it is less relevant when using Mobile VR headsets, due to the combination of lenses and device field of view (DFOV).
- Image sharpness. Regardless on the indoor and outdoor panoramas, both phones scored very similar results (see **Figure 116**). This suggests that the use of different phones to visualize the same environment did not alter the level of realism perceived in terms of image sharpness.
- Image not pixeled and pixel edges. Both phones achieved very high scores regardless of the viewed panorama (see **Figure 117** and **Figure 118**). This result was unexpected because LG G3 has higher pixel density and resolution than the iPhone 6 (see **Table 17**). I believe that using a Mobile VR headset the display resolution is alone less influent on the image visual perception, and that it should be considered in combination with the lenses and the device field of view (DFOV).

- Realism compared to real life. Both phones achieved good realism compared to real life (see **Figure 127**). I believe this is due to the good impression Mobile VR provided to the test users for remote observation, regardless of the phone used.
- Realism compared to Photos and Videos. Both phones achieved high scores for realism compared to photos and videos. This suggests that regardless of the display used, Mobile VR can provide higher levels of realism than photos or videos.
- Realism of objects deformation and natural elements. Both phones achieved good scores (see **Figure 129**).
- Realism in terms of emotions. Both phones achieved good scores (see **Figure 131**).
- Influence of colours over emotions. Both phones achieved high scores (see **Figure 132**).

6.5.2 Comfort

6.5.2.1 Outdoor vs Indoor

Discomfort. According to **Figure 134** most of the discomfort was perceived more with Lachea panorama. Indeed, "very much" discomfort for Lachea scored 7 points regardless of the phone, while Monello scored only 1 point for "very much". I believe this is due to stitching errors on Lachea (see subsection 5.1.2), which appeared more visible with the screen of the LG G3 (5 points over 7) than with the one of iPhone 6 (2 points over 7). It can be assumed that the LG G3 showed more clearly the stitching errors of Lachea panorama and induced more discomfort on test users. This suggests that the more stitching errors are present in a panorama, the more uncomfortable the remote observation will result.

Max duration of viewing before discomfort occurs. Results show that the only combination of panorama and phone that guaranteed a long comfortable time of remote observation was the Lachea with the iPhone 6 (see **Figure 136**). I think this is because Lachea had more details to explore than Monello and gave more "holiday" feeling to test users, which in turn could have induced more relax and comfort.

No significant difference was reported for:

- Eye strain or visual fatigue. Both Lachea and Monello achieved an acceptable level of comfort, regardless of the phone used (see **Figure 135**).

6.5.2.2 iPhone 6 vs LG G3

Discomfort. According to **Figure 134** the Lachea viewed on LG G3 caused the highest discomfort to test users, compared to Lachea viewed on iPhone 6. It can be assumed that the LG G3 showed more clearly the stitching errors of Lachea panorama (see subsection 5.1.2) and induced more discomfort on test users. This suggests that when stitching errors or distortions are presented in the panorama, a display with specifications like the iPhone 6 is recommended.

Max duration of viewing before discomfort occurs. Results show that the only combination of panorama and phone that guaranteed a long comfortable time of viewing was the Lachea with the iPhone 6 (see **Figure 136**). I believe this is because the iPhone 6 shows less details of the panorama and might have masked with its lower resolution some of the stitching errors of Lachea

No significant difference was reported for:

- Eye strain or visual fatigue. Both iPhone 6 and LG G3 achieved an acceptable level of comfort, regardless of the panorama viewed (see **Figure 135**).

6.5.3 Presence

6.5.3.1 Outdoor vs Indoor

Sense of presence. Both panoramas regardless of the phone achieved very good scores (see **Figure 137**). However, **Figure 138** reported that Monello with LG G3 performed the highest scores for "feeling of being there" inside the panorama. Furthermore, the trend suggests that generally Monello provided more sense of presence than Lachea. This suggests that indoor environments with scene features like the ones of Monello cave can provide higher sense of presence. I believe this is due to:

- The higher 3D impression provided by objects very close to the viewer and the small close-range bounded spaces of the panorama, which could recall the boundaries of the viewer when wearing the Mobile VR headset;
- The simplicity of the panorama, which might have helped proper depth perception, like suggested by previous researches (see paragraph 4.2.1.1).

Induced emotions. According to **Figure 143**, Lachea induced slightly more relax and joy than Monello. This was expected because I think that the Lachea panorama recalls more the idea of "being on holiday" thanks to the beach and the sea. Furthermore, the 20% of test users experienced "anxiety" with the Monello cave, 75% of which using LG G3. This emotion was not provided by Lachea panorama. "Happiness" was almost equally induced by both panoramas, with slightly lower scores for Monello when using the LG G3.

Furthermore, when asking test users to estimate induced emotions through the 6-Likart scale, results reported that both Lachea and Monello achieved positive emotions (see **Figure 145**), very high enjoyment (see **Figure 146**), very high happiness (see **Figure 147**), and high relax (see **Figure 150**) without significant differences.

Results also reported a significant difference for sadness with Lachea island compared to Monello, but only when using the LG G3 (see **Figure 148**). This suggests that when panoramas with wide ranges of colours are visualized on a screen like the LG G3 there is almost complete absence of sadness compared to all other scenarios. However, it is not possible to assume that outdoor environments provide less sadness compared to indoor environments because for iPhone 6 the two panoramas achieved similar results.

Results also reported that when using LG G3 there was a significant difference between Lachea and Monello for scariness, with the latter having higher scores (see **Figure 149**). This suggests that the use of a display like the LG G3 to visualize panoramas with scene features like the Monello panorama (see subsection 5.1.2) can amplify induced scariness on viewers. I believe this is because the LG G3 larger device field of view might have enhanced the sense of presence (which resulted in highest scores for Monello LG G3, see **Figure 138**) and activated more emotions (including fear). This confirms previous researches proving that presence and emotions activate each other [319] [333] [334] [322].

Finally, results reported a strong significant difference between Lachea and Monello in terms of "feeling like going on holiday", with Lachea having the best performances regardless of the phones used (see **Figure 151**). This was expected as the Lachea panorama provides a summer-like feeling thanks to the presence of the beach, sunny weather, green vegetation, and wide range of colours (see subsection 5.1.2). This is in line with previous researches, which proved that bright colours rise mainly positive emotional associations, whilst dark colours mainly negative emotional associations [335].

No significant difference was reported for:

- Tunnel vision. No significant difference was reported between Lachea and Monello when using the same phones (see **Figure 139**). This suggests that the content of the scene has only a marginal influence than display specifications on tunnel vision.
- Isolation. Both panoramas achieved a very good isolation from the surrounding real environment (see **Figure 141**). This suggests that the only factor responsible of the isolation for the viewer is the Mobile VR headset in use.
- Influence of isolation over emotions. **Figure 142** reported no significant difference between Lachea and Monello, which scored quite high results in terms of influence of isolation over emotions. This is justified because both panoramas also achieved similar isolation scores, regardless of the phone used.
- Influence of emotions over sense of presence. Test users expressed the belief that emotions contribute positively to the achieved sense of presence, regardless of the panorama showed (see **Figure 152**). This result is also confirmed in literature [319] [333] [334] [322].
- Emotions caused by coming back to real life after the VR experience. In **Figure 153** no significant difference was found between Lachea and Monello panorama, with both scoring an average of no emotions perceived. However, the standard deviation regardless of the phone used was quite high, meaning that the response to the VR experience was quite diverse. This shows how much the subjective perception of each test user is still very difficult to predict in VR. Further works on the psychological emotional response of test users could be done in the future.

6.5.3.2 iPhone 6 vs LG G3

Sense of presence. Both phones regardless of the panorama achieved very good scores (see **Figure 137**). This suggests that the bigger size of the LG G3 display, its higher pixel density, and its higher resolution did not significantly affect the overall achieved sense of presence of test users.

Tunnel vision. **Figure 139** clearly shows significant differences between iPhone 6 and LG G3, regardless of the viewed panorama. In detail, iPhone 6 caused more tunnel vision due to its smaller screen size, which in turn offered tinier device field of view compared to the one of the LG G3. This suggests that to avoid tunnel vision larger screens and with specifications similar to the LG G3 should be used, regardless of the content showed on the panorama. However, **Figure 140** reported that test users did not mind too much the effect of tunnel vision, which did not alter considerably the sense of presence. This result is in contrast with previous researches [39], which stated that a narrow device field of view with tunnel vision reduce immersion in VR.

Induced emotions. According to results in **Figure 143** "surprise" was mostly experienced when viewing Lachea using the LG G3 (40%) and Monello using iPhone 6 (45%). Curiously, this emotion was almost not provided at all by Lachea viewed on iPhone 6. I believe this is because test users appreciated more the large screen and pixel density of the LG G3, and were more surprised by the visual effect achieved with that setup.

Furthermore, when asking test users to estimate induced emotions through the 6-Likart scale, results reported that both iPhone 6 and LG G3 achieved positive emotions (see **Figure 145**), very high enjoyment (see **Figure 146**), very high happiness (see **Figure 147**), low levels of scariness (see **Figure 149**), high relax (see **Figure 150**), and similar scores for feeling like going on holiday when using the same panorama (see **Figure 151**) without significant differences.

Results also reported a significant difference for sadness when using the LG G3 to visualize panoramas like the Lachea (see **Figure 148**). Indeed, while Lachea on LG G3 induced significantly less sadness than Monello using the same phone, the two panoramas had no significant difference when using the iPhone 6. This suggests that a display like the one of the LG G3 can provide more positive feelings to the viewer when showing panoramas rich in colours and details compared to the almost monochromatic scenes of Monello panorama. Therefore, such a display can be used to significantly reduce sadness when visualizing colourful remote environments.

No significant difference was reported for:

- Isolation. Both phones achieved a very good isolation from the surrounding real environment (see **Figure 141**). This suggests that the only factor responsible of the isolation for the viewer is the Mobile VR headset in use.
- Influence of isolation over emotions. **Figure 142** reported no significant difference between iPhone 6 and LG G3: for both the influence of isolation over emotions was estimated quite high. I believe this is justified because both phones also achieved similar isolation scores. Furthermore, this is in line with previous researches that reported an increased sense of presence when more isolation is provided [316] [336]. More presence indeed activates more emotions [319] [333] [334] [322].
- Influence of emotions over sense of presence. Test users expressed the belief that emotions contribute positively to the achieved sense of presence, regardless of the phone used in Mobile VR (see **Figure 152**). This result is also confirmed in literature [319] [333] [334] [322].
- Emotions caused by coming back to real life after the VR experience. In **Figure 153** no significant difference was found between iPhone 6 and LG G3, with both scoring an average of no emotions perceived. However, the standard deviation regardless of the panorama viewed was quite high, meaning that the response to the VR experience was quite diverse. This shows how much the subjective perception of each test user is still very difficult to predict in VR. Further works on the psychological emotional response of test users could be done in the future.

6.5.4 Depth Perception

6.5.4.1 Outdoor vs Indoor

Distance perception. Results report that the Monello panorama achieved higher scores than the Lachea in terms of distance estimation (see **Figure 155**). This was expected as Monello presents objects closer to the viewer, and distance estimation errors might have occurred less. Indeed, with Lachea panorama far objects appeared almost flat, while in Monello this effect was reduced, especially when using the iPhone 6.

Lights and shadows contribution to 3D perception. Results reported that Monello panorama achieved slightly higher scores than Lachea panorama, regardless of the phone used (see **Figure 157**). This might be explained by the strong shadows and artificial illumination of the Monello cave, which presented also more depth planes due to the dissimilar rocks surrounding the viewer. Furthermore, I believe that the monochromatic appearance of the scene helped the viewer to focus more on lights and shadows, performing better space awareness. However, the overall contribution to 3D was quite high, which was expected as lights and shadows are considered to be reliable monocular depth cues [23] [337] [338] [339]**.**

Distorted 3D perception in far or close objects. Results reported slightly less distortions with Monello panorama when using the iPhone 6 (see **Figure 158**). This was expected as the scene presented fewer stitching errors than the Lachea panorama.

Initial time needed before perceiving 3D. Results in **Figure 160** reported that the Monello panorama was able to provide 3D perception slightly faster than the Lachea panorama, especially when using iPhone 6. Furthermore, Lachea forced several test users (15% with iPhone 6, 20% with LG G3) to continuously get used to the 3D. This amount of test users was not reported when viewing the Monello panorama. This might be due to more stitching errors presented in the Lachea outdoor panorama (see subsection 5.1.2).

No significant difference was reported for:

- 3D depth impression. Both panoramas performed equally with very good scores (see **Figure 154**). This result was quite unexpected because I believed that scene features of the Monello panorama would have induced more 3D impression due to objects presented closer to the viewer. However, this suggests that the overall depth perception was good regardless of the panorama and phone used. I suspect that the Mobile VR headset lenses configuration (e.g. distance from the eyes, interpupillary distance) had a major responsibility as it did not change during all the trials.
- Color contribution to 3D perception. Results in **Figure 156** reported that there was no significant difference between Lachea and Monello panorama, and both scored very high ratings.
- Influence of 3D perception over emotions. Results in **Figure 159** reported that both panoramas achieved very high scores. I believe this is because the majority of the test users was positively affected by the possibility to visualize an existing place in 3D through Mobile VR.

6.5.4.2 iPhone 6 vs LG G3

Distance perception. Results reported that the iPhone 6 achieved higher scores than the LG G3 for distance estimation, especially when visualizing the Monello panorama (75% of test users, see **Figure 155**). This might suggest that for indoor environments with scene features like the Monello panorama a smaller display with specifications like the iPhone 6 improves distance perception. I believe this might also be influenced by the graphical field of view (GFOV) chosen for the panorama.

Color contribution to 3D perception. In **Figure 156** test users considered the contribution of colours "very much" on their 3D perception more on LG G3 (40% indoor and outdoor) than on iPhone 6 (15% indoor, 25% outdoor). I think that the LG G3 larger device field of view affected slightly more the 3D perception of the scene compared to iPhone 6. In literature I found an interesting debate [340] on the influence of device field of view and graphical field of view on distance perception: some studies reported no influence [341] [342], others a considerable one [343] [344] [345] (which agree with my results).

Lights and shadows contribution to 3D perception. Results reported that there is not much difference between iPhone 6 and LG G3, except for when the iPhone 6 is used to visualize the Lachea panorama (see **Figure 157**). I believe this is because the Lachea presents less strong shadows than the Monello panorama, and the lower resolution and pixel density of the iPhone 6 made them even less noticeable for test users. This suggests that the combination of scene with few shadows and low resolution / low pixel density display reduces 3D performances when relying on lights and shadows.

Distorted 3D perception in far or close objects. Results reported that the iPhone 6 caused more distortions in far objects (30%) than the LG G3 (15%) when viewing the Lachea panorama (see **Figure 158**). By contrast, the LG G3 caused more distortions in near objects (25%) than the iPhone 6 (10%) when viewing the Lachea panorama. I suspect this is due to the different device and graphical field of views of the two devices, which might have introduced alteration on disparity values of the scene. Furthermore, the iPhone 6 provided the lowest amount of distortions when viewing the Monello panorama. I believe this is due to both the lower amount of stitching errors of the Monello and the lower resolution of the iPhone 6, which masked errors more than the LG G3.

Initial time needed before perceiving 3D. Results in **Figure 160** reported that the iPhone 6 was able to provide 3D perception slightly faster than the LG G3, regardless of the panorama viewed. Indeed, Lachea (75%) and Monello (80%) were perceived in 3D by test users immediately with iPhone 6. I believe this is because the iPhone 6 presented lower resolution, which partially masked possible stitching errors better and provided a slightly simpler scene to the viewer. This assumption is justified by my literature analysis, which proved that simpler scenes can benefit 3D perception (see paragraph 4.2.1.1).

No significant difference was reported for:

- 3D depth impression. Both phones performed equally with very good scores (see **Figure 154**). This result was not expected because I thought that the higher resolution of the LG G3 would have increased the stereoscopic resolution as well, improving depth perception. However, it should be reminded that test users answered this question subjectively, and that some of them might have not noticed any difference in resolution due to myopia or astigmatism, resulting in similar 3D depth impression for all phone-panorama configurations. Furthermore, it can be assumed that the Mobile VR headset lenses setup (e.g. distance from the eyes, interpupillary distance) had a major responsibility, as it did not change during all the trials.

- Influence of 3D perception over emotions. Results in **Figure 159** reported that both phones achieved very high scores. I believe this is because most of the test users were positively affected by the possibility to visualize an existing place in 3D through Mobile VR.

6.6 Guidelines

Based on the previous sections' results and analysis, a Guidelines section is provided also for this user study to summarize the main outcomes in terms of concise guidelines for system designers. The guidelines are laid down looking at the outcome of the user study against the four elements that have been considered as relevant contributors to VR remote visual observation: realism, comfort, presence and depth-perception.

- **How to improve realism?**
  - Use high resolution displays to show panoramas rich in details. However, when simple scenes are presented, display resolution is less influent on realism.
  - Reduce blurred vision, especially due to stitching errors. The viewer should enjoy a clear observation of the whole scene, like it would be in real life.
  - Use display with high pixel density and resolution when the scene is very rich in details, especially when it features wide colours spectrum and high levels of vividness. This is because higher resolution enables more colour shades, which make the panorama more realistic.
  - Prioritize the use of natural illumination and natural environments, which can provide better realism in terms of image brightness, especially with high resolution and high pixel density. This also applies to panoramas with high vividness.
  - When a scene contains lots of colours, light reflections, and natural illumination (e.g. landscape, forest, beach), image contrast covers an important role for the realistic perception of the environment.
  - Inside a Mobile VR headset, screen size, resolution, and pixel density might be more important in terms of realistic image brightness reproduction than the actual colour reproduction skills of the display.
  - When visualizing a panorama on different screen sizes, make sure to adapt the graphical field of view of the VR viewer so that object sizes and 3D keep their realistic appearance.
  - Avoid tunnel vision by using large screen sizes or appropriate lenses for the Mobile VR headset. However, if the screen is small and the lenses introduce excessive visual distortions, realism might be low despite the absence of tunnel vision.

- **How to improve comfort?**
  - Avoid stitching errors and blurred vision. When stitching errors are present, lower resolution, smaller screen size, and lower pixel density might alleviate their negative effect.
  - Prioritize relaxing environments to the ones that induce fear and claustrophobia.

- **How to improve presence?**
    - Prioritize environments offering objects that are close to the viewer without excessive disparity (strong visible 3D) and simple design to provide less cue conflicts.
    - Avoid tunnel vision. However, if tunnel vision cannot be removed, a close-range bounded environment could recall the boundaries of the HMD and thus enhance the viewer's sense of presence.
    - Prioritize environments that can induce strong emotions to the viewer (e.g. feeling of going on holiday, enjoyment, happiness, relax). Within this purpose, remember that:
        - Wide colour spectrum and high pixel density can reduce sadness;
        - Dark environments with almost monochromatic colours combined with high resolution and high pixel density can induce more easily fear and scariness;
        - Nice weather, natural environments, open spaces, sunlight, sea, water reflections can easily induce good feelings.
        - Bright colours rise mainly positive emotional associations;
        - Dark colours rise mainly negative emotional associations.

- **How to improve depth perception?**
    - Prioritize objects closer to the viewer to deliver stronger depth perception. However, remember to avoid excessive disparity to allow proper depth reconstruction of the scene.
    - When strong depth cues are provided, distance estimation is improved as the viewer can have more objects as reference for space and depth perception.
    - Strong shadows might help to improve depth perception through more visible monocular cues, which are beneficial specially to solve occurring binocular cue conflicts.
    - Prioritize more depth planes on the scene to help the viewer perceive relative distances between objects more accurately and to enhance achieved space awareness.
    - Avoid stitching errors, vertical parallax, and possible perspective mismatches between left-eye and right-eye panoramas. Less stitching errors can also reduce the time needed for the viewer to see the scene in 3D.
    - Choose the appropriate graphical field of view (GFOV) of the Mobile VR viewer, according to the screen size used inside the HMD. This also depends on the camera used to capture the environment, and on the setup adopted to generate the 3D panorama via software.
    - If low resolution displays are used, monocular cues (e.g. lights and shadows) should be intensified so that the viewer will still be able to combine them with binocular cues and better estimate scene depth.
    - Choose corrective lenses when the viewer is affected by astigmatism or myopia, to keep disparity and depth cues visible enough for an appropriate space awareness of the remote environment.

6.7 Conclusion

In this chapter the design and outcomes of the second usability evaluation proposed by this Thesis were presented.

The focus of this user study was on the influence of the *display* specifications and the characteristics of the observed *environments* over realism, comfort, sense of presence, and depth perception. To perform the study, an iPhone 6 and an LG G3 were used inside a Mobile VR headset and assessed in combination with an indoor and an outdoor environment.

The variables defined for this analysis, data on test users, and information on the chosen statistical tools to calculate results were presented.

Results were compared with what was found in the previous systematic review and were used to devise new guidelines to optimize the VR system and improve the visual perception of the remote environments.

CHAPTER 7. USABILITY EVALUATION - EYE-ADAPTED HDR VIEWING

This chapter describes the third phase of experimentation related to the role of "eye-adaptation driven HDR viewing". The experimentation aims at further investigate previously assessed elements within a specific VR headset technology: eye-tracked headsets, to drive HDR viewing modality of an environment. The chapter starts by introducing proposed evaluation's design (section 7.1), and it continues by presenting implementation (section 7.2), procedure (see section 7.3), user study's results (section 7.3.1), results analysis (section 7.4.6), and guidelines (section 7.6).

7.1 Evaluation Design

The purpose of this usability study is to understand if the use of dynamic HDR and eye adaptation techniques in virtual reality with real pictures can improve sense of presence, comfort, depth perception, and realism when observing remote environments. To do it, *HDR modality* and *environment* are investigated.

7.1.1 HDR Modality

We distinguish three modalities for HDR visualization in virtual reality:

1. **Static HDR**. A static 3D panorama photo is showed in virtual reality with no change in colours and lights regardless of the head rotation and eyes position of the viewer. This modality is the commonly used by Mobile VR headsets to visualize photos depicting real existing places without the use of CGI.

2. **Dynamic Head Tracking (HT) HDR**. A 3D panoramic HDR picture is dynamically converted into LDR according to the viewer's head rotation but regardless of eyes position. The resulting LDR picture shows the best lighting and colours conditions around the centre of the virtual reality screen, changing exposure values accordingly to emulate the behaviour of the human eye pupil when watching strong light sources. For example, if in the panorama a strong light comes from a window and the outside appears completely white, the generated LDR picture will reduce exposure levels to replace the white of the outside area with a non-overexposed version of that photo when the viewer's head is pointing the window. When viewer's head points somewhere else inside the room, exposure levels will be increased again to replace the underexposed areas of the room with an optimal exposition to enhance colours and scene details.

3. **Dynamic Eye Tracking (ET) HDR**. The same behaviour of the dynamic head tracking HDR is reproduced, but this time considering the viewer's eyes position too. This means that the area that the viewer's eyes are pointing rather than the centre of the screen will be updated, to show optimal lighting conditions and enhance scene colour details accordingly. This modality requires extra computation to track the viewer's eyes in order to make the virtual vision as believable as possible.

7.1.2 Environment

To comply with requirements of the proposed experiment objectives (see subsection 3.4.1) I decide to choose two environments at the **University of Hertfordshire**: an indoor hall (**Hutton Hub**, see **Figure 163**) and an outdoor area (Outside **Main Reception Entrance**, see **Figure 164**). The two environments are chosen because of their characteristics; the most relevant of them are listed in **Table 19**.

HUTTON HUB	MAIN RECEPTION
Indoor	Outdoor
Artificial lights from the roof, and sunlight coming from large windows	Only natural lights, with sunlight covered by cloudy sky and reflections on buildings
Tables and chairs close to the viewer	Viewer quite far from objects and people
Few dark areas	Many dark areas
Scene with posters, adverts, and wooden surfaces	Scene with trees, garden, rocks, and reflective surfaces on several buildings

Table 19. *Differences between Hutton Hub panorama and Main Reception Entrance panorama.*

Figure 163 *- Hutton Hub HDR static panorama.*

Figure 164 *- Main Reception HDR static panorama.*

7.1.3 Usability variables

For the aim of the investigation, the following set of independent and dependent variables is considered.

Independent variables:

- **HDR modality used**. This refers to static HDR, dynamic head tracking HDR, and dynamic eye tracking HDR.

Dependent variables:

- **Realism in terms of Presence**
  - Sense of presence;
  - Isolation;
  - Feeling like observing a real place;
- **Realism in terms of Emotions**
  - Likeliness;
  - Happiness;
  - Disappointment;
  - Desire to visit the remote environment in real life;
  - Positive mood change after viewing in VR;
- **Realism in terms of Virtual Vision vs Real Life**
  - Motivation to use VR for remote observation;
  - Natural eye adaptation;
  - Virtual lights vs Real Life lights;
  - Color realism;
  - Realism of Darkness;
  - Realism of Brightness;
  - Virtual vision compared to Real vision;
- **Comfort**
  - Level of comfort;
  - Time before feeling discomfort;
  - Perceptual issues;
- **Depth perception**
  - 3D accuracy;
  - Lights and Shadows contribution to 3D;
  - Realistic space perception;
  - Perceived size of the environment.

7.1.4 Participants

The complete set of test users consists of 24 people, of which 8 females and 16 males.

The following additional information on test users are collected to investigate possible effects of personal differences on dependent variables:

- **Personal information**
    - Age
    - Gender
    - Use of glasses in real life
    - Eye problems such as astigmatism, myopia, presbyopia, hypermetropia
    - Previous experience with computer games
    - Previous experience with 3D displays or 3D digital images
    - Previous experience with virtual reality
    - Interpupillary distance

7.1.5 Statistical tools

A within subject design is adopted. In detail, the following assumptions are made to perform several ***within-subject paired t-test*** [310], in line with literature guidelines [311] [312]. Furthermore, multiple t-tests in Excel are preferred to the MANOVA, to obtain further details on possible significant differences. Therefore, all possible coupling combinations for *HDR modalities* and *environment* are considered as follows:

1. Indoor vs Outdoor when using Static HDR;
2. Indoor vs Outdoor when using Dynamic HDR with Head Tracking;
3. Indoor vs Outdoor when using Dynamic HDR with Eye Tracking;
4. Static HDR vs Dynamic HDR with Head Tracking, regardless of viewed environment;
5. Static HDR vs Dynamic HDR with Eye Tracking, regardless of viewed environment;
6. Dynamic HDR with Head Tracking vs Dynamic HDR with Eye Tracking, regardless of viewed environment.

For each of the six combinations, the p-value resulting from each *paired t-test* performed on each question is calculated.

7.2 Implementation

One of the challenges for this user study is to capture a real environment producing a 3D HDR panorama with all relative visual information to correctly process lights of the scene in virtual reality (see subsection 7.2.1). Furthermore, the second challenge is to implement an HDR system (see subsection 7.2.2) able to:

- Dynamically convert an HDR panorama into LDR fast enough and with good approximation;
- Track viewer's head rotations;
- Track viewer's eyes positions with minimum delays;
- Calculate the right amount of exposure to show on the LDR resulting panorama.

7.2.1 How to capture a real environment for dynamic HDR

The problem of the technique previously used with the Fuji 3D camera is that while the camera captures multiple pictures of the surrounding environment light conditions change. Therefore, it is not ideal to use the same acquisition method to get HDR panoramas from multiple exposures.

To solve this problem, I find more suitable the use of a 3D 360 high resolution camera, able to capture the environment faster and more accurately. Within this purpose, I decide to use an **insta360 Pro camera** [346] (see **Figure 165**), as it can capture 3D 360 8K panoramas in RAW format and at different exposures in a few seconds.

Figure 165 *- insta360 Pro camera.*

Furthermore, the software "Insta360 Stitcher" (see **Figure 166**), which is provided with the camera, is used to generate HDR and RAW panoramas from the captured pictures.

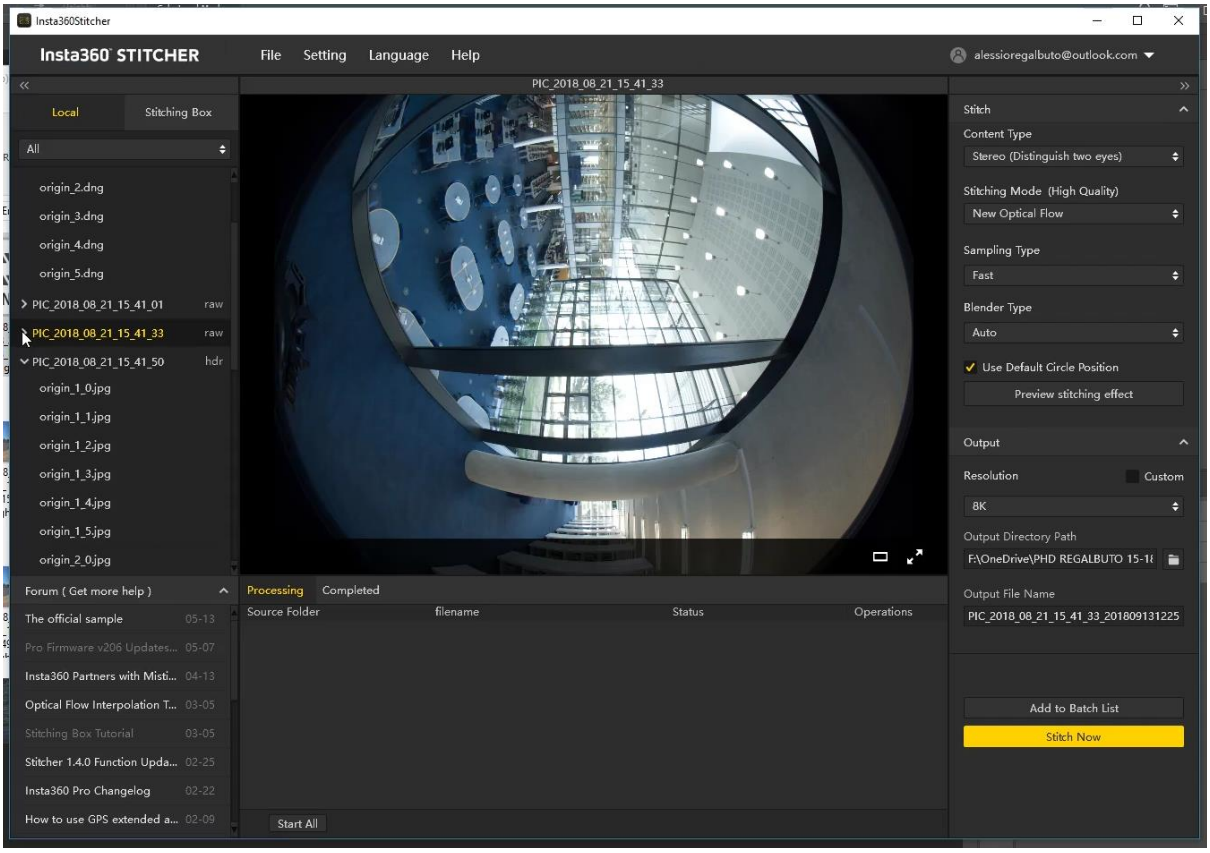


Figure 166 *- Insta360 Stitcher, used to create 3D panoramas from captured pictures of the 6 lenses of the insta360 Pro camera.*

7.2.2 How to implement a system for static and dynamic HDR

The main requirement to emulate eye adaptation for a dynamic HDR panorama visualization is the use of an eye tracking device. Within this purpose, I decide to use the FOVE VR headset (see **Figure 167**), which offers optical tracking capability through infrared cameras behind the HMD lenses (see **Figure 168**).

Figure 167 *- FOVE HMD with eye tracking.*

Figure 168 *- FOVE HMD (view of lenses and eye tracking integrated system).*

7.2.2.1 The problem of the floating-point buffer

I decide to use Unity game engine [16] to create the proposed HDR visualization system. This engine has a built-in eye-adaptation functionality that works with HDR panoramas represented by 32-bit color channels on 32-bits floating-point display buffers. However, the FOVE VR display is unable to handle floating-point images, thus unable to directly use Unity built-in eye adaptation with 32-bit color values.

To solve this problem, I propose to create a system that manually changes the exposure of an HDR skybox and dynamically converts it to an LDR panorama to be shown on the HMD at runtime. To calculate exposure value changes of the LDR panorama, an approximate estimation of the light sources within the scene is required, to reduce light intensity when the user faces them and increase it when facing dark areas. This mechanism attempts to adjust over-exposed and under-exposed areas of the panorama, inducing the impression of the eye-adapted HDR viewing. To do it, I decide to create a grayscale ground-truth panorama that highlights light sources in white and dark areas in black.

It should be noted that such ground-truth panorama is an approximation and aims only at providing convincing exposure changes to induce the eye-adaptation sensation on the viewers. The aim is not focused on precise exposure values, due to limitations of the used hardware and visualization system.

7.2.2.2 Generating approximate ground-truth illumination maps

To generate a ground-truth panorama for the designed system, I initially thought to manually select the areas where light sources of the scene are approximately located. However, I was able to find an alternative way to speed up the process by using Adobe Photoshop.

More specifically, I first generate an LDR tone-mapped version of the original HDR panorama, by converting 32-bit color channels to 16-bit (see **Figure 169** and **Figure 170**).

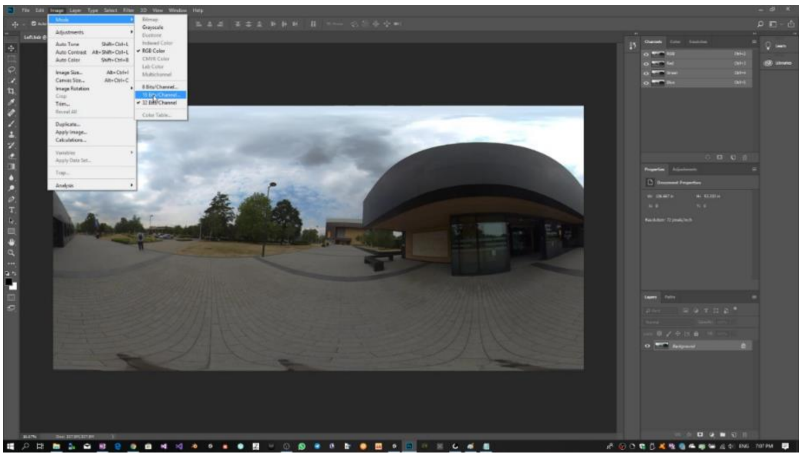


Figure 169 *- Step 1. Converting HDR panorama from 32-bit color channels to 16-bit.*

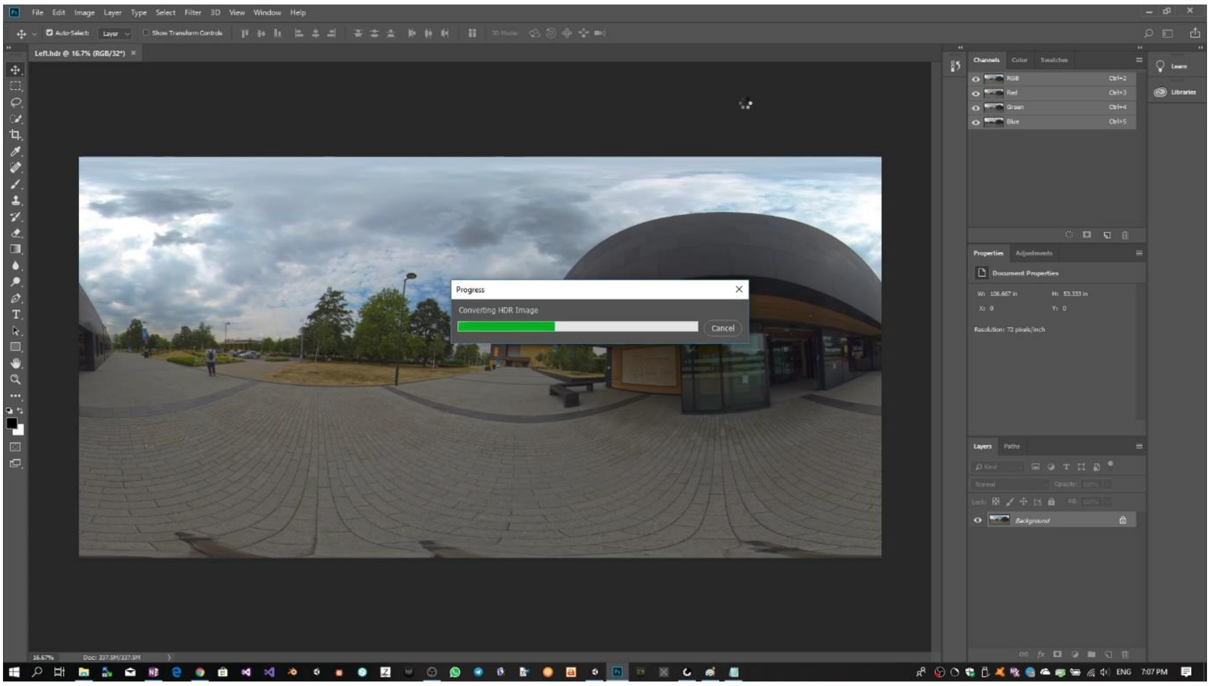


Figure 170 *- Step 2. HDR conversion to 16-bit.*

Once generated an LDR panorama, I load the RGB channel as a selection by using the shortcut CTRL + click on the RGB channel (see top right of **Figure 171**).

Despite the approximation, this results in the selection of the areas of the panorama that present illuminated zones of the scene. Then I create an alpha channel from this selection (see **Figure 172**).

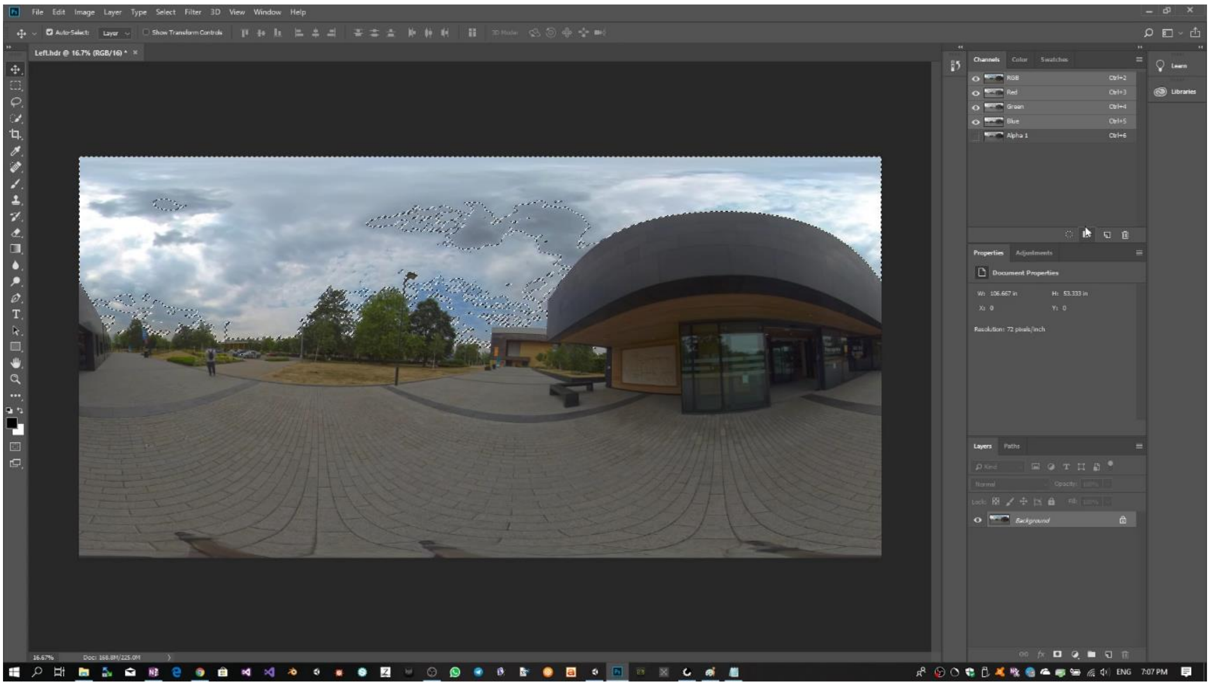


Figure 171 *- Step 3. Load RGB channel as a selection (CTRL + click on RGB channel) and create alpha channel from the selection.*

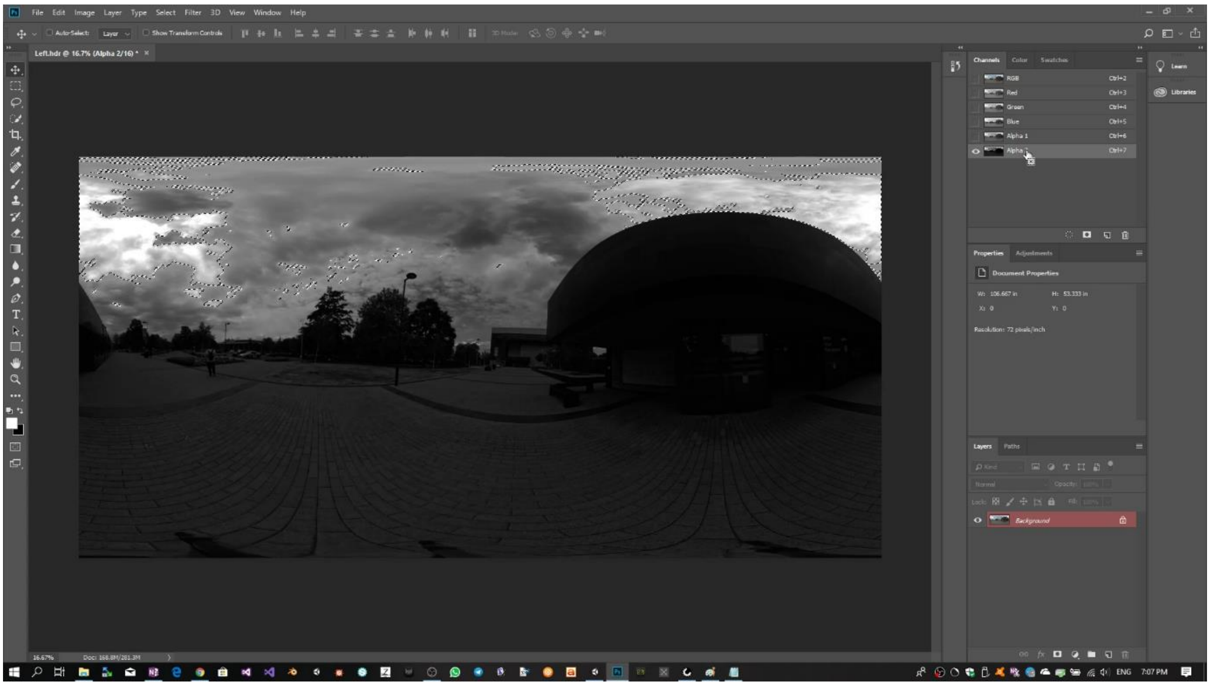


Figure 172 - Step 4. Making under-exposed areas look black and producing new Alpha layer. (CTRL + ALT + SHIFT + click on alpha channels).

Once created the alpha channel layer containing information from the previous selection, I create an intersection of this channel with itself, to make the difference between bright and dark areas even more visible. To do it, I use the shortcut CTRL + ALT + SHIFT + click on the generated alpha channel and copy the resulting selection on a new channel (see **Figure 173**). Repeating this operation multiple times, I obtain different approximated ground-truth panoramas that can be used for the HDR system.

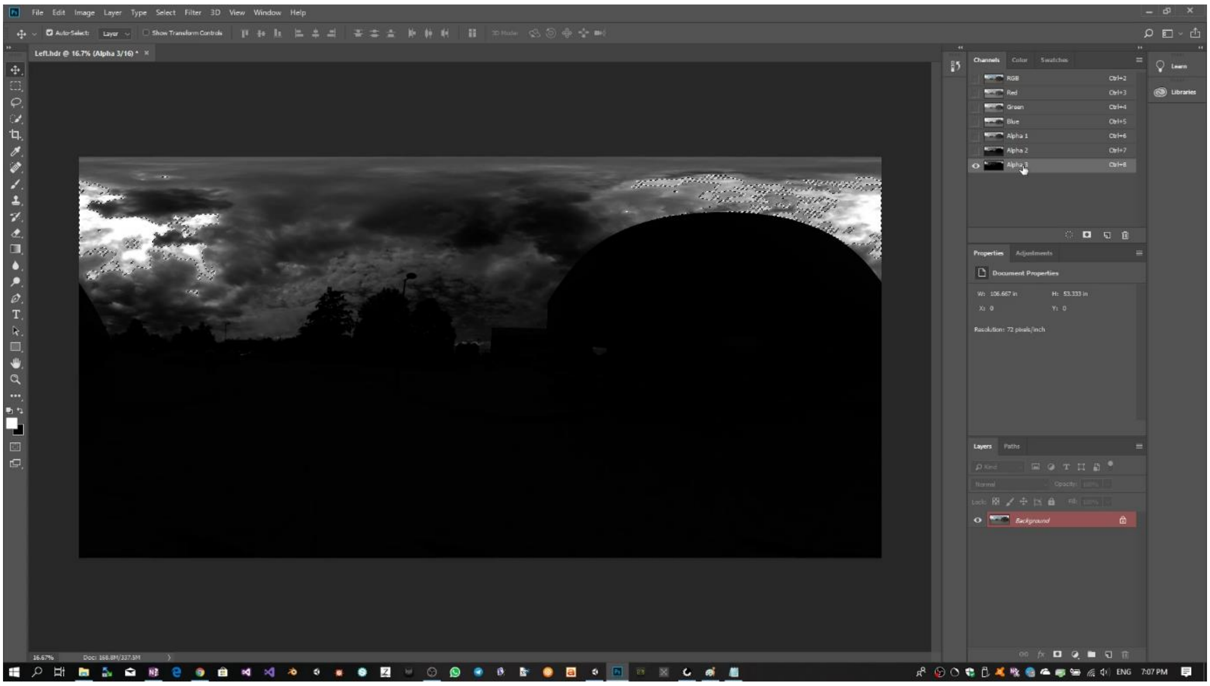


Figure 173 - Step 5. Repeating step 4 for many of the Alpha layers, to leave only details of bright areas in grayscale.

7.2.2.3 Creating the HDR system in Unity

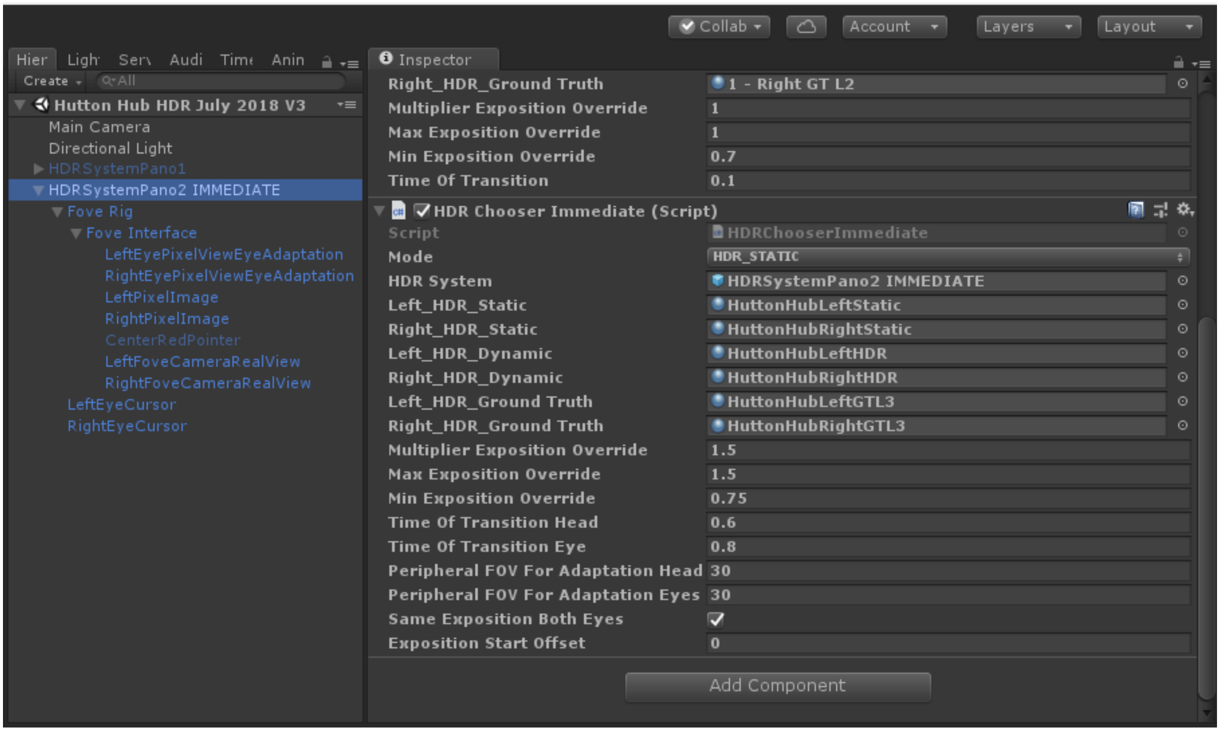


***Figure 174** - Snapshot of the Unity HDR modality chooser script, containing references to static HDR skyboxes, dynamic HDR skyboxes, and ground truth (GT) exposure skyboxes.*

Figure 174 shows a snapshot of my HDR system inside Unity. Furthermore, it presents the script to choose the desired HDR modality: *static HDR (HDR_STATIC)*, *dynamic HDR according to head rotations (HDR_DYNAMIC_HT)*, or *dynamic HDR according to eye's position (HDR_DYNAMIC_ET)*.

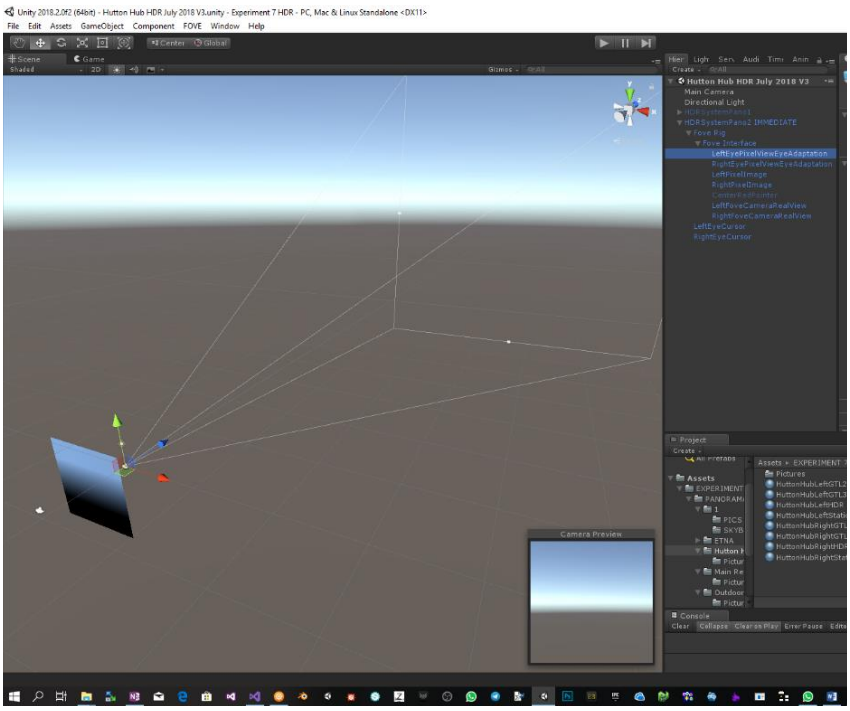


***Figure 175** - Snapshot of the layout of the virtual cameras used to evaluate the change of exposure of the ground truth panorama when the user moves head or eyes.*

Figure 175 shows a snapshot of the layout of the developed HDR system. The squared plane represents a render texture that takes information from the ground-truth panorama and shows the average approximated light intensity of the gazed area pointed by the viewer. I use it for debugging purposes, to better calibrate the system.

A closer look to the game objects of the scene is shown in **Figure 176**. In particular, *LeftEyePixelViewEyeAdaptation and LeftEyePixelViewEyeAdaptation* are two cameras that follow head or eye movements to evaluate the new light intensity of the gazed area by using the ground-truth panorama as a reference. *LeftPixelImage* and *RightPixelImage* are two squared planes that are used for debugging purposes to show what area of the ground-truth panorama is visualized by the two eye-adaptation cameras. *LeftFoveCameraRealView* and *RightFoveCameraRealView* are the virtual cameras of the FOVE interface, which visualize the converted LDR skybox and display it on the VR headset.

LeftEyeCursor and *RightEyeCursor* are two game objects used to display a small coloured cube on the estimated tracked position of the gazed area of the panorama. I consider this information to have an estimation of the accuracy of the optical tracking calibration procedure for each test user.

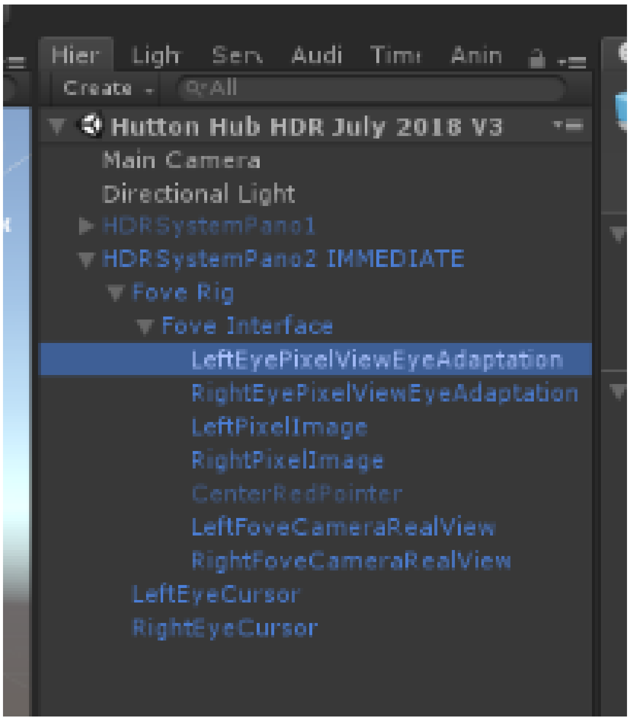


Figure 176 - Snapshot of the structure of the proposed HDR system in Unity.

Figure 177 presents a snapshot of the configuration that I decide to use for the FOVE virtual cameras, to handle the representation of the scene in LDR mode (note the “Allow HDR” checkbox disabled, to avoid the problems caused by the limitations of the FOVE display buffer).

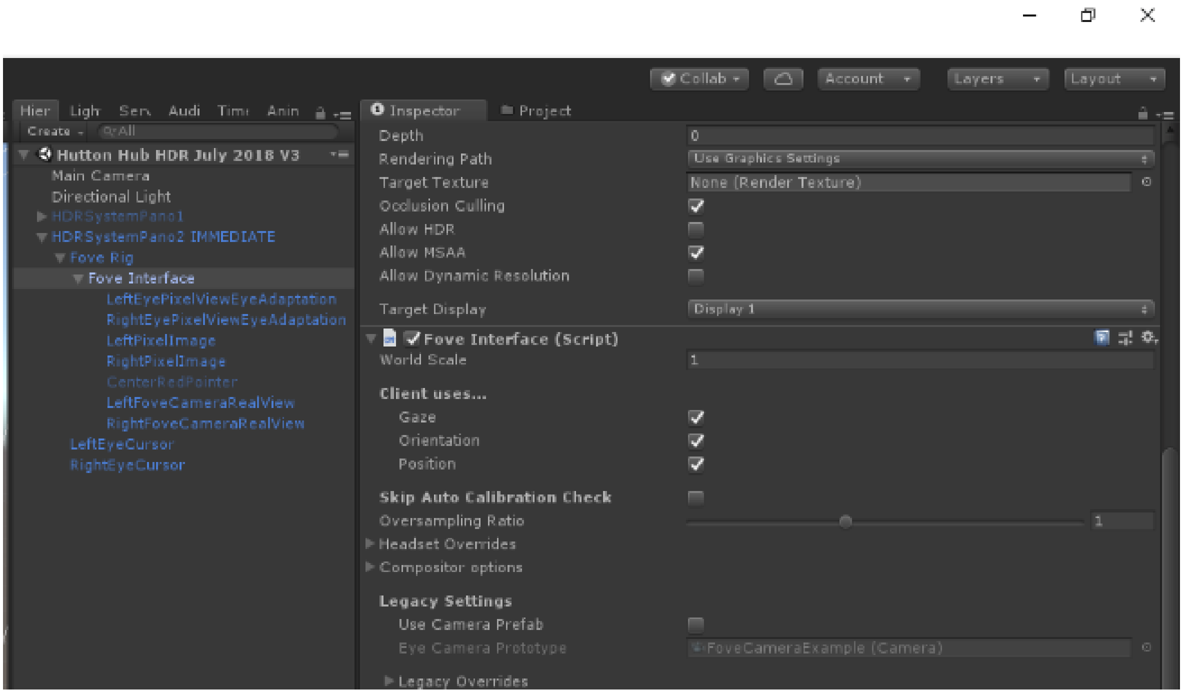


Figure 177 - Snapshot of the FOVE Interface configuration, with the HDR System V3 script, which controls the change in exposure of the real view according to the light determined by the eye adaptation cameras.

Figure 178 and **Figure 179** present snapshots of the skyboxes used by the HDR system. In particular, ground truth panorama (GT), static HDR panorama, and dynamic HDR panorama are shown.

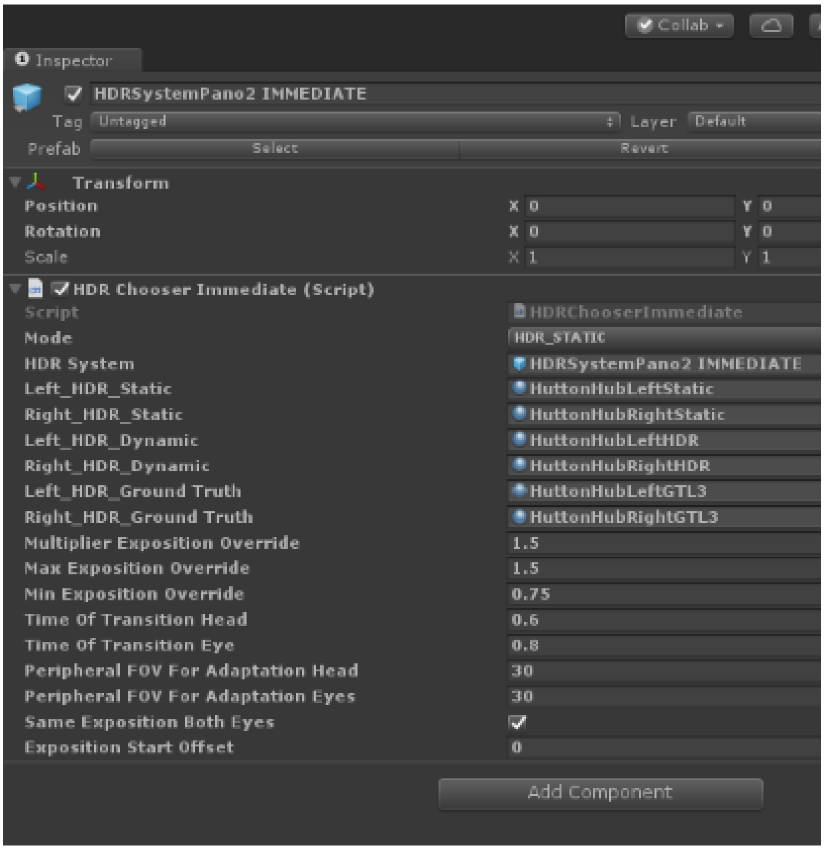


Figure 178 – Snapshot showing the references to the used skyboxes for the HDR system.

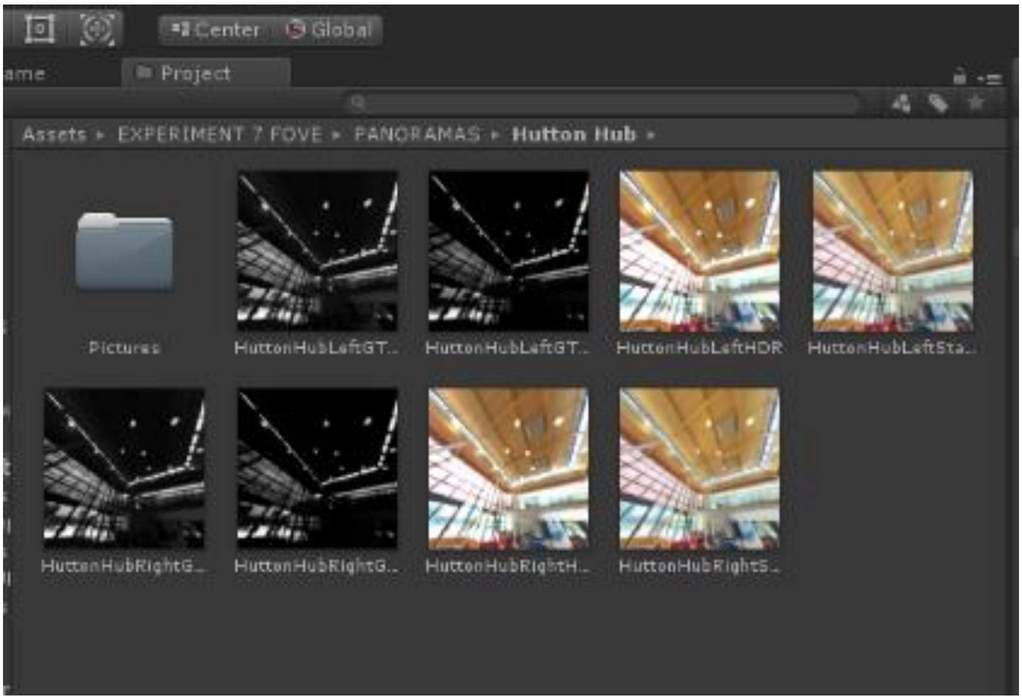


Figure 179 *- Snapshot showing the set of skyboxes for the HDR system, which are ground truth (GT), static HDR, and dynamic HDR.*

7.2.2.4 Testing the system

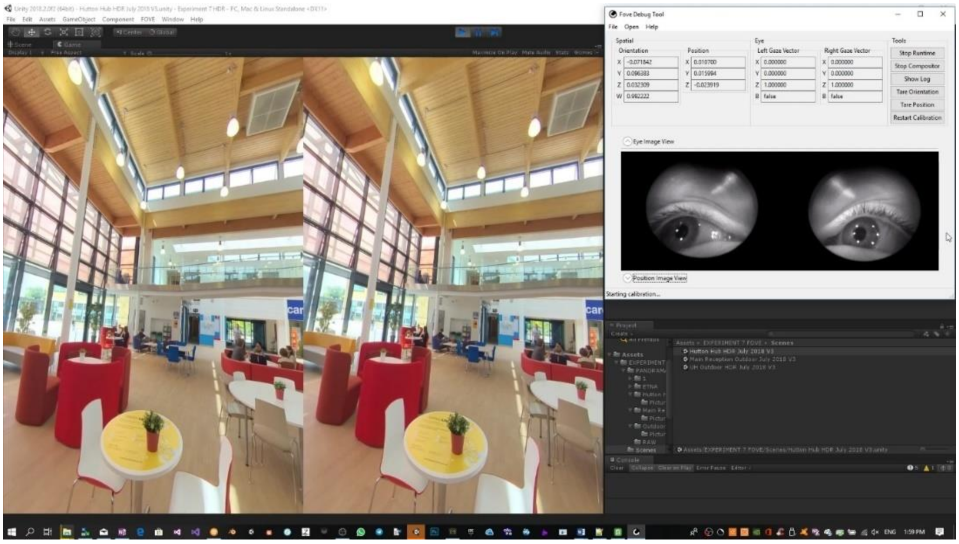


Figure 180 *– Unity implemented system combined with the eye tracking of the FOVE HMD, with snapshot of the viewer's eyes observing the Hutton Hub panorama.*

Figure 180 and **Figure 181** show snapshots of the test users visualizing the indoor and outdoor panoramas with the presented HDR modalities. In figure, details from the stream of the eye-tracking infrared cameras are shown, to check the correct monitoring of the viewer's eye for the dynamic eye-adapted HDR viewing of the scenes.

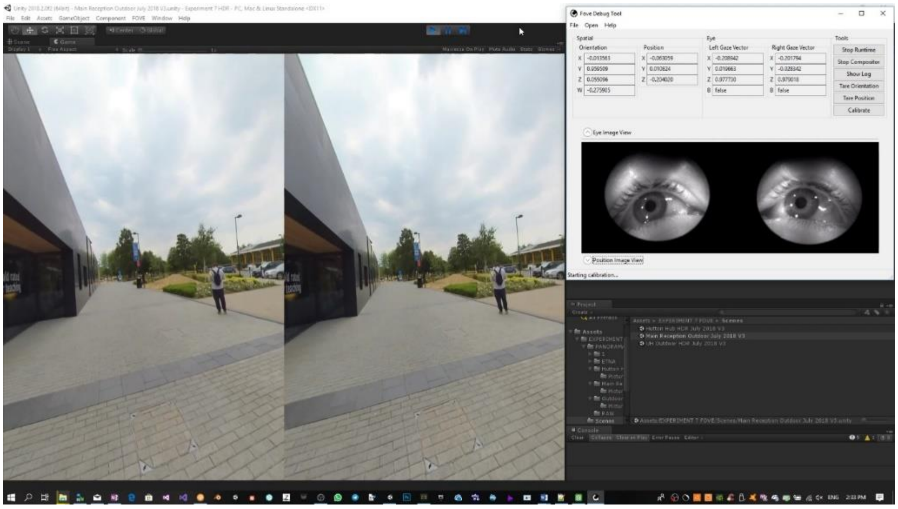


Figure 181 *- Unity developed system combined with eye tracking of FOVE, with Main Reception panorama.*

7.3 Procedure

To design the procedure of this user study it was not possible to reuse the results from the pilot tests of previous user studies as this evaluation uses different hardware and a new configuration of the virtual environment visualization. Therefore, a new pilot test was performed to choose the final optimal procedure to use.

7.3.1 Chosen Procedure

The following steps are taken to perform this usability evaluation:

1. The test user answers personal information questionnaire and agrees to the ethical forms on Google Forms;
2. More information on the purpose and procedure of this usability are explained, and the questionnaire on HDR modalities is read to increase awareness and to understand what aspects of the panorama should be carefully observed;
3. The test user wears a virtual reality headset that provides also eye tracking capabilities. Once adjusted, a calibration procedure is performed to detect eyes positions and movements as accurately as possible;
4. Once eyes calibration is completed, an outdoor or indoor environment is presented in virtual reality, starting with static HDR modality. Once the viewer explored every area of the panorama and feels ready, the modality is switched to HDR dynamic with head tracking. Once the new modality is completely explored, the viewer requests to switch modality and the HDR dynamic with eye tracking is presented.
5. To allow a better comparison of the modalities, all modalities are shown again in the same order, starting with the static HDR, then dynamic with head tracking, then dynamic with eye tracking.
6. Before completing the evaluation, a blue small cube is shown to the user. The cube continuously follows the gaze of the viewer according to data from the eye tracking system. The viewer is then requested to judge from 0 to 6 the accuracy of the position of the blue cube compared to the actual point of the screen that the eyes are watching.
7. Once that the headset has been removed, questions on Google Forms are digitally answered. Then, the same procedure is repeated for the other indoor or outdoor panorama.

7.3.2 VR Viewing Scheduling

Test user ID	Indoor panorama	Outdoor panorama
#1	First Trial	Second Trial
#2	Second Trial	First Trial
#3	First Trial	Second Trial
#4	Second Trial	First Trial
#5	First Trial	Second Trial
#6	Second Trial	First Trial
#7	First Trial	Second Trial
#8	Second Trial	First Trial
#9	First Trial	Second Trial
#10	Second Trial	First Trial
#11	First Trial	Second Trial
#12	Second Trial	First Trial
#13	First Trial	Second Trial
#14	Second Trial	First Trial
#15	First Trial	Second Trial
#16	Second Trial	First Trial
#17	First Trial	Second Trial
#18	Second Trial	First Trial
#20	First Trial	Second Trial
#21	Second Trial	First Trial
#22	First Trial	Second Trial
#23	Second Trial	First Trial
#24	First Trial	Second Trial

***Table 20**. Scheduling for the HDR usability evaluation. All participants were alternated to reduce possible biased answers.*

Table 20 shows the proposed scheduling for the visualization of the two HDR panoramas.

7.4 Results

For clarity, I will refer to static HDR as **S-HDR**, to dynamic HDR with head tracking as **HDR-HT**, and to dynamic HDR with eye tracking as **HDR-ET**.

7.4.1 Realism in terms of Presence

7.4.1.1 Sense of presence

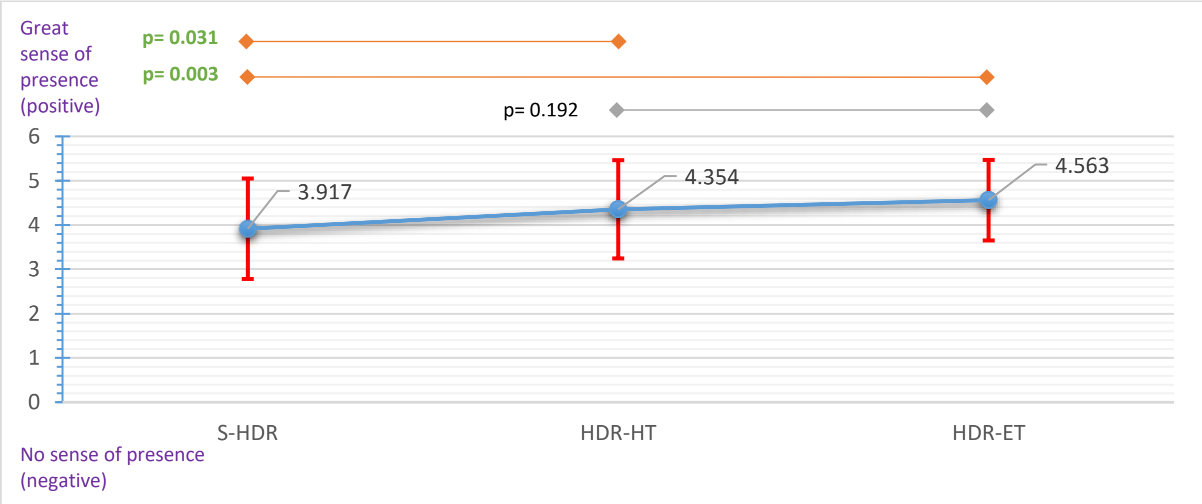


***Figure 182** - Significant differences for sense of presence between S-HDR and both HDR-HT and HDR-ET (See Appendix B – question HR1.1) [min = 0, Max = +6].*

Results reported no significant difference between indoor and outdoor environments when using the same HDR modality, with all achieving high scores. However, dynamic HDR-HT presented significant higher levels of sense of presence compared to S-HDR. Furthermore, also dynamic HDR-ET scored significantly higher ratings than S-HDR regardless of the panorama viewed (see **Figure 182**). No significant difference occurred between dynamic HDR-HT and dynamic HDR-ET.

7.4.1.2 Isolation

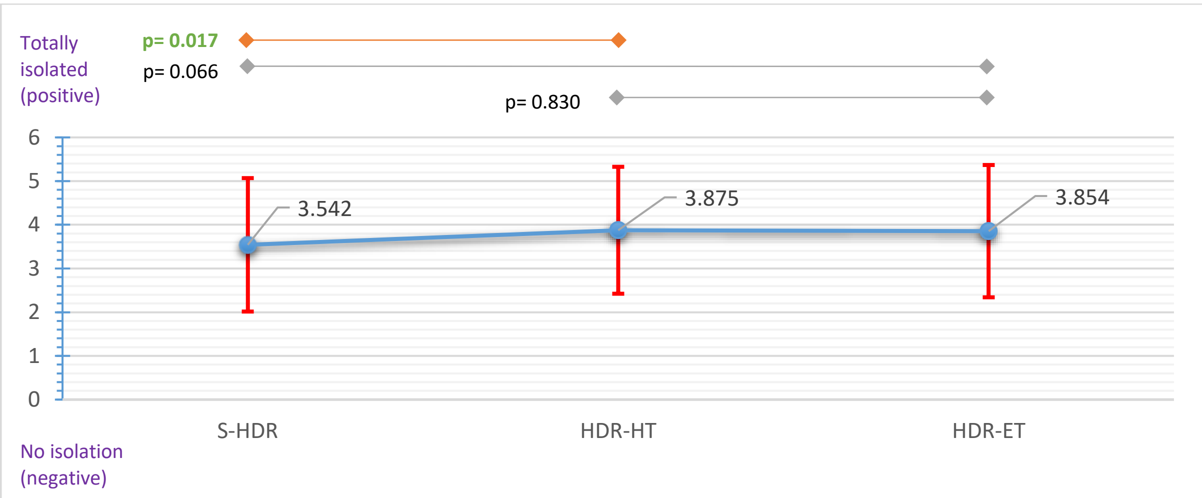


***Figure 183** - Significant differences for isolation between S-HDR and HDR-HT (See Appendix B – question HR1.2) [min = 0, Max = +6].*

No significant difference between indoor and outdoor environments was reported when using same HDR modality, with both achieving quite high scores. However, a significant difference resulted between S-HDR and HDR-HT, regardless of viewed panorama (see **Figure 183**).

7.4.1.3 Feeling like observing a real place

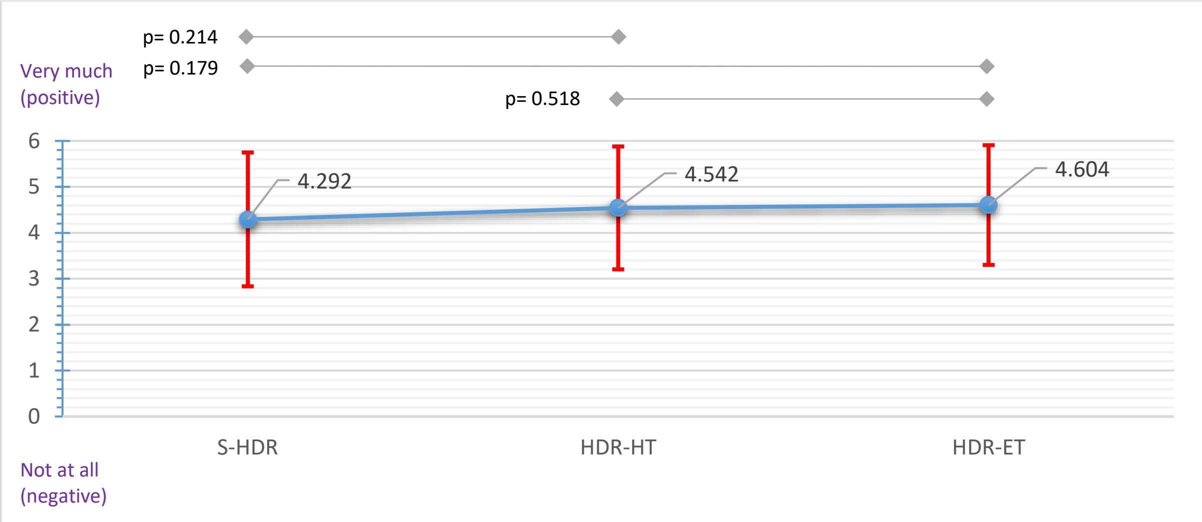


Figure 184 *- No significant difference resulted between the three HDR modalities, regardless of the panorama viewed. With all modalities the virtual place looked very similar to the real one (See Appendix B – question HR1.3) [min = 0, Max = +6].*

No significant difference was reported between indoor and outdoor, regardless of the HDR modality used, with both achieving very high scores. Furthermore, no significant difference between HDR modalities resulted, regardless of the panorama viewed (see **Figure 184**).

7.4.2 Realism in terms of Emotions

7.4.2.1 Likeliness

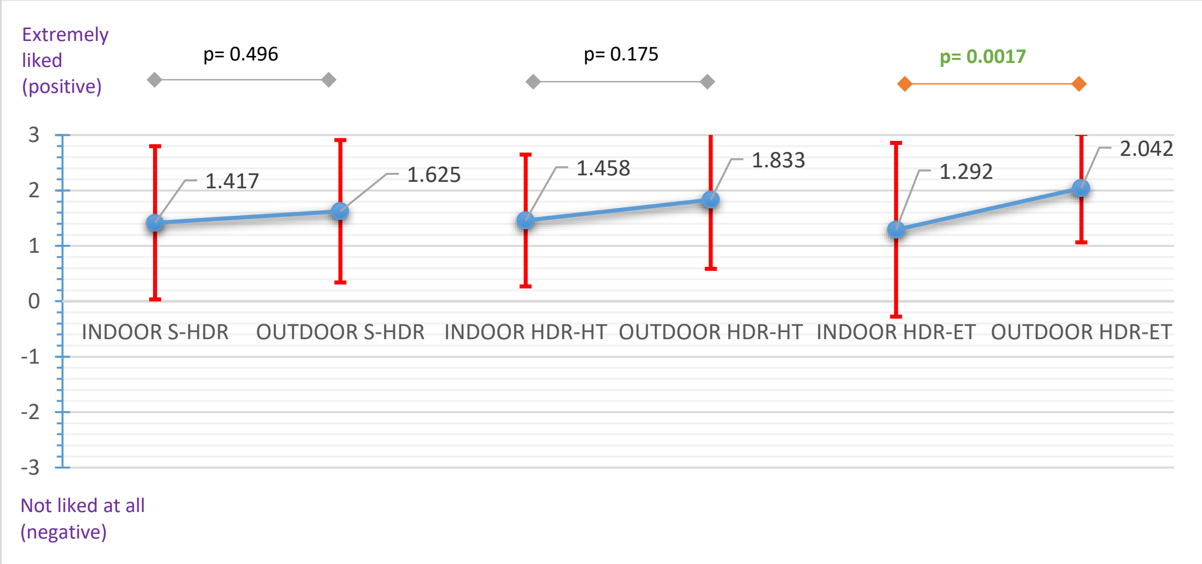


Figure 185 *- Significant difference between indoor and outdoor when using HDR-ET (See Appendix B – question HR2.1) [min = -3, Max = +3].*

When comparing indoor and outdoor using the same HDR modality, results reported that HDR-ET scored significantly higher ratings with the outdoor environment (see **Figure 185**).

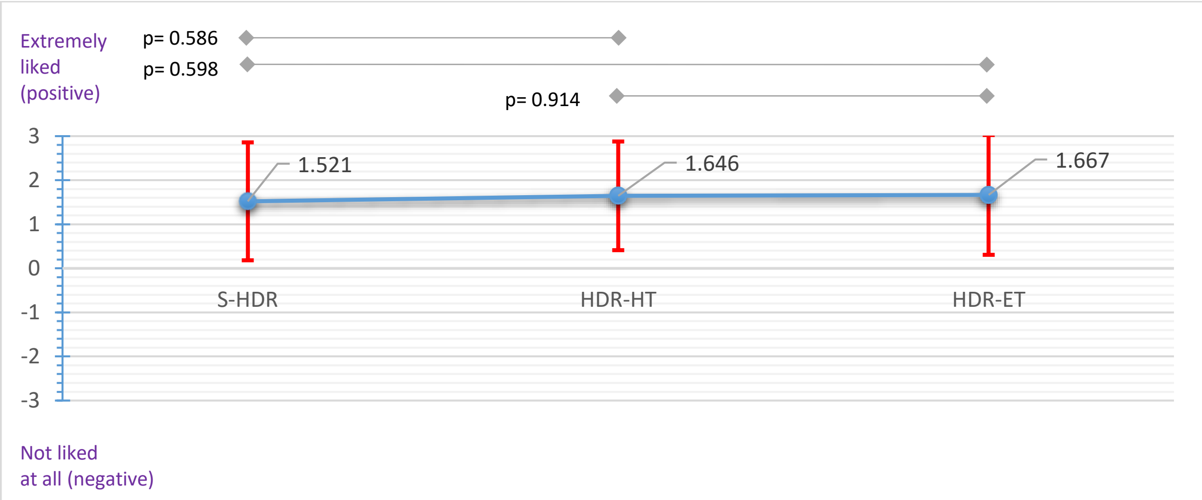


Figure 186 *- No significant difference resulted for likeliness between the three HDR modalities, regardless of the panorama viewed (See Appendix B – question HR2.1) [min = -3, Max = +3].*

Furthermore, no significant difference was reported between HDR modalities regardless of the panorama viewed (see **Figure 186**).

7.4.2.2 Happiness

No significant difference was reported between indoor and outdoor when using the same HDR modality, with scores quite high (averages between 1.208 and 1.708).

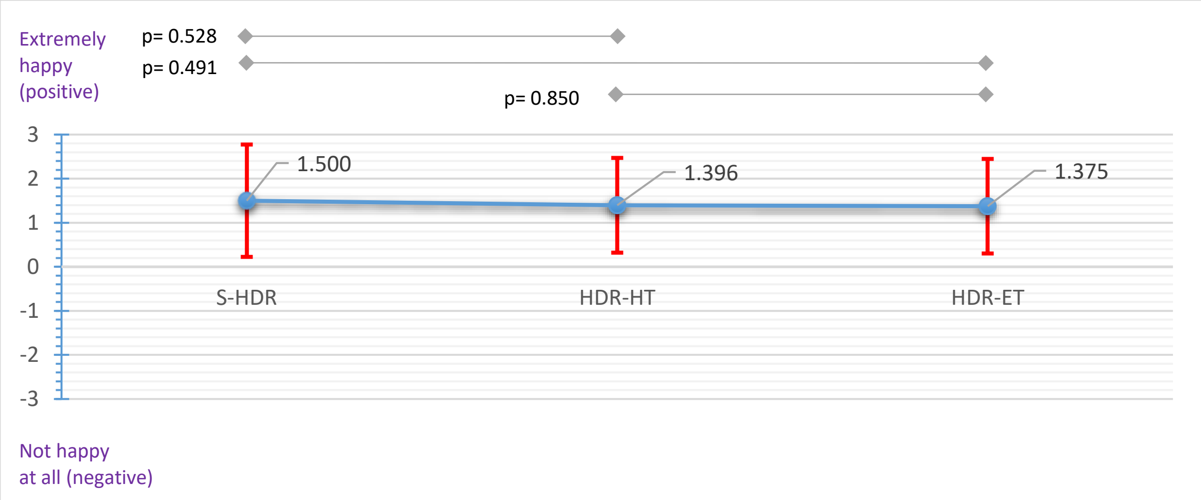


Figure 187 - No significant difference between HDR modalities in terms of happiness (See Appendix B – question HR2.2) [min = -3, Max = +3].

Furthermore, results show no significant difference between HDR modalities, regardless of the panorama viewed (see **Figure 187**).

7.4.2.3 Disappointment

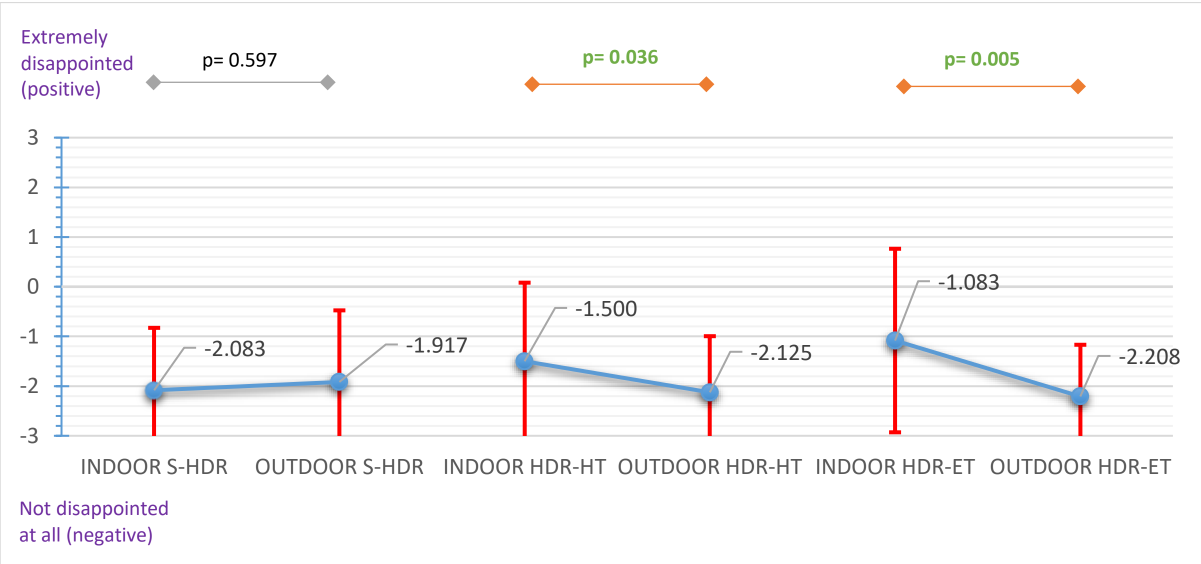


Figure 188 - Significant difference in terms of disappointment, between indoor and outdoor when using HDR-HT and HDR-ET, both with higher scores for indoor (See Appendix B – question HR2.3) [min = -3, Max = +3].

Results reported significantly different values between indoor and outdoor when using the same HDR modality. Test users rated indoor significantly more disappointing than outdoor when HDR-HT and HDR-ET were used (see **Figure 188**).

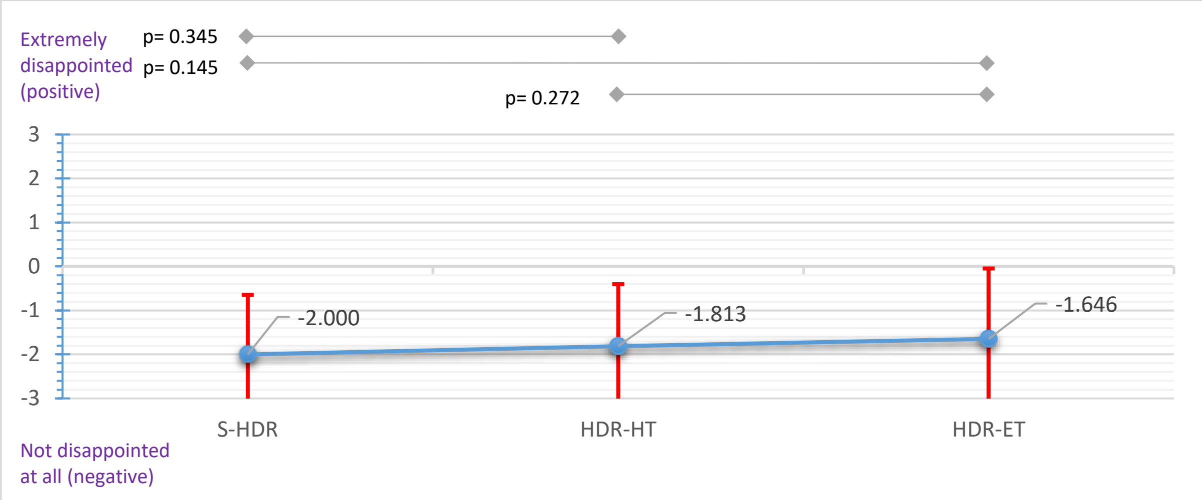


Figure 189 - No significant difference resulted between HDR modalities when viewing the same panorama (See Appendix B – question HR2.3) [min = -3, Max = +3].

Furthermore, no significant difference resulted in terms of disappointment between all HDR modalities when the same panorama was viewed (see **Figure 189**).

7.4.2.4 Desire to visit the remote environment in real life

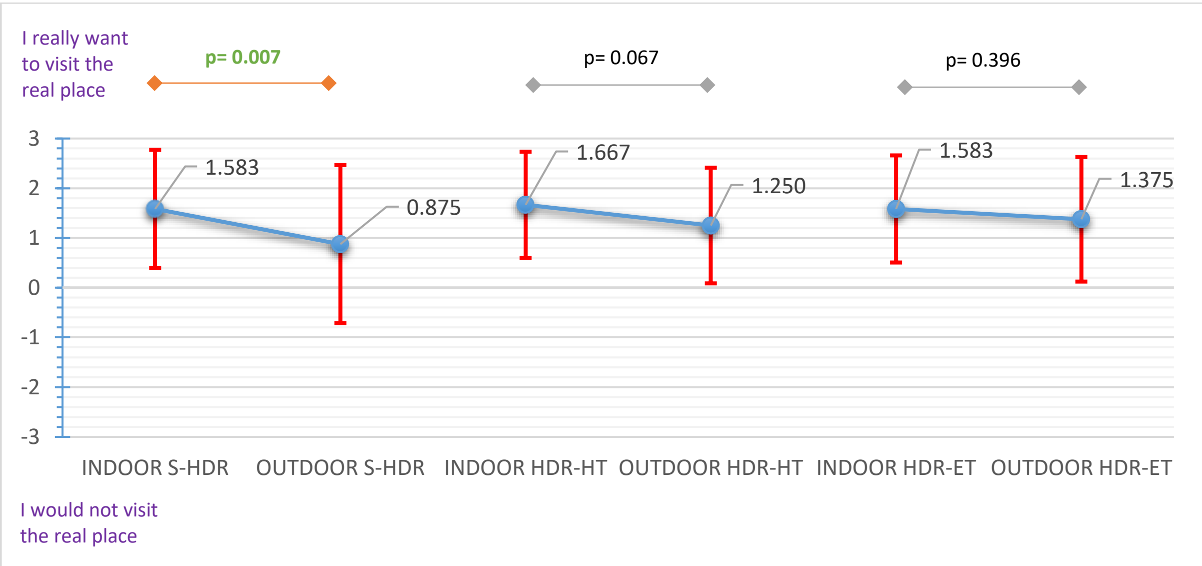


Figure 190 - Significant difference between indoor and outdoor when using S-HDR (See Appendix B – question HR2.4) [min = -3, Max = +3].

Results reported that a significant difference exists between indoor and outdoor only when using S-HDR, with indoor achieving higher scores (see **Figure 190**).

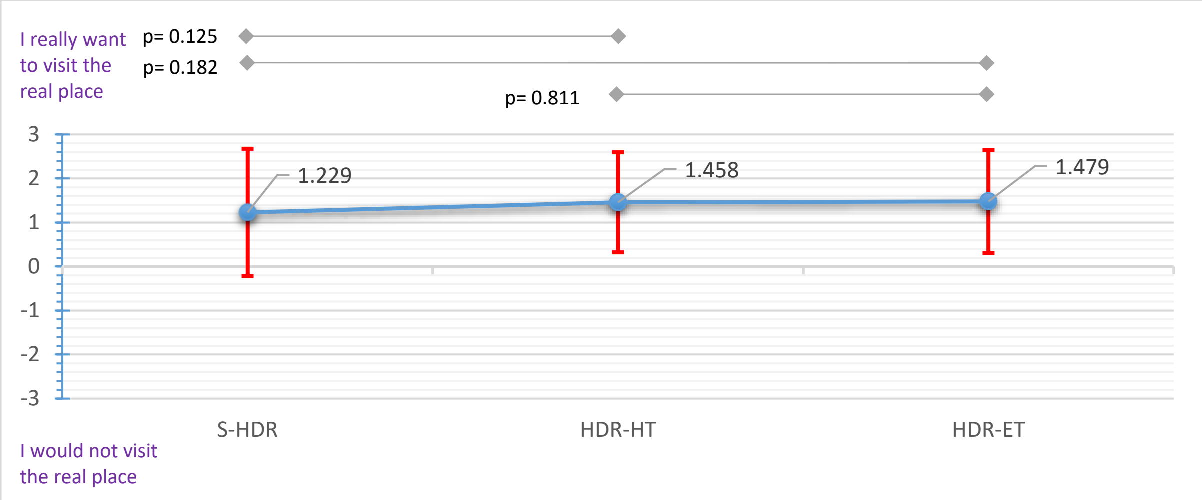


Figure 191 *- No significant difference was reported in terms of desire to visit the real place, for all HDR modalities when using same panorama (See Appendix B – question HR2.4) [min = -3, Max = +3].*

Furthermore, results reported no significant difference between all HDR modalities when using the same panorama (see **Figure 191**).

7.4.2.5 Positive mood change after viewing in VR

Results reported no significant difference between indoor and outdoor when using the same HDR modality, with all results highlighting a very slight positive mood change.

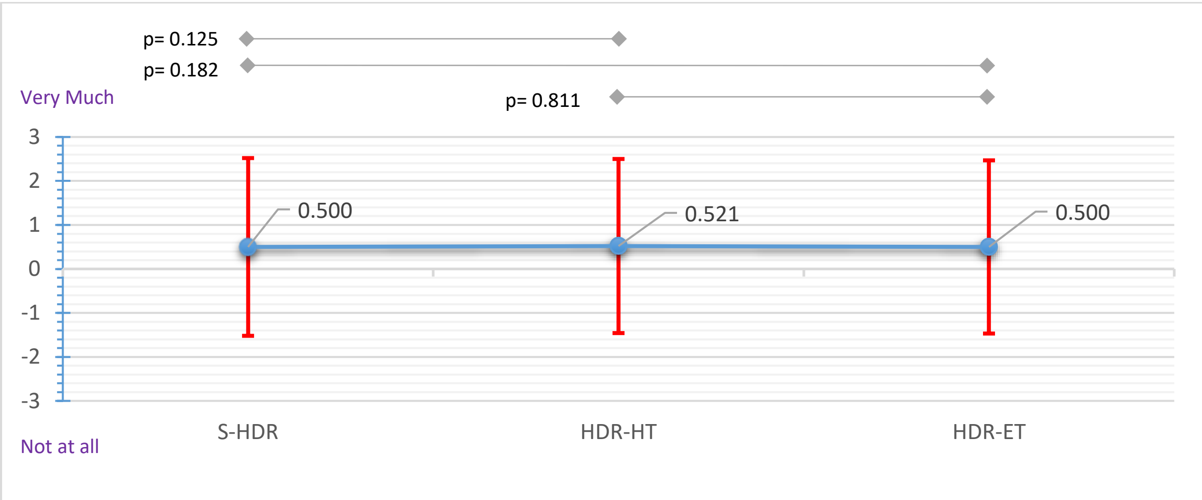


Figure 192 *- No significant difference in terms of positive mood change, between HDR modalities when observing the same panorama (See Appendix B – question HR2.5) [min = -3, Max = +3].*

Furthermore, no significant difference was reported in terms of positive mood change when observing the same panorama with all HDR modalities (see **Figure 192**).

7.4.2.6 Motivation to use VR for remote observation

Results reported no significant difference between indoor and outdoor when using the same HDR modality, with all results highlighting a very high motivation in the use of VR for remote observation.

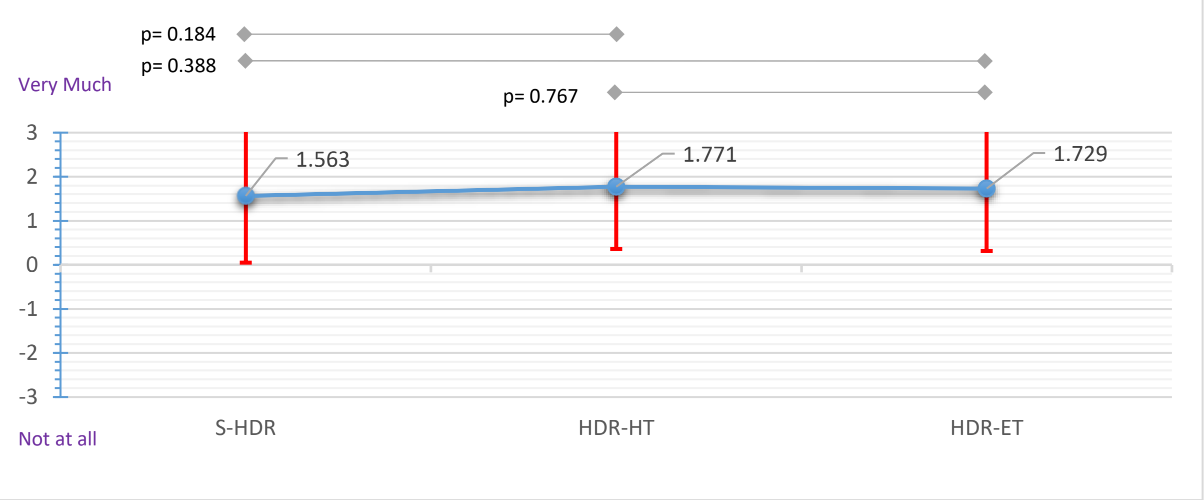


Figure 193 *- No significant difference in terms of motivation to use VR for remote observation, between all HDR modalities when observing the same panorama (See Appendix B – question HR2.6) [min = -3, Max = +3].*

Furthermore, no significant difference was found between HDR modalities when viewing the same panorama (see **Figure 193**).

7.4.3 Realism in terms of Virtual Vision vs Real Life

7.4.3.1 Natural eye adaptation

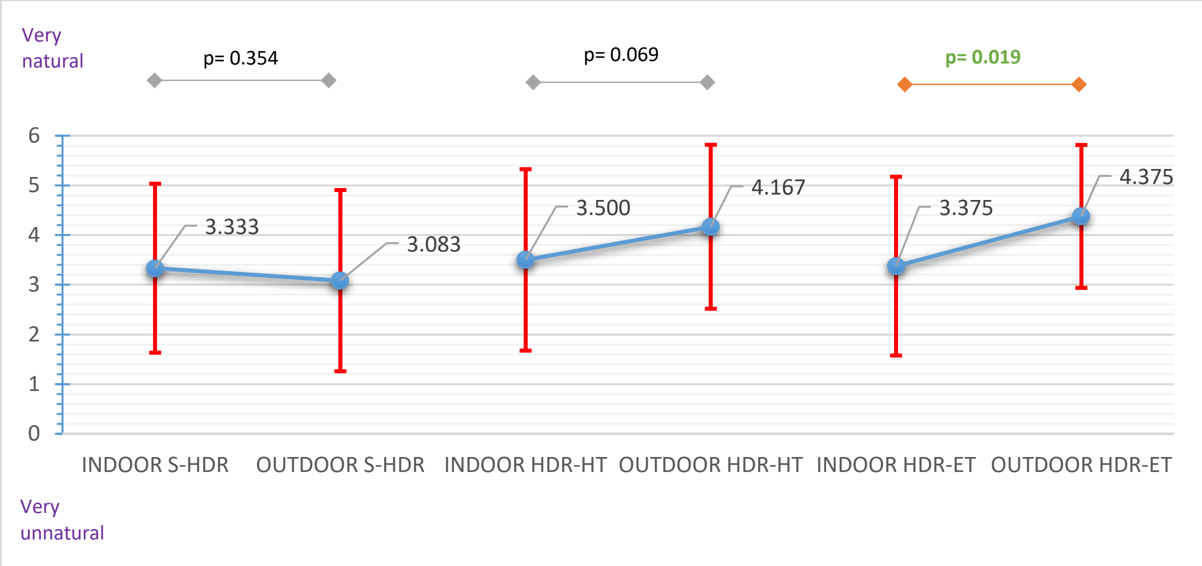


Figure 194 *- Significant difference in terms of natural eye adaptation between indoor and outdoor when using HDR-ET (See Appendix B – question HR3.1) [min = 0, Max = 6].*

Results showed a significant difference between indoor and outdoor when using HDR-ET, with outdoors getting the highest scores (see **Figure 194**).

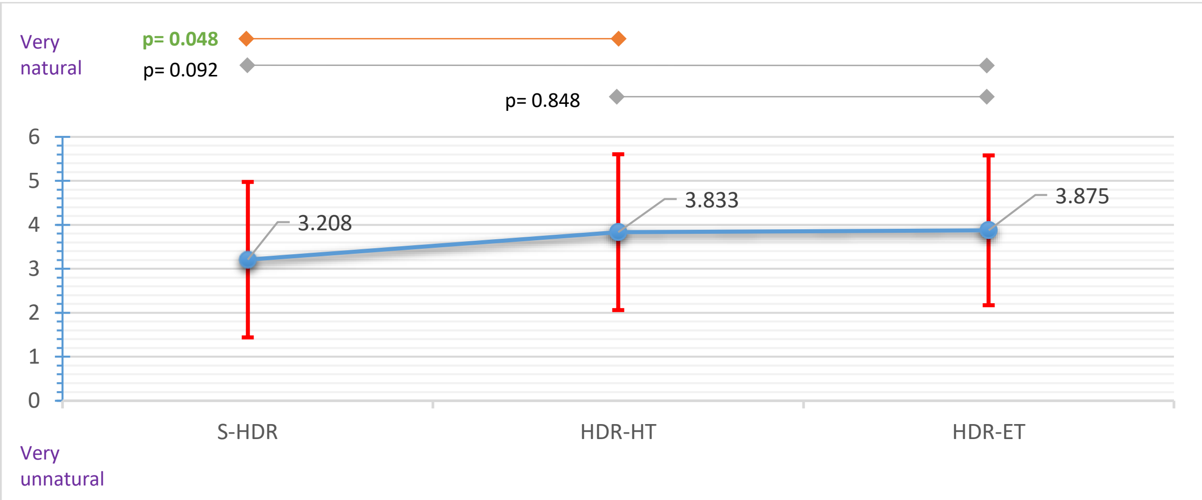


Figure 195 *- Significant difference between S-HDR and HDR-HT when viewing same panorama (See Appendix B – question HR3.1) [min = 0, Max = 6].*

Furthermore, results reported a significant difference between S-HDR and HDR-HT regardless of the panorama viewed, with the latter performing better in terms of natural eye adaptation (see **Figure 195**).

7.4.3.2 Virtual lights vs Real Life lights

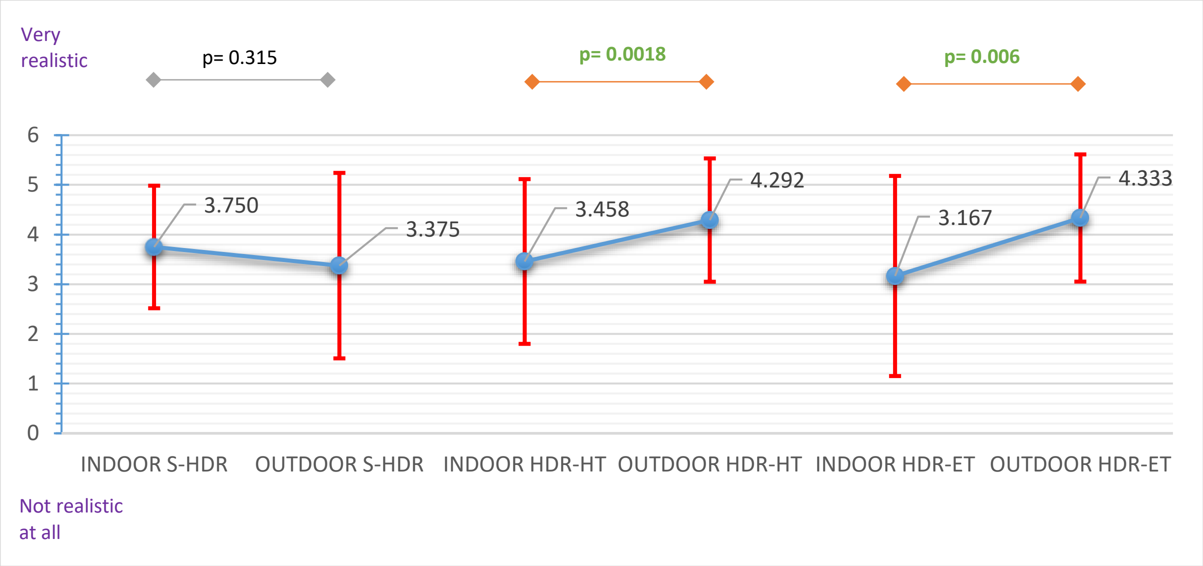


***Figure 196** - Significant differences for realism of virtual lights, between indoor and outdoor when using the same HDR-HT or HDR-ET modality, with outdoor performing better (See Appendix B – question HR3.2) [min = 0, Max = 6].*

Results reported significant differences for realism of virtual lights, between indoor and outdoor when using the same HDR-HT or HDR-ET modality, with outdoor achieving better realism for lights in the remote observed environment than indoor (see **Figure 196**).

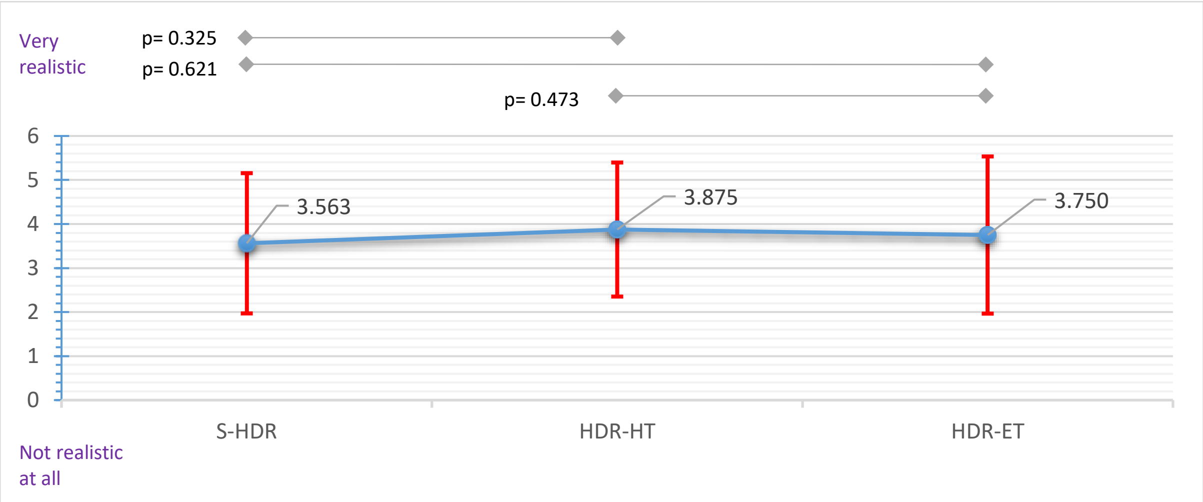


***Figure 197** - No significant difference for realism of virtual lights, between HDR modalities when viewing the same panorama (See Appendix B – question HR3.2) [min = 0, Max = 6].*

Furthermore, no significant difference for realism of virtual lights was found between HDR modalities when viewing the same panorama, with all of them achieving quite high scores (see **Figure 197**).

7.4.3.3 Color realism

No significant difference was found between indoor and outdoor when using the same HDR modality, with both panoramas achieving very high scores (with averages between 3.917 and 5.083).

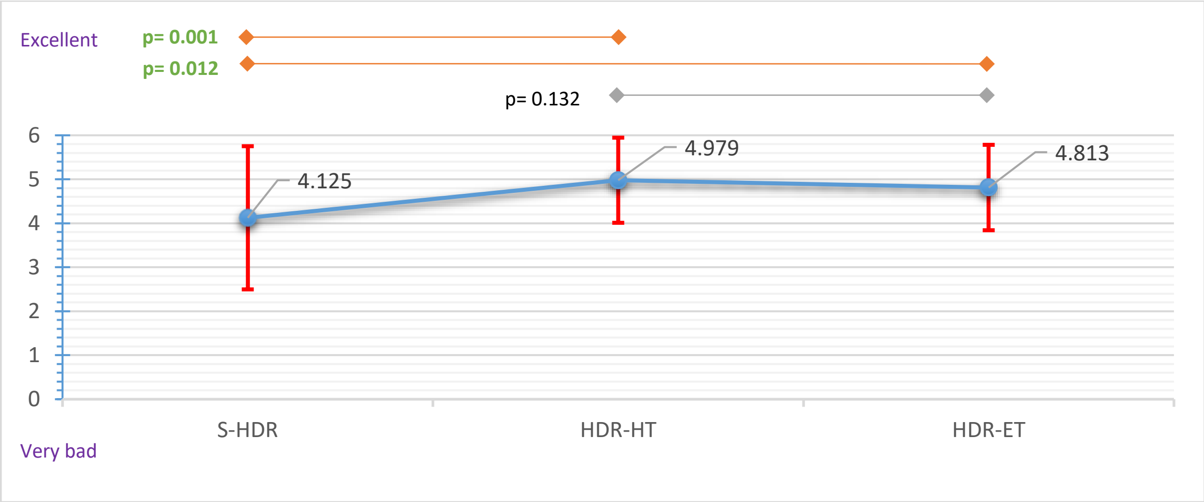


Figure 198 *- Significant difference between S-HDR and both HDR-HT and HDR-ET, with S-HDR performing worse (See Appendix B – question HR3.3) [min = 0, Max = 6].*

Furthermore, significant differences were found between S-HDR and both HDR-HT and HDR-ET modalities when viewing the same panorama, with S-HDR performing worse than the other two (see **Figure 198**).

7.4.3.4 Realism of Darkness

No significant difference was found between indoor and outdoor when using the same HDR modality, with both panoramas achieving very high scores (with averages between 3.292 and 4.167).

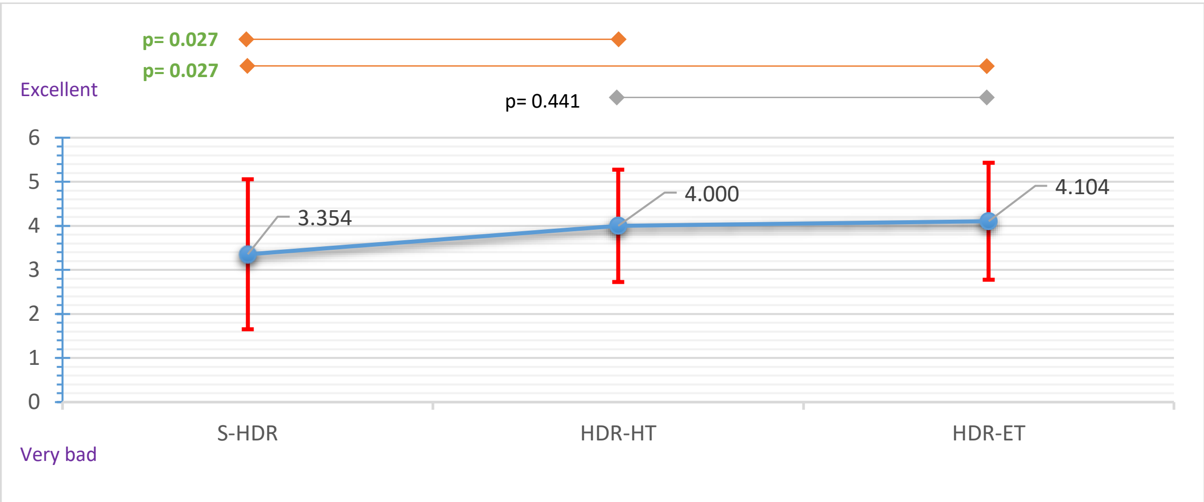


Figure 199 *- Significant differences for darkness realism between S-HDR and both HDR-HT and HDR-ET (See Appendix B – question HR3.4) [min = 0, Max = 6].*

Furthermore, results reported significant differences for realism of darkness between S-HDR and both HDR-HT and HDR-ET when viewing the same environment, with S-HDR performing worse than the other two (see **Figure 199**).

7.4.3.5 Realism of Brightness

No significant difference was found between indoor and outdoor when using the same HDR modality, with both panoramas achieving very high scores (with averages between 3.583 and 4.833).

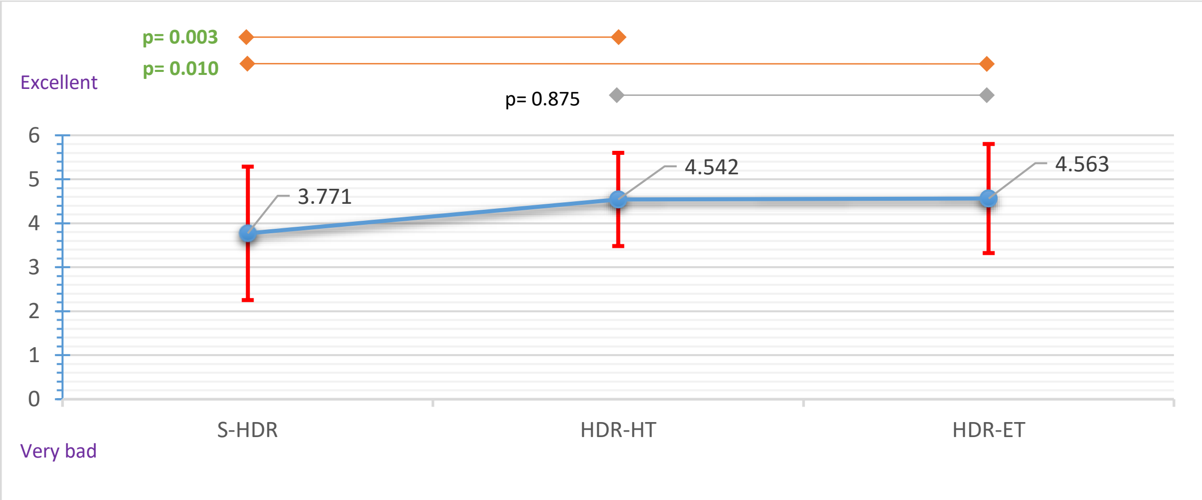


Figure 200 *- Significant difference between S-HDR and both HDR-HT and HDR-ET when observing the same panorama (See Appendix B – question HR3.5) [min = 0, Max = 6].*

Furthermore, significant differences were found between S-HDR and both HDR-HT and HDR-ET, with S-HDR performing worse than the other two, when observing the same environment (see **Figure 200**).

7.4.3.6 Virtual vision compared to Real vision

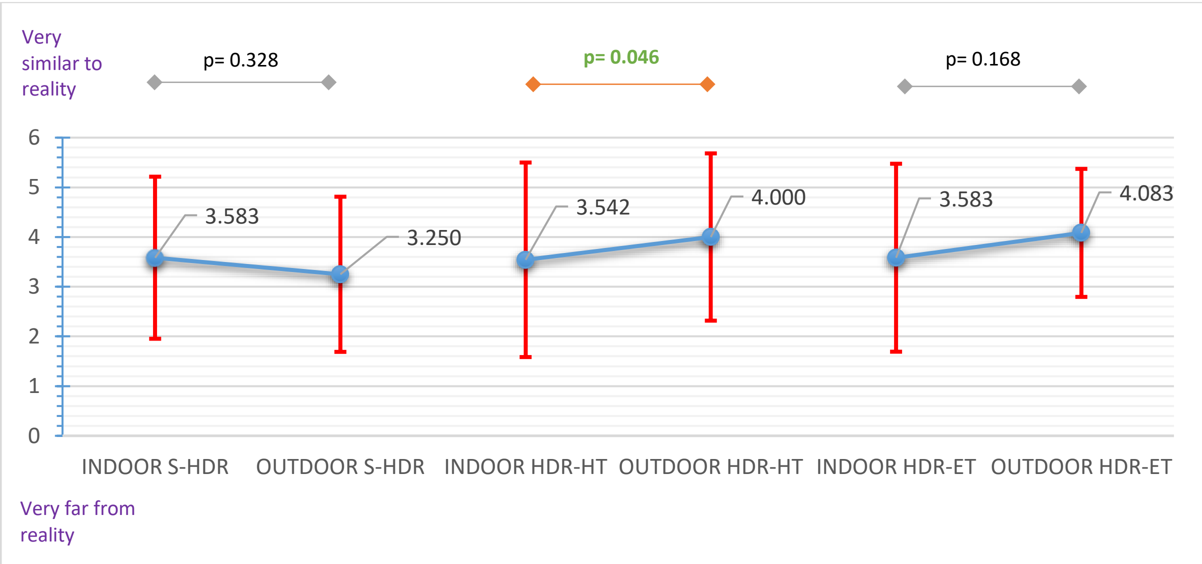


Figure 201 *- Significant difference in terms of realism of the virtual vision compared to real life between indoor and outdoor (See Appendix B – question HR3.6) [min = 0, Max = 6].*

Results showed a significant difference in terms of realism of the virtual vision compared to real life between indoor and outdoor when using same HDR modality. In detail, outdoor performed significantly better than indoor when using HDR-HT (see **Figure 201**).

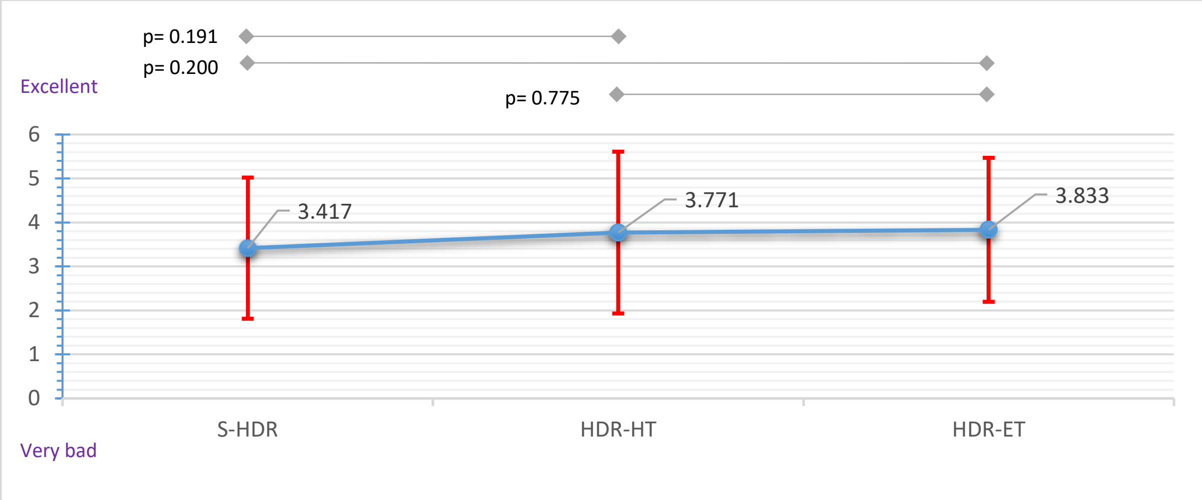


Figure 202 *- No significant difference between HDR modalities when viewing the same panorama, in terms of realism of virtual vision compared to real life vision (See Appendix B – question HR3.6) [min = 0, Max = 6].*

Furthermore, no significant difference was found between HDR modalities when viewing the same panorama, with all achieving high scores (see **Figure 202**).

7.4.4 Comfort

7.4.4.1 Level of comfort

No significant difference was found between indoor and outdoor when using the same HDR modality, with both panoramas achieving very high scores (with averages between 4.042 and 4.792).

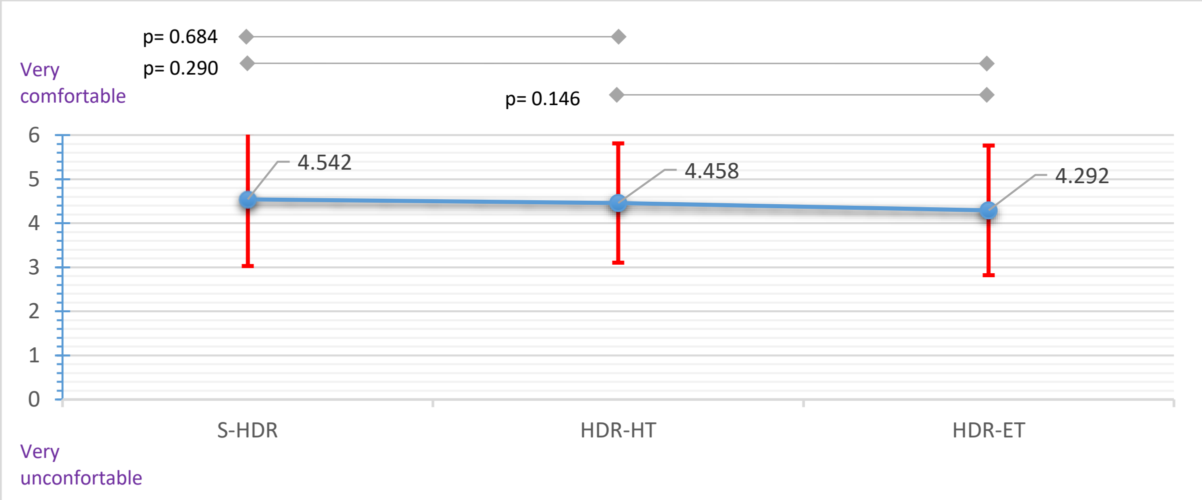


Figure 203 *- No significant difference for comfort between HDR modalities when viewing the same panorama (See Appendix B – question HC1.1) [min = 0, Max = 6].*

Furthermore, no significant difference was found between HDR modalities for comfort when viewing the same panorama, with all achieving very high scores (see **Figure 203**).

7.4.4.2 Time before feeling discomfort

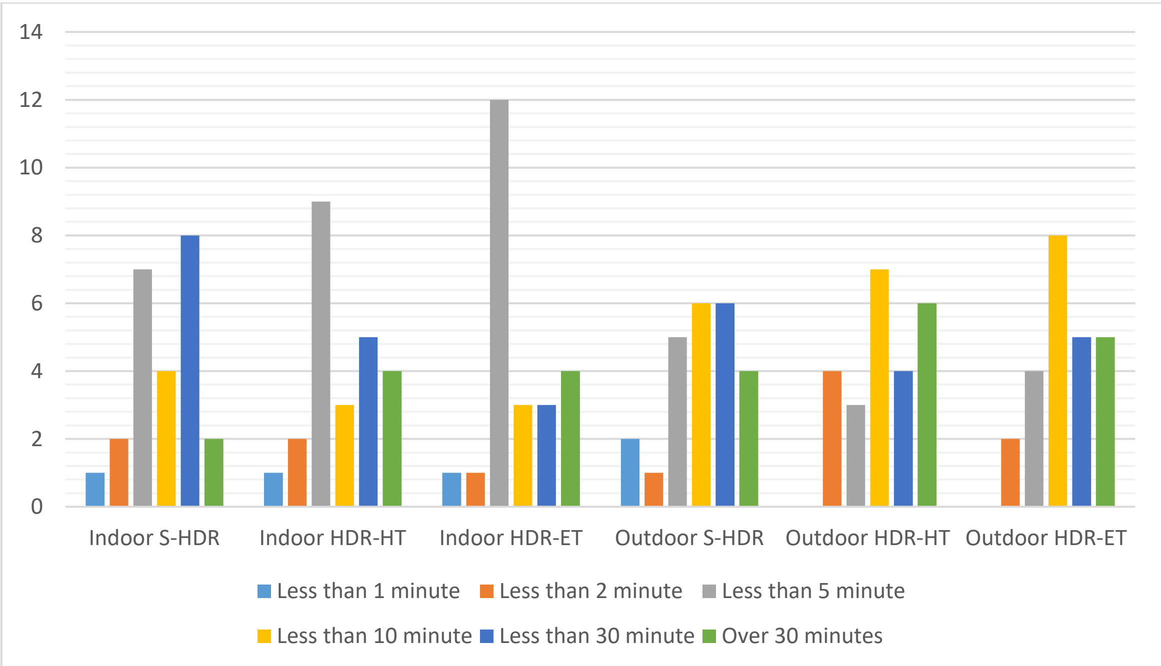


Figure 204 *- Number of test users that chose each answer to question <For how long you think you can watch the panorama before getting tired, using this setup?> (See Appendix B – question HC1.2)*

Results on time before feeling discomfort for each HDR modality and panorama are reported in **Figure 204**.

7.4.4.3 Perceptual issues

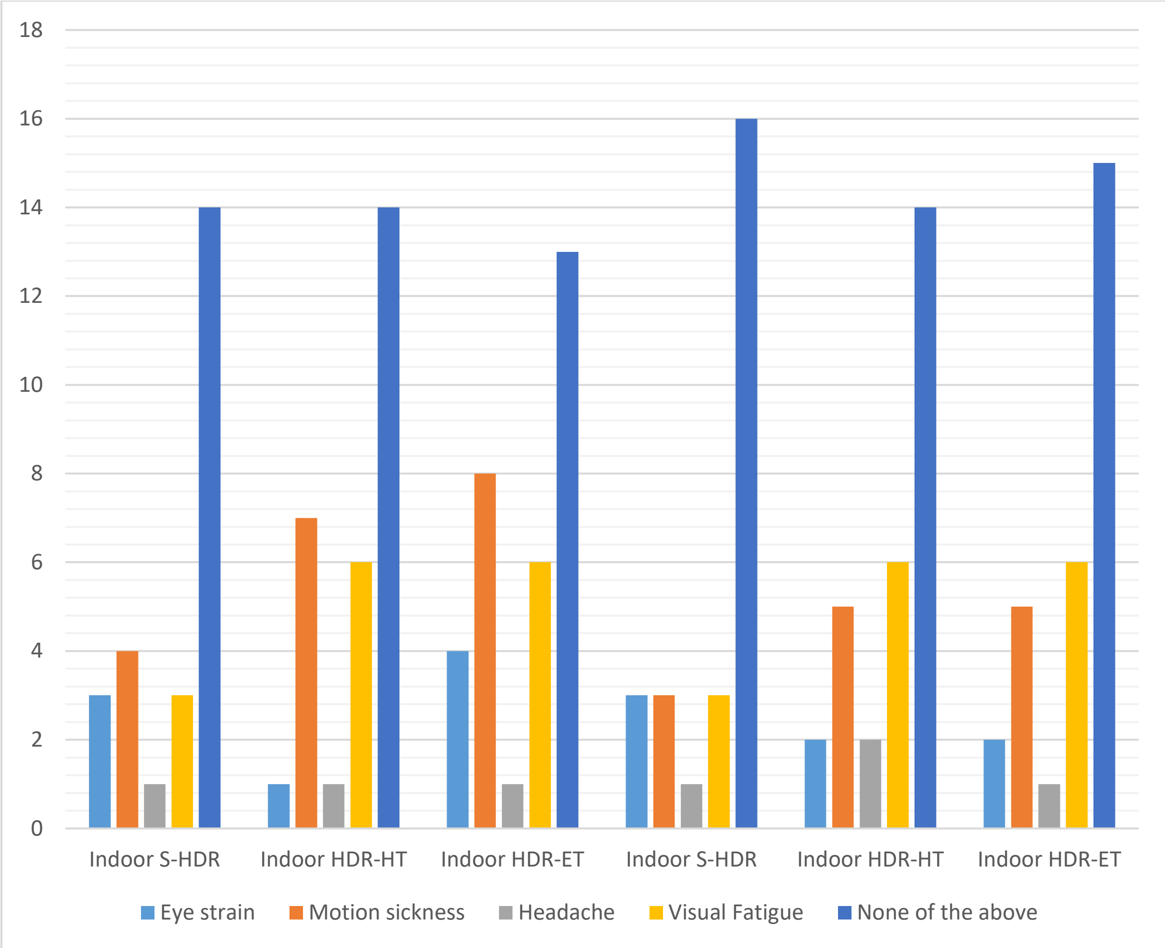


Figure 205 *- Number of test users that chose each answer to question <Did you perceive any of the following for each setup? Check all the perceptual issues that apply>*
(See Appendix B – question HC1.3)

Results on occurred perceptual issues while using Mobile VR for each HDR modality and observed panorama are reported in **Figure 205**.

7.4.5 Depth perception

7.4.5.1 3D accuracy

No significant difference was found between indoor and outdoor when using the same HDR modality, with both panoramas achieving very high scores (with averages between 4.125 and 4.750).

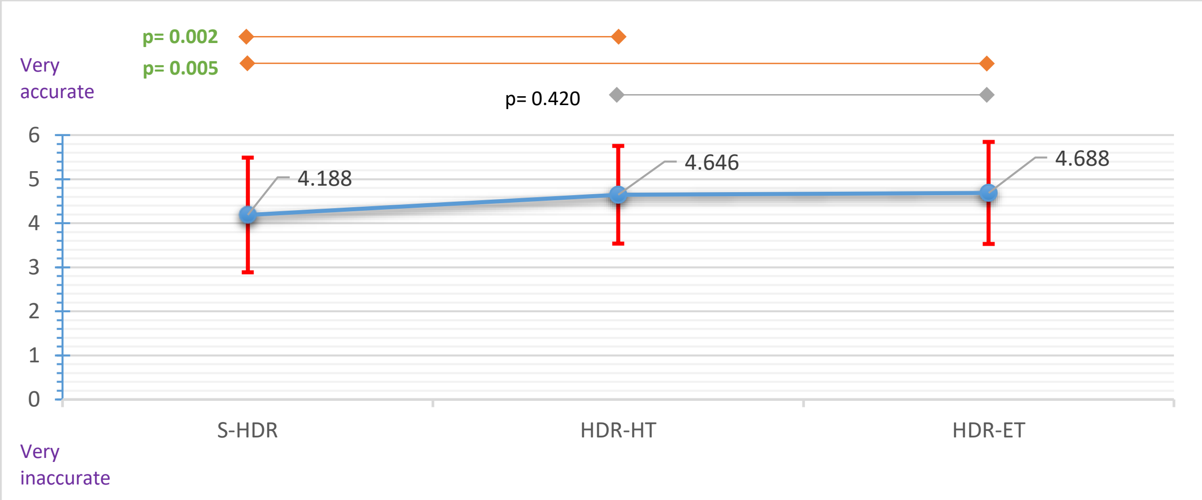


Figure 206 *- Significant differences for 3D accuracy between S-HDR and both HDR-HT and HDR-ET when viewing the same panorama (See Appendix B – question HD1.1) [min = 0, Max = 6].*

Furthermore, significant differences between S-HDR and both HDR-HT and HDR-ET were found, with S-HDR performing worse than the other two in terms of 3D accuracy (see **Figure 206**).

7.4.5.2 Lights and Shadows contribution to 3D

No significant difference was found between indoor and outdoor when using the same HDR modality, with both panoramas achieving very high scores (with averages between 0.583 and 1.917).

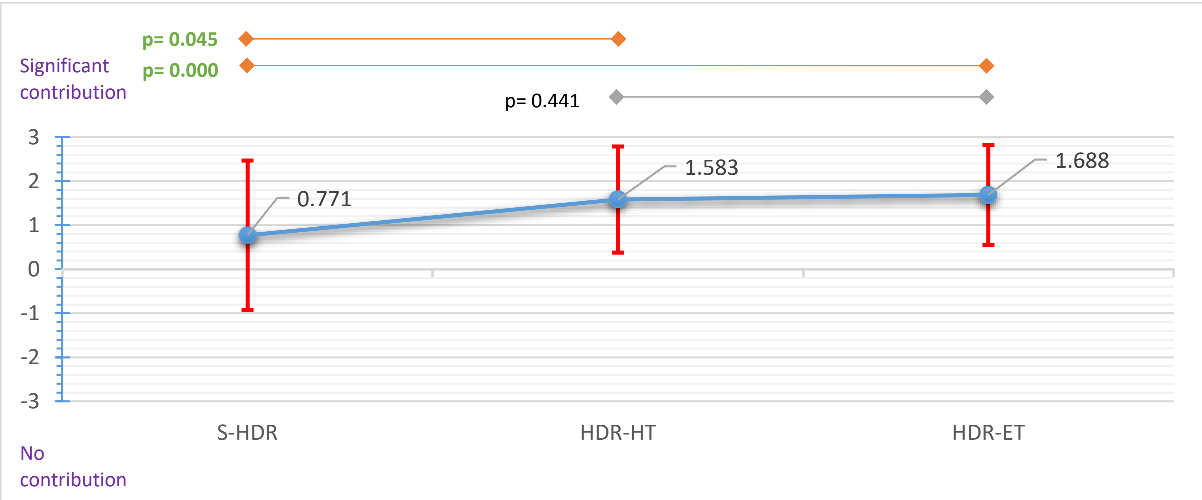


Figure 207 *- Significant differences between S-HDR and both HDR-HT and HDR-ET in terms of lights and shadows contribution to 3D, when viewing same panorama (See Appendix B – question HD1.2) [min = -3, Max = +3].*

Furthermore, significant differences were found between S-HDR and both HDR-HT and HDR-ET in terms of lights and shadows contribution to 3D when viewing same panorama, with S-HDR performing worse than the other two (see **Figure 207**).

7.4.5.3 Realistic space perception

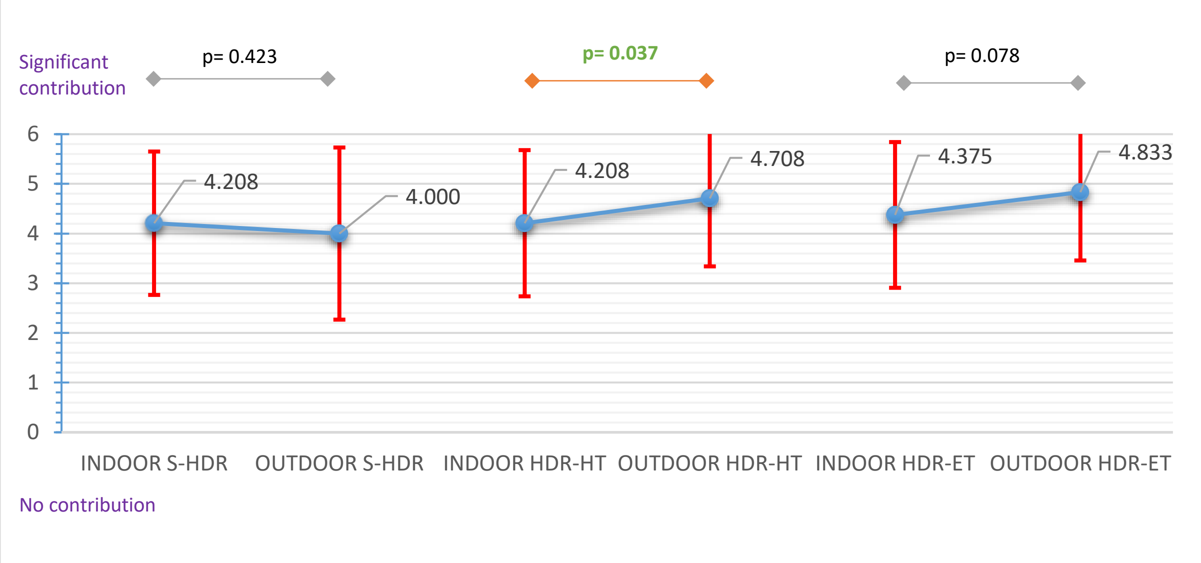


Figure 208 *- Significant difference between indoor and outdoor in terms of realistic space perception, when using the same HDR modality (See Appendix B – question HD1.3) [min = 0, Max = 6].*

Results showed that indoor and outdoor are significantly different when using HDR-HT modality, with outdoor achieving higher scores (see **Figure 208**).

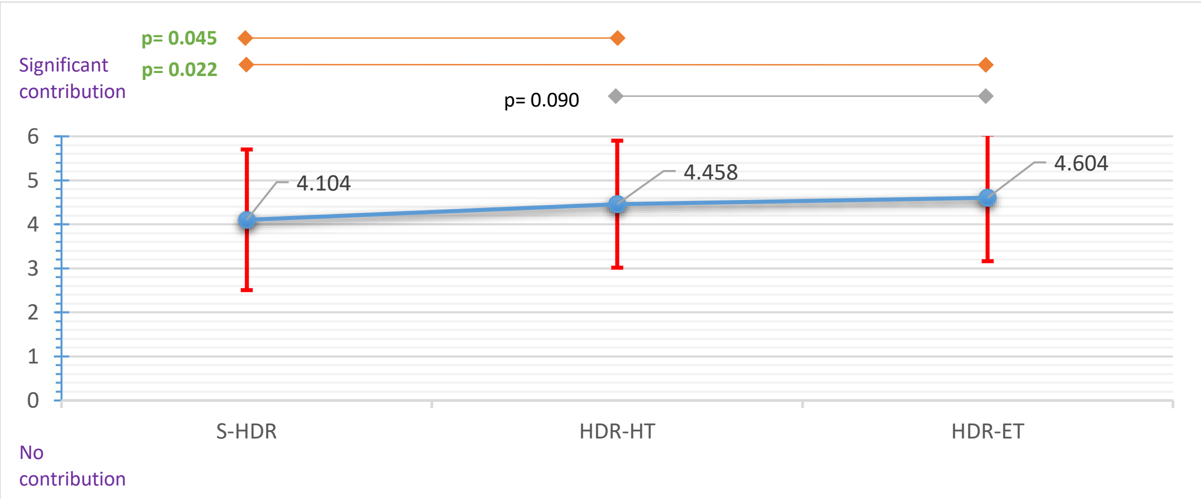


Figure 209 *- Significant difference between S-HDR and both HDR-HT and HDR-ET in terms of realistic space perception, when viewing the same panorama (See Appendix B – question HD1.3) [min = 0, Max = 6].*

7.4.5.4 Perceived size of the environment

No significant difference was found between indoor and outdoor when using the same HDR modality, with both panoramas achieving very high scores (with averages between 4.167 and 4.708).

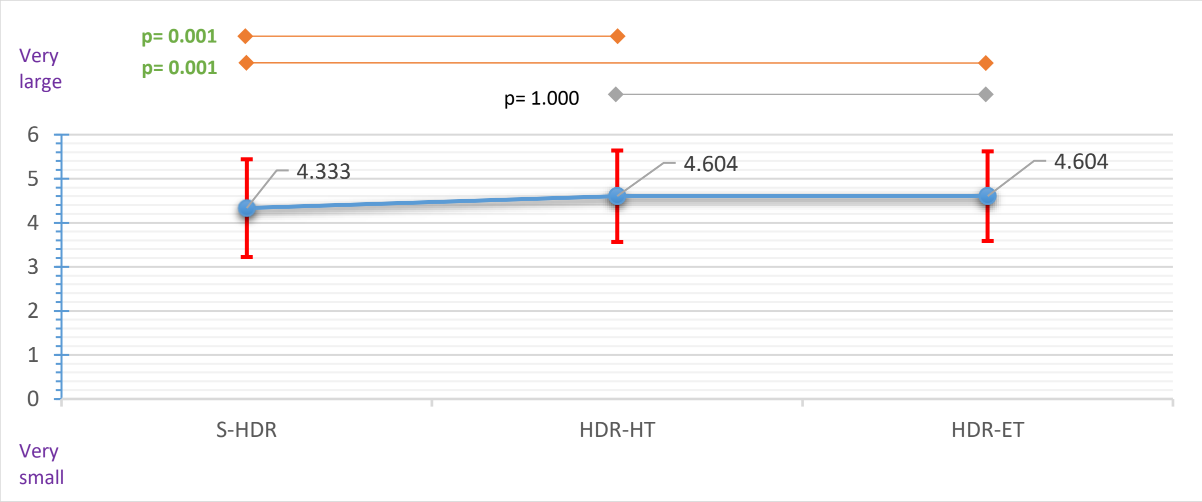


Figure 210 *- Significant difference between S-HDR and both HDR-HT and HDR-ET in terms of perceived size of environment (See Appendix B – question HD1.4) [min = 0, Max = 6].*

Furthermore, results showed significant differences between S-HDR and both HDR-HT and HDR-ET in terms of perceived size of the environment, when viewing the same panorama. S-HDR showed the room slightly smaller than the other two HDR modalities (see **Figure 210**).

7.4.6 Final comments reported by test users

What follows presents some additional information provided by test users at the end of the user study. This represents useful knowledge towards future works and better result analysis.

Hutton Hub (indoor panorama)

Test users reported that small dark areas of the panorama tended to overexpose the whole image with both HDR-HT and HDR-ET. This was probably caused by the implementation, which was unable to capture accurately enough the real environmental lighting conditions, resulting in unbalanced light changes and excessive light intensity differences. In some cases, this caused the bizarre feeling of being able to control the lights of the scene (e.g. make the sun light appear or disappear) just rotating the head and looking around.

Furthermore, the speed of the light change was perceived differently by test users, who reported both too slow and too fast changes. I believe this is due to their fast head rotations and to possible delays caused by the computation of the virtual scene observed through the HMD.

HDR-ET looked more consistent. However, several users reported less comfort while watching the indoor panorama, compared to when viewing the outdoor environment. I believe this is because some inaccuracies were masked by the large space of the outdoor, which seemed to perform better for light changes than the indoor when using HDR-ET.

Even though colours were reported vivid with S-HDR, many test users agreed that HDR-ET and HDR-HT reproduced them even better thanks to the dynamic lighting conditions. Furthermore, when looking outside the windows many test users reported that S-HDR was not clear, while HDR-HT and HDR-ET provided more details thanks to the believable change of lighting and exposure levels.

Main Reception (outdoor panorama)

Many test users reported that S-HDR lacked in dynamism for colour reproduction. By contrast, dynamic HDR modalities worked better, especially because of the better natural lighting of the outdoor environment.

HDR-ET was very realistic, especially for the light changes shown when looking from the main reception to the sky, and vice versa. This is reasonable as the areas of the scene covered by the sky and by the main reception building were quite large and reduced possible inaccuracies in light changes.

Furthermore, some test users reported that if HDR-ET was slower in light changes it would have performed better than the HDR-HT.

For future works, one of the test users suggested to further investigate the performance of dynamic HDR with outdoor environment at night. This might be very interesting because a few test users perceived more realism when passing from dark to light areas, which is the perfect condition for scenes having night illumination.

7.5 Analysis

For clarity, I will refer to static HDR as **S-HDR**, to dynamic HDR with head tracking as **HDR-HT**, and to dynamic HDR with eye tracking as **HDR-ET**.

7.5.1 Realism in terms of Presence

7.5.1.1 Indoor vs Outdoor using same HDR modality

No significant difference was found in terms of:

- Sense of Presence. Both indoor and outdoor when using the same HDR modality provided similarly high level of presence. This suggests presence was not substantially influenced by the environmental characteristics when using HDR panoramas and same HDR modality to visualize panoramas.
- Isolation. No difference between indoor and outdoor was reported when using the same HDR modality. This suggests environmental characteristics do not have a substantial impact on achieved isolation when using HDR panoramas and same HDR modality.
- Feeling like observing a real place. No difference was reported between indoor and outdoor. This was expected as both panoramas were quite realistic without substantial differences thanks to the capture method adopted for the implementation, which was improved compared to previous user studies.

7.5.1.2 HDR modalities when viewing same panorama

Sense of Presence. S-HDR modality performed worse than HDR-HT and HDR-ET, regardless of the panorama viewed (see **Figure 182**). This suggests that dynamic HDR better emulates human vision when watching over-exposed or under-exposed areas of an environment, resulting in higher sense of presence and natural vision.

Dynamic HDR also allows the viewer to naturally interact with visualized light conditions. According to previous studies, this interaction benefits sense of presence [316], which confirms the obtained result. Furthermore, I believe dynamic HDR provides enhanced pictorial realism [316], increased environmental details [347], and greater visual realism [348] [316] [336], which were all reported to improve presence.

Finally, no significant difference was found between HDR-HT and HDR-HT. This suggests that eye tracking offered no additional value to the sense of presence that was already achieved by HDR-HT.

Isolation. A significant difference was reported only between S-HDR and HDR-HT (see **Figure 183**). This result was unexpected as I believed isolation was similarly perceived regardless of both panorama and HDR modality. This suggests that HDR-HT is more suitable to provide better isolation.

No significant difference was found in terms of:

- Feeling like observing a real place. HDR modalities performed very similarly, scoring high results (see **Figure 184**). This was not expected as I believed that HDR-HT and HDR-ET were more suitable to achieve a better human eye emulation. I can assume that this result was due to some implementation issues that might have negatively influenced the viewer's opinion (e.g. lights too strong, darkness too fast, etc. For more information, the reader is suggested to see subsection 7.4.6).

7.5.2 Realism in terms of Emotions

7.5.2.1 Indoor vs Outdoor using same HDR modality

Likeliness. Outdoor performed significantly better than indoor when using the HDR-ET modality (see **Figure 185**). This was expected as eye tracking and light changes might have performed better with the larger spaces of the outdoor panorama, which reduced the risk of light inaccuracies.

Disappointment. Indoor performed significantly worse in terms of disappointment achieving higher scores than outdoor when using HDR-HR and HDR-ET (see **Figure 188**). I believe this is due to implementation issues, which might have caused excessive light changes within very small areas of the panorama due to the presence of the windows and pillars (see subsection 7.4.6).

Desire to visit the remote environment in real life. Indoor performed significantly better than outdoor when using S-HDR (see **Figure 190**). This was not reported with the other two HDR modalities. I believe this was caused by the environmental characteristics of the outdoor panorama, which presented larger spaces and less 3D objects close to the viewer.

No significant differences were found in terms of:

- Happiness. Indoor and outdoor achieved quite high scores. This suggests that test users were almost equally happy about viewing both panoramas, regardless of the HDR modality used.
- Positive mood change after viewing in VR. The change in mood was very low for both panoramas when using the same HDR modality.
- Motivation to use VR for remote observation. Both panoramas reported very high motivation to use VR, when using the same HDR modality.

7.5.2.2 HDR modalities when viewing same panorama

No significant difference was found in terms of:

- Likeliness. All modalities achieved high scores (see **Figure 186**).
- Happiness. All modalities achieved high scores (see **Figure 187**).
- Disappointment. All modalities achieved low scores (see **Figure 189**).
- Desire to visit the remote environment in real life. All modalities achieved quite high scores (see **Figure 191**).
- Positive mood change after viewing in VR. All modalities reported minimal mood changes (see **Figure 192**).
- Motivation to use VR for remote observation. All modalities achieved high scores (see **Figure 193**).

7.5.3 Realism in terms of Virtual Vision vs Real Life

7.5.3.1 Indoor vs Outdoor using same HDR modality

Natural eye adaptation. Outdoor performed significantly better than indoor only when using HDR-ET (see **Figure 194**). I believe this is because of the better performances of light changes and eye tracking environment due to larger spaces and less inaccuracy errors of the outdoor panorama.

Virtual lights vs Real Life lights. Outdoor performed significantly better than indoor when using HDR-HT or HDR-ET (see **Figure 196**). I assume this is again due to the larger spaces of the panorama, which reduced possible excessive light changes and inaccuracies on eye tracking. Furthermore, indoor seemed to perform a little better than outdoor when using S-HDR, but the difference was not significant.

Virtual vision compared to Real vision. Outdoor performed significantly better than indoor when using HDR-HT (see **Figure 201**). I believe this is due to the excessive light changes experienced when observing the indoor environment, which presented more light changes in very small portions of the panorama.

No significant difference was reported in terms of:

- Color realism. Both indoor and outdoor achieved very high scores when using same HDR modality.
- Realism of Darkness. Both indoor and outdoor achieved very high scores when using same HDR modality.
- Realism of Brightness. Both indoor and outdoor achieved very high scores when using same HDR modality.

7.5.3.2 HDR modalities when viewing same panorama

Natural eye adaptation. S-HDR performed significantly worse than HDR-HT and HDR-ET (see **Figure 195**). This is justified by the static illumination, which did not emulate any eye adaptation at all by showing always same colours and exposure levels.

Color realism. S-HDR performed significantly worse than HDR-HT and HDR-ET (see **Figure 198**). This result was expected, as I believe that the dynamic adaptation of lights and colours allowed the viewer to visualize more details of the scene (e.g. real colours of the space behind the window instead of a white blurry picture, real colours of dark areas instead of black uniform picture).

Realism of Darkness. S-HDR performed significantly worse than HDR-HT and HDR-ET (see **Figure 199**). This was expected as dynamic light adaptation and exposure changes allowed the viewer to visualize more details of dark areas instead of black unclear pictures.

Realism of Brightness. S-HDR performed significantly worse than HDR-HT and HDR-ET (see **Figure 200**). This suggests that also to visualize bright areas the benefit of light adaptation and exposure changes was higher than static HDR pictures.

No significant difference was reported in terms of:

- Virtual lights vs Real Life lights. All modalities performed well (see **Figure 197**).
- Virtual vision compared to Real vision. All modalities performed well (see **Figure 202**).

7.5.4 Comfort

7.5.4.1 Indoor vs Outdoor using same HDR modality

Time before feeling discomfort. (See **Figure 204**) When using S-HDR, outdoor panorama was a little more comfortable since slightly more test users reported they would be able to watch the panorama for over 30 minutes without experiencing discomfort.

When using HDR-HT or HDR-ET, outdoor was much more comfortable and people could watch it for longer times. I assume this was again due to excessive light changes for the indoor panorama, which presented more inaccuracies than the outdoor.

Perceptual issues. Results suggest a slight trend of test users to perceive motion sickness more with the indoor panorama than with the outdoor (see **Figure 205**). I believe this is due to the excessive light changes in the indoor panorama, which occurred more frequently than in the outdoor panorama.

No significant difference was reported in terms of:

- Level of comfort. Both panoramas achieved very high scores when using the same HDR modality.

7.5.4.2 HDR modalities when viewing same panorama

Time before feeling discomfort. When looking to comfort occurring only after 30 minutes, S-HDR achieved less scores than HDR-HT and HDR-ET (see **Figure 204**). However, I suspect that many test users were influenced by the panorama characteristics, due to the excessive light changes and the visual differences between indoor and outdoor. For this reason, it is not possible to distinguish the effect of the HDR modality independently from the panorama viewed, and to identify a clear trend.

Perceptual issues. I noticed a slight trend showing that S-HDR provided less motion sickness than HDR-HT and HDR-ET (see **Figure 205**). I assume this is due to the amount of light changes with the dynamic HDR modalities, which might have caused more realism but to the expenses of comfort.

No significant difference was reported in terms of:

- Level of comfort. All HDR modalities achieved very high scores when viewing the same panorama (see **Figure 203**).

7.5.5 Depth perception

7.5.5.1 Indoor vs Outdoor using same HDR modality

Realistic space perception. Outdoor performed significantly better than indoor when using HDR-HT (see **Figure 208**). I assume this is due to the excessive changes of lights, that sometimes on the indoor panorama disoriented the viewers, resulting in confusing depth perception and space perception.

No significant difference was found in terms of:

- 3D accuracy. Both panoramas achieved high scores when using the same HDR modality.
- Lights and Shadows contribution to 3D. Both panoramas achieved very high scores when using the same HDR modality.
- Perceived size of the environment. Both panoramas achieved very high scores when using the same HDR modality.

7.5.5.2 HDR modalities when viewing same panorama

3D accuracy. S-HDR performed significantly worse than HDR-HT and HDR-ET (see **Figure 206**). I believe this is because dynamic HDR revealed more details than the S-HDR, resulting in more disparity details and better depth perception.

Lights and Shadows contribution to 3D. S-HDR performed significantly worse than HDR-HT and HDR-ET (see **Figure 207**). I think that the reason is the lack of details shown by the S-HDR compared to the dynamic light adaptation of HDR-HT and HDR-ET, which inevitably showed more lights and shadows on the scene.

Realistic space perception. S-HDR performed significantly better than HDR-HT and HDR-ET (see **Figure 209**). I believe that the reason is the larger amount of details provided to the viewer thanks to the dynamic lights and exposure adaptations of the scene.

Perceived size of the environment. Results report that the size of the panorama when using S-HDR appeared significantly smaller than when using HDR-HT and HDR-ET (see **Figure 210**). I suspect this was due to the dynamic light and exposure adaptation, which altered the space perception of the viewers. However, it is not possible to understand from test whether this was beneficial, as the perceived size of the room was not measurable compared to the actual size of the real environment and it was only a subjective impression.

7.6 Guidelines

Like for the previous two user studies, for this experiment a Guidelines section is provided to summarize main outcomes in terms of concise recommendations for system designers. As before the guidelines are laid down after the four elements that have been considered as relevant contributors to remote visual observation. For this experiment these are as before: realism, comfort, presence and depth-perception. Furthermore, emotions have also been considered since the role they played in the previous user study's results was significant.

- **How to improve Realism in terms of Presence?**
    - Use HDR-HT and HDR-ET to improve the sense of presence by providing eye adaptation to the viewer. This guarantees better visual perception of environmental details and colours, regardless of over-exposed and under-exposed areas (which are dynamically corrected in real-time).
    - Use HDR-HT to provide better isolation, which in turn enhances sense of presence.

- **How to improve Realism in terms of Emotions?**
    - Prefer HDR-ET for outdoor environments, especially when natural illumination is provided within the scene.
    - Avoid disappointments by offering a correct light adaptation when using HDR-HT and HDR-ET. When excessive changes occur within small areas of the panorama, and too much light or darkness are shown, the user feels like being able to turn on or off lights of the scene, which is not realistic and causes negative emotions.
    - Prioritize 3D perception in the panoramas, as it helps the viewer to enjoy the experience and activate higher levels of presence.

- **How to improve Realism in terms of Virtual Vision vs Real Vision?**
    - Offer accurate and natural eye adaptation to the viewer, by controlling the amount of light in the scene when using HDR-HT and HDR-ET. It is suggested to depict open spaces that present large portions of the panorama having different light conditions. This helps to avoid inaccuracies and excessive light changes when small movements are performed by head or eyes.
    - Avoid S-HDR for colour realism and prioritize HDR-HT or HDR-ET. This is because dynamic HDR provides more details and natural colours of over-exposed and under-exposed areas of the panorama, emulating the behaviour of human eyes.
    - HDR-HT and HDR-ET perform significantly better than S-HDR in terms of darkness and brightness reproductions. These are suggested for a more realistic observation of remote environments.

- **How to improve Comfort?**
    - S-HDR can be very comfortable when light conditions of the panorama are extreme and eye adaptation would imply too many fast changes of exposure with HDR-HT and HDR-ET. However, when light changes are not too extreme and excessive intensity is avoided, HDR-HT and HDR-ET are much more comfortable than S-HDR, as they better emulate the human eye behaviour when looking light sources or dark areas.

    - Avoid too many light changes with HDR-HT and HDR-ET to avoid motion sickness and visual fatigue. Instead, prioritize smooth and slow changes and large areas of brightness or darkness.

- **How to improve Depth Perception?**
    - HDR-HT and HDR-ET are preferred, as they reveal more details of the scene and increase details of 3D disparity values as well.
    - Prioritize scenes that use large areas of lights and shadows, so that HDR-HT and HDR-ET will provide their best performances for 3D perception.
    - Use HDR-HT and HDR-ET for better space awareness of the surrounding environment.
    - There might be a difference in perceived size of the environment when using S-HDR rather than HDR-HT and HDR-ET. Dynamic HDR might provide better distance estimations.

7.7 Conclusion

In this chapter, the design and outcomes of the third usability evaluation proposed by this thesis were presented.

The aim of this last usability evaluation was to assess the use of eye tracking and HDR panoramic pictures in VR and to analyse possible visual improvements when using eye-adapted HDR viewing over static HDR panoramic pictures.

After a brief explanation of the different HDR modalities this user study investigated, the design of the usability evaluation was discussed. In detail, a static HDR panorama, a head-tracked dynamic HDR panorama, and an eye-tracked dynamic HDR panorama have been compared when observing an indoor and an outdoor environment.

Furthermore, implementation details on the HDR 3D panorama acquisition with the *insta360 Pro* camera, on the developed HDR visualization system in *Unity 3D*, and on the use of the *FOVE VR* headset with eye tracking capabilities were presented.

Results measured the influence of each *HDR modality* and each observed *environment* over realism, sense of presence, comfort, and depth perception. These were then discussed and compared with what was found in the previous systematic review. Among them, the most important outcome demonstrated that both head-tracked and eye-tracked HDR dynamic 3D panoramas in VR can benefit the visual perception of test users. This was not found in any of the previous analysed researches of the systematic review, and represents new unique and innovative knowledge produced by this thesis.

Finally, new guidelines based on the outcome of the user study were presented, based on results of this final usability evaluation.

CHAPTER 8. CONCLUSIONS AND FUTURE RESEARCH

8.1 Summary

The research presented in this work aimed at understanding the role of technology setup and camera-display parameters in providing convincing visual realism, for reducing the perceptual gap between direct and indirect observation of real places when using Mobile VR headsets.

To achieve this goal, a comprehensive systematic investigation was proposed through three development phases:

- *Systematic Review for Learning*. A systematic literature investigation was proposed to collect outcomes from the state-of-the-art, in the context of 3D acquisition and visualization systems, applied to Virtual Reality for remote observation of existing environments.
- *Build Knowledge through Assessments*. Two user studies were presented, to assess the impact of *familiarity* with the visualized scene, observed *environment*, and *display* on realism, comfort, presence, and depth perception for VR remote observation of existing places.
- *Learning from Usability Results*. A focused evaluation was proposed after the lesson learned from the previous user studies' results, to assess the benefits of eye-adapted HDR viewing in VR for indirect panoramic observation of existing places.

All user studies' results were evaluated in terms of achieved *realism*, *depth perception*, *comfort*, and *sense of presence,* and presented in the form of guidelines for VR system designers and content creators.

The structure of this thesis has been organically presented to the reader, with the intent to facilitate the understanding of the discussed topics. In particular, Chapter 2 provided the reader with important definitions and background knowledge; Chapter 3 presented the aims and motivation of this investigation with a research plan to achieve proposed goals; Chapter 4 concerned the systematic review; Chapter 5 reported on the influence of familiarity and environment through the first user study; Chapter 6 reported on the impact of display and environment through the second user study; Chapter 7 reported on eye-adapted HDR viewing in VR through the third proposed user study.

For convenience, the section below groups the most relevant overall conclusions from the outcomes of this research, with the aim of delivering simple guidelines for better performances in remote visual observation of real places through Virtual Reality headsets.

8.2 Contributions

Some of the most important outcomes of the proposed user studies on remote observation of real places through VR can be summarized as follows in **Table 21**, **Table 22**, **Table 23**, and **Table 24**.

HOW TO IMPROVE REALISM	
FAMILIARITY	- Use Mobile VR rather than photos and videos, especially when viewers already know the place;
DISPLAY	- Use very large horizontal size of display, especially when people view the remote environment for their first time. This also avoids tunnel vision, which has a detrimental impact on realism. - High resolution displays should be used when the panorama is rich in details. However, with simple scenes resolution is not as relevant. - High pixel density is beneficial especially when the scene presents wide colours spectrum and high levels of vividness. - Image contrast is relevant when the scene presents lots of colours, light reflections, and natural illumination (e.g. landscape, forest, beach). - Colours reproduction capabilities of the display are less important than screen size, resolution, and pixel density when using Mobile VR.
ENVIRONMENT & HDR	- Give preference to natural features (e.g. trees, sea, rocks, sunlight) as they make the environment more realistic and believable. - Reduce stitching errors and blurred vision. - Prioritize the use of natural illumination and vivid scenes. - Make sure to adapt the graphical field of view of the panorama to the size of the screen used, as it influences the realistic perception of surrounding spaces. - Give preference to dynamic HDR over static HDR when colours realism is considered very important: eye adaptation allows viewers to appreciate more authentic lights and colours of the remote environment.

***Table 21**. Summary of the most important guidelines to improve realism, resulting from all proposed user studies.*

HOW TO IMPROVE COMFORT	
FAMILIARITY	- When the viewer knows the place, differences between real and virtual environments might be detectable more easily, and possible stitching errors might become more uncomfortable to the viewer.
DISPLAY	- Avoid excessive display brightness as it might cause visual fatigue to viewers.
ENVIRONMENT & HDR	- Keep the scene simple, so that viewers will avoid uncomfortable vision in 3D. - Avoid stitching errors for the observed panoramic scenes. - Prioritize relaxing environments to the ones that induce fear and claustrophobia. - Use static HDR if light changes are too frequent with dynamic HDR, as they might cause more visual fatigue and motion sickness despite the increased level of realism. - Use dynamic HDR when light changes are smooth and not frequent, and the scene presents large areas of brightness or darkness. This is because eye-adapted HDR viewing reproduces more realistically the human visual system, causing less discomfort and more enjoyment to the viewer.

Table 22. *Summary of the most important guidelines to improve comfort, resulting from all proposed user studies.*

<table>
<tr><th colspan="2">HOW TO IMPROVE PRESENCE</th></tr>
<tr><td>FAMILIARITY</td><td>- People that have real life memories related to the visualized place can achieve higher sense of presence due to the emotional involvement, which activates presence.</td></tr>
<tr><td>DISPLAY</td><td>- Enhance immersion (e.g. provide also audio or tactile feedback together with the 3D display) as higher immersion causes more emotions, which in turn enable more sense of presence.
- Be sure to guarantee a good level of isolation from the place where the viewer stands in real life, to avoid any distraction from the virtual experience.
- Try to avoid tunnel vision, especially with outdoor environments.</td></tr>
<tr><td>ENVIRONMENT & HDR</td><td>- Enhance emotions (e.g. feeling of going on holiday, enjoyment, happiness, relax) as they activate sense of presence. Within this purpose, remember that:
o Wide colour spectrum and high pixel density can reduce sadness;
o Dark environments with almost monochromatic colours combined with high resolution and high pixel density can induce more easily fear and scariness;
o Nice weather, natural environments, open spaces, sunlight, sea, water reflections can easily induce good feelings.
o Bright colours rise mainly positive emotional associations;
o Dark colours rise mainly negative emotional associations.
- Prioritize environments offering objects that are close to the viewer without excessive disparity (strong visible 3D), and simple design to provide less cue conflicts.
- When tunnel vision occurs, indoor close-range bounded environments should be preferred as they could recall the boundaries of the HMD and thus enhance the viewer's sense of presence.
- Use dynamic HDR as eye adaptation provides better visual perception of environmental details and colours, regardless of the over-exposed and under-exposed areas (which are dynamically corrected in real-time).
- Use HDR with head tracking to provide better isolation, which in turns enhances the sense of presence.</td></tr>
</table>

***Table 23**. Summary of the most important guidelines to improve presence, resulting from all proposed user studies.*

HOW TO IMPROVE DEPTH PERCEPTION	
FAMILIARITY	- Depth perception of people that have no familiarity with the place and try Mobile VR for the first time should be enhanced (e.g. providing more monocular cues such as lights and shadows, texture cues, relative size cues).
DISPLAY	- Consider an appropriate time to allow viewers to adjust IPD and distance from the screen accordingly, so that they can enjoy an optimal HMD setup and better 3D perception. - If low resolution displays are used, monocular cues (e.g. lights and shadows) should be intensified so that the viewer will still be able to combine them with binocular cues and better estimate scene depth. - Choose corrective lenses when the viewer is affected by astigmatism or myopia, to keep disparity and depth cues visible enough for an appropriate space awareness of the remote environment.
ENVIRONMENT & HDR	- Use visual aids (e.g. arrows, grids, or virtual avatars) to allow the viewer to have points of references that can help better evaluate the surrounding spaces. - Provide monocular cues (e.g. lights and shadows, texture cues, relative sizes), so that viewers can solve possible binocular cues conflicts and better perceive depths and distances of the scene. - Prioritize objects that are close to the viewers but avoiding excessive disparity. - Prioritize scene designs having several and large depth intervals. This will guarantee more virtual spaces between background and foreground objects, showing more distinctly their position within the 3D space, improving relative distance estimations. - Use simple scenes to allow the user to perceive shapes and objects in 3D with low occurrence of visual conflicts. - Avoid stitching errors, vertical parallax, and possible perspective mismatches between left-eye and right-eye panoramas. Less stitching errors can also reduce the time needed for the viewer to see the scene in 3D. - Choose the appropriate graphical field of view (GFOV) of the Mobile VR viewer, according to the screen size used inside the HMD. This also depends on the camera used to capture the environment, and on the setup adopted to generate the 3D panorama. - Prioritize the use of dynamic HDR over static HDR, as it reveals more details of the scene, increases 3D disparity details, and provides better space awareness. This applies especially to large areas of lights and shadows.

Table 24. *Summary of the most important guidelines to improve depth perception, resulting from all proposed user studies.*

8.3 Future Works

I strongly believe that my research can be used as a possible starting point to further investigate the role of camera-display configurations and virtual reality devices, to deliver more convincing realistic remote observation of existing places.

Therefore, I recommend the reader with the following research topics, for potential future assessments and user studies:

- **Use of professional 3D panoramic cameras** to test the effect of higher resolution photography (e.g. 8K or above) on remote observation in virtual reality. I expect this investigation to find better levels of realism, less 3D distortions, but more challenges in terms of visual delays, due to the higher amount of details to show on display, which calls for more powerful graphic cards and devices to visualize them. This might be interesting to devise techniques to reduce motion sickness caused by delays and conflicting sensory stimuli.

- **Assessments on 3D panoramic Videos in Virtual Reality**. Due to limitations of the professional 3D panoramic cameras that I could have used for my research, there is still room for further investigation on the effect of different 3D video camera layouts, frame rates, video resolutions, and interactive visual aids when showing 3D panoramic movies in virtual reality for remote observation of existing places. This might open new possibilities to hotspot-to-hotspot navigations in 3D panoramas and enable a new frontier for exploring places with natural, accurate and interactive remote navigation.

- **Assessments on dynamic HDR Videos**. After the many advantages that I discovered in this investigation through the use of eye-adapted HDR viewing for still panoramas in virtual reality, I expect that further investigation could be carried out for dynamic HDR videos and interactive media. This might present further challenges like handling moving objects, minimum distances to avoid stitching errors according to the used camera layout, and methods to acquire at high frequencies different exposed panoramas to correctly process light changes for HDR eye adaptation.

- **Assessments on Real-Time 3D panoramic streaming.** I believe that a lot of advantages might be offered by teleoperated navigation systems able to stream in real-time robots' point of view for enhanced remote operations. Within this purpose, networking problems, communication delays, video resolution performances, and intuitive interactions might represent the biggest challenges.

- **Photogrammetry using 3D and 360 cameras.** Even if today photogrammetry mainly relies on standard cameras and very few algorithms exist to process equirectangular panoramic pictures from 360 cameras, I believe that modern professional cameras might offer a powerful mean to reduce the required number of shots to capture entire existing environments with good results. However, I still expect challenges like the distortions introduced by large field of views of 360 cameras, the problem of taking pictures of the environment avoiding the tripod and unwanted objects on the scene, and the possible use of moving platforms to allow 3D 360 cameras to be automatically transported through rooms and places with image stabilization capabilities.

REFERENCES

[1] The New Penguin English Dictionary, Robert Allen, 2000.

[2] W. R. Sherman and A. B. Craig, Understanding virtual reality: Interface, application, and design, Elsevier, 2002.

[3] S. Johnston, "Virtual Reality as New Media is Revolutionary, Int J Adv Technol 8:182," 2017. [Online]. Available: https://www.omicsonline.org/open-access/virtual-reality-as-new-media-is-revolutionary-0976-4860-1000182.php?aid=88502.

[4] M. L. Heilig, Sensorama simulator, Google Patents, 1962.

[5] M. L. Heilig, "Stereoscopic-television apparatus for individual use". Patent US Patent 2,955,156, 4 October 1960.

[6] C. Cruz-Neira, D. J. Sandin, T. A. DeFanti, R. V. Kenyon and J. C. Hart, "The CAVE: audio visual experience automatic virtual environment," *Communications of the ACM,* vol. 35, no. 6, pp. 64--73, 1992.

[7] C. Thinus-Blanc and F. Gaunet, "Representation of space in blind persons: vision as a spatial sense?," *Psychological bulletin,* vol. 121, no. 1, p. 20, 1997.

[8] E. Foulke, "Perception, cognition and the mobility of blind pedestrians," *Spatial abilities: Development and physiological foundations,* pp. 55--76, 1982.

[9] H. L. Pick, D. H. Warren and J. C. Hay, "Sensory conflict in judgments of spatial direction," *Perception & Psychophysics,* vol. 6, no. 4, pp. 203--205, 1969.

[10] Oculus VR, "Oculus Rift," Facebook, [Online]. Available: https://www.oculus.com/. [Accessed 24 August 2018].

[11] H. Corporation, "HTC VIVE," HTC Corporation, [Online]. Available: https://www.vive.com/eu/. [Accessed 24 August 2018].

[12] S. Corp, "StarVR," StarVR Corp, [Online]. Available: https://www.starvr.com/. [Accessed 24 August 2018].

[13] Samsung, "Samsung Gear VR with Controller - The Official Samsung Galaxy Site," [Online]. Available: https://www.samsung.com/global/galaxy/gear-vr/.

[14] B. S.Fisher, 1991.

[15] K. Ukai and P. A. Howarth, "Visual fatigue caused by viewing stereoscopic motion images: Background, theories, and observations," *Displays,* vol. 29, no. 2, pp. 106--116, 2008.

[16] U. Technologies, "Unity 3D Game Engine," [Online]. Available: https://unity3d.com/.

[17] I. NaturalPoint, "OptiTrack Motion Tracking Camera Systems," [Online]. Available: http://optitrack.com/products/flex-3/.

[18] Nvidia, “Nvidia 3D Vision,” Nvidia, [Online]. Available: http://www.nvidia.co.uk/object/buy-wireless-3d-glasses-uk.html. [Accessed 17 July 2018].

[19] S. Livatino, “Coherent touchless interaction with steroscopic 3D images”. UK Patent GB2533777A, 24 12 2014.

[20] J. E. Cutting and P. M. Vishton, “Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth,” in *Perception of space and motion*, Elsevier, 1995, pp. 69--117.

[21] W. B. Thompson, P. Willemsen, A. A. Gooch, S. H. Creem-Regehr, J. M. Loomis and A. C. Beall, “Does the quality of the computer graphics matter when judging distances in visually immersive environments?,” *Presence: Teleoperators \& Virtual Environments,* vol. 13, no. 5, pp. 560--571, 2004.

[22] C. J. Lin, B. H. Woldegiorgis, D. Caesaron and L.-Y. Cheng, “Distance estimation with mixed real and virtual targets in stereoscopic displays,” *Displays,* vol. 36, pp. 41--48, 2015.

[23] S. Winkler and D. Min, “Stereo/multiview picture quality: Overview and recent advances,” *Signal Processing: Image Communication,* vol. 28, no. 10, pp. 1358--1373, 2013.

[24] Pixabay. [Online]. Available: https://pixabay.com/en/railway-rails-track-train-2439189/.

[25] A. Alimbetov, Pexels, [Online]. Available: https://www.pexels.com/photo/architecture-buildings-business-city-303159/.

[26] I. Pereira, Pexels, [Online]. Available: https://www.pexels.com/photo/arrows-close-up-dark-energy-394377/.

[27] Pixabay. [Online]. Available: https://pixabay.com/en/skyline-dubai-skyscrapers-city-2018079/.

[28] F. Paggiaro, “Pexels,” [Online]. Available: https://www.pexels.com/photo/shadow-depth-of-field-photography-of-human-index-finger-793060/.

[29] C. Larson, “Pexels,” [Online]. Available: https://www.pexels.com/photo/saraha-desert-1123567/.

[30] M. W. McGreevy, “Virtual reality and planetary exploration,” in *Virtual Reality*, Elsevier, 1993, p. 46–65.

[31] X. Tang and H. Yamada, “Tele-operation construction robot control system with virtual reality technology,” *Procedia Engineering,* vol. 15, pp. 1071--1076, 2011.

[32] Y. Kato, “A remote navigation system for a simple tele-presence robot with virtual reality,” in *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference*, IEEE, 2015, pp. 4524--4529.

[33] S. Livatino, G. Muscato and F. Privitera, “Stereo viewing and virtual reality technologies in mobile robot teleguide,” *IEEE Transactions on Robotics,* vol. 25, no. 6, pp. 1343--1355, 2009.

[34] S. Shaikh, K. Bokde, A. Ingale and B. Tekwani, “Virtual Tourism,” 2018.

[35] NASA, “Pixabay,” [Online]. Available: https://pixabay.com/en/mars-mars-rover-space-travel-robot-67522/.

[36] B. Series, “Subjective methods for the assessment of stereoscopic 3DTV systems,” 2015.

[37] M. Slater, “A note on presence terminology,” *Presence connect,* vol. 3, no. 3, pp. 1--5, 2003.

[38] D. A. Bowman and R. P. McMahan, “Virtual reality: how much immersion is enough?,” *Computer,* vol. 40, no. 7, 2007.

[39] S.-Y. Lu, M. Shpitalni and R. Gadh, “Virtual and augmented reality technologies for product realization,” *CIRP Annals,* vol. 48, no. 2, pp. 471--495, 1999.

[40] A. Mitiche and J.-M. Létang, “Stereokinematic analysis of visual data in active convergent stereoscopy,” *Robotics and autonomous systems,* vol. 25, no. 1-2, pp. 43--71, 1998.

[41] Panasonic, “How focal length affects viewing angle,” [Online]. Available: http://av.jpn.support.panasonic.com/support/global/cs/dsc/knowhow/knowhow12.html.

[42] H. Kim and J. Lee, “A study on the possibility of implementing a real-time stereoscopic 3D rendering TV system,” *Displays,* vol. 40, pp. 24--34, 2015.

[43] H. J. Pragnell, The London Panoramas of Robert Barker and Thomas Girtin, circa 1800, Topographical Society, 1968.

[44] O. Faugeras, Panoramic vision: sensors, theory, and applications, Springer Science & Business Media, 2013, pp. 5,6.

[45] R. Barker, Artist, *London from the Roof of Albion Mills.* [Art]. 1792.

[46] S. Amic, K. C. Galitz, L. D. Cars, D. Lobstein, B. Mottin, T. Galifot, B. Tillier and G. Courbet, Gustave Courbet, Hatje Cantz, 2008.

[47] CUTGANA, “Visual 3D Fruition Project,” [Online]. Available: http://www.cutgana.unict.it/VirtualTours3D/.

[48] D. G. Lowe, “Object recognition from local scale-invariant features,” in *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, vol. 2, IEEE, 1999, pp. 1150--1157.

[49] D. G. Lowe, “Method and apparatus for identifying scale invariant features in an image and use of same for locating an object in an image”. Patent US Patent 6,711,293, 23 March 2004.

[50] J. Xing and Z. Miao, “An improved algorithm on image stitching based on SIFT features,” in *Innovative Computing, Information and Control, 2007. ICICIC'07. Second International Conference on*, IEEE, 2007, pp. 453--453.

[51] Y. Li, Y. Wang, W. Huang and Z. Zhang, "Automatic image stitching using sift," in *Audio, Language and Image Processing, 2008. ICALIP 2008. International Conference on*, IEEE, 2008, pp. 568--571.

[52] Z. Min, Z. Jiguo and X. Xusheng, "Panorama stitching based on sift algorithm and levenberg-marquardt optimization," *Physics Procedia,* vol. 33, pp. 811--818, 2012.

[53] J. Illingworth and J. Kittler, "A survey of the Hough transform," *Computer vision, graphics, and image processing,* vol. 44, no. 1, pp. 87--116, 1988.

[54] M. M. Pandya, N. G. Chitaliya and S. R. Panchal, "Accurate image registration using SURF algorithm by increasing the matching points of images," *International Journal of Computer Science and Communication Engineering,* vol. 2, no. 1, pp. 14--19, 2013.

[55] L. Juan and G. Oubong, "SURF applied in panorama image stitching," in *Image Processing Theory Tools and Applications (IPTA), 2010 2nd International Conference on*, IEEE, 2010, pp. 495--499.

[56] E. Rublee, V. Rabaud, K. Konolige and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *Computer Vision (ICCV), 2011 IEEE international conference on*, IEEE, 2011, pp. 2564--2571.

[57] E. Adel, M. Elmogy and H. Elbakry, "Image stitching system based on ORB feature-based technique and compensation blending," *International Journal of Advanced Computer Science and Applications,* vol. 9, no. 6, 2015.

[58] S. T. Suen, E. Y. Lam and K. K. Wong, "Photographic stitching with optimized object and color matching based on image derivatives," *Optics Express,* vol. 15, no. 12, pp. 7689--7696, 2007.

[59] L. Juan and O. Gwun, "A comparison of sift, pca-sift and surf," *International Journal of Image Processing (IJIP),* vol. 3, no. 4, pp. 143--152, 2009.

[60] P. Ostiak, "Implementation of HDR panorama stitching algorithm," in *Proceedings of the 10th CESCG Conference*, Citeseer, 2006, pp. 24--26.

[61] Y. Xiong and K. Pulli, "Fast panorama stitching for high-quality panoramic images on mobile phones," *IEEE Transactions on Consumer Electronics,* vol. 56, no. 2, 2010.

[62] A. Regalbuto, S. Livatino, K. Edwards and I. Mporas, "Mobile VR Headset Usability Evaluation of 2D and 3D Panoramic Views Captured with Different Cameras," in *International Conference on Interactive Collaborative Robotics*, Springer, 2017, pp. 191--200.

[63] B. Lane, "Stereoscopic displays," in *Proc. SPIE*, vol. 367, 1982, pp. 20--32.

[64] C. Wheatstone, "XVIII. Contributions to the physiology of vision.—Part the first. on some remarkable, and hitherto unobserved, phenomena of binocular vision," *Philosophical transactions of the Royal Society of London,* vol. 128, pp. 371--394, 1838.

[65] G. Kim, Designing virtual reality systems, Springer, 2005.

[66] J. Aggergaard, "Unsplash," [Online]. Available: https://unsplash.com/photos/38Hg7GMTogo.

[67] A. Vallainc, "Unsplash," [Online]. Available: https://unsplash.com/photos/yFcIxuArJEQ.

[68] Nvidia, "GEFORGE GTX 1070 & 1070 Ti," [Online]. Available: https://www.nvidia.com/en-gb/geforce/products/10series/geforce-gtx-1070/.

[69] Nvidia, "GEFORGE GTX 1080," [Online]. Available: https://www.nvidia.com/en-gb/geforce/products/10series/geforce-gtx-1080/.

[70] F. Steinicke, "The science and fiction of the ultimate display," in *Being really virtual*, Springer, 2016, pp. 19--32.

[71] E. Reinhard, W. Heidrich, P. Debevec, S. Pattanaik, G. Ward and K. Myszkowski, High dynamic range imaging: acquisition, display, and image-based lighting, Morgan Kaufmann, 2010.

[72] M. A. Robertson, S. Borman and R. L. Stevenson, "Estimation-theoretic approach to dynamic range enhancement using multiple exposures," *Journal of Electronic Imaging,* vol. 12, no. 2, pp. 219--229, 2003.

[73] G. Krawczyk, K. Myszkowski and D. Brosch, "HDR tone mapping," in *High-Dynamic-Range (HDR) Vision*, Springer, 2007, pp. 147--178.

[74] Adobe, "About high dynamic range images," [Online]. Available: https://helpx.adobe.com/uk/photoshop/using/high-dynamic-range-images.html. [Accessed 30 October 2018].

[75] FOVE, Inc., "FOVE Eye Tracking Virtual Reality Headset," FOVE, Inc., [Online]. Available: https://www.getfove.com. [Accessed 8 8 2018].

[76] S. Livatino and K. Christina, "Simple Guidelines for Testing VR Applications," INTECH Open Access Publisher, 2008.

[77] D. MacIsaac, "An introduction to action research," *Retrieved January,* vol. 20, p. 2002, 1995.

[78] R. Winter, Learning from experience: Principles and practice in action-research, Falmer Pr, 1989.

[79] X. Faulkner, Usability Engineering, Palgrave Macmillan, 2000.

[80] J. Nielsen, Usability Engineering, Morgan Kaufmann, 1993.

[81] J. &. M. R. Nielsen, Usability Inspection Methods, 1994.

[82] J. Rubin, Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests (Wiley Technical Communications Library), Wiley, 1994.

[83] R. Fisher, The design of experiments, 5th ed., Oliver & Boyd,, 1935/1951, p. 13.

[84] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, P. Group and others, “Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement,” *PLoS medicine,* vol. 6, no. 7, p. e1000097, 2009.

[85] “Science Direct,” Elsevier, [Online]. Available: http://www.sciencedirect.com/.

[86] “Web of Science,” Clarivate Analytics (United States), [Online]. Available: https://clarivate.com/products/web-of-science/.

[87] S. Livatino and F. Privitera, “3D visualization technologies for teleguided robots,” in *Proceedings of the ACM symposium on Virtual reality software and technology*, ACM, 2006, pp. 240--243.

[88] J. Häkkinen, T. Kawai, J. Takatalo, T. Leisti, J. Radun, A. Hirsaho and G. Nyman, “Measuring stereoscopic image quality experience with interpretation based quality methodology,” in *Image Quality and System Performance V*, vol. 6808, International Society for Optics and Photonics, 2008, p. 68081B.

[89] J. R. Wilson, “Effects of participating in virtual environmentsa review of current knowledge,” *Safety science,* vol. 23, no. 1, pp. 39--51, 1996.

[90] R. Vlad, P. Ladret and A. Guérin, “Three factors that influence the overall quality of the stereoscopic 3D content: image quality, comfort, and realism,” in *Image Quality and System Performance X*, vol. 8653, International Society for Optics and Photonics, 2013, p. 865309.

[91] T. Piantanida, D. Boman and J. Gille, “Human perceptual issues and virtual reality,” *Virtual Reality Systems,* vol. 1, no. 1, pp. 43--52, 1993.

[92] J. M. Moshell, B. S. Blau, B. Knerr, D. R. Lampton and J. P. Bliss, “A research testbed for virtual environment training applications,” in *Virtual Reality Annual International Symposium, 1993., 1993 IEEE*, IEEE, 1993, pp. 83--89.

[93] M. S. Banks, J. C. Read, R. S. Allison and S. J. Watt, “Stereoscopy and the human visual system,” *SMPTE motion imaging journal,* vol. 121, no. 4, pp. 24--43, 2012.

[94] K. Templin, P. Didyk, T. Ritschel, K. Myszkowski and H.-P. Seidel, “Highlight microdisparity for improved gloss depiction,” *ACM Transactions on Graphics (TOG),* vol. 31, no. 4, p. 92, 2012.

[95] A. Blake and H. Bülthoff, “Does the brain know the physics of specular reflection?,” *Nature,* vol. 343, no. 6254, p. 165, 1990.

[96] A. Hurlbert, B. Cumming and A. Parker, “Recognition and perceptual use of specular reflections,” *Investigative Ophthalmology and Visual Science,* vol. 32, no. 4, p. 105, 1991.

[97] Y. Sakano and H. Ando, “Effects of head motion and stereo viewing on perceived glossiness,” *Journal of Vision,* vol. 10, no. 9, pp. 15--15, 2010.

[98] G. Wendt, F. Faul and R. Mausfeld, “Highlight disparity contributes to the authenticity and strength of perceived glossiness,” *Journal of Vision,* vol. 8, no. 1, pp. 14--14, 2008.

[99] M. S. Landy, L. T. Maloney, E. B. Johnston and M. Young, "Measurement and modeling of depth cue combination: In defense of weak fusion," *Vision research,* vol. 35, no. 3, pp. 389--412, 1995.

[100] S. Reichelt, R. Häussler, G. Fütterer and N. Leister, "Depth cues in human visual perception and their realization in 3D displays," in *Three-Dimensional Imaging, Visualization, and Display 2010 and Display Technologies and Applications for Defense, Security, and Avionics IV*, vol. 7690, International Society for Optics and Photonics, 2010, p. 76900B.

[101] B. Masia, G. Wetzstein, P. Didyk and D. Gutierrez, "A survey on computational displays: Pushing the boundaries of optics, computation, and perception," *Computers & Graphics,* vol. 37, no. 8, pp. 1012--1038, 2013.

[102] B. Rogers and M. Graham, "Similarities between motion parallax and stereopsis in human depth perception," *Vision research,* vol. 22, no. 2, pp. 261--270, 1982.

[103] M. Hogervorst, M. Bradshaw and R. Eagle, "Spatial frequency tuning for 3-D corrugations from motion parallax," *Vision Research,* vol. 40, no. 16, pp. 2149--2158, 2000.

[104] M. F. Bradshaw and B. J. Rogers, "The interaction of binocular disparity and motion parallax in the computation of depth," *Vision research,* vol. 36, no. 21, pp. 3457--3468, 1996.

[105] M. Mikkola, S. Jumisko-Pyykkö, D. Strohmeier, A. Boev and A. Gotchev, "Stereoscopic depth cues outperform monocular ones on autostereoscopic display," *IEEE Journal of Selected Topics in Signal Processing,* vol. 6, no. 6, pp. 698--709, 2012.

[106] F. Speranza, W. J. Tam, R. Renaud and N. Hur, "Effect of disparity and motion on visual comfort of stereoscopic images," in *Stereoscopic displays and virtual reality systems XIII*, vol. 6055, International Society for Optics and Photonics, 2006, p. 60550B.

[107] N. Hiruma and T. Fukuda, "Accommodation response to binocular stereoscopic TV images and their viewing conditions," *SMPTE,* vol. 102, no. 12, pp. 1137-1144, December 1993.

[108] T. Miyashita and T. Uchida, "Cause of fatigue and its improvement in stereoscopic displays," *Proceedings of the SID,* vol. 31, pp. 249-254, 1990.

[109] M. Wöpking, "Viewing comfort with stereoscopic pictures: An experimental study on the subjective effects of disparity magnitude and depth of focus," *Journal of the society for information display,* vol. 3, no. 3, pp. 101--103, 1995.

[110] J. P. McIntire, P. R. Havig and E. E. Geiselman, "Stereoscopic 3D displays and human performance: A comprehensive review," *Displays,* vol. 35, no. 1, pp. 18--26, 2014.

[111] G. R. Jones, D. Lee, N. S. Holliman and D. Ezra, "Controlling perceived depth in stereoscopic images," in *Stereoscopic Displays and Virtual Reality Systems VIII*, vol. 4297, International Society for Optics and Photonics, 2001, pp. 42--54.

[112] A. J. Woods, T. Docherty and R. Koch, “Image distortions in stereoscopic video systems,” in *Stereoscopic displays and applications IV*, vol. 1915, International Society for Optics and Photonics, 1993, pp. 36--49.

[113] S. P. Williams and R. V. Parrish, “New computational control techniques and increased understanding for stereo 3-D displays,” in *Stereoscopic Displays and Applications*, vol. 1256, International Society for Optics and Photonics, 1990, pp. 73--83.

[114] C. Ware, C. Gobrecht and M. A. Paton, “Dynamic adjustment of stereo display parameters,” *IEEE transactions on systems, man, and cybernetics-part A: systems and humans,* vol. 28, no. 1, pp. 56--65, 1998.

[115] V. S. Grinberg, G. W. Podnar and M. Siegel, “Geometry of binocular imaging,” in *Stereoscopic Displays and Virtual Reality Systems*, vol. 2177, International Society for Optics and Photonics, 1994, pp. 56--66.

[116] D. Kim, S. Choi and K. Sohn, “Depth adjustment for stereoscopic images and subjective preference evaluation,” *Journal of Electronic Imaging,* vol. 20, no. 3, p. 033011, 2011.

[117] P. Didyk, T. Ritschel, E. Eisemann, K. Myszkowski and H.-P. Seidel, “A perceptual model for disparity,” in *ACM Transactions on Graphics (TOG)*, vol. 30, ACM, 2011, p. 96.

[118] J. Li, M. Barkowsky and P. Le Callet, “Visual discomfort of stereoscopic 3D videos: Influence of 3D motion,” *Displays,* vol. 35, no. 1, pp. 49--57, 2014.

[119] S. Yano, M. Emoto and T. Mitsuhashi, “Two factors in visual fatigue caused by stereoscopic HDTV images,” *Displays,* vol. 25, no. 4, pp. 141--150, 2004.

[120] M. Lambooij, W. A. IJsselsteijn and I. Heynderickx, “Visual discomfort of 3D TV: Assessment methods and modeling,” *Displays,* vol. 32, no. 4, pp. 209--218, 2011.

[121] T. Bando, A. Iijima and S. Yano, “Visual fatigue caused by stereoscopic images and the search for the requirement to prevent them: A review,” *Displays,* vol. 33, no. 2, pp. 76--83, 2012.

[122] S.-H. Cho and H.-B. Kang, “An Assessment of Visual Discomfort Caused by Motion-in-Depth in Stereoscopic 3D Video,” in *BMVC*, 2012, pp. 1--10.

[123] I. P. Howard and B. J. Rogers, Binocular vision and stereopsis, New York: Oxford University Press, USA, 1995.

[124] J. Li, M. Barkowsky, J. Wang and P. Le Callet, “Study on visual discomfort induced by stimulus movement at fixed depth on stereoscopic displays using shutter glasses,” in *Digital Signal Processing (DSP), 2011 17th International Conference*, IEEE, 2011, pp. 1--8.

[125] B. Mendiburu, 3D movie making: stereoscopic digital cinema from script to screen, Focal press, 2012.

[126] G. Sun and N. Holliman, “Evaluating methods for controlling depth perception in stereoscopic cinematography,” in *Stereoscopic Displays and Applications XX*, vol. 7237, International Society for Optics and Photonics, 2009, p. 72370I.

[127] J. P. Wann, S. Rushton and M. Mon-Williams, “Natural problems for stereoscopic depth perception in virtual environments,” *Vision research,* vol. 35, no. 19, pp. 2731--2736, 1995.

[128] S. Koppal, C. L. Zitnick, M. Cohen, S. B. Kang, B. Ressler and A. Colburn, “A viewer-centric editor for 3D movies,” *IEEE Computer Graphics and Applications,* vol. 31, no. 1, pp. 20--35, 2011.

[129] M. Emoto, T. Niida and F. Okano, “Repeated vergence adaptation causes the decline of visual functions in watching stereoscopic television,” *Journal of display technology,* vol. 1, no. 2, p. 328, 2005.

[130] E. Peli, “The visual effects of head-mounted display (HMD) are not distinguishable from those of desk-top computer display,” *Vision research,* vol. 38, no. 13, pp. 2053--2066, 1998.

[131] A. OOHIRA and M. OCHIAI, “Influence on visual function by a stereoscopic TV programme with binocular liquid crystal shutter and Hi-Vision TV display,” *Ergonomics,* vol. 39, no. 11, pp. 1310--1314, 1996.

[132] S. Pastoor, “Human factors of 3D imaging: results of recent research at Heinrich-Hertz-Institut Berlin,” in *Proc. 2nd International Display Workshop IDW'95,(Oct.) Hamamatsu, Japan*, vol. 3, 1995, pp. 69--72.

[133] R. North and D. B. Henson, “Adaptation to prism-induced heterophoria in subjects with abnormal binocular vision or asthenopia,” *American journal of optometry and physiological optics,* vol. 58, no. 9, pp. 746--752, 1981.

[134] B. Winn, B. Gilmartin, D. L. Sculfor and J. C. Bamford, “Vergence adaptation and senescence.,” *Optometry and Vision Science,* vol. 71, pp. 1-4, 1994.

[135] J. F. Norman, T. E. Dawson and A. K. Butler, “The effects of age upon the perception of depth and 3-D shape from differential motion and binocular disparity,” *Perception,* vol. 29, no. 11, pp. 1335--1359, 2000.

[136] J. F. Norman, H. F. Norman, A. E. Craft, C. L. Walton, A. N. Bartholomew, C. L. Burton, E. Y. Wiesemann and C. E. Crabtree, “Stereopsis and aging,” *Vision research,* vol. 48, no. 23-24, pp. 2456--2465, 2008.

[137] S. Yano, S. Ide, T. Mitsuhashi and H. Thwaites, “A study of visual fatigue and visual comfort for 3D HDTV/HDTV images,” *Displays,* vol. 23, no. 4, pp. 191--201, 2002.

[138] T. Shibata, J. Kim, D. M. Hoffman and M. S. Banks, “The zone of comfort: Predicting visual discomfort with stereo displays,” *Journal of vision,* vol. 11, no. 8, pp. 11--11, 2011.

[139] T. Oskam, A. Hornung, H. Bowles, K. Mitchell and M. H. Gross, “OSCAM-optimized stereoscopic camera control for interactive 3D.,” *ACM Trans. Graph.,* vol. 30, no. 6, pp. 189--1, 2011.

[140] C. M. Schor and C. W. Tyler, “Spatio-temporal properties of Panum's fusional area,” *Vision Research,* vol. 21, no. 5, pp. 683--692, 1981.

[141] J. Kuze and K. Ukai, “Subjective evaluation of visual fatigue caused by motion images,” *Displays,* vol. 29, no. 2, pp. 159--166, 2008.

[142] W. J. Tam, F. Speranza, S. Yano, K. Shimono and H. Ono, “Stereoscopic 3D-TV: visual comfort,” *IEEE Transactions on Broadcasting,* vol. 57, no. 2, pp. 335--346, 2011.

[143] M. Lambooij, M. Fortuin, I. Heynderickx and W. IJsselsteijn, “Visual discomfort and visual fatigue of stereoscopic displays: A review,” *Journal of Imaging Science and Technology,* vol. 53, no. 3, pp. 30201--1, 2009.

[144] W. J. Tam, F. Speranza, C. Vázquez, R. Renaud and N. Hur, “Visual comfort: stereoscopic objects moving in the horizontal and mid-sagittal planes,” in *Stereoscopic Displays and Applications XXIII*, vol. 8288, International Society for Optics and Photonics, 2012, p. 828813.

[145] W. Chen, J. Fournier, M. Barkowsky and P. Le Callet, “New stereoscopic video shooting rule based on stereoscopic distortion parameters and comfortable viewing zone,” in *stereoscopic displays and applications XXII*, vol. 7863, International Society for Optics and Photonics, 2011, p. 78631O.

[146] M. Lang, A. Hornung, O. Wang, S. Poulakos, A. Smolic and M. Gross, “Nonlinear disparity mapping for stereoscopic 3D,” in *ACM Transactions on Graphics (TOG)*, vol. 29, ACM, 2010, p. 75.

[147] R. Neuman, Interviewee, *Chief Stereographer, Disney Animation Studios.* [Interview]. 2009.

[148] I. Tsirlin, L. M. Wilcox and R. S. Allison, “The effect of crosstalk on the perceived depth from disparity and monocular occlusions,” *IEEE Transactions on broadcasting,* vol. 57, no. 2, pp. 445--453, 2011.

[149] S. J. Daly, R. T. Held and D. M. Hoffman, “Perceptual issues in stereoscopic signal processing,” *IEEE Transactions on broadcasting,* vol. 57, no. 2, pp. 347--361, 2011.

[150] V. De Silva, A. Fernando, S. Worrall, H. K. Arachchi and A. Kondoz, “Sensitivity analysis of the human visual system for depth cues in stereoscopic 3-D displays,” *IEEE Transactions on Multimedia,* vol. 13, no. 3, pp. 498--506, 2011.

[151] C. Hendrix and W. Barfield, “Relationship between monocular and binocular depth cues for judgements of spatial information and spatial instrument design,” *Displays,* vol. 16, no. 3, pp. 103--113, 1995.

[152] S.-W. Jung and S.-J. Ko, “Depth sensation enhancement using the just noticeable depth difference,” *IEEE Transactions on Image Processing,* vol. 21, no. 8, pp. 3624--3637, 2012.

[153] H. Van Hateren, “A cellular and molecular model of response kinetics and adaptation in primate cones and horizontal cells,” *Journal of vision,* vol. 5, no. 4, pp. 5--5, 2005.

[154] A. B. Watson, A. J. Ahumada and J. E. Farrell, “Window of visibility: a psychophysical theory of fidelity in time-sampled visual motion displays,” *JOSA A,* vol. 3, no. 3, pp. 300--307, 1986.

[155] D. Kelly, “Motion and vision. II. Stabilized spatio-temporal threshold surface,” *Josa,* vol. 69, no. 10, pp. 1340--1349, 1979.

[156] B. Girod, “Motion compensation: visual aspects, accuracy, and fundamental limits,” in *Motion analysis and image sequence processing*, Springer, 1993, pp. 125--152.

[157] J. Laird, M. Rosen, J. Pelz, E. Montag and S. Daly, “Spatio-velocity CSF as a function of retinal velocity using unstabilized stimuli,” in *Human Vision and Electronic Imaging XI*, vol. 6057, International Society for Optics and Photonics, 2006, p. 605705.

[158] Y. Nojiri, H. Yamanoue, S. Ide, S. Yano and F. Okana, “Parallax distribution and visual comfort on stereoscopic HDTV,” in *Proc. IBC*, vol. 373, 2006.

[159] M. Lambooij, M. Murdoch, W. A. IJsselsteijn and I. Heynderickx, “The impact of video characteristics and subtitles on visual comfort of 3D TV,” *Displays,* vol. 34, no. 1, pp. 8--16, 2013.

[160] D. B. Henson and B. G. Dharamshi, “Oculomotor adaptation to induced heterophoria and anisometropia,” *Investigative ophthalmology & visual science,* vol. 22, no. 2, pp. 234--240, 1982.

[161] L. M. Wilcox, R. S. Allison, J. Helliker, B. Dunk and R. C. Anthony, “Evidence that viewers prefer higher frame-rate film,” *ACM Transactions on Applied Perception (TAP),* vol. 12, no. 4, p. 15, 2015.

[162] J. W. Miller, “Study of visual acuity during the ocular pursuit of moving test objects. II. Effects of direction of movement, relative movement, and illumination,” *JOSA,* vol. 48, no. 11, pp. 803--808, 1958.

[163] M. T. Lambooij, W. A. IJsselsteijn and I. Heynderickx, “Visual discomfort in stereoscopic displays: a review,” in *Stereoscopic Displays and Virtual Reality Systems XIV*, vol. 6490, International Society for Optics and Photonics, 2007, p. 64900I.

[164] J. Li, M. Barkowsky and P. Le Callet, “The influence of relative disparity and planar motion velocity on visual discomfort of stereoscopic videos,” in *International Workshop on Quality of Multimedia Experience QoMEX*, 2011, pp. pp--1.

[165] Y. J. Jung, S.-i. Lee, H. Sohn, H. Park and Y. M. Ro, “Visual comfort assessment metric based on salient object motion information in stereoscopic video,” *Journal of Electronic Imaging,* vol. 21, no. 1, p. 011008, 2012.

[166] J. Häkkinen, T. Kawai, J. Takatalo, T. Leisti, J. Radun, A. Hirsaho and G. Nyman, “Measuring Stereoscopic Image Quality Experience with Interpretation,” in *Image Quality and System Performance V*, vol. 6808, International Society for Optics and Photonics, 2008, p. 68081B.

[167] S.-P. Du, B. Masia, S.-M. Hu and D. Gutierrez, “A metric of visual comfort for stereoscopic motion,” *ACM Transactions on Graphics (TOG),* vol. 32, no. 6, p. 222, 2013.

[168] A. Jain and J. Konrad, “Crosstalk in automultiscopic 3-D displays: Blessing in disguise?,” in *Stereoscopic Displays and Virtual Reality Systems XIV*, vol. 6490, International Society for Optics and Photonics, 2007, p. 649012.

[169] P. J. Seuntiëns, L. M. Meesters and W. A. IJsselsteijn, “Perceptual attributes of crosstalk in 3D images,” *Displays,* vol. 26, no. 4-5, pp. 177--183, 2005.

[170] D. M. Hoffman, A. R. Girshick, K. Akeley and M. S. Banks, “Vergence--accommodation conflicts hinder visual performance and cause visual fatigue,” *Journal of vision,* vol. 8, no. 3, pp. 33--33, 2008.

[171] H. K. Hatada, Y. Ishigure, S. Suyama, H. Takada, M. Date, K. Ishikawa and Toyohiko, “Effect on Depth Perception by a Blur in a Depth-fused 3-D Display,” *The Journal of The Institute of Image Information and Television Engineers,* vol. 60, no. 3, pp. 431-438, 2006.

[172] J. A. Marshall, C. A. Burbeck, D. Ariely, J. P. Rolland and K. E. Martin, “Occlusion edge blur: a cue to relative visual depth,” *JOSA A,* vol. 13, no. 4, pp. 681--688, 1996.

[173] G. Mather, “The use of image blur as a depth cue,” *Perception,* vol. 26, no. 9, pp. 1147--1158, 1997.

[174] G. Mather and D. R. Smith, “Depth cue integration: stereopsis and image blur,” *Vision research,* vol. 40, no. 25, pp. 3501--3506, 2000.

[175] G. Mather and D. R. Smith, “Blur discrimination and its relation to blur-mediated depth perception,” *Perception,* vol. 31, no. 10, pp. 1211--1219, 2002.

[176] R. P. O'Shea, D. G. Govan and R. Sekuler, “Blur and contrast as pictorial depth cues,” *Perception,* vol. 26, no. 5, pp. 599--612, 1997.

[177] B. Masia, G. Wetzstein, C. Aliaga, R. Raskar and D. Gutierrez, “Display adaptive 3D content remapping,” *Computers & Graphics,* vol. 37, no. 8, pp. 983--996, 2013.

[178] L. Lipton, Foundations of the stereoscopic cinema: a study in depth, Van Nostrand Reinhold, 1982.

[179] S. Heinzle, P. Greisen, D. Gallup, C. Chen, D. Saner, A. Smolic, A. Burg, W. Matusik and M. Gross, “Computational stereo camera system with programmable control loop,” in *ACM Transactions on Graphics (TOG)*, vol. 30, ACM, 2011, p. 94.

[180] C. Nitschke, A. Nakazawa and H. Takemura, “Display-camera calibration using eye reflections and geometry constraints,” *Computer Vision and Image Understanding,* vol. 115, no. 6, pp. 835--853, 2011.

[181] T. Bonfort, P. Sturm and P. Gargallo, “General specular surface triangulation,” in *Asian Conference on Computer Vision*, Springer, 2006, pp. 872--881.

[182] N. Funk and Y.-H. Yang, “Using a raster display for photometric stereo,” in *Proceedings of the IEEE Canadian Conference on Computer and Robot Vision*, IEEE, 2007, pp. 201--207.

[183] Y. Francken, C. Hermans and P. Bekaert, “Screen-camera calibration using gray codes,” in *Computer and Robot Vision, 2009. CRV'09. Canadian Conference on*, IEEE, 2009, pp. 155--161.

[184] M. Tarini, H. P. Lensch, M. Goesele and H.-P. Seidel, “3D acquisition of mirroring objects using striped patterns,” *Graphical Models,* vol. 67, no. 4, pp. 233--259, 2005.

[185] Y. Francken, C. Hermans, T. Cuypers and P. Bekaert, “Fast normal map acquisition using an LCD screen emitting gradient patterns,” in *Computer and Robot Vision, 2008. CRV'08. Canadian Conference on*, IEEE, 2008, pp. 189--195.

[186] M. Sharma, S. Chaudhury, B. Lall and M. Venkatesh, “A flexible architecture for multi-view 3DTV based on uncalibrated cameras,” *Journal of Visual Communication and Image Representation,* vol. 25, no. 4, pp. 599--621, 2014.

[187] B. Ward, S. B. Kang and E. Bennett, “Depth director: A system for adding depth to movies,” *IEEE Computer Graphics and Applications,* vol. 31, no. 1, pp. 36--48, 2011.

[188] W. A. IJsselsteijn, H. de Ridder and J. Vliegen, “Subjective evaluation of stereoscopic images: effects of camera parameters and display duration,” *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 10, no. 2, pp. 225--233, 2000.

[189] P. Milgram and M. Krüger, “Adaptation effects in stereo due to on-line changes in camera configuration,” in *Stereoscopic Displays and Applications III*, vol. 1669, International Society for Optics and Photonics, 1992, pp. 122--135.

[190] A. V. Bhavsar and A. N. Rajagopalan, “Depth Estimation with a Practical Camera,” in *BMVC*, 2009, pp. 1--11.

[191] “The Physics Factbook - Focal Length Of A Human Eye,” [Online]. Available: https://hypertextbook.com/facts/2002/JuliaKhutoretskaya.shtml. [Accessed 30 July 2018].

[192] R. Serway, R. Beichner and J. Jewett, Physics for Scientists and Engineers, Saunders golden sunburst series ed., Saunders College Publishing, 2000.

[193] J. Cameron, J. Skofronick and R. Grant, Physics of the Body, Second Edition. Madison, WI: Medical Physics Publishing ed., Medical Physics Pub., 1999.

[194] L. Grossweiner, The Science of Phototherapy, Taylor & Francis, 1994.

[195] University of Virginia, “Phys 531, Lecture 11, Survey of Optical Systems [pdf],” 7 October 2004. [Online]. Available: http://galileo.phys.virginia.edu/classes/531.cas8m.fall04/l11.pdf. [Accessed 30 July 2018].

[196] “Human eye,” [Online]. Available: https://en.wikipedia.org/wiki/Human_eye#cite_note-SavinoDanesh-Meyer2012-7. [Accessed 30 July 2018].

[197] P. J. Savino and H. V. Danesh-Meyer, Color Atlas and Synopsis of Clinical Ophthalmology--Wills Eye Institute--Neuro-Ophthalmology, Lippincott Williams & Wilkins, 2012, p. 12.

[198] S. Ryan, A. Schachat, C. Wilkinson, D. Hinton, S. Sadda and P. Wiedemann, Retina E-Book, Elsevier Health Sciences, 2012, p. 342.

[199] W. Trattler, P. Kaiser and N. Friedman, Review of Ophthalmology E-Book: Expert Consult - Online and Print, Elsevier Health Sciences, 2012, p. 255.

[200] I. P. Howard, Seeing in depth, Vol. 1: Basic mechanisms., University of Toronto Press, 2002.

[201] G. Smith and D. A. Atchison, THE EYE AND VISUAL OPTICAL INSTRUMENTS. Edition en anglais, Cambridge University Press, 1997.

[202] W. Chen, J. Fournier, M. Barkowsky and P. Le Callet, “New requirements of subjective video quality assessment methodologies for 3DTV,” in *Video Processing and Quality Metrics 2010 (VPQM)*, 2010.

[203] M. Lambooij, W. IJsselsteijn and I. Heynderickx, “Stereoscopic displays and visual comfort: a review,” *Einhoven University of Technology Einhoven Univeristy of Technology Philips Research, SPIE Newsroom 2 April,* 2007.

[204] ITU, “Subjective Assessment of Stereoscope Television Pictures,” in *RECOMMENDATION ITU-R BT.1438*, 2000.

[205] A. S. Percival, “The relation of convergence to accommodation and its practical bearing,” *Ophthal. Rev.,* vol. 11, pp. 313--328, 1892.

[206] Y. Nojiri, H. Yamanoue, A. Hanazato and F. Okano, “Measurement of parallax distribution and its application to the analysis of visual comfort for stereoscopic HDTV,” in *Stereoscopic Displays and Virtual Reality Systems X*, vol. 5006, International Society for Optics and Photonics, 2003, pp. 195--206.

[207] A. K. Moorthy, C.-C. Su, A. Mittal and A. C. Bovik, “Subjective evaluation of stereoscopic image quality,” *Signal Processing: Image Communication,* vol. 28, no. 8, pp. 870--883, 2013.

[208] D. L. MacAdam, “Stereoscopic perceptions of size, shape, distance and direction,” *Journal of the Society of Motion Picture and Television Engineers,* vol. 62, no. 4, pp. 271--293, 1954.

[209] F. Solari, M. Chessa, M. Garibotti and S. P. Sabatini, “Natural perception in dynamic stereoscopic augmented reality environments,” *Displays,* vol. 34, no. 2, pp. 142--152, 2013.

[210] D. A. Southard, “Transformations for stereoscopic visual simulation,” *Computers & Graphics,* vol. 16, no. 4, pp. 401--410, 1992.

[211] D. Kim and K. Sohn, “Visual fatigue prediction for stereoscopic image,” *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 21, no. 2, pp. 231--236, 2011.

[212] H. Yamanoue, “The relation between size distortion and shooting conditions for stereoscopic images,” *SMPTE journal,* vol. 106, no. 4, pp. 225--232, 1997.

[213] J. Baker, "Generating images for a time-multiplexed stereoscopic computer graphics system," *True Three-Dimensional Imaging Techniques & Display Technologies,* vol. 761, pp. 44--53, 1987.

[214] L. F. Hodges and E. T. Davis, "Geometric considerations for stereoscopic virtual environments," *Presence: Teleoperators & Virtual Environments,* vol. 2, no. 1, pp. 34--43, 1993.

[215] S. Kitrosser, "Photography in the service of stereoscopy," *Journal of Imaging Science and Technology,* vol. 42, no. 4, pp. 295--300, 1998.

[216] W. Robinett and R. Holloway, "The visual display transformation for virtual reality," *Presence: Teleoperators & Virtual Environments,* vol. 4, no. 1, pp. 1--23, 1995.

[217] D. A. Southard, "Viewing model for virtual environment displays," *Journal of Electronic Imaging,* vol. 4, no. 4, pp. 413--421, 1995.

[218] Z. Wartell, L. F. Hodges and W. Ribarsky, "Balancing fusion, image depth and distortion in stereoscopic head-tracked displays," in *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, ACM Press/Addison-Wesley Publishing Co., 1999, pp. 351--358.

[219] H. Yamanoue, "The differences between toed-in camera configurations and parallel camera configurations in shooting stereoscopic images," in *Multimedia and Expo, 2006 IEEE International Conference on*, IEEE, 2006, pp. 1701--1704.

[220] H. Yamanoue, M. Okui and F. Okano, "Geometrical analysis of puppet-theater and cardboard effects in stereoscopic HDTV images," *IEEE Transactions on circuits and systems for video technology,* vol. 16, no. 6, pp. 744--752, 2006.

[221] K. Masaoka, A. Hanazato, M. Emoto, H. Yamanoue, Y. Nojiri and F. Okano, "Spatial distortion prediction system for stereoscopic images," *Journal of Electronic Imaging,* vol. 15, no. 1, p. 013002, 2006.

[222] R. Allison, "Analysis of the influence of vertical disparities arising in toed-in stereoscopic cameras," *Journal of Imaging Science and Technology,* vol. 51, no. 4, pp. 317--327, 2007.

[223] "Committee Draft of ISO/IEC 23002–3 Auxiliary Video Data Representations," in *ISO/IEC JTC 1/SC 29/WG 11*, 2006.

[224] D. Astruc, S. Cazin, E. Cid, O. Eiff, L. Lacaze, P. Robin, F. Toublanc and I. Cáceres, "A stereoscopic method for rapid monitoring of the spatio-temporal evolution of the sand-bed elevation in the swash zone," *Coastal Engineering,* vol. 60, pp. 11--20, 2012.

[225] C. Kim, A. Hornung, S. Heinzle, W. Matusik and M. Gross, "Multi-perspective stereoscopy from light fields," *ACM Transactions on Graphics (TOG),* vol. 30, no. 6, p. 190, 2011.

[226] F. Devernay, S. Duchêne and A. Ramos-Peon, "Adapting stereoscopic movies to the viewing conditions using depth-preserving and artifact-free novel view synthesis," in

Stereoscopic Displays and Applications XXII, vol. 7863, International Society for Optics and Photonics, 2011, p. 786302.

[227] M. Obrist, D. Wurhofer, T. Meneweger, T. Grill and M. Tscheligi, "Viewing experience of 3DTV: An exploration of the feeling of sickness and presence in a shopping mall," *Entertainment computing,* vol. 4, no. 1, pp. 71--81, 2013.

[228] H. E. Tasli and A. A. Alatan, "User assisted disparity remapping for stereo images," *Signal Processing: Image Communication,* vol. 28, no. 10, pp. 1374--1389, 2013.

[229] T. Yan, R. W. Lau, Y. Xu and L. Huang, "Depth mapping for stereoscopic videos," *International Journal of Computer Vision,* vol. 102, no. 1-3, pp. 293--307, 2013.

[230] H. Azari, I. Cheng, K. Daniilidis and A. Basu, "Optimal pixel aspect ratio for enhanced 3D TV visualization," *Computer Vision and Image Understanding,* vol. 116, no. 1, pp. 38--53, 2012.

[231] B. T. Backus, D. J. Fleet, A. J. Parker and D. J. Heeger, "Human cortical activity correlates with stereoscopic depth perception," *Journal of neurophysiology,* vol. 86, no. 4, pp. 2054--2068, 2001.

[232] I. Cheng and A. Basu, "Perceptually optimized 3-D transmission over wireless networks," *IEEE Transactions on Multimedia,* vol. 9, no. 2, pp. 386--396, 2007.

[233] L.-F. Cheong, "Depth perception under motion and stereo with implications for 3D TV," in *3DTV Conference, 2007*, IEEE, 2007, pp. 1--4.

[234] Y. Pan, I. Cheng and A. Basu, "Quality metric for approximating subjective evaluation of 3-D objects," *IEEE Transactions on Multimedia,* vol. 7, no. 2, pp. 269--279, 2005.

[235] L. Cheong, C. Fermüller and Y. Aloimonos, "Effects of errors in the viewing geometry on shape estimation," *Computer Vision and Image Understanding,* vol. 71, no. 3, pp. 356--372, 1998.

[236] T. Brodský, C. Fermüller and Y. Aloimonos, "Structure from motion: Beyond the epipolar constraint," *International Journal of Computer Vision,* vol. 37, no. 3, pp. 231--258, 2000.

[237] N. Ayache, Artificial vision for mobile robots: stereo vision and multisensory perception, Mit Press, 1991.

[238] T. Shibata, S. Kurihara, T. Kawai, T. Takahashi, T. Shimizu, R. Kawada, A. Ito, J. Häkkinen, J. Takatalo and G. Nyman, "Evaluation of stereoscopic image quality for mobile devices using interpretation based quality methodology," in *Stereoscopic Displays and Applications XX*, vol. 7237, International Society for Optics and Photonics, 2009, p. 72371E.

[239] T. Mitsuhashi, "A study of the relationship between scanning specifications and picture quality," NHK Technical Research Laboratories, 1980.

[240] B. Javidi and F. Okano, "Three-dimensional television, video, and display technologies," Springer Science & Business Media, 2002.

[241] T. D. Basha, Y. Moses and S. Avidan, “Stereo seam carving a geometrically consistent approach,” *IEEE transactions on pattern analysis and machine intelligence,* vol. 35, no. 10, pp. 2513--2525, 2013.

[242] G. Humphreys and P. Hanrahan, “A distributed graphics system for large tiled displays,” in *Proceedings of the conference on Visualization'99: celebrating ten years*, IEEE Computer Society Press, 1999, pp. 215--223.

[243] R. Raskar, M. S. Brown, R. Yang, W. C. Chen, G. Welch, H. Towles, B. Seales and H. Fuchs, “Multi-projector displays using camera-based registration,” in *Proceedings of the conference on Visualization'99: celebrating ten years*, IEEE Computer Society Press, 1999, pp. 161--168.

[244] A. Majumder and R. Stevens, “Perceptual photometric seamlessness in projection-based tiled displays,” *ACM Transactions on Graphics (TOG),* vol. 24, no. 1, pp. 118--139, 2005.

[245] M. Brown, A. Majumder and R. Yang, “Camera-based calibration techniques for seamless multiprojector displays,” *IEEE Transactions on Visualization and Computer Graphics,* vol. 11, no. 2, pp. 193--206, 2005.

[246] A. Majumder, Z. He, H. Towles and G. Welch, “Achieving color uniformity across multi-projector displays,” in *Visualization 2000. Proceedings*, IEEE, 2000, pp. 117--124.

[247] A. Boev, R. Bregovic and A. Gotchev, “Visual-quality evaluation methodology for multiview displays,” *Displays,* vol. 33, no. 2, pp. 103--112, 2012.

[248] J. D. Pfautz, “Depth perception in computer graphics,” University of Cambridge, Computer Laboratory, 2002.

[249] H. Pan, X.-F. Feng and S. Daly, “LCD motion blur modeling and analysis,” in *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, vol. 2, IEEE, 2005, pp. II--21.

[250] “STARVR,” Starbreeze Studios and Acer, [Online]. Available: https://www.starvr.com/. [Accessed 31 July 2018].

[251] P. Jackson, Interviewee, *Peter Jackson Responds to 'Hobbit' Footage Critics, Explains 48-Frames Strategy.* [Interview]. 24 April 2012.

[252] E. Engstrom, “A study of television image characteristics: Part Two: Determination of frame frequency for television in terms of flicker characteristics,” *Proceedings of the Institute of Radio Engineers,* vol. 23, no. 4, pp. 295--310, 1935.

[253] T. Kurita, “Moving Picture Quality Improvement for Hold-type AM-LCDs,” in *SID Symposium Digest of Technical Papers*, vol. 32, Wiley Online Library, 2001, pp. 986--989.

[254] D. M. Hoffman, V. I. Karasev and M. S. Banks, “Temporal presentation protocols in stereoscopic displays: Flicker visibility, perceived motion, and perceived depth,” *Journal of the Society for Information Display,* vol. 19, no. 3, pp. 271--297, 2011.

[255] Y. Kuroki, “Improvement of 3D visual image quality by using high frame rate,” *Journal of the Society for Information Display,* vol. 20, no. 10, pp. 566--574, 2012.

[256] J. Liang, C. Shaw and M. Green, “On temporal-spatial realism in the virtual reality environment,” in *Proceedings of the 4th annual ACM symposium on User interface software and technology*, ACM, 1991, pp. 19--25.

[257] A. Mazikowski and J. Lebiedź, “Image projection in immersive 3D visualization laboratory,” *Procedia Computer Science,* vol. 35, pp. 842--850, 2014.

[258] M. Barkowsky, P. Campisi, P. Le Callet and V. Rizzo, “Crosstalk measurement and mitigation for autostereoscopic displays,” in *Three-Dimensional Image Processing (3DIP) and Applications*, vol. 7526, International Society for Optics and Photonics, 2010, p. 75260R.

[259] F. L. Kooi and A. Toet, “Visual comfort of binocular and 3D displays,” *Displays,* vol. 25, no. 2-3, pp. 99--108, 2004.

[260] A. J. Woods and C. R. Harris, “Comparing levels of crosstalk with red/cyan, blue/yellow, and green/magenta anaglyph 3D glasses,” in *Stereoscopic displays and applications XXI*, vol. 7524, International Society for Optics and Photonics, 2010, p. 75240Q.

[261] M. Salmimaa and T. Järvenpää, “3-D crosstalk and luminance uniformity from angular luminance profiles of multiview autostereoscopic 3-D displays,” *Journal of the Society for Information Display,* vol. 16, no. 10, pp. 1033--1040, 2008.

[262] S. Pastoor, “Wahrnehmungsgrenzen fu¨r stationa¨res u¨bersprechen,” Internal Report HHI BMFT-Vorhaben: 01 BK 101.

[263] A. Hanazato, M. Okui, H. Yamanoue and I. Yuyama, “Evaluation of crosstalk in stereoscopic display,” *Proceedings 3D Image Conference,* p. 258–263, 1999.

[264] L. Lipton, “Factors affecting" ghosting" in time-multiplexed piano-stereoscopic CRT display systems,” in *True Three-Dimensional Imaging Techniques & Display Technologies*, vol. 761, International Society for Optics and Photonics, 1987, pp. 75--79.

[265] L. M. Wilcox and J. A. Stewart, “Determinants of perceived image quality: ghosting vs. brightness,” in *Stereoscopic Displays and Virtual Reality Systems X*, vol. 5006, International Society for Optics and Photonics, 2003, pp. 263--269.

[266] A. Hanazato, “Evaluation of Cross Talk in Stereoscopic Display,” in *3D conference, 1999*, 1999.

[267] S. Pala, R. Stevens and P. Surman, “Optical cross-talk and visual comfort of a stereoscopic display used in a real-time application,” in *Stereoscopic Displays and Virtual Reality Systems XIV*, vol. 6490, International Society for Optics and Photonics, 2007, p. 649011.

[268] M. Lambooij, M. Fortuin and I. Heynderickx, “Visual discomfort associated with 3D displays,” in 5th Int. Workshop Video Process. Quality Metrics for Consum. Electron,” in *Dept. Of Electrical and Electronic Eng., Yonsei University, Seoul, South Korea*, vol. 7237, Citeseer, 2010.

[269] E. Chino, K. Tajiri, H. Kawakami, H. Ohira, K. Kamijo, H. Kaneko, S. Kato, Y. Ozawa, T. Kurumisawa, K. Inoue and others, “Development of Wide-Color-Gamut Mobile Displays with

Four-Primary-Color LCDs," in *SID symposium digest of technical papers*, vol. 37, Wiley Online Library, 2006, pp. 1221--1224.

[270] S. Ueki, K. Nakamura, Y. Yoshida, T. Mori, K. Tomizawa, Y. Narutaki, Y. Itoh and K. Okamoto, "Five-Primary-Color 60-Inch LCD with Novel Wide Color Gamut and Wide Viewing Angle," in *SID Symposium Digest of Technical Papers*, vol. 40, Wiley Online Library, 2009, pp. 927--930.

[271] H.-C. Cheng, I. Ben-David and S.-T. Wu, "Five-Primary-Color LCDs," *Journal of Display Technology,* vol. 6, no. 1, pp. 3--7, 2010.

[272] Y.-C. Yang, K. Song, S. Rho, N.-S. Rho, S. Hong, K. B. Deul, M. Hong, K. Chung, W. Choe, S. Lee and others, "Development of Six Primary-Color LCD," in *SID Symposium Digest of Technical Papers*, vol. 36, Wiley Online Library, 2005, pp. 1210--1213.

[273] H. Sugiura, H. Kaneko, S. Kagawa, M. Ozawa, J. Someya, H. Tanizoe, H. Ueno and T. Kimura, "Improved Six-Primary-Color 23-in. WXGA LCD using Six-Color LEDs," in *SID Symposium Digest of Technical Papers*, vol. 37, Wiley Online Library, 2006, pp. 1126--1129.

[274] H. Sugiura, H. Kaneko, S. Kagawa, J. Someya and H. Tanizoe, "Six-primary-color lcd monitor using six-color leds with an accurate calibration system," in *Color Imaging XI: Processing, Hardcopy, and Applications*, vol. 6058, International Society for Optics and Photonics, 2006, p. 60580H.

[275] T. Ajito, T. Obi, M. Yamaguchi and N. Ohyama, "Expanded color gamut reproduced by six-primary projection display," in *Projection Displays 2000: Sixth in a Series*, vol. 3954, International Society for Optics and Photonics, 2000, pp. 130--138.

[276] S. Roth, I. Ben-David, M. Ben-Chorin, D. Eliav and O. Ben-David, "Wide gamut, high brightness multiple primaries single panel projection displays," in *SID Symposium Digest of Technical Papers*, vol. 34, Wiley Online Library, 2003, pp. 118--121.

[277] M. S. Brennesholtz, S. C. McClain, S. Roth and D. Malka, "A Single Panel LCOS Engine with a Rotating Drum and a Wide Color Gamut," in *SID Symposium Digest of Technical Papers*, vol. 36, Wiley Online Library, 2005, pp. 1814--1817.

[278] S. Roth and W. Caldwell, "Four primary color projection display," in *SID Symposium Digest of Technical Papers*, vol. 36, Wiley Online Library, 2005, pp. 1818--1821.

[279] J. Morovic and M. R. Luo, "The fundamentals of gamut mapping: A survey," *Journal of Imaging Science and Technology,* vol. 45, no. 3, pp. 283--290, 2001.

[280] *IEC 61996-2-2. Multimedia systems and equipment - color measurement and management - Part 2-2: Color management—extended RGB color space scRGB,* 2003.

[281] *IEC 61966-2-4 First edition. Multimedia systems and equipment - Color measurement and management: Part 2-4: Color management - Extendedgamut YCC colour space for video applications-xvYCC,* 2006.

[282] W. H. Encyclopedia, "Visual acuity," *VA,* 1965.

[283] B. Sajadi, M. Gopi and A. Majumder, "Edge-guided resolution enhancement in projectors via optical pixel sharing," *ACM Transactions on Graphics (TOG),* vol. 31, no. 4, p. 79, 2012.

[284] B. Sajadi, D. Qoc-Lai, A. H. Ihler, M. Gopi and A. Majumder, "Image enhancement in projectors via optical pixel shift and overlay," in *IEEE International Conference on Computational Photography (ICCP)*, IEEE, 2013, pp. 1--10.

[285] Z.-I. Hara and N. Shiramatsu, "Improvement in the picture quality of moving pictures for matrix displays," *Journal of the Society for Information Display,* vol. 8, no. 2, pp. 129--137, 2000.

[286] A. Basu, "Optimal discretization for stereo reconstruction," *Pattern Recognition Letters,* vol. 13, no. 11, pp. 813--820, 1992.

[287] Q.-H. Wang, X.-F. Li, L. Zhou, A.-H. Wang and D.-H. Li, "Cross-talk reduction by correcting the subpixel position in a multiview autostereoscopic three-dimensional display based on a lenticular sheet," *Applied optics,* vol. 50, no. 7, pp. B1--B5, 2011.

[288] R. Hayes, "3-D movies: A history and filmography of stereoscopic cinema," 1989.

[289] L. Lipton, "Composition for electrostereoscopic displays," in *Stereo computer graphics*, Princeton University Press, 1993, pp. 11--25.

[290] P. Howarth, "Virtual Reality: an occupational health hazard of the future," in *RCN Occupational Health Nurses Forum*, 1994.

[291] S.-n. Yang, T. Schlieski, B. Selmins, S. C. Cooper, R. A. Doherty, P. J. Corriveau and J. E. Sheedy, "Stereoscopic viewing and reported perceived immersion and symptoms," *Optometry and vision science,* vol. 89, no. 7, pp. 1068--1080, 2012.

[292] A. Y. Nee, S. Ong, G. Chryssolouris and D. Mourtzis, "Augmented reality applications in design and manufacturing," *CIRP Annals-manufacturing technology,* vol. 61, no. 2, pp. 657--679, 2012.

[293] J. Juang, W. Hung and S. Kang, "SimCrane 3D+: A crane simulator with kinesthetic and stereoscopic vision," *Advanced Engineering Informatics,* vol. 27, no. 4, pp. 506--518, 2013.

[294] C.-H. Chang, C.-K. Liang and Y.-Y. Chuang, "Content-aware display adaptation and editing for stereoscopic images," in *ACM SIGGRAPH 2011 Posters*, ACM, 2011, p. 30.

[295] P. Didyk, T. Ritschel, E. Eisemann, K. Myszkowski, H.-P. Seidel and W. Matusik, "A luminance-contrast-aware disparity model and applications," *ACM Transactions on Graphics (TOG),* vol. 31, no. 6, p. 184, 2012.

[296] P. Didyk, T. Ritschel, E. Eisemann, K. Myszkowski and H.-P. Seidel, "Apparent stereo: The cornsweet illusion can enhance perceived depth," in *Human Vision and Electronic Imaging XVII*, vol. 8291, International Society for Optics and Photonics, 2012, p. 82910N.

[297] M. F. Bradshaw and B. J. Rogers, "Sensitivity to horizontal and vertical corrugations defined by binocular disparity," *Vision research,* vol. 39, no. 18, pp. 3049--3056, 1999.

[298] H. Wu and K. Rao, “Perceptual video quality metrics - A review,” in *Digital video image quality and perceptual coding*, CRC press, 2005, pp. 191--216.

[299] E. Montag and M. Fairchild, “Fundamentals of human vision and vision modeling,” *Digital video image quality and perceptual coding,* 2006.

[300] F. Xiao, J. M. DiCarlo, P. B. Catrysse and B. A. Wandell, “High dynamic range imaging of natural scenes,” in *Color and Imaging Conference*, vol. 2002, Society for Imaging Science and Technology, 2002, pp. 337--342.

[301] R. Mantiuk, S. Daly and L. Kerofsky, “Display adaptive tone mapping,” in *ACM Transactions on Graphics (TOG)*, vol. 27, ACM, 2008, p. 68.

[302] W. Brown, “Avatar: Stereoscopic cinema, gaseous perception and darkness,” *Animation,* vol. 7, no. 3, pp. 259--271, 2012.

[303] E. Reinhard, W. Heidrich, P. Debevec, S. Pattanaik, G. Ward and K. Myszkowski, High dynamic range imaging: acquisition, display, and image-based lighting, Morgan Kaufmann, 2010.

[304] A. Gorea and C. W. Tyler, “New look at Bloch’s law for contrast,” *JOSA A,* vol. 3, no. 1, pp. 52--61, 1986.

[305] Y. Nojiri, H. Yamanoue, A. Hanazato, M. Emoto and F. Okano, “Visual comfort/discomfort and visual fatigue caused by stereoscopic HDTV viewing,” in *Stereoscopic Displays and Virtual Reality Systems XI*, vol. 5291, International Society for Optics and Photonics, 2004, pp. 303--314.

[306] J. Schild and M. Masuch, “Designing stereoscopic information visualization for 3D-TV: What can we learn from S3D gaming?,” in *Stereoscopic Displays and Applications XXIII*, vol. 8288, International Society for Optics and Photonics, 2012, p. 82882P.

[307] J. Freeman and S. E. Avons, “Focus group exploration of presence through advanced broadcast services,” in *Human Vision and Electronic Imaging V*, vol. 3959, International Society for Optics and Photonics, 2000, pp. 530--540.

[308] W. Barfield, D. Zeltzer, T. Sheridan and M. Slater, “Presence and performance within virtual environments,” *Virtual environments and advanced interface design,* pp. 473--513, 1995.

[309] M. Slater and M. Usoh, “Presence in immersive virtual environments,” in *Virtual Reality Annual International Symposium*, IEEE, 1993, pp. 90--96.

[310] H. Hsu and P. A. Lachenbruch, “Paired t test,” *Wiley Encyclopedia of Clinical Trials,* 2008.

[311] J. C. De Winter, “Using the Student's t-test with extremely small sample sizes,” *Practical Assessment, Research \& Evaluation,* vol. 18, no. 10, 2013.

[312] C. A. Boneau, “The effects of violations of assumptions underlying the t test.,” *Psychological bulletin,* vol. 57, no. 1, p. 49, 1960.

[313] CUTGANA, “3D Tour Project for Cutgana - Dr. Salvatore Livatino & Alessio Regalbuto,” [Online]. Available: http://www.cutgana.unict.it/VirtualTours3D/. [Accessed 17 September 2018].

[314] Google, “Google Forms,” [Online]. Available: https://www.google.co.uk/forms/about/.

[315] M. Slater, D. P. Pertaub and A. Steed, “Public speaking in virtual reality: Facing an audience of avatars,” *IEEE Computer Graphics and Applications,* vol. 19, no. 2, pp. 6--9, 1999.

[316] M. J. Schuemie, P. Van Der Straaten, M. Krijn and C. A. Van Der Mast, “Research on presence in virtual reality: A survey,” *CyberPsychology & Behavior,* vol. 4, no. 2, pp. 183--201, 2001.

[317] J. Prothero and H. Hoffman, “Widening the field of view increases the sense of presence within immersive virtual environments,” *Tech. Rep.,* 1995.

[318] H. G. Hoffman, T. Richards, B. Coda, A. Richards and S. R. Sharar, “The illusion of presence in immersive virtual reality during an fMRI brain scan,” *CyberPsychology & Behavior,* vol. 6, no. 2, pp. 127--131, 2003.

[319] J. Diemer, G. W. Alpers, H. M. Peperkorn, Y. Shiban and A. Mühlberger, “The impact of perception and presence on emotional reactions: a review of research in virtual reality,” *Frontiers in psychology,* vol. 6, p. 26, 2015.

[320] M. F. Macedonio, T. D. Parsons, R. A. Digiuseppe, B. A. Weiderhold and A. A. Rizzo, “Immersiveness and physiological arousal within panoramic video-based virtual reality,” *Cyberpsychology & Behavior,* vol. 10, no. 4, pp. 508--515, 2007.

[321] B. Herbelin, F. Vexo and D. Thalmann, “Sense of presence in virtual reality exposures therapy,” in *Proceedings of the 1st International Workshop on Virtual Reality Rehabilitation, Lausanne, Switzerland*, Citeseer, 2002.

[322] M. P. Huang and N. E. Alessi, “Presence as an emotional experience,” *Studies in health technology and informatics,* pp. 148--153, 1999.

[323] R. S. Renner, B. M. Velichkovsky and J. R. Helmert, “The perception of egocentric distances in virtual environments-a review,” *ACM Computing Surveys (CSUR),* vol. 46, no. 2, p. 23, 2013.

[324] B. J. Mohler, S. H. Creem-Regehr, W. B. Thompson and H. H. Bülthoff, “The effect of viewing a self-avatar on distance judgments in an HMD-based virtual environment,” *Presence: Teleoperators and Virtual Environments,* vol. 19, no. 3, pp. 230--242, 2010.

[325] B. Ries, V. Interrante, M. Kaeding and L. Phillips, “Analyzing the effect of a virtual avatar's geometric and motion fidelity on ego-centric spatial perception in immersive virtual environments,” in *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology*, ACM, 2009, pp. 59--66.

[326] M. Slater, M. Usoh and Y. Chrysanthou, “The influence of dynamic shadows on presence in immersive virtual environments,” in *Virtual Environments’ 95*, Springer, 1995, pp. 8--21.

[327] [Online]. Available: https://www.anandtech.com/show/8554/the-iphone-6-review/8. [Accessed 17 September 2018].

[328] [Online]. Available: https://www.anandtech.com/show/8169/the-lg-g3-review/2. [Accessed 17 September 2018].

[329] J. Freeman and S. E. Avons, “Focus group exploration of presence through advanced broadcast services,” vol. 3959, pp. 530--539, 2000.

[330] M. Slater, V. Linakis, M. Usoh, R. Kooper and G. Street, “Immersion, presence, and performance in virtual environments: An experiment with tri-dimensional chess,” vol. 163, p. 72, 1996.

[331] M. Krijn, P. M. Emmelkamp, R. Biemond, C. d. W. de Ligny, M. J. Schuemie and C. A. van der Mast, “Treatment of acrophobia in virtual reality: The role of immersion and presence,” *Behaviour research and therapy,* vol. 42, no. 2, pp. 229--239, 2004.

[332] R. Pausch, D. Proffitt and G. Williams, “Quantifying immersion in virtual reality,” in *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, ACM Press/Addison-Wesley Publishing Co., 1997, pp. 13--18.

[333] M. F. Macedonio, T. D. Parsons, R. A. Digiuseppe, B. A. Weiderhold and A. A. Rizzo, “Immersiveness and physiological arousal within panoramic video-based virtual reality,” *Cyberpsychology \& Behavior,* vol. 10, no. 4, pp. 508--515, 2007.

[334] B. K. Wiederhold, R. Davis and M. D. Wiederhold, “The effects of immersiveness on physiology.,” 1998.

[335] M. Hemphill, “A note on adults' color--emotion associations,” *The Journal of genetic psychology,* vol. 157, no. 3, pp. 275--280, 1996.

[336] B. G. Witmer and M. J. Singer, “Measuring presence in virtual environments: A presence questionnaire,” *Presence,* vol. 7, no. 3, pp. 225--240, 1998.

[337] C. Hendrix and W. Barfield, “Relationship between monocular and binocular depth cues for judgements of spatial information and spatial instrument design,” *Displays,* vol. 16, no. 3, pp. 103--113, 1995.

[338] K. Berbaum, D. Tharp and K. Mroczek, “Depth perception of surfaces in pictures: Looking for conventions of depiction in Pandora's box,” *Perception,* vol. 12, no. 1, pp. 5--20, 1983.

[339] M. Ferre, R. Aracil and M. A. Sanchez-Uran, “Stereoscopic human interfaces,” *IEEE Robotics & Automation Magazine,* vol. 15, no. 4, 2008.

[340] P. Willemsen, M. B. Colton, S. H. Creem-Regehr and W. B. Thompson, “The effects of head-mounted display mechanical properties and field of view on distance judgments in virtual environments,” *ACM Transactions on Applied Perception (TAP),* vol. 6, no. 2, p. 8, 2009.

[341] J. M. Knapp and J. M. Loomis, “Limited field of view of head-mounted displays is not the cause of distance underestimation in virtual environments,” *Presence: Teleoperators & Virtual Environments,* vol. 13, no. 5, pp. 572--577, 2004.

[342] R. Messing and F. Durgin, “Space perception and cues to distance in virtual reality,” in *Proceedings of the 1st Symposium on Applied perception in graphics and visualization*, ACM, 2004, p. 176.

[343] P. B. Kline and B. G. Witmer, “Distance perception in virtual environments: Effects of field of view and surface texture at near distances,” in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 40, SAGE Publications Sage CA: Los Angeles, CA, 1996, pp. 1112--1116.

[344] B. Wu, T. L. Ooi and Z. J. He, “Perceiving distance accurately by a directional process of integrating ground information,” *Nature,* vol. 428, no. 6978, p. 73, 2004.

[345] B. G. Witmer and W. J. Sadowski Jr, “Nonvisually guided locomotion to a previously viewed target in real and virtual environments,” *Human factors,* vol. 40, no. 3, pp. 478--488, 1998.

[346] I. TM, “Insta360 Pro - Transport your audience: VR in 8K,” Insta360 TM, [Online]. Available: https://www.insta360.com/product/insta360-pro. [Accessed 8 8 2018].

[347] T. B. Sheridan, “Musings on telepresence and virtual presence,” *Presence: Teleoperators & Virtual Environments,* vol. 1, no. 1, pp. 120--126, 1992.

[348] M. Slater, P. Khanna, J. Mortensen and I. Yu, “Visual realism enhances realistic response in an immersive virtual environment,” *IEEE computer graphics and applications,* vol. 29, no. 3, 2009.

[349] F. Steinicke, G. Bruder, K. Hinrichs and A. Steed, “Gradual transitions and their effects on presence and distance estimation,” *Computers & Graphics,* vol. 34, no. 1, pp. 26--33, 2010.

[350] M. Meehan, S. Razzaque, M. C. Whitton and F. P. Brooks Jr, “Effect of latency on presence in stressful virtual environments,” in *IEEE Virtual Reality, 2003. Proceedings.*, IEEE, 2003, pp. 141-148.

[351] C. Hendrix and W. Barfield, “Presence within virtual environments as a function of visual display parameters,” *Presence: Teleoperators & Virtual Environments,* vol. 5, no. 3, pp. 274--289, 1996.

[352] J. E. Cutting, “How the eye measures reality and virtual reality,” *Behavior Research Methods, Instruments, & Computers,* vol. 29, no. 1, pp. 27--36, 1997.

[353] J. W. Kelly, L. S. Donaldson, L. A. Sjolund and J. B. Freiberg, “More than just perception--action recalibration: Walking through a virtual environment causes rescaling of perceived space,” *Attention, Perception, & Psychophysics,* vol. 75, no. 7, pp. 1473--1485, 2013.

[354] P. Willemsen, A. A. Gooch, W. B. Thompson and S. H. Creem-Regehr, “Effects of stereo viewing conditions on distance perception in virtual environments,” *Presence: Teleoperators and Virtual Environments,* vol. 17, no. 1, pp. 91--101, 2008.

[355] R. G. Eggleston, W. P. Janson and K. A. Aldrich, “Virtual reality system effects on size-distance judgements in a virtual environment,” in *Virtual Reality Annual International Symposium, 1996., Proceedings of the IEEE 1996*, IEEE, 1996, pp. 139--146.

[356] V. Interrante, B. Ries and L. Anderson, “Distance perception in immersive virtual environments, revisited,” in *Virtual Reality Conference*, IEEE, 2006, pp. 3--10.

[357] S. R. Ellis and B. M. Menges, “Judgments of the distance to nearby virtual objects: Interaction of viewing conditions and accommodative demand,” *Presence: Teleoperators & Virtual Environments,* vol. 6, no. 4, pp. 452--460, 1997.

[358] R. T. Surdick, E. T. Davis, R. A. King and L. F. Hodges, “The perception of distance in simulated visual displays: A comparison of the effectiveness and accuracy of multiple depth cues across viewing distances,” *Presence: Teleoperators & Virtual Environments,* vol. 6, no. 5, pp. 513--531, 1997.

[359] G. Bruder, A. Pusch and F. Steinicke, “Analyzing effects of geometric rendering parameters on size and distance estimation in on-axis stereographics,” in *Proceedings of the ACM Symposium on Applied Perception*, ACM, 2012, pp. 111--118.

[360] F. Kellner, B. Bolte, G. Bruder, U. Rautenberg, F. Steinicke, M. Lappe and R. Koch, “Geometric calibration of head-mounted displays and its effects on distance estimation,” *IEEE Transactions on Visualization & Computer Graphics,* no. 4, pp. 589--596, 2012.

[361] S. A. Kuhl, W. B. Thompson and S. H. Creem-Regehr, “Minification influences spatial judgments in virtual environments,” in *Proceedings of the 3rd symposium on Applied perception in graphics and visualization*, ACM, 2006, pp. 15--19.

[362] S. A. Kuhl, W. B. Thompson and S. H. Creem-Regehr, “HMD calibration and its effects on distance judgments,” *ACM Transactions on Applied Perception (TAP),* vol. 6, no. 3, p. 19, 2009.

[363] R. Zhang, A. Nordman, J. Walker and S. A. Kuhl, “Minification affects verbal-and action-based distance judgments differently in head-mounted displays,” *ACM Transactions on Applied Perception (TAP),* vol. 9, no. 3, p. 14, 2012.

[364] B. Bodenheimer, J. Meng, H. Wu, G. Narasimham, B. Rump, T. P. McNamara, T. H. Carr and J. J. Rieser, “Distance estimation in virtual and real environments using bisection,” in *Proceedings of the 4th symposium on Applied perception in graphics and visualization*, ACM, 2007, pp. 35--40.

[365] J. S. Lappin, A. L. Shelton and J. J. Rieser, “Environmental context influences visually perceived distance,” *Perception & psychophysics,* vol. 68, no. 4, pp. 571--581, 2006.

[366] J. K. Witt, J. K. Stefanucci, C. R. Riener and D. R. Proffitt, “Seeing beyond the target: Environmental context affects distance perception,” *Perception,* vol. 36, no. 12, pp. 1752--1768, 2007.

[367] M. Leyrer, S. A. Linkenauger, H. H. Bülthoff, U. Kloos and B. Mohler, “The influence of eye height and avatars on egocentric distance estimates in immersive virtual environments,” in

Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization, ACM, 2011, pp. 67--74.

[368] R. Ratan and B. S. Hasler, “Exploring self-presence in collaborative virtual teams.,” *PsychNology Journal,* vol. 8, no. 1, 2010.

[369] M. Slater and M. Usoh, “Body centred interaction in immersive virtual environments,” *Artificial life and virtual reality,* vol. 1, no. 1994, pp. 125--148, 1994.

[370] M. Slater, M. Usoh and A. Steed, “Depth of presence in virtual environments,” *Presence: Teleoperators & Virtual Environments,* vol. 3, no. 2, pp. 130--144, 1994.

[371] L. Phillips, V. Interrante, M. Kaeding, B. Ries and L. Anderson, “Correlations between physiological response, gait, personality, and presence in immersive virtual environments,” *Presence,* vol. 21, no. 2, pp. 119--141, 2012.

[372] V. Interrante, B. Ries, J. Lindquist, M. Kaeding and L. Anderson, “Elucidating factors that can facilitate veridical spatial perception in immersive virtual environments,” *Presence: Teleoperators and Virtual Environments,* vol. 17, no. 2, pp. 176--198, 2008.

[373] B. J. Mohler, H. H. Bülthoff, W. B. Thompson and S. H. Creem-Regehr, “A full-body avatar improves egocentric distance judgments in an immersive virtual environment,” in *Proceedings of the 5th symposium on Applied perception in graphics and visualization*, ACM, 2008, p. 194.

[374] L. Phillips, B. Ries, M. Kaeding and V. Interrante, “Avatar self-embodiment enhances distance perception accuracy in non-photorealistic immersive virtual environments,” 2010.

[375] S. H. Creem-Regehr, P. Willemsen, A. A. Gooch and W. B. Thompson, “The influence of restricted viewing conditions on egocentric distance perception: Implications for real and virtual indoor environments,” *Perception,* vol. 34, no. 2, pp. 191--204, 2005.

[376] A. Naceri and R. Chellali, “The effect of isolated disparity on depth perception in real and virtual environments,” in *Virtual Reality Short Papers and Posters (VRW), 2012 IEEE*, IEEE, 2012, pp. 107--108.

[377] A. Tellegen and G. Atkinson, “Openness to absorbing and self-altering experiences ("absorption"), a trait related to hypnotic susceptibility.,” *Journal of abnormal psychology,* vol. 83, no. 3, p. 268, 1974.

[378] N. I. Harway, “Judgment of distance in children and adults.,” *Journal of Experimental Psychology,* vol. 65, no. 4, p. 385, 1963.

[379] J. F. Norman, A. M. Clayton, C. F. Shular and S. R. Thompson, “Aging and the perception of depth and 3-D shape from motion parallax.,” *Psychology and aging,* vol. 19, no. 3, p. 506, 2004.

[380] G. Bruder, F. A. Sanz, A.-H. Olivier and A. Lécuyer, “Distance estimation in large immersive projection systems, revisited,” in *Virtual Reality (VR), 2015 IEEE*, IEEE, 2015, pp. 27--32.

[381] K. Takahashi, T. Meilinger, K. Watanabe and H. H. Bülthoff, “Psychological influences on distance estimation in a virtual reality environment,” *Frontiers in human neuroscience,* vol. 7, p. 580, 2013.

[382] S. Cole, E. Balcetis and D. Dunning, “Affective signals of threat increase perceived proximity,” *Psychological science,* vol. 24, no. 1, pp. 34--40, 2013.

[383] Y. J. Xiao and J. J. Van Bavel, “See your friends close and your enemies closer: Social identity and identity threat shape the representation of physical distance,” *Personality and Social Psychology Bulletin,* vol. 38, no. 7, pp. 959--972, 2012.

[384] D. R. Morse and B. B. Cohen, “Desensitization using meditation-hypnosis to control" needle" phobia in two dental patients.,” *Anesthesia progress,* vol. 30, no. 3, p. 83, 1983.

[385] D. B. Shabani and W. W. Fisher, “Stimulus fading and differential reinforcement for the treatment of needle phobia in a youth with autism,” *Journal of applied behavior analysis,* vol. 39, no. 4, pp. 449--452, 2006.

[386] N. Liberman, Y. Trope and E. Stephan, “Psychological distance,” *Social psychology: Handbook of basic principles,* vol. 2, pp. 353--383, 2007.

[387] L. M. Wilcox, R. S. Allison, S. Elfassy and C. Grelik, “Personal space in virtual reality,” *ACM Transactions on Applied Perception (TAP),* vol. 3, no. 4, pp. 412--428, 2006.

[388] S. Schnall, “Embodiment in affective space: Social influences on spatial perception,” *Spatial dimensions of social thought,* vol. 18, p. 129, 2011.

[389] D. R. Proffitt, M. Bhalla, R. Gossweiler and J. Midgett, “Perceiving geographical slant,” *Psychonomic bulletin & review,* vol. 2, no. 4, pp. 409--428, 1995.

[390] D. R. Proffitt, “Embodied perception and the economy of action,” *Perspectives on psychological science,* vol. 1, no. 2, pp. 110--122, 2006.

[391] M. Veltkamp, H. Aarts and R. Custers, “Perception in the service of goal pursuit: Motivation to attain goals enhances the perceived size of goal-instrumental objects,” *Social Cognition,* vol. 26, no. 6, pp. 720--736, 2008.

[392] E. Balcetis and D. Dunning, “Wishful seeing: More desired objects are seen as closer,” *Psychological science,* vol. 21, no. 1, pp. 147--152, 2010.

[393] A. L. Alter and E. Balcetis, “Fondness makes the distance grow shorter: Desired locations seem closer because they seem more vivid,” *Journal of Experimental Social Psychology,* vol. 47, no. 1, pp. 16--21, 2011.

[394] T. Iachini, Y. Coello, F. Frassinetti and G. Ruggiero, “Body space in social interactions: a comparison of reaching and comfort distance in immersive virtual reality,” *PloS one,* vol. 9, no. 11, p. e111511, 2014.

[395] Y. Delevoye-Turrell, A. Bartolo and Y. Coello, “Motor representations and the perception of space (in N. Gangopadhyay, M. Madary, F. Spicer, Perception, action, and consciousness: Sensorimotor Dynamics and Two Visual Systems ed.): Oxford University press,” 2010.

[396] Y. Delevoye-Turrell, C. Vienne and Y. Coello, “Space boundaries in schizophrenia,” *Social Psychology,* 2011.

[397] D. P. Kennedy, J. Gläscher, J. M. Tyszka and R. Adolphs, “Personal space regulation by the human amygdala,” *Nature neuroscience,* vol. 12, no. 10, p. 1226, 2009.

[398] M. A. Dosey and M. Meisels, “Personal space and self-protection,” *Journal of personality and social psychology,* vol. 11, no. 2, p. 93, 1969.

[399] R. Messing and F. H. Durgin, “Distance perception and the visual horizon in head-mounted displays,” *ACM Transactions on Applied Perception (TAP),* vol. 2, no. 3, pp. 234--250, 2005.

[400] M. Teghtsoonian and R. Teghtsoonian, “Scaling apparent distance in natural indoor settings,” *Psychonomic Science,* vol. 16, no. 6, pp. 281--283, 1969.

[401] R. Teghtsoonian and M. Teghtsoonian, “Scaling apparent distance in a natural outdoor setting,” *Psychonomic Science,* vol. 21, no. 4, pp. 215--216, 1970.

[402] B. Williams, T. Rasor and G. Narasimham, “Distance perception in virtual environments: a closer look at the horizon and the error,” in *Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization*, ACM, 2009, pp. 7--10.

[403] L. Phillips, B. Ries, V. Interrante, M. Kaeding and L. Anderson, “Distance perception in NPR immersive virtual environments, revisited,” in *Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization*, ACM, 2009, pp. 11--14.

[404] F. Steinicke, G. Bruder, K. Hinrichs, M. Lappe, B. Ries and V. Interrante, “Transitional environments enhance distance perception in immersive virtual reality systems,” in *Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization*, ACM, 2009, pp. 19--26.

[405] J. M. Plumert, J. K. Kearney, J. F. Cremer and K. Recker, “Distance perception in real and virtual environments,” *ACM Transactions on Applied Perception (TAP),* vol. 2, no. 3, pp. 216--233, 2005.

[406] P. Willemsen and A. A. Gooch, “Perceived egocentric distances in real, image-based, and traditional virtual environments,” in *Virtual Reality, 2002. Proceedings. IEEE*, IEEE, 2002, pp. 275--276.

[407] G. Riva, F. Mantovani, C. S. Capideville, A. Preziosa, F. Morganti, D. Villani, A. Gaggioli, C. Botella and M. Alcañiz, “Affective interactions using virtual reality: the link between presence and emotions,” *CyberPsychology & Behavior,* vol. 10, no. 1, pp. 45--56, 2007.

[408] M. Price and P. Anderson, “The role of presence in virtual reality exposure therapy,” *Journal of anxiety disorders,* vol. 21, no. 5, pp. 742--751, 2007.

[409] J. Bailey, J. N. Bailenson, A. S. Won, J. Flora and K. C. Armel, "Presence and memory: immersive virtual reality effects on cued recall," in *Proceedings of the International Society for Presence Research Annual Conference*, Citeseer, 2012, pp. 24--26.

[410] H. Q. Dinh, N. Walker, L. F. Hodges, C. Song and A. Kobayashi, "Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments," in *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)*, IEEE, 1999, pp. 222--228.

[411] C. Heeter, "Reflections on real presence by a virtual person," *Presence: Teleoperators \& Virtual Environments,* vol. 12, no. 4, pp. 335--345, 2003.

[412] W. Wirth, T. Hartmann, S. Böcking, P. Vorderer, C. Klimmt, H. Schramm, T. Saari, J. Laarni, N. Ravaja, F. R. Gouveia and others, "A process model of the formation of spatial presence experiences," *Media psychology,* vol. 9, no. 3, pp. 493--525, 2007.

[413] M. Slater and M. Usoh, "Representations systems, perceptual position, and presence in immersive virtual environments," *Presence: Teleoperators \& Virtual Environments,* vol. 2, no. 3, pp. 221--233, 1993.

[414] M. Lombard and T. Ditton, "At the heart of it all: The concept of presence," *Journal of Computer-Mediated Communication,* vol. 3, no. 2, 1997.

[415] J. Steuer, "Defining virtual reality: Dimensions determining telepresence," *Journal of communication,* vol. 42, no. 4, pp. 73--93, 1992.

[416] R. B. Welch, T. T. Blackmon, A. Liu, B. A. Mellers and L. W. Stark, "The effects of pictorial realism, delay of visual feedback, and observer interactivity on the subjective sense of presence," *Presence: Teleoperators & Virtual Environments,* vol. 5, no. 3, pp. 263--273, 1996.

[417] M. Meehan, "An Objective Surrogate for Presence: Physiological Response. Presence 2000-3rd International Workshop on Presence," 2000.

[418] J. D. Prothero, H. G. Hoffman, D. E. Parker, T. A. Furness III and M. J. Wells, "Foreground/background manipulations affect presence," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 39, SAGE Publications Sage CA: Los Angeles, CA, 1995, pp. 1410--1414.

[419] T. Schubert, F. Friedmann and H. Regenbrecht, "The experience of presence: Factor analytic insights," *Presence: Teleoperators & Virtual Environments,* vol. 10, no. 3, pp. 266--281, 2001.

[420] M. Schuemie, M. Bruynzeel, L. Drost, M. Brinckman, G. De Haan, P. Emmelkamp and C. Van der Mast, "Treatment of acrophobia in virtual reality: A pilot study," in *Conference Proceedings Euromedia*, 2000, pp. 271--275.

[421] H. T. Regenbrecht, T. W. Schubert and F. Friedmann, "Measuring the sense of presence and its relations to fear of heights in virtual environments," *International Journal of Human-Computer Interaction,* vol. 10, no. 3, pp. 233--249, 1998.

[422] M. Slater and S. Wilbur, “A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments,” *Presence: Teleoperators & Virtual Environments,* vol. 6, no. 6, pp. 603--616, 1997.

[423] W. Barfield and C. Hendrix, “The effect of update rate on the sense of presence within virtual environments,” *Virtual Reality,* vol. 1, no. 1, pp. 3--15, 1995.

[424] C. Hendrix and W. Barfield, “The sense of presence within auditory virtual environments,” *Presence: Teleoperators & Virtual Environments,* vol. 5, no. 3, pp. 290--301, 1996.

[425] R. M. Baños, C. Botella, A. Garcia-Palacios, H. Villa, C. Perpiñá and M. Alcaniz, “Presence and reality judgment in virtual environments: a unitary construct?,” *CyberPsychology & Behavior,* vol. 3, no. 3, pp. 327--335, 2000.

[426] A. A. Rizzo, M. Wiederhold and J. G. Buckwalter, “Basic issues in the use of virtual environments for mental health applications,” *Studies in health technology and informatics,* pp. 21--42, 1998.

[427] M. Usoh, E. Catena, S. Arman and M. Slater, “Using presence questionnaires in reality,” *Presence: Teleoperators & Virtual Environments,* vol. 9, no. 5, pp. 497--503, 2000.

[428] G. Fontaine, “The experience of a sense of presence in intercultural and international encounters,” *Presence: Teleoperators & Virtual Environments,* vol. 1, no. 4, pp. 482--490, 1992.

[429] T. Kunnapas, “Distance perception as a function of available visual cues.,” *Journal of experimental psychology,* vol. 77, no. 4, p. 523, 1968.

[430] J. W. Philbeck and J. M. Loomis, “Comparison of two indicators of perceived egocentric distance under full-cue and reduced-cue conditions.,” *Journal of Experimental Psychology: Human Perception and Performance,* vol. 23, no. 1, p. 72, 1997.

[431] S. Uno and M. Slater, “The sensitivity of presence to collision response,” in *Virtual Reality Annual International Symposium, 1997., IEEE 1997*, IEEE, 1997, pp. 95--103.

[432] W. IJsselsteijn, H. d. Ridder, J. Freeman, S. E. Avons and D. Bouwhuis, “Effects of stereoscopic presentation, image motion, and screen size on subjective and objective corroborative measures of presence,” *Presence: Teleoperators & Virtual Environments,* vol. 10, no. 3, pp. 298--311, 2001.

[433] L. J. Hettinger and M. W. Haas, Virtual and adaptive environments: Applications, implications, and human performance issues, CRC Press, 2003.

[434] B. T. Backus, M. S. Banks, R. van Ee and J. A. Crowell, “Horizontal and vertical disparity, eye position, and stereoscopic slant perception,” *Vision research,* vol. 39, no. 6, pp. 1143--1170, 1999.

[435] W. Barfield and C. Rosenberg, *Motion parallax as a variable for perspective display design,* Department of Industrial Engineering, FU-20, University of Washington, , 1995.

[436] S. J. Watt, K. Akeley, M. O. Ernst and M. S. Banks, “Focus cues affect perceived depth,” *Journal of vision,* vol. 5, no. 10, pp. 7--7, 2005.

[437] W. Barfield, D. Zeltzer, T. Sheridan and M. Slater, “Presence and performance within virtual environments,” *Virtual environments and advanced interface design,* pp. 473--513, 1995.

[438] B. Series, “Subjective methods for the assessment of stereoscopic 3DTV systems,” 2015.

[439] M. Slater and M. Usoh, “Presence in immersive virtual environments,” in *Virtual Reality Annual International Symposium, 1993., 1993 IEEE*, IEEE, 1993, pp. 90--96.

[440] S. Winkler and D. Min, “Stereo/multiview picture quality: Overview and recent advances,” *Signal Processing: Image Communication,* vol. 28, no. 10, pp. 1358--1373, 2013.

[441] M. Wöpking, “Viewing comfort with stereoscopic pictures: An experimental study on the subjective effects of disparity magnitude and depth of focus,” *Journal of the society for information display,* vol. 3, no. 3, pp. 101--103, 1995.

[442] M. Wöpking, “Viewing comfort with stereoscopic pictures: An experimental study on the subjective effects of disparity magnitude and depth of focus,” *Journal of the society for information display,* vol. 3, no. 3, pp. 101--103, 1995.

APPENDIX A. Pilot tests and related procedures

Pilot tests - list of questions used in modality 1 and modality 2

Which VR Panorama did you see?	☐ **Island Lachea** ☐ **Monello Cave**
Phone	☐ **iPhone 6** ☐ **LG** ☐ ________________
VR Headset	☐ **VR Shinecon** ☐ ________________________
How much you felt **immersed**?	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4
Distorted 3D? If yes, which objects?	☐ Yes, near objects ☐ Yes, far objects ☐ Yes, near and far objects ☐ No distortion
Depth perceived **on display's borders** is **equal to** the one perceived on **centre of display**? ☐ Other ____________________	☐ Yes ☐ No, better depth perceived on border ☐ No, better depth perceived on centre
How much **time** before you got used to **depth perception without strain**?	☐ I perceived depth immediately in a natural way ☐ I strained my eyes at the beginning only ☐ I strained my eyes at the beginning all the time
How natural are **colours**?	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4
How much **lights and shadows** improved 3D perception?	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4
Eye discomfort while watching 3D? ____________________________________	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4
How long could you **watch 3D without discomfort**?	☐ < 1 min ☐ < 5 min ☐ < 10 min ☐ < 30 min ☐ > 30 min
Evaluate objects' **depth and shape** in 3D	☐ Far objects appeared flat ☐ Close objects appeared flat ☐ Both Far and close objects appeared flat ☐ All objects appeared in 3D well
How much **distances** appeared **real and correct**?	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4
How much **movement** made the experience more **realistic**? (from 0 to 5)	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4

How much the **panorama** was **realistic thanks to 3D**? (from 0 to 5)	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4
How much did you **like the experience**? (from 0 to 5)	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4
Evaluate **quality** of **3D depth perception**	☐ -3 ☐ -2 ☐ -1 ☐ 0 ☐ 1 ☐ 2 ☐ 3
How much you felt **"present in place"** inside the virtual panorama?	☐ -3 ☐ -2 ☐ -1 ☐ 0 ☐ 1 ☐ 2 ☐ 3
How much VR can help to visit remote place realistically?	☐ -3 ☐ -2 ☐ -1 ☐ 0 ☐ 1 ☐ 2 ☐ 3
Evaluate **comfort** (eyes strain, headache, nausea, fatigue)	☐ -3 ☐ -2 ☐ -1 ☐ 0 ☐ 1 ☐ 2 ☐ 3
Evaluate **how much** you felt **isolated from the real world** using the VR headset	☐ -3 ☐ -2 ☐ -1 ☐ 0 ☐ 1 ☐ 2 ☐ 3
Evaluate **how much realistic VR is compared to real life**	☐ -3 ☐ -2 ☐ -1 ☐ 0 ☐ 1 ☐ 2 ☐ 3
Evaluate **how much realistic VR is compared to photos and video**	☐ -3 ☐ -2 ☐ -1 ☐ 0 ☐ 1 ☐ 2 ☐ 3
How many distortions aggravated objects and natural elements?	☐ -3 ☐ -2 ☐ -1 ☐ 0 ☐ 1 ☐ 2 ☐ 3
How much you think **a display horizontally larger** would **improve realism** in VR?	☐ -3 ☐ -2 ☐ -1 ☐ 0 ☐ 1 ☐ 2 ☐ 3
How much you **felt** you could **touch objects**?	☐ -3 ☐ -2 ☐ -1 ☐ 0 ☐ 1 ☐ 2 ☐ 3

Comparative analysis

Which tour was more realistic? (feeling of being there)	☐ Island Lachea (Outdoor) ☐ Monello Cave (Indoor) ☐ Both were equally realistic ☐ None of them was realistic
Evaluate the accuracy of movements in this VR experience	☐ Too many delays in my movement ☐ Acceptable accuracy ☐ Almost responsive movement ☐ Very responsive movement ☐ Extremely realistic movement

Pilot tests – list of questions used with modality 3

Test User ID: _________ Panorama Viewed: ______________________________

Background		
Questions	Phone1	Phone2
Did you perceive 3D immediately? Y/N		

Realism (similar to actual viewing)		
Questions	Phone1	Phone2
How realistic the scene appears to you? From 0 to 6		
Do the objects in the scene appear distorted? From 0 to 6		
Do you think there is a difference in the level of realism between the edge and the centre of the screen? From 0 to 6		
Do you think colours make this experience more realistic? From 0 to 6		
Do you think lights and shadows make this experience more realistic? From 0 to 6		
How would you rate the level of realism in terms of conveyed sensations and emotions? From 0 to 6		
Do you think that your head's movement enhanced the level of realism of the experience? From 0 to 6		
Do you think that colours appear natural? From 0 to 6		
Do you think that brightness (light intensity) appear natural? From 0 to 6		
Do you think that contrast (contrast between colours) appear natural? From 0 to 6		
Do you think that sharpness (focus) appear natural? From 0 to 6		
Can you see pixels on the screen? From 0 to 6		

Presence (feeling of being there)		
Questions	Phone1	Phone2
How much do you feel "present" in the environment shown? From 0 to 6		
Do you think colours make you feel more present in the scene? From 0 to 6		
Do you think lights and shadows make you feel more present in the scene? From 0 to 6		

Comfort		
Questions	Phone1	Phone2
Do you feel any pain or strain while watching? From 0 to 6		
How long do you think you could watch? (< 1 min, < 5 min, < 10 min, < 30 min, > 30 min)		

Depth Perception		
Questions	Phone1	Phone2
How would you rate the achieved 3D depth impression? from 0 to 6		

Some of the objects' depth in the scene may appear increased or reduced compared to their actual depth. Is the apparent depth of objects decreased (flattened) than their actual depth? Y/N		
If yes, how much depth is reduced? 20% 40% 60% 80% 100%		
Is the apparent depth of objects increased than their actual depth? Y/N		
If yes, how much depth is increased? 20% 40% 60% 80% 100%		
Do you have the feeling that you could touch the closer objects? Y/N		
Do you think 3D depth makes the panorama more realistic? Y/N		
If yes, how much? From 0 to 6		

Phone1 vs Phone2		
Questions	Phone1	Phone2
Which tour seems the most realistic? (Phone1 / Phone2 / the same realism) from 0 to 6		
Which tour is more realistic in terms of colours and lighting? (Phone1 / Phone2 / the same realism) from 0 to 6		
Which tour made you feel more "present"? (Phone1 / Phone2 / the same realism) from 0 to 6		
Which tour appeared more comfortable? (Phone1 / Phone2 / the same realism) from 0 to 6		

Distance evaluation – UH Hutton Hub

Phone1

Object	FOV 90	FOV 80	FOV 70	FOV 60	Real life
White pillar					
Table					
Girl					

Phone2

Object	FOV 90	FOV 80	FOV 70	FOV 60	Real life
White pillar					
Table					
Girl					

APPENDIX B. Usability evaluations questionnaires

Usability evaluation – Display and Environment – questionnaire

Realism

The impact of the two independent variables (i.e. Phone type, panorama showed) over the level of realism perceived by test users is analysed. The following is the list of questions related to this analysis:

Descriptive nominal qualitative data

HR1.1 Is the realism of the 3D effect perceived at the screen borders similar to that perceived at the screen centre? [Yes, similar 3D effects / No, the 3D effect is better at borders / No, the 3D effect at borders is more distorted / Other]

Descriptive numerical qualitative data

R2.1 How would you rate the achieved level of realism compared to actually being in the real place? [-3, +3]

R2.2 How would you rate the achieved level of realism compared to that provided by photographs and Videos? [-3, +3]

R2.3 How would you rate the achieved level of realism in terms of deformation of observed objects/nature elements? [-3, +3]

R2.4 Do you believe that a display with larger horizontal size can increase the level of realism? [-3, +3]

R2.5 How would you rate the level of realism in terms of conveyed sensations and emotions? [-3, +3]

R2.6 How would you rate the overall quality of the panorama images seen through the VR display, in terms of image definition? [0,6]

R2.7 How sharp (not blurred, focused) the panorama images appear on the VR display you used to watch the panorama? [0, 6]

R2.8 Do you see pixel edges (i.e. areas of the image that look like a mosaic image) when using the VR display to watch the 3D panorama? [0, 6]

R2.9 How much do you think the pixel edges affect the realism of the panorama that you viewed? [0, 6]

R2.10 How much do you think images appeared vivid on the VR display you used to watch the 3D panorama? [0, 6]

R2.11 How much do you think the vividness of the display affects the realism of the panorama that you viewed? [0, 6]

R2.12 How bright do you think the display you used to see the panorama is? [0, 6]

R2.13 How much do you think the brightness of the display affects the realism of the panorama that you viewed? [0, 6]

R2.14 How would you rate the intensity of colours of the VR display you used to watch the panorama? [0,6]

R2.15 How much do you think color intensity affects the realism of the panorama that you viewed? [0, 6]

R2.16 How would you rate the contrast (difference in luminance or color that makes an object distinguishable) of images on the VR display you used to watch the panorama? [0, 6]

R2.17 How much do you think contrast affects the realism of the panorama that you viewed? [0, 6]

R2.18 How much do you think the colours of the panorama affected the emotions you felt while viewing it with the VR display? [0, 6]

Comfort

The impact of the two independent variables (i.e. Phone type, panorama showed) over the level of comfort perceived by test users is analysed. The following is the list of questions related to this analysis:

Descriptive nominal qualitative data

C1.1 Did you experience discomfort during 3D viewing of the panorama? [Very little / Little / Quite some / Much / Very Much]

C1.2 How long do you think you can watch the 3D panorama before getting tired? [Less than 1 minute / Less than 5 minutes / Less than 10 minutes / Less than 30 minutes / More than 30 minutes]

Descriptive numerical qualitative data

C2.1 How would you rate the suitability of the VR setup to the specific application? [-3, +3]

C2.2 How would you rate the comfort experience (in terms of eye strain, headache, nausea, tiredness)? [-3, +3]

Presence

The impact of the two independent variables (i.e. Phone type, panorama showed) over the *sense of presence* ("feeling of being there") perceived by test users is analysed. The following is the list of questions related to this analysis:

Descriptive nominal qualitative data

P1.1 How much did you feel "immersed" into the panorama? [I did not feel immersed / A little / Quite some / Much / Completely]

P1.2 Select all the emotions that this panorama with this VR display made you feel [Happiness, Sadness, Relax, Fear, Anger, Joy, Surprise, Anxiety, Indifference, Other not listed].

Descriptive numerical qualitative data

P2.1 How would you rate the overall sense of presence achieved in the remote place (feeling of being there)? [-3, +3]

P2.2 How would you rate the general sense of isolation from the surrounding environment? [-3, +3]

P2.3 How much did you notice black borders of the display you used to watch the panorama? [0, 6]

P2.4 Did the black borders affect your sense of presence (feeling of being there) in the panorama that you viewed? [0, 6]

P2.5 How did you feel while watching this panorama? (Bad = 0 / Good = 6) [0, 6]

P2.6 How much did you enjoy watching this panorama? [0, 6]

P2.7 How much happy did you feel watching this panorama? [0, 6]

P2.8 How much sad did you feel watching this panorama? [0, 6]

P2.9 How much scared did you feel watching this panorama? [0, 6]

P2.10 How much relaxed did you feel watching this panorama? [0, 6]

P2.11 How much this panorama makes you feel like going on holidays? [0, 6]

P2.12 How much do you believe that the isolation provided by this VR setup made you feel more emotions? [0, 6]

P2.13 How much do you think the achieved emotions influenced your sense of presence (feeling of being there)? [0, 6]

P2.14 How happy did you feel to come back to reality after watching this panorama? [0, 6]

Depth Perception

The impact of the two independent variables (i.e. Phone type, panorama showed) over the depth perception of test users is analysed. The following is the list of questions related to this analysis:

Descriptive nominal qualitative data

D1.1 Did you perceive a distorted 3D? [Yes, in near objects / Yes, in far objects / No, no distortion]

D1.2 Did you perceive the 3D immediately or your eyes needed some time to get used to it? [Yes, I perceived 3D immediately / No, some time was needed at the beginning / No, I had to continuously get used to it]

D1.3 How much do you think that colours have contributed to the perception of your 3D? [Very little / Little / Quite some / Much / Very Much]

D1.4 How much do you think that lights and shadows have contributed to the perception of your 3D? [Very little / Little / Quite some / Much / Very Much]

D1.5 How did you perceive distances in 3D? [Near and far objects were indistinguishable, Far objects were flat, Near objects were flat, Distances were realistically reproduced, I felt like I was able to touch objects near me]

Descriptive numerical qualitative data

D2.1 How would you rate the perceived 3D depth impression? [-3, +3]

D2.2 How much the achieved 3D perception affected the emotions you felt while viewing it with the VR display? [0, 6]

Usability evaluation – HDR – questionnaire

Realism in terms of Presence

Sense of presence

HR1.1 On a scale of 0 to 6, how great was the perceived "sense of presence" (feeling of being there) into the shown environment? (0 = no sense of presence, 6 = great sense of presence)

Isolation

HR1.2 On a scale of 0 to 6, how isolated from reality did you feel with this setup? (0 = no isolation, 6 = totally isolated)

Feeling like observing a real place

HR1.3 On a scale of 0 to 6, how much this setup made you feel you are observing a real place? (0 = not at all, 6 = very much)

Realism in terms of Emotions

Likeliness

HR2.1 On a scale of -3 to +3, how much did you like this setup to observe a remote environment? (-3 = not liked at all, +3 = extremely liked)

Happiness

HR2.2 On a scale of -3 to +3, how happy did you feel? (-3 = not happy at all, +3 = extremely happy)

Disappointment

HR2.3 On a scale of -3 to +3, how disappointed did you feel? (-3 = not disappointed at all, +3 = extremely disappointed)

Desire to visit the remote environment in real life

HR2.4 On a scale of -3 to +3, how much would you like to visit the place shown in the panorama? (-3 = I would not visit the real place, +3 = I really want to visit the real place)

Positive mood change after viewing in VR

HR2.5 On a scale of -3 to +3, has your mood positively changed after viewing this panoramic view? (-3 = not at all, +3 = very much)

Motivation to use VR for remote observation

HR2.6 On a scale of -3 to +3, do you feel more motivated now to use virtual reality to remotely observe places? (-3 = not at all, +3 = very much)

Realism in terms of Virtual Vision vs Real Life

Natural eye adaptation

HR3.1 On a scale of 0 to 6, how natural was the adaptation of the eyes to the amount of light in the scene? (0 = very unnatural, 6 = very natural)

Virtual lights vs Real Life lights

HR3.2 On a scale of 0 to 6, how realistic was the reproduction of lights compared to real life? (0 = not realistic at all, 6 = very realistic)

Color realism

HR3.3 On a scale of 0 to 6, how do you rate colours of the scene, in terms of realism? (0 = very bad, 6 = excellent)

Realism of Darkness

HR3.4 Focus on an area of the panorama where you expect to be very dark. On a scale of 0 to 6, how much this setup reproduced darkness realistically? (0 = very bad, 6 = excellent)

Realism of Brightness

HR3.5 Focus on an area of the panorama where you expect to be very light. On a scale of 0 to 6, how much this setup reproduced lights realistically? (0 = very bad, 6 = excellent)

Virtual vision compared to Real vision

HR3.6 On a scale of 0 to 6, how do you rate the virtual vision of the scene in this setup compared to the natural behaviour of your eyes? (0 = Very far from reality, 6 = Very similar to reality)

Comfort

Level of comfort

HC1.1 On a scale of 0 to 6, how comfortable this setup was for your eyes? (0 = Very uncomfortable, 6 = Very comfortable)

Time before feeling discomfort

HC1.2 For how long you think you can watch the panorama before getting tired, using this setup? [Less than 1 minute / Less than 2 minutes / Less than 5 minutes / Less than 10 minutes / Less than 30 minutes / Over 30 minutes]

Perceptual issues

HC1.3 Did you perceive any of the following for each setup? Check all the perceptual issues that apply. [Eye strain / Motion sickness / Headache / Visual fatigue / None of the above]

Depth perception

3D accuracy

HD1.1 On a scale of 0 to 6, how much do you believe the 3D perception of the objects of the scene was accurate? (0 = very inaccurate, 6 = very accurate)

Lights and Shadows contribution to 3D

HD1.2 On a scale of -3 to +3, how much do you believe that lights and shadows positively contributed to the perception of your 3D, using this setup? (-3 = no contribution, +3 = significant contribution)

Realistic space perception

HD1.3 On a scale of 0 to 6, how much this setup helped you to perceive the space around you realistically? (0 = no contribution, 6 = significant contribution)

Perceived size of the environment

HD1.4 On a scale of 0 to 6, how large the space of the environment appears to you, using this setup? (0 = very small, 6 = very large)

APPENDIX C. Unity 3D projects: images from practical activities

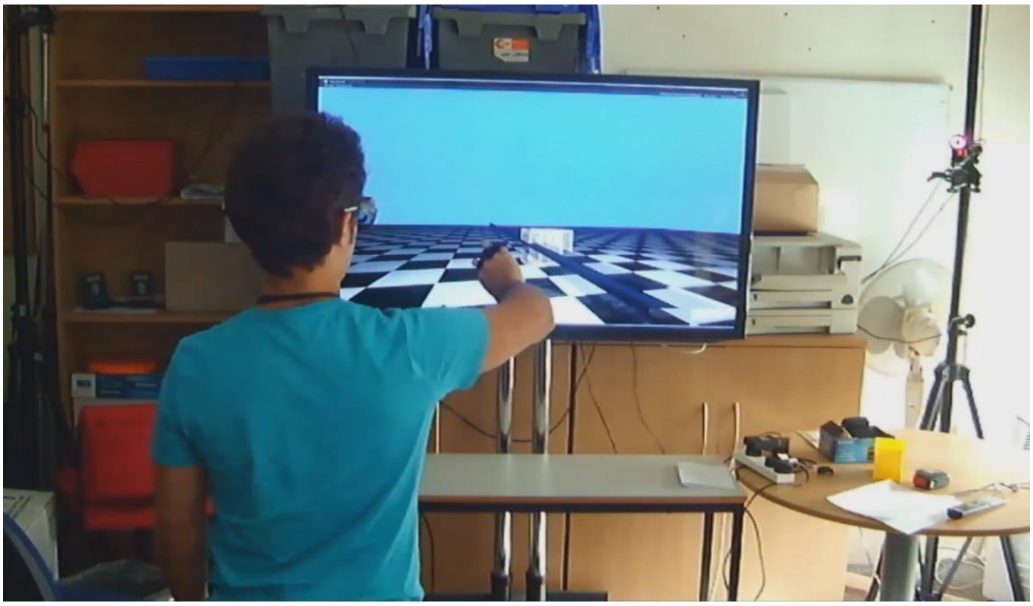


Figure 211 *- Moving a virtual object touching it with my real hand, using OptiTrack motion tracking. This was useful to understand the principles of stereoscopy and the formula contained in Dr Livatino's filed invention [19].*

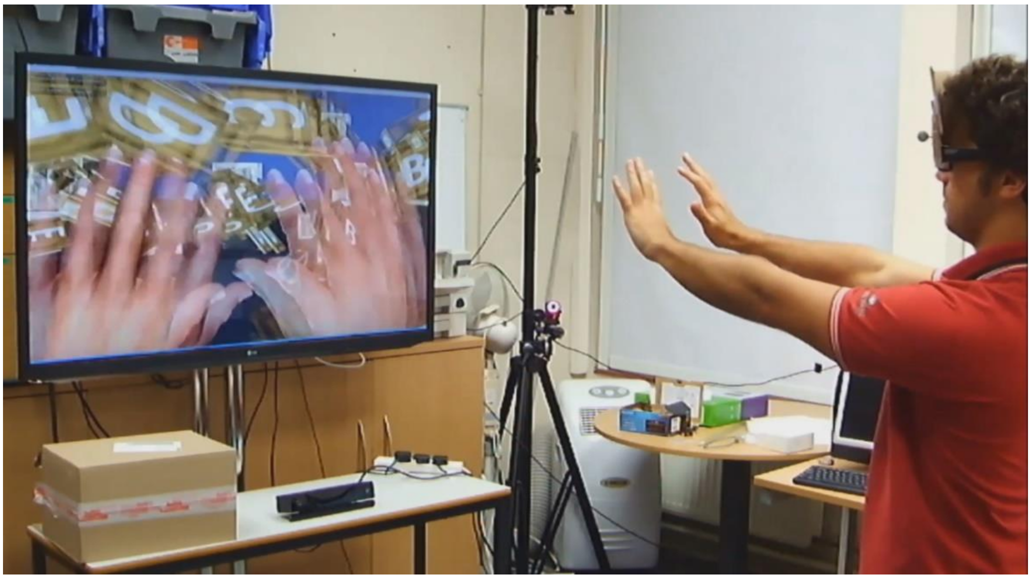


Figure 212 *- Testing 3D visualization and interaction with virtual objects using Leap Motion and OptiTrack motion tracking.*

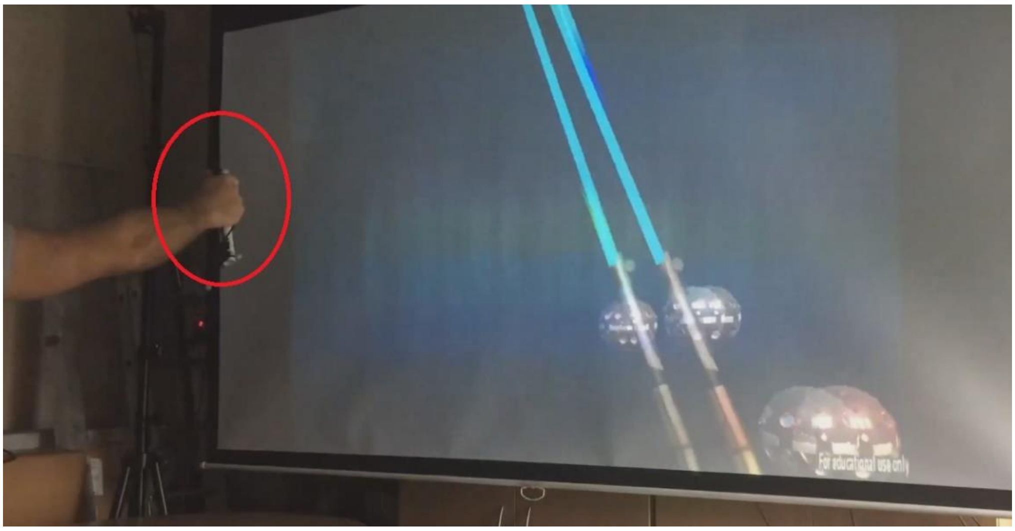


Figure 213 - *Using the formula from [19] to precisely map a virtual object (lightsaber of the Star Wars' demo) on a real object (highlighted by the red circle on my hand).*

Figure 211, **Figure 212**, **Figure 213** and **Figure 214** show some of the practical tests I developed in Unity 3D to better understand induced depth perception through 3D displays and interactive digital systems. In detail, this was done to better understand the principles of 3D realistic visualization, the role of camera and display parameters, and different ways to interact with virtual environments (e.g. using OptiTrack motion camera system, Leap Motion, Kinect, VR HMDs).

Figure 214 - *Testing my 3D VR archery game, developed to practice with OptiTrack, Unity 3D, virtual 3D camera and display parameters.*

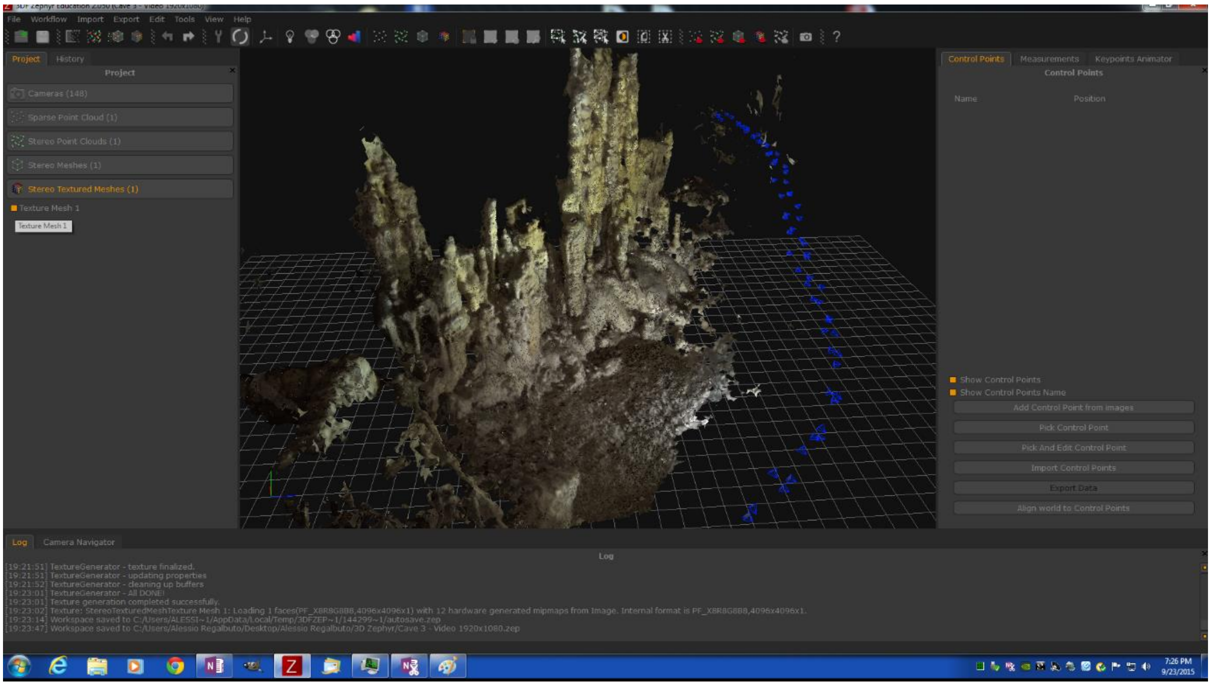


Figure 215 *- Screenshot from the 3D reconstruction of the stones of the Monello's Cave using 3D Zephyr.*

Figure 216 *- Detail of the Laplacian Smoother Filter applied to the reconstructed point cloud in 3D Zephyr.*

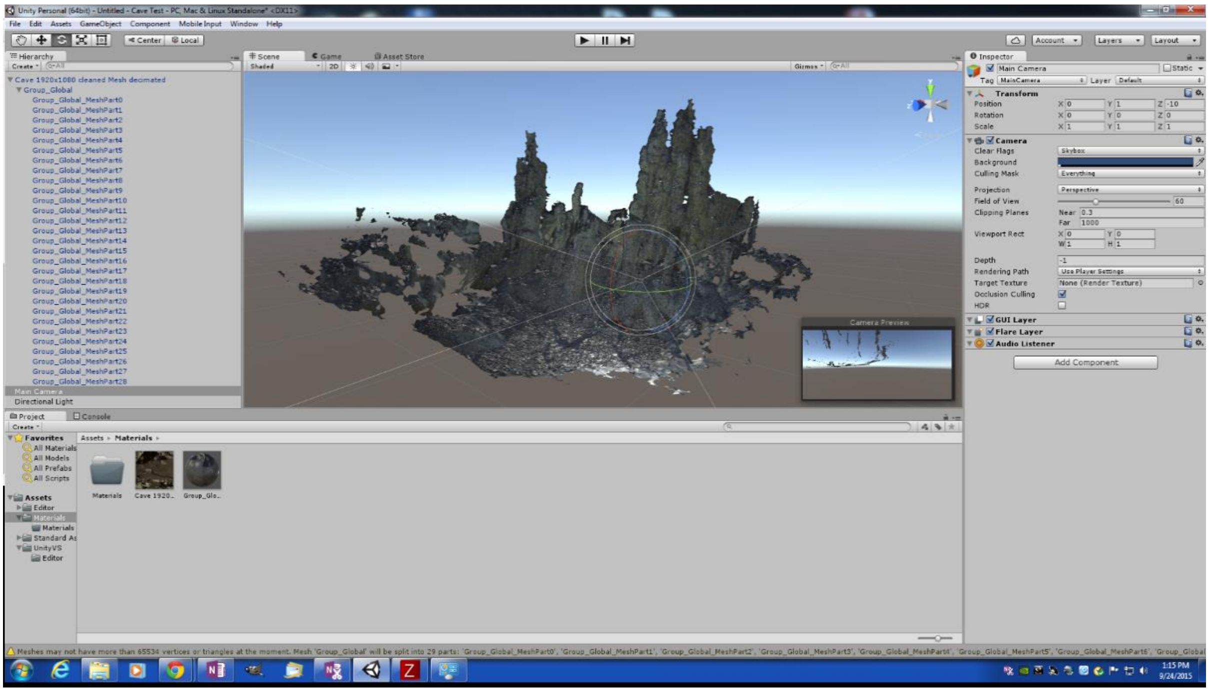


Figure 217 - *Generated mesh from 3D Zephyr and imported inside Unity 3D.*

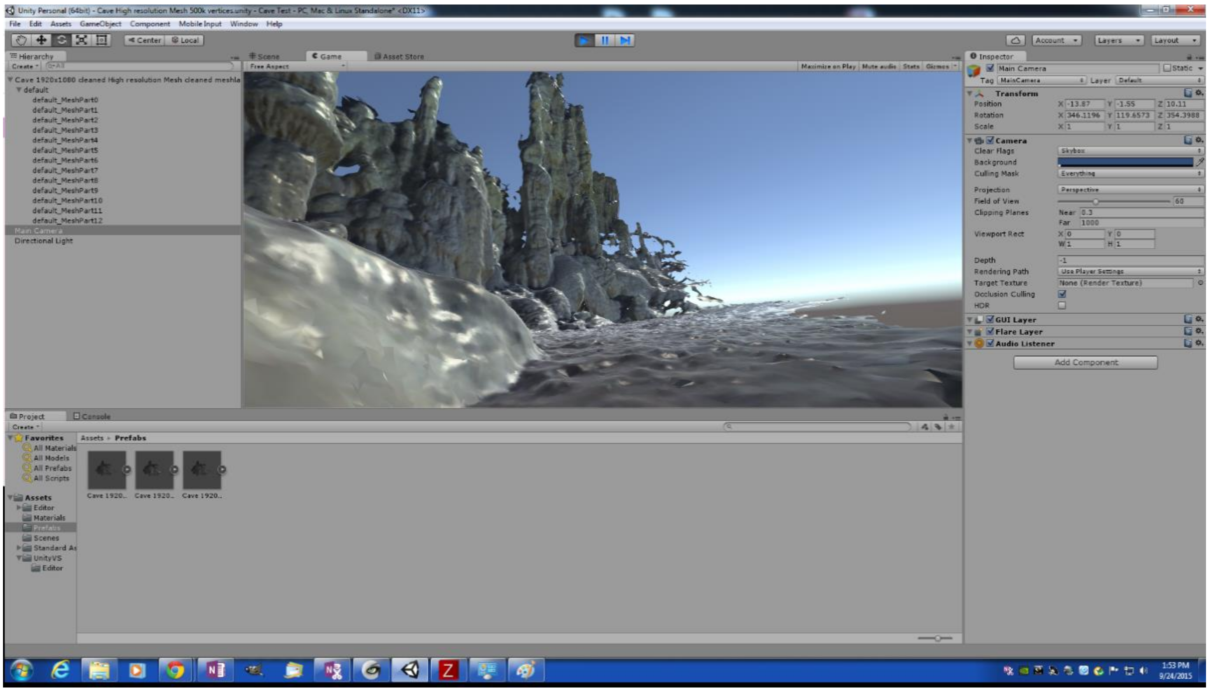


Figure 218 - *Screenshot of the moving camera in the Unity 3D executed scene.*

Figure 219 - Screenshot of the moving camera in the Unity 3D executed scene.

Figure 215, **Figure 216**, **Figure 217**, **Figure 218** and **Figure 219** show my first attempts to capture the 3D model of a real existing place (Monello's cave in Sicily) using photogrammetry. This was helpful to understand one of the valid approaches that can be used to digitalize an existing place in Virtual Reality.

However, despite the good quality of the result, it was decided to use 3D panoramic photos instead of photogrammetry to conduct the proposed investigation of this thesis.

APPENDIX D. Source codes and scripts

Edited XML and code to convert 2D to 3D panoramas generated by Panotour

Original code produced by Panotour for the 2D hotspot

```
<image type="CUBE" multires="true" baseindex="0" tilesize="512" devices="!androidstock|webgl">
 <level tiledimagewidth="8599" tiledimageheight="8599">
  <front url="_10left_10/0/4/%v_%u.jpg"/>
  <right url="_10left_10/1/4/%v_%u.jpg"/>
  <back url="_10left_10/2/4/%v_%u.jpg"/>
  <left url="_10left_10/3/4/%v_%u.jpg"/>
  <up url="_10left_10/4/4/%v_%u.jpg"/>
  <down url="_10left_10/5/4/%v_%u.jpg"/>
 </level>
 <level tiledimagewidth="4299" tiledimageheight="4299">
  <front url="_10left_10/0/3/%v_%u.jpg"/>
  <right url="_10left_10/1/3/%v_%u.jpg"/>
  <back url="_10left_10/2/3/%v_%u.jpg"/>
  <left url="_10left_10/3/3/%v_%u.jpg"/>
  <up url="_10left_10/4/3/%v_%u.jpg"/>
  <down url="_10left_10/5/3/%v_%u.jpg"/>
 </level>
 <level tiledimagewidth="2149" tiledimageheight="2149">
  <front url="_10left_10/0/2/%v_%u.jpg"/>
  <right url="_10left_10/1/2/%v_%u.jpg"/>
  <back url="_10left_10/2/2/%v_%u.jpg"/>
  <left url="_10left_10/3/2/%v_%u.jpg"/>
  <up url="_10left_10/4/2/%v_%u.jpg"/>
  <down url="_10left_10/5/2/%v_%u.jpg"/>
 </level>
 <level tiledimagewidth="1074" tiledimageheight="1074">
  <front url="_10left_10/0/1/%v_%u.jpg"/>
  <right url="_10left_10/1/1/%v_%u.jpg"/>
  <back url="_10left_10/2/1/%v_%u.jpg"/>
  <left url="_10left_10/3/1/%v_%u.jpg"/>
  <up url="_10left_10/4/1/%v_%u.jpg"/>
  <down url="_10left_10/5/1/%v_%u.jpg"/>
 </level>
 <level tiledimagewidth="537" tiledimageheight="537">
  <front url="_10left_10/0/0/%v_%u.jpg"/>
  <right url="_10left_10/1/0/%v_%u.jpg"/>
  <back url="_10left_10/2/0/%v_%u.jpg"/>
  <left url="_10left_10/3/0/%v_%u.jpg"/>
  <up url="_10left_10/4/0/%v_%u.jpg"/>
  <down url="_10left_10/5/0/%v_%u.jpg"/>
 </level>
</image>
<image type="CUBE" devices="androidstock.and.no-webgl">
 <front url="_10left_10/mobile/0.jpg"/>
 <right url="_10left_10/mobile/1.jpg"/>
 <back url="_10left_10/mobile/2.jpg"/>
 <left url="_10left_10/mobile/3.jpg"/>
 <up url="_10left_10/mobile/4.jpg"/>
 <down url="_10left_10/mobile/5.jpg"/>
</image>
```

Edited code to support in a single hotspot both left and right panoramas in 3D

```
<image type="CUBE" multires="true" baseindex="0" tilesize="512" devices="!androidstock|webgl" stereo="true"
stereolabels="l|r">
  <level tiledimagewidth="8599" tiledimageheight="8599">
   <front url="_10%t_10/0/4/%v_%u.jpg"/>
   <right url="_10%t_10/1/4/%v_%u.jpg"/>
   <back url="_10%t_10/2/4/%v_%u.jpg"/>
   <left url="_10%t_10/3/4/%v_%u.jpg"/>
   <up url="_10%t_10/4/4/%v_%u.jpg"/>
   <down url="_10%t_10/5/4/%v_%u.jpg"/>
  </level>
  <level tiledimagewidth="4299" tiledimageheight="4299">
   <front url="_10%t_10/0/3/%v_%u.jpg"/>
   <right url="_10%t_10/1/3/%v_%u.jpg"/>
   <back url="_10%t_10/2/3/%v_%u.jpg"/>
   <left url="_10%t_10/3/3/%v_%u.jpg"/>
   <up url="_10%t_10/4/3/%v_%u.jpg"/>
   <down url="_10%t_10/5/3/%v_%u.jpg"/>
  </level>
  <level tiledimagewidth="2149" tiledimageheight="2149">
   <front url="_10%t_10/0/2/%v_%u.jpg"/>
   <right url="_10%t_10/1/2/%v_%u.jpg"/>
   <back url="_10%t_10/2/2/%v_%u.jpg"/>
   <left url="_10%t_10/3/2/%v_%u.jpg"/>
   <up url="_10%t_10/4/2/%v_%u.jpg"/>
   <down url="_10%t_10/5/2/%v_%u.jpg"/>
  </level>
  <level tiledimagewidth="1074" tiledimageheight="1074">
   <front url="_10%t_10/0/1/%v_%u.jpg"/>
   <right url="_10%t_10/1/1/%v_%u.jpg"/>
   <back url="_10%t_10/2/1/%v_%u.jpg"/>
   <left url="_10%t_10/3/1/%v_%u.jpg"/>
   <up url="_10%t_10/4/1/%v_%u.jpg"/>
   <down url="_10%t_10/5/1/%v_%u.jpg"/>
  </level>
  <level tiledimagewidth="537" tiledimageheight="537">
   <front url="_10%t_10/0/0/%v_%u.jpg"/>
   <right url="_10%t_10/1/0/%v_%u.jpg"/>
   <back url="_10%t_10/2/0/%v_%u.jpg"/>
   <left url="_10%t_10/3/0/%v_%u.jpg"/>
   <up url="_10%t_10/4/0/%v_%u.jpg"/>
   <down url="_10%t_10/5/0/%v_%u.jpg"/>
  </level>
 </image>
 <image type="CUBE" devices="androidstock.and.no-webgl" stereo="true" stereolabels="l|r">
  <front url="_10%t_10/html5/0.jpg"/>
  <right url="_10%t_10/html5/1.jpg"/>
  <back url="_10%t_10/html5/2.jpg"/>
  <left url="_10%t_10/html5/3.jpg"/>
  <up url="_10%t_10/html5/4.jpg"/>
  <down url="_10%t_10/html5/5.jpg"/>
 </image>
```

XmlSerializerPano class

```
using PanoramaConverter.Model;
using System;
using System.Collections.Generic;
using System.IO;
using System.Linq;
using System.Text;
using System.Text.RegularExpressions;
using System.Threading.Tasks;
using System.Windows.Media.Imaging;
using System.Xml.Serialization;

namespace PanoramaConverter.Utils {
  public static class XmlSerializerPano {

    public static string CurrentTourPath;
    public static string UpdatedPanoramaCode;

    public static BitmapImage LoadPanoramaPreview() {
      try {
        return new BitmapImage(new Uri(System.IO.Path.GetDirectoryName(CurrentTourPath) + "/thumbnail.png"));
      } catch (Exception ex) {
        Console.WriteLine(ex.Message);
        return null;
      }
    }

    public static krpano GenerateKrpanoFromXML(string xmlContent) {
      XmlSerializer ser = new XmlSerializer(typeof(krpano));
      using (StringReader sr = new StringReader(xmlContent)) {
        return (krpano)ser.Deserialize(sr);
      }
    }

    public static string GenerateXmlFromKrpano<Krpano>(Krpano panoToGenerate) {
      XmlSerializer xmlSerializer = new XmlSerializer(panoToGenerate.GetType());

      using (StringWriter textWriter = new StringWriter()) {
        xmlSerializer.Serialize(textWriter, panoToGenerate);
        return textWriter.ToString();
      }
    }

    public static String GenerateUpdatedXML(string xmlContent) {
      Regex rexBegin = new Regex(@"<?xml[\s\S]*?<scene");
      var resBegin = rexBegin.Match(xmlContent);
      string begin = resBegin.Value.Substring(0, resBegin.Value.Length - 6);

      Regex rexLast = new Regex(@"<?xml[\s\S]*?<scene");
      var resLast = rexBegin.Match(xmlContent);
      String ending = xmlContent.Substring(xmlContent.LastIndexOf("</scene>"));

      Regex rexInner = new Regex(@"<scene[\s\S]*?<\/scene>");
      MatchCollection results = rexInner.Matches(xmlContent);

      StringBuilder replacement = new StringBuilder("<?" + begin);
```

```
        foreach(Match result in results) {
          XmlSerializer ser = new XmlSerializer(typeof(Scene));
          using (StringReader sr = new StringReader(result.Value)) {
            string toAdd = UpdateScene((Scene)ser.Deserialize(sr));
            toAdd = toAdd.Replace("<?xml version=\"1.0\" encoding=\"utf-16\"?>\r\n", "");
            toAdd = toAdd.Replace("xmlns:xsi=\"http://www.w3.org/2001/XMLSchema-instance\"
xmlns:xsd=\"http://www.w3.org/2001/XMLSchema\" ", "");
            replacement.Append(toAdd);
          }
        }

        replacement.Append(ending.Replace("</scene>", ""));
        UpdatedPanoramaCode = replacement.ToString();
        return replacement.ToString();
      }

      public static string GenerateXmlFromScene(Scene panoToGenerate) {
        XmlSerializer xmlSerializer = new XmlSerializer(panoToGenerate.GetType());

        using (StringWriter textWriter = new StringWriter()) {
          xmlSerializer.Serialize(textWriter, panoToGenerate);
          return textWriter.ToString();
        }
      }

      public static String UpdateScene(Scene scene) {
        //Conversion...
        string directoryPath = System.IO.Path.GetDirectoryName(CurrentTourPath);
        List<string> directories = new List<string>();

        foreach (string folder in Directory.GetDirectories(directoryPath)) {
          directories.Add(System.IO.Path.GetFileName(folder));
        }

        foreach (Model.Image img in scene.Image) {
          string folderNameWithoutNumber = img.Front.Url.Split('_')[0];
          string leftNumber = img.Front.Url.Split('_')[1].Split('/')[0];
          string rightNumber = string.Empty;

          //Getting other number from current folder
          var res = directories.Where(x => x.Split('_')[0] == folderNameWithoutNumber);
          foreach (string result in res) {
            if (!result.Split('/')[0].Contains(leftNumber)) {
              rightNumber = result.Split('_')[1].Split('/')[0];
              continue;
            }
          }

          img.Stereo = "true";
          img.Stereolabels = leftNumber + "|" + rightNumber;
          img.Front.Url = img.Front.Url.Replace(leftNumber, "%t");
          img.Right.Url = img.Right.Url.Replace(leftNumber, "%t");
          img.Back.Url = img.Back.Url.Replace(leftNumber, "%t");
          img.Left.Url = img.Left.Url.Replace(leftNumber, "%t");
          img.Up.Url = img.Up.Url.Replace(leftNumber, "%t");
          img.Down.Url = img.Down.Url.Replace(leftNumber, "%t");
        }
        return GenerateXmlFromScene(scene);
      }
    }
}
```

PanoramaConverter.Model namespace classes

```
using System;
using System.Collections.Generic;
using System.Linq;
using System.Text;
using System.Threading.Tasks;
using System.Xml.Serialization;

namespace PanoramaConverter.Model {
  [XmlRoot(ElementName = "autorotate")]
  public class Autorotate {
    [XmlAttribute(AttributeName = "horizon")]
    public string Horizon { get; set; }
    [XmlAttribute(AttributeName = "tofov")]
    public string Tofov { get; set; }
    [XmlAttribute(AttributeName = "waittime")]
    public string Waittime { get; set; }
    [XmlAttribute(AttributeName = "speed")]
    public string Speed { get; set; }
  }

  [XmlRoot(ElementName = "panoview")]
  public class Panoview {
    [XmlAttribute(AttributeName = "h")]
    public string H { get; set; }
    [XmlAttribute(AttributeName = "v")]
    public string V { get; set; }
    [XmlAttribute(AttributeName = "fov")]
    public string Fov { get; set; }
    [XmlAttribute(AttributeName = "hmin")]
    public string Hmin { get; set; }
    [XmlAttribute(AttributeName = "hmax")]
    public string Hmax { get; set; }
    [XmlAttribute(AttributeName = "vmin")]
    public string Vmin { get; set; }
    [XmlAttribute(AttributeName = "vmax")]
    public string Vmax { get; set; }
    [XmlAttribute(AttributeName = "fovmax")]
    public string Fovmax { get; set; }
  }

  [XmlRoot(ElementName = "view")]
  public class View {
    [XmlAttribute(AttributeName = "fisheye")]
    public string Fisheye { get; set; }
    [XmlAttribute(AttributeName = "limitview")]
    public string Limitview { get; set; }
    [XmlAttribute(AttributeName = "hlookatmin")]
    public string Hlookatmin { get; set; }
    [XmlAttribute(AttributeName = "hlookatmax")]
    public string Hlookatmax { get; set; }
    [XmlAttribute(AttributeName = "vlookatmin")]
    public string Vlookatmin { get; set; }
    [XmlAttribute(AttributeName = "vlookatmax")]
    public string Vlookatmax { get; set; }
    [XmlAttribute(AttributeName = "maxpixelzoom")]
    public string Maxpixelzoom { get; set; }
    [XmlAttribute(AttributeName = "fovtype")]
    public string Fovtype { get; set; }
    [XmlAttribute(AttributeName = "fovmax")]
    public string Fovmax { get; set; }
    [XmlAttribute(AttributeName = "fov")]
```

```
        public string Fov { get; set; }
        [XmlAttribute(AttributeName = "hlookat")]
        public string Hlookat { get; set; }
        [XmlAttribute(AttributeName = "vlookat")]
        public string Vlookat { get; set; }
    }

    [XmlRoot(ElementName = "preview")]
    public class Preview {
        [XmlAttribute(AttributeName = "url")]
        public string Url { get; set; }
        [XmlAttribute(AttributeName = "type")]
        public string Type { get; set; }
        [XmlAttribute(AttributeName = "striporder")]
        public string Striporder { get; set; }
    }

    [XmlRoot(ElementName = "front")]
    public class Front {
        [XmlAttribute(AttributeName = "url")]
        public string Url { get; set; }
    }

    [XmlRoot(ElementName = "right")]
    public class Right {
        [XmlAttribute(AttributeName = "url")]
        public string Url { get; set; }
    }

    [XmlRoot(ElementName = "back")]
    public class Back {
        [XmlAttribute(AttributeName = "url")]
        public string Url { get; set; }
    }

    [XmlRoot(ElementName = "left")]
    public class Left {
        [XmlAttribute(AttributeName = "url")]
        public string Url { get; set; }
    }

    [XmlRoot(ElementName = "up")]
    public class Up {
        [XmlAttribute(AttributeName = "url")]
        public string Url { get; set; }
    }

    [XmlRoot(ElementName = "down")]
    public class Down {
        [XmlAttribute(AttributeName = "url")]
        public string Url { get; set; }
    }

    [XmlRoot(ElementName = "image")]
    public class Image {
        [XmlElement(ElementName = "front")]
        public Front Front { get; set; }
        [XmlElement(ElementName = "right")]
        public Right Right { get; set; }
        [XmlElement(ElementName = "back")]
        public Back Back { get; set; }
        [XmlElement(ElementName = "left")]
        public Left Left { get; set; }
```

```
        [XmlElement(ElementName = "up")]
        public Up Up { get; set; }
        [XmlElement(ElementName = "down")]
        public Down Down { get; set; }
        [XmlAttribute(AttributeName = "stereo")]
        public string Stereo { get; set; }
        [XmlAttribute(AttributeName = "stereolabels")]
        public string Stereolabels { get; set; }
        [XmlAttribute(AttributeName = "type")]
        public string Type { get; set; }
        [XmlAttribute(AttributeName = "devices")]
        public string Devices { get; set; }
    }

    [XmlRoot(ElementName = "scene")]
    public class Scene {
        [XmlElement(ElementName = "autorotate")]
        public Autorotate Autorotate { get; set; }
        [XmlElement(ElementName = "panoview")]
        public Panoview Panoview { get; set; }
        [XmlElement(ElementName = "view")]
        public View View { get; set; }
        [XmlElement(ElementName = "preview")]
        public Preview Preview { get; set; }
        [XmlElement(ElementName = "image")]
        public List<Image> Image { get; set; }
        [XmlAttribute(AttributeName = "name")]
        public string Name { get; set; }
        [XmlAttribute(AttributeName = "heading")]
        public string Heading { get; set; }
        [XmlAttribute(AttributeName = "thumburl")]
        public string Thumburl { get; set; }
        [XmlAttribute(AttributeName = "backgroundsound")]
        public string Backgroundsound { get; set; }
        [XmlAttribute(AttributeName = "backgroundsoundloops")]
        public string Backgroundsoundloops { get; set; }
        [XmlAttribute(AttributeName = "haslocalsounds")]
        public string Haslocalsounds { get; set; }
        [XmlAttribute(AttributeName = "haspolygons")]
        public string Haspolygons { get; set; }
        [XmlAttribute(AttributeName = "titleid")]
        public string Titleid { get; set; }
        [XmlAttribute(AttributeName = "descriptionid")]
        public string Descriptionid { get; set; }
        [XmlAttribute(AttributeName = "multires")]
        public string Multires { get; set; }
        [XmlAttribute(AttributeName = "planar")]
        public string Planar { get; set; }
        [XmlAttribute(AttributeName = "full360")]
        public string Full360 { get; set; }
        [XmlAttribute(AttributeName = "video")]
        public string Video { get; set; }
        [XmlAttribute(AttributeName = "seen")]
        public string Seen { get; set; }
    }
}
```

Code for Unity custom HDR system

HDRChooserImmediate class

```
using System.Collections;
using System.Collections.Generic;
using UnityEngine;

public class HDRChooserImmediate : MonoBehaviour {

    public enum HDRMode { HDR_STATIC, HDR_HEAD_TRACKING, HDR_EYE_TRACKING }
    public HDRMode Mode;

    public GameObject HDRSystem;
    public Material Left_HDR_Static, Right_HDR_Static, Left_HDR_Dynamic, Right_HDR_Dynamic;
    public Material Left_HDR_GroundTruth, Right_HDR_GroundTruth;
    public float MultiplierExpositionOverride = 1F;
    public float MaxExpositionOverride = 1F;
    public float MinExpositionOverride = 0.61F;
    public float TimeOfTransitionHead = 0.8F;
    public float TimeOfTransitionEye = 0.01F;
    public float PeripheralFOVForAdaptationHead = 30F;
    public float PeripheralFOVForAdaptationEyes = 9F;
    public bool SameExpositionBothEyes = true;
    public float ExpositionStartOffset = 0F;

    private Transform FoveInterface, //To activate / deactivate scripts
        LeftEvaluator, RightEvaluator, //To activate / deactivate lookAt
        LeftRealEye, RightRealEye,  //To change Skybox to static / dynamic
        LeftEyeCursor, RightEyeCursor, CenterRedPointer; //Pointer and cursors for Dynamic visualization
    private bool initNeeded = true;

    void Start() {
        if (HDRSystem != null) {
            FoveInterface = HDRSystem.transform.Find("Fove Rig").transform.Find("Fove Interface");
            LeftEvaluator = FoveInterface.transform.Find("LeftEyePixelViewEyeAdaptation");
            RightEvaluator = FoveInterface.transform.Find("RightEyePixelViewEyeAdaptation");
        }
    }

    void Update() {
        if (initNeeded) {
            LeftRealEye = FoveInterface.transform.Find("FOVE Eye (Left)");
            RightRealEye = FoveInterface.transform.Find("FOVE Eye (Right)");
            LeftEvaluator.GetComponent<Skybox>().material = Left_HDR_GroundTruth;
            RightEvaluator.GetComponent<Skybox>().material = Right_HDR_GroundTruth;
            LeftEyeCursor = HDRSystem.transform.Find("LeftEyeCursor");
            RightEyeCursor = HDRSystem.transform.Find("RightEyeCursor");
            CenterRedPointer = FoveInterface.Find("CenterRedPointer");
            initNeeded = false;
        }

        //Update if I want to correct exposition of both eyes simultaneously or independently
        FoveInterface.GetComponent<HDRSystemV3>().sameExpositionInBothEyes = SameExpositionBothEyes;


        switch (Mode) {
            case HDRMode.HDR_EYE_TRACKING:
                FoveInterface.GetComponent<HDRSystemV3>().enabled = true;
                FoveInterface.GetComponent<HDRSystemV3>().LeftGroundTruth = Left_HDR_GroundTruth;
                FoveInterface.GetComponent<HDRSystemV3>().RightGroundTruth = Right_HDR_GroundTruth;
                FoveInterface.GetComponent<HDRSystemV3>().LeftAdaptedView = Left_HDR_Dynamic;
                FoveInterface.GetComponent<HDRSystemV3>().RightAdaptedView = Right_HDR_Dynamic;
                FoveInterface.GetComponent<HDRSystemV3>().expositionMultiplicationFactor = MultiplierExpositionOverride;
                FoveInterface.GetComponent<HDRSystemV3>().maxViewableExposition = MaxExpositionOverride;
                FoveInterface.GetComponent<HDRSystemV3>().minViewableExposition = MinExpositionOverride;
                FoveInterface.GetComponent<HDRSystemV3>().smoothTime = TimeOfTransitionEye;
```

```
            //Setting peripheral adaptation FOV or both the cameras of adapted view
FoveInterface.GetComponent<HDRSystemV3>().UpdatePeripheralFOVForAdaptation(PeripheralFOVForAdaptationEyes);
            FoveInterface.GetComponent<HDRSystemV3>().expositionStartOffset = ExpositionStartOffset;
            LeftEvaluator.GetComponent<LookAtTarget>().enabled =
            RightEvaluator.GetComponent<LookAtTarget>().enabled = true;
            LeftRealEye.GetComponent<Skybox>().material = Left_HDR_Dynamic;
            RightRealEye.GetComponent<Skybox>().material = Right_HDR_Dynamic;
            CenterRedPointer.gameObject.SetActive(false);
            LeftEyeCursor.gameObject.SetActive(true);
            RightEyeCursor.gameObject.SetActive(true);
            break;
          case HDRMode.HDR_HEAD_TRACKING:
            FoveInterface.GetComponent<HDRSystemV3>().enabled = true;
            FoveInterface.GetComponent<HDRSystemV3>().LeftGroundTruth = Left_HDR_GroundTruth;
            FoveInterface.GetComponent<HDRSystemV3>().RightGroundTruth = Right_HDR_GroundTruth;
            FoveInterface.GetComponent<HDRSystemV3>().LeftAdaptedView = Left_HDR_Dynamic;
            FoveInterface.GetComponent<HDRSystemV3>().RightAdaptedView = Right_HDR_Dynamic;
            FoveInterface.GetComponent<HDRSystemV3>().expositionMultiplicationFactor = MultiplierExpositionOverride;
            FoveInterface.GetComponent<HDRSystemV3>().maxViewableExposition = MaxExpositionOverride;
            FoveInterface.GetComponent<HDRSystemV3>().minViewableExposition = MinExpositionOverride;
            FoveInterface.GetComponent<HDRSystemV3>().smoothTime = TimeOfTransitionHead;


            //Setting peripheral adaptation FOV or both the cameras of adapted view
            FoveInterface.GetComponent<HDRSystemV3>().UpdatePeripheralFOVForAdaptation(PeripheralFOVForAdaptationHead);
            FoveInterface.GetComponent<HDRSystemV3>().expositionStartOffset = ExpositionStartOffset;
            LeftEvaluator.GetComponent<LookAtTarget>().enabled =
            RightEvaluator.GetComponent<LookAtTarget>().enabled = false;
            LeftEvaluator.transform.localRotation =
            RightEvaluator.transform.localRotation = Quaternion.Euler(0, 0, 0);
            LeftRealEye.GetComponent<Skybox>().material = Left_HDR_Dynamic;
            RightRealEye.GetComponent<Skybox>().material = Right_HDR_Dynamic;
            CenterRedPointer.gameObject.SetActive(true);
            LeftEyeCursor.gameObject.SetActive(false);
            RightEyeCursor.gameObject.SetActive(false);
            break;
          case HDRMode.HDR_STATIC:
            FoveInterface.GetComponent<HDRSystemV3>().enabled = false;
            LeftEvaluator.GetComponent<LookAtTarget>().enabled =
            RightEvaluator.GetComponent<LookAtTarget>().enabled = false;
            LeftRealEye.GetComponent<Skybox>().material = Left_HDR_Static;
            RightRealEye.GetComponent<Skybox>().material = Right_HDR_Static;
            CenterRedPointer.gameObject.SetActive(false);
            LeftEyeCursor.gameObject.SetActive(false);
            RightEyeCursor.gameObject.SetActive(false);
            break;
        }

        if (Input.GetKeyUp(KeyCode.Space)) {
          if (Mode == HDRMode.HDR_EYE_TRACKING)
            Mode = 0;
          else
            Mode = Mode + 1;
        }
      }
    }
```

HDRSystemV3 class

```
using System.Collections;
using System.Collections.Generic;
using UnityEngine;

public class HDRSystemV3 : MonoBehaviour {

    public static float CurrentExposure;
    public float expositionMultiplicationFactor = 10f;

    //Skyboxes to be used
    public Material LeftGroundTruth, RightGroundTruth, LeftAdaptedView, RightAdaptedView;

    //These are the actual pictures that I will project on the planes that the FOVE will see
    public Camera leftEyeExposition, rightEyeExposition;

    //These are the single pixels 1x1 textures I will use with the HDR adapting cameras as target render, to evaluate the exposition and
    //set it on the leftEyeExposition and rightEyeExposition skyboxes at runtime.
    public Camera leftPixelEyeExposition, rightPixelEyeExposition;
    public GameObject FoveInterface;
    public float leftCurrentExposition = 1f, rightCurrentExposition = 1f;

    public float maxViewableExposition = 2.5F;
    public float minViewableExposition = -5F;

    Camera leftEyeFOVE, rightEyeFOVE;
    Texture2D leftPixelEyeExposition_texture, rightPixelEyeExposition_texture;
    private bool initNeeded = true;

    //Smoothing
    public float smoothTime = 0.3F;
    private float yVelocity = 0.0F;

    //Same exposition of eyes
    public bool sameExpositionInBothEyes = true;

    public float expositionStartOffset = 0F;

    // Use this for initialization
    void Start() {

    }

    // Update is called once per frame
    void Update() {
        if (initNeeded) {
            UpdateFoveLayersEyesSkyboxesFOVE();
            SetupRenderTextures();
            initNeeded = false;
        }

        //Evaluating exposure
        UpdateRenderTextures();
        CalculateAverageBrightness();

        //Updating real view exposure
        UpdateEyeSkyboxesExpositionValue();
    }
```

```
    #region METHODS
    /// <summary>
    /// Checks if my system is ready to operate or if some Gameobjects / Cameras are missing from the parameters on the script. The missing ones have to be initialized from Unity GUI Editor.
    /// </summary>
    /// <returns></returns>
    public bool isHDRSystemAlessioInitialized() {
        if (FoveInterface == null) {
            Debug.LogError("FOVE INTERFACE NOT SET ON THE HDRSystemAlessio script!");
        }
        if (leftEyeFOVE == null) {
            Debug.LogError("leftEyeFOVE NOT SET ON THE HDRSystemAlessio script!");
        }
        if (rightEyeFOVE == null) {
            Debug.LogError("rightEyeFOVE NOT SET ON THE HDRSystemAlessio script!");
        }
        if (leftPixelEyeExposition == null) {
            Debug.LogError("leftPixelEyeExposition NOT SET ON THE HDRSystemAlessio script!");
        }
        if (rightPixelEyeExposition == null) {
            Debug.LogError("rightPixelEyeExposition NOT SET ON THE HDRSystemAlessio script!");
        }
        if (leftEyeExposition == null) {
            Debug.LogError("leftEyeExposition NOT SET ON THE HDRSystemAlessio script!");
        }
        if (rightEyeExposition == null) {
            Debug.LogError("rightEyeExposition NOT SET ON THE HDRSystemAlessio script!");
        }
        return !(FoveInterface == null || leftEyeFOVE == null || rightEyeFOVE == null || leftPixelEyeExposition == null || rightPixelEyeExposition == null || leftEyeExposition == null || rightEyeExposition == null);
    }

    public void SetupSkyBoxes() {
        if (LeftAdaptedView != null && leftEyeFOVE != null) {
            leftEyeFOVE.GetComponent<Skybox>().material = LeftAdaptedView;
        }
        if (RightAdaptedView != null && rightEyeFOVE != null) {
            rightEyeFOVE.GetComponent<Skybox>().material = RightAdaptedView;
        }
        if (LeftGroundTruth != null && leftPixelEyeExposition != null) {
            leftPixelEyeExposition.GetComponent<Skybox>().material = LeftGroundTruth;
        }
        if (RightGroundTruth != null && rightPixelEyeExposition != null) {
            rightPixelEyeExposition.GetComponent<Skybox>().material = RightGroundTruth;
        }
    }

    public void UpdateFoveLayersEyesSkyboxesFOVE() {
        if (FoveInterface == null) return;

        Transform leftEye = FoveInterface.transform.Find("FOVE Eye (Left)");
        Transform rightEye = FoveInterface.transform.Find("FOVE Eye (Right)");

        if (leftEyeFOVE == null && leftEye != null) {
            leftEyeFOVE = leftEye.gameObject.GetComponent<Camera>();
        }

        if (rightEyeFOVE == null && rightEye != null) {
            rightEyeFOVE = rightEye.gameObject.GetComponent<Camera>();
        }

        //Setting Skyboxes for ground truth and adapted view
        SetupSkyBoxes();
    }
```

```
    public void SetupRenderTextures() {
        if (!isHDRSystemAlessioInitialized()) return;

        leftPixelEyeExposition_texture = new Texture2D(leftPixelEyeExposition.targetTexture.width, leftPixelEyeExposition.targetTexture.height, TextureFormat.ARGB32, true);
        rightPixelEyeExposition_texture = new Texture2D(rightPixelEyeExposition.targetTexture.width, rightPixelEyeExposition.targetTexture.height, TextureFormat.ARGB32, true);
    }

    public void UpdatePeripheralFOVForAdaptation(float newPeripheralFOVInDegrees) {
        if (leftPixelEyeExposition != null) {
            leftPixelEyeExposition.fieldOfView = newPeripheralFOVInDegrees;
        }
        if (rightEyeExposition != null) {
            rightEyeExposition.fieldOfView = newPeripheralFOVInDegrees;
        }
    }

    public void UpdateRenderTextures() {
        UpdateTexture(leftPixelEyeExposition.targetTexture, true);
        UpdateTexture(rightPixelEyeExposition.targetTexture, false);
    }

    void CalculateAverageBrightness() {
        if (!isHDRSystemAlessioInitialized()) return;

        //2x2 texture, I do the average
        leftCurrentExposition = expositionMultiplicationFactor /
            (Brightness(leftPixelEyeExposition_texture.GetPixel(0, 0)) +
            Brightness(leftPixelEyeExposition_texture.GetPixel(0, 1)) +
            Brightness(leftPixelEyeExposition_texture.GetPixel(1, 0)) +
            Brightness(leftPixelEyeExposition_texture.GetPixel(1, 1))) / 4f;

        if (sameExpositionInBothEyes) {
            rightCurrentExposition = leftCurrentExposition;
        } else {
            //2x2 texture, I do the average
            rightCurrentExposition = expositionMultiplicationFactor /
                (Brightness(rightPixelEyeExposition_texture.GetPixel(0, 0)) +
                Brightness(rightPixelEyeExposition_texture.GetPixel(0, 1)) +
                Brightness(rightPixelEyeExposition_texture.GetPixel(1, 0)) +
                Brightness(rightPixelEyeExposition_texture.GetPixel(1, 1))) / 4f;
        }
    }

    private float Brightness(Color c) {
        return Mathf.Sqrt(
           c.r * c.r * .241F +
           c.g * c.g * .691F +
           c.b * c.b * .068F);
    }

    public void UpdateEyeSkyboxesExpositionValue() {
        if (!isHDRSystemAlessioInitialized()) return;

        //Smoothing the exposure value
        float newexposureLeft = Mathf.Max(Mathf.Min(Mathf.SmoothDamp(LeftAdaptedView.GetFloat("_Exposure"), leftCurrentExposition + expositionStartOffset, ref yVelocity, smoothTime), maxViewableExposition), minViewableExposition);
        float newexposureRight = sameExpositionInBothEyes? newexposureLeft : Mathf.Max(Mathf.Min(Mathf.SmoothDamp(RightAdaptedView.GetFloat("_Exposure"), rightCurrentExposition + expositionStartOffset, ref yVelocity, smoothTime), maxViewableExposition), minViewableExposition);

        LeftAdaptedView.SetFloat("_Exposure", newexposureLeft);
        RightAdaptedView.SetFloat("_Exposure", newexposureRight);
    }
```

```
    public void UpdateTexture(RenderTexture rTex, bool isLeft) {
        if (!isHDRSystemAlessioInitialized()) return;
        RenderTexture.active = rTex;
        if (isLeft) {
            leftPixelEyeExposition_texture.ReadPixels(new Rect(0, 0, rTex.width, rTex.height), 0, 0);
            leftPixelEyeExposition_texture.Apply();
        } else {
            rightPixelEyeExposition_texture.ReadPixels(new Rect(0, 0, rTex.width, rTex.height), 0, 0);
            rightPixelEyeExposition_texture.Apply();
        }
    }
    #endregion
}
```

APPENDIX E. Additional References

The following table gives more information on the reference numbers used in **Figure 57**. Each row shows on the first column the reference number reported by **Figure 57**, and on the second column all corresponding references related to that reference number.

REFERENCES OF DIAGRAM – REFERENCES OF THIS THESIS	
1	Transitional environments improve distance estimates. Furthermore, a wider field of view contributes to sense of presence [349].
2	Steinicke et al. [349] reported studies showing that low latency contributes to sense of presence [350].
3	Steinicke et al. [349] reported studies showing that dynamic shadows of objects contribute to sense of presence [351] [349].
4	Field of view has an influence on scene perception [352].
5	User interaction with feedbacks in space in VR enhances distance perception reducing errors [353].
7	Kelly et al. [353] reported studies showing that low graphics reduce distance estimation [21].
8	Kelly et al. [353] reported studies showing that Inaccurate S3D might alter distance estimation, but within action space (2-30m) it is likely not the source of distance compression reported in previous virtual environment studies [354].
9	Kelly et al. [353] reported several researches on the influence of reduced field of view on distance estimation [341] [340].
10	Investigation on other possible causes on distance compression in VR [355].
11	Assessment on egocentric distance perception in a high fidelity, low latency, immersive virtual environment [356].
12	Willemsen et al. [340] reported that accommodation and convergence are absolute egocentric cues, but individually, do not have much direct effect beyond personal space (i.e., out to about 2 m) [20]. At distances up to 2 m, accommodation and convergence have been shown to be important cues that influence space perception in virtual environments [357] [358].
13	Review on many of the potential factors influencing the perception of egocentric distances in Virtual Environments [323].
14	Renner et al. [323] reported that it has already been shown that greater visual realism enhances presence and physiological responses [348]. Furthermore, if the GFOV settings do not correspond to the DFOV of the HMD, images are minified or magnified, which was repeatedly shown to influence participants' distance estimates [359] [360] [361] [362] [363].
15	Renner et al. [323] reviewed studies showing better matching task performance in a virtual outdoor environment than in a virtual indoor environment [364].
16	Renner et al. [323] reported results of Lappin et al. [365] and Witt et al. [366], who showed in real environments an effect of environmental context (which is not yet fully understood) over distance estimation.
17	Renner et al. [323] reported studies of Mohler et al. [324], which showed that both a tracked and a static avatar improved distance estimates even if the avatar was dislocated.
18	Renner et al. [323] reported studies of Leyrer et al. [367], which in their experiment found that a tracked avatar improved distance estimates significantly only if ownership was controlled (i.e., the participants' feeling that the avatar was located at the same location as their body and the feeling that the avatar was their own body). Thus, an avatar might improve distance estimates only if the user accepts it as the representation of his or her own body and not if the user sees it as an object. The feeling that a virtual body is one's own body is also described as self-presence [368] and is thought to enhance the sense of presence [369] [370].

19	Renner et al. [323] reported that Interrante et al. [356] used familiar objects inside a virtual environment environment, and a virtual replica of the room that the participants had seen before. This led to the interesting finding that participants did not underestimate virtual distances when the virtual environment was an exact replica of the real environment they were currently standing in. This finding was replicated by Phillips et al. [371].
20	Renner et al. [323] reported that Interrante et al. [372] made another experiment on different sizes of replicas of real environments and concluded that participants were not able to use metric information from the real environment to calibrate their blind walking estimates. Instead, the good estimates might be due to a higher sense of presence.
21	Renner et al. [323] reported a number of authors have argued that a higher sense of presence might improve distance perception (e.g., Interrante et al. [356]; Mohler et al. [373]; Phillips et al. [374]; Ries et al. [325]; Steinicke et al. [349]).
22	Renner et al. [323] reported that many authors tested for gender differences but found none when evaluating distance estimations in Virtual Environments (e.g. Creem-Regehr et al. [375]; Interrante et al. [356]; Naceri and Chellali [376]).
23	Renner et al. [323] reported that Phillips et al. [371] correlated the scores of several personality questionnaires that assess personality traits, which are hypothesized to be associated with a higher sense of presence, with distance estimates. They found a significant correlation with the trait absorption, indicating that participants who were more open to absorbing and self-altering experiences [377] underestimated distances to a greater extent.
24	Renner et al. [323] reported that age needs to be considered if the participants are children. This is in line with research in real environments showing differing spatial perception in children and older observers (e.g. Harway [378]; Norman et al. [379]). Furthermore, they reported that the level of experience with the used virtual reality hardware system or virtual reality in general might have an influence.
25	Results from the experiment of Bruder et al. [380] report that both screen distance and parallax have a strong asymmetric effect on distance judgments: an increased distance underestimation was found for positive parallax conditions, while less distance overestimation for negative and zero.
26	Lappin et al. [365] found an effect of environmental context on distance estimation in virtual environments. However, the experiment by Bodenheimer et al. [364] reported opposite results. Bodenheimer et al. believed this was because all outdoor experiments were done in the afternoon, so their subjects may not have been able to see properly with the sun shining in their face.
27	Bodenheimer et al. [364] reported studies by Wu et al. [344], who showed that a vertical FOV of 21◦ or less leads to an underestimation of distance.
28	Takahashi et al. [381] studied psychological influences on distance estimation in a virtual reality environment.
29	Takahashi et al. [381] reported studies showing that: threatening objects (e.g., a living tarantula) are perceived as closer (Cole et al. [382]); a location related to a rival group (e.g. Fenway Park for a Yankees fan) is imagined as nearer when accompanied by a feeling of threat (Xiao and Van Bavel [383]); pointy objects tend to evoke aversion (Morse and Cohen [384]; Shabani and Fisher [385]); distance perception modulation might be related to the violation of personal space (Liberman et al. [386]); objects in a virtual environment are felt as intrusive (Wilcox et al. [387]); intrusive cones would be perceived as closer when they violate the observers' personal space (Schnall [388]).
30	Takahashi et al. [381] reported studies showing that: hills steeper after 1-h run (Proffitt et al. [389]; Proffitt [390]); a glass of water appears larger when thirsty (Veltkamp et al. [391]); mental and bodily states modify spatial perception (Proffitt [390]); desired objects are felt as nearer or closer (Balcetis and Dunning [392]; Alter and Balcetis [393]).

31	Iachini et al. [394] showed in their experiment that an emotional influence existed in their viewers and affected distances.
32	Iachini et al. [394] reported studies from literature showing that environmental properties, emotional state, and dangerousness of situation have an influence on reachability judgements [395] [396] [397] [398].
33	Knapp et al. [341] conducted studies on the effect of a limited field of view of HMDs on distance underestimation in virtual environments.
34	Knapp et al. [341] reported studies showing that limiting vertical FOV when the head is stationary produced a significant underestimation [343]. Furthermore, they reported in their experiment different results, and justified them with the need to distinguish between "instantaneous FOV," as set by the simulated HMD, and "effective field of regard," which is determined by the instantaneous FOV sweeping out a larger region of space as the head moves.
35	Messing et al. [399] conducted a study investigating relations between distance perception and the visual horizon in HMDs.
36	Messing et al. [399] reported studies showing that the choice of wireframe or photorealistic scenes has not a primary influence over distance estimation [21] [353].
37	Messing et al. [399] reported results from Teghtsoonian et al. [400] [401] showing that verbal estimates of distances in indoor spaces were expansive (had exponents greater than 1), while those in outdoor spaces were slightly compressive (exponents slightly less than 1).
38	Williams et al. [402] investigated the possible influence of the angle of declination (which is formed when the height of the viewer is different from the height of the camera used to visualize the virtual environment) on distance estimation.
39	Williams et al. [402] reported a previous study [402] showing that people were better at judging distances if they had virtual feet. They found that this was especially evident for distances that were within a few meters.
40	Phillips et al. [403] reported that despite in literature photorealism seems not to affect distance perception, their results show that for places already known or in which the viewer is located (e.g. high-fidelity realistic replica of the environment) distance underestimation is reduced, especially when compared to non-photorealistic rendering (NPR drawing style).
41	Steinicke et al. [404] reported that transitional elements like doors from Reality to Virtual Environment can improve distance estimation. This was proved via blind walking experiments. Furthermore, the sense of presence is enhanced via transitional environments too.
42	Messing et al. [342] reported same results in 3 experiments: the first proving that wearing an HMD with restricted FOV compresses distance estimation, the second that distance perception indoor is expanded but compressed outdoors (which replicates the same results of the experiment of Teghtsoonians [400] [401]), the third that a manipulation of the angle of declination below the horizon can alter distance estimation.
43	Plumert et al. [405] reported that distance perception may be better in virtual environments involving large-screen immersive displays than in those involving head-mounted displays (HMDS).
44	Renner et al. [323] reported studies showing that greater visual realism enhances presence and physiological responses (Slater et al. [348]).
45	Plumert et al. [405] reported that According to Wu et al. [344] a restricted vertical FOV (21◦ or less) leads to underestimation of distance in the real environment. Likewise, Witmer and Sadowski [345]suggest that the reduced vertical FOV in HMDs may degrade convergent linear perspective and relative size as cues to distance.
46	Plumert et al. [405] reported that Knapp and Loomis [341] recently found that a reduced vertical FOV similar to that experienced with HMDs had no impact on blindfolded walking in the real environment. Likewise, Creem-Rehehr et al. [375] found that not being able to see

	the area around one's feet (a typical feature of HMDs) does not impair people's ability to perceive distance in the real environment. Together, these findings suggest that the reduced vertical FOV in commonly used HMDs does not account for underestimation of distance in virtual environments using HMDs. However, this does not rule out the possibility that reduced vertical FOV in combination with other aspects of virtual environments contributes to underestimation of distance in HMDs.
48	Willemsen et al. [406] consider real-time rendering and immersive display technologies as possible causes of underestimated distance in VE. They proved that distance estimation compression in Virtual Environments is due to the display device, not to the graphics of the panorama (photographic vs traditional rendered). This is also proved by their experiment, which shows that display has a major effect on underestimated distances.
49	Results from Willemsen et al. [340] do indicate that there is a reliable effect of underestimation when viewing the real world with the mock HMD suggesting that mechanical aspects of HMDs account for some of the distance compression effects found in virtual environment research.
50	Diemer et al. [319] reported that presence is the activator of emotions in VR. Furthermore, high quality HMD and simple scenes provide more emotions. Finally, more immersion results in increased presence. This might suggest that more immersion causes more emotions, but literature offers conflicting theories.
51	Riva et al. [407] reported that a circular interaction between presence and emotions exists, and that affective content has a strong influence on the sense of presence.
52	Macedonio et al. [320] reported that individual variables of users such as absorption and hypnotisability influence the achieved sense of presence. This suggests that individual differences may moderate presence.
53	Price and Anderson [408] reported in their experiment that individuals with a phobia reproduced in the VE felt more present.
54	Bailey et al. [409] reported results from a study by Dinh et al. [410], who suggest that increased levels of sensory input (e.g. olfactory feedback) can increase the levels of overall presence (in a virtual office) and the memory of the virtual environment. Furthermore, they reported other researchers, who suggested that individuals' personalities, past experiences, and mental abilities for imagination are important factors for understanding presence (Heeter [411]; Wirth et al. [412]).
55	Schuemie et al. [316] reported in their survey that Slater et al. [370] stress the participant's sense of "being there" in the virtual environment, and point out that a high sense of presence in a VE requires a simultaneous low level of presence in the real world and vice versa. Furthermore, as Slater et al. [315] note, a key result of presence is that a person remembers the VE as a place rather than a set of pictures. However, it is still unclear whether higher measured presence causes stronger emotional responses in a VE or the other way around. Witmer and Singer [336] determined several factors that are thought to contribute to a sense of presence: • Control factors, the amount of control the user had on events in the VE. • Sensory factors, the quality, number and consistency of displays. • Distraction factors, the degree of distraction by objects and events in the real world. • Realism factors, the degree of realism of the portrayed VE. Furthermore, they found three factors, which did not perfectly match the original factors, mentioned above. These factors, which regrouped items from the original factors, were labelled: • Involved/Control — the control and responsiveness of a VE, and how involving a VE is.

	• Natural — the naturalness of interactions and control of locomotion, and the consistency of a VE. • Interface Quality — the amount of interference or distraction from task performance, and the participant's ability to concentrate on the tasks. In terms of causes of presence according to literature, the following results were reported: <u>1. Results by Slater and Usoh</u> [413] • High quality, high resolution information. • Consistency across all displays. • Interaction with environment. • Virtual body, the representation of the user's body in the VE. • Effect of action should be anticipated. <u>2. Results by Witmer and Singer</u> [336] • Control factors, the control the users have. • Sensory factors, the richness of the displayed information and consistency across displays. • Distraction factors, how much the user is distracted from the VE. • Realism factors, pictorial and social realism of the VE. 3. <u>Results by Sheridan</u> [347]: • Extent of sensory information. • Control of relation of sensors to environment. • Ability to modify physical environment. 4. <u>Results by Lombard and Ditton</u> [414]: • The form in which the information is presented. • The content of the information. • User characteristics. 5. <u>Results by Steuer</u> [415]: • Vividness refers to the ability of a technology to produce a sensorially rich mediated environment. • Interactivity refers to the degree to which users of a medium can influence the form or content of the mediated environment. • User characteristics refers to the individual differences in users. Furthermore, Welch et al. [416] reported a significant effect for pictorial realism. In terms of measuring presence, studies by Meehan [417] reported a correlation was found between skin conductance and presence as measured using the questionnaire by Slater and colleagues. These results tend to be supported by the findings of Wiederhold et al. [334], who performed a within subject experiment with five subjects, one diagnosed as being afraid of flying. During exposure to an airplane simulator on either a screen or using an HMD, skin conductance was found to be significantly higher for the HMD condition, which also generated the highest presence ratings on a presence questionnaire.
56	Hoffman et al. [318] reported that displaying images close to the patient's eyes helps give patients the illusion of "presence," the sensation that they are actually inside the computer-generated environment, interacting with virtual objects, instead of merely watching the virtual world on a distant computer screen [418]. Furthermore, increasing the field-of-view of a VR display has also been shown to increase presence.

57	Slater et al. [326] reported that shadows increase VIVIDNESS of the visual display, which enhance immersion, which enhances the sense of presence. Indeed, they speculate that the additional information provided by shadows about the movements of the VB in relationship to the surfaces of the VE can enhance this degree of association, and hence the degree of presence. Furthermore, even if they were unable to test this in their experiment, they still consider the proposition that shadows, increasing the degree of vividness of the visual displays, will enhance the sense of presence. The existence of a relationship between dynamic shadows and the sense of presence is not obvious and is motivated by the idea that presence is (amongst other things) a function of immersion, and immersion requires vividness.
58	Krijn et al. [331] reported that immersion (as the objective qualification of the VR equipment) is of influence on presence (Schubert et al. [419]).
59	Krijn et al. [331] reported some studies that did find a linear relationship between presence and anxiety experienced (Schuemie et al. [420]).
60	Krijn et al. [331] reported that other studies like the one they performed found no relationship between presence and anxiety (Regenbrecht et al. [421]).
61	Slater et al. [422] reported that the impact of the display aspects (Inclusive, Surrounding, Extensive, Vivid) is mediated through two filters: the application or task context and the perceptual requirements of the individual.
62	Slater et al. [422] reported that a study by Barfield and Hendrix [423] examined the influence on reported presence of display update rate. They found that there was such an influence that presence generally increased with increasing update rate, but that the reported presence was approximately constant between about 15Hz and 20Hz.
63	Slater et al. [422] reported that Hendrix and Barfield [351] discovered that stereopsis and a wider geometric field of view are positively correlated with the achieved sense of presence.
64	Slater et al. [422] reported that in the study by Welch et. al. [416], delay in visual feedback was another independent factor. A higher level of presence was reported under the condition of minimal delay, and this was a more important factor than the level of pictorial realism. Hendrix and Barfield [351] found that head-tracking significantly increased the reported sense of presence.
65	Slater et al. [422] reported that experimental studies by Hendrix and Barfield [424] examined the impact of sound on subjective presence. They discovered that spatialized sound leads to a higher reported presence than both no sound and non-spatialized sound.
66	Baños et al. [425] reported that according to Witmer and Singer [336] and Rizzo et al. [426] two factors are necessary to experience presence: involvement and immersion. Involvement has been defined by Witmer and Singer as “a psychological state experienced as a consequence of focusing one’s energy and attention on a coherent set of stimuli or meaningfully related activities and events”. It suggests being a function of the user’s internal characteristics, depending on variables such as interest and motivation.
67	Pausch et al. [332] proved that HMD allows better task performance compared to Desktop PC.
68	Usoh et al. [427] presented presence questionnaires and reported previous studies discussing factors influencing presence. Among them: - **High resolution**; - **Consistency of the displayed environment** across all sensory modalities. - **The possibility of the individual being able to navigate through and interact with objects in the environment**, including interaction with other actors which may spontaneously react to the individual;

	- **The individual's virtual body**, their self-representation within the environment, should be similar in appearance or functionality to the individual's own body, and respond appropriately to the movements of their head, eyes, and limbs; - **The connection between individual's actions and effects of those actions** should be simple enough for the individual to quickly learn.
69	Herbelin et al. [321] reported that kids are very fast in taking off from reality when telling stories, whereas adults - and more patients who reject a fearful situation- need to be more stimulated. Therefore, there is an essential need for appropriate mediations of the stimuli to enter an artificial presence state in virtual reality environments, especially for adults.
70	Herbelin et al. [321] reported studies by Huang and Alessi [322] that already supposed sense of presence and emotions to be strongly linked: "Like emotions, presence is continuously changing and dynamic".
71	Witmer and Singer [336] reported that according to Fontaine [428], presence seems to be a matter of focus. Focus occurs when one directs attention toward something. This focus is continually shifting in everyday life, as is obvious from the amount of presence required in performing everyday tasks like commuting. They also discuss the problem of the degree of control: in general, the more control a person has over the task environment or in interacting with the VE, the greater the experience of presence. Furthermore, on isolation they suggest that devices able to isolate users from their actual physical environment may increase presence in a VE. Finally, they discuss on scene realism suggesting that presence should increase as a function of VE scene realism (as governed by scene content, texture, resolution, light sources, field of view (FOV), dimensionality, etc.), and that scene realism does not require real-world content but refers to the connectedness and continuity of the stimuli being experienced.
72	Witmer and Singer [336] reported a study by Prothero and Hoffman [317] showing that using an eye mask to limit the field of view reduces the amount of presence.
73	Witmer and Singer [336] reported considerations by Sheridan [347] on "environmental richness": the greater the extent of sensory information transmitted to appropriate sensors of the observer, the stronger the sense of presence will be.
74	Renner et al. [323] reported that the lack of pictorial cues was proved to encourage distance underestimation. From research in real environments, it is known that when depth cues are reduced the accuracy of distance perception declines (Kunnapas [429]; Philbeck and Loomis [430]) and severe distortions of distance and size perception can occur.
75	Bailey et al. [409] reported other researchers, who suggested that individuals'personalities, past experiences, and mental abilities for imagination are important factors for understanding presence (Heeter [411]; Wirth et al. [412]).
76	Steinicke et al. [349] reported that a wider field of view and realistic physical simulations [431] contribute to a user's sense of presence.
77	Steinicke et al. [349] reported that stereoscopic display contributes to the sense of presence [432].
78	Hettinger et al. [433] reported in their book discussed "Immersion", "Sense of Immersion", and the relation with presence. When the sense of immersion (feeling of being surrounded by a computer-generated environment) is coupled with realistic visual imagery and high degree of interactivity in VR, an enhanced sense of presence is achieved.

Table 25. References of the diagram in ***Figure 57***, mapped with the references of this thesis.

APPENDIX F. Further discussions

Issues and Comments on the 3D panorama implementations

Vertical rotation vs horizontal rotation

An interesting topic has been the choice to capture the panoramic spherical pictures by rotating the camera tracing vertical circles or horizontal circles. **Figure 220** shows the two different techniques compared and the results obtained.

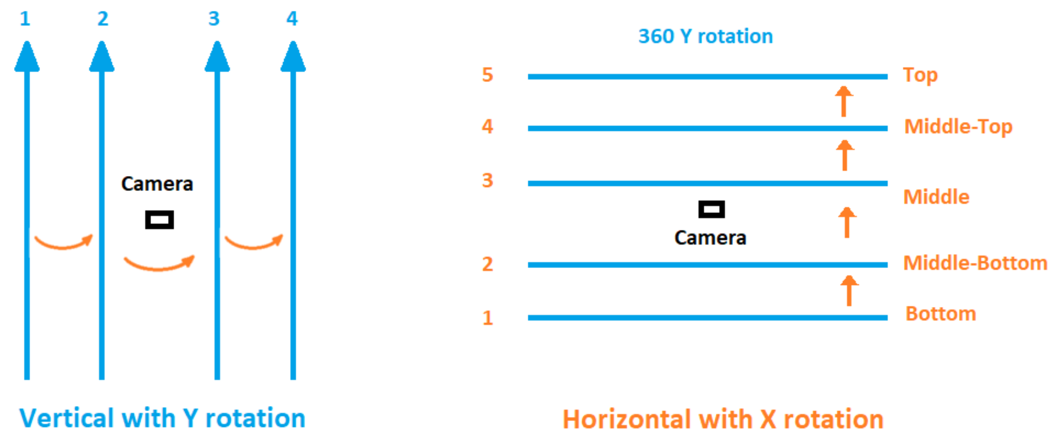


Figure 220 *- Two different approaches to capture the panorama. On the left, pictures are taken from bottom to top before rotating the camera horizontally. On the right, the camera is rotated horizontally for 360 degrees before moving vertically.*

Results are very similar, even in low light environments like the Monello's cave. However, the problem inside the cave was that it was completely dark, and a small light was used to enlighten each portion of the scene captured by the camera. The question was on whether the stitching with *Autopano* was going to have problems because of the different illumination during rotation of camera and light, and if the first or the second technique was going to work better than the other.

The verdict was that, even using a single light moving with the camera, the stitching was successful in both types of rotation, giving the same good result without problems.

The following **Figure 221** and **Figure 222** are examples of the two panoramas taken using the two different techniques, with no *InPaint* auto completion applied to the roof or to the floor of the scene.

Figure 221 *- Panorama taken using vertical acquisition with Y axis rotation.*

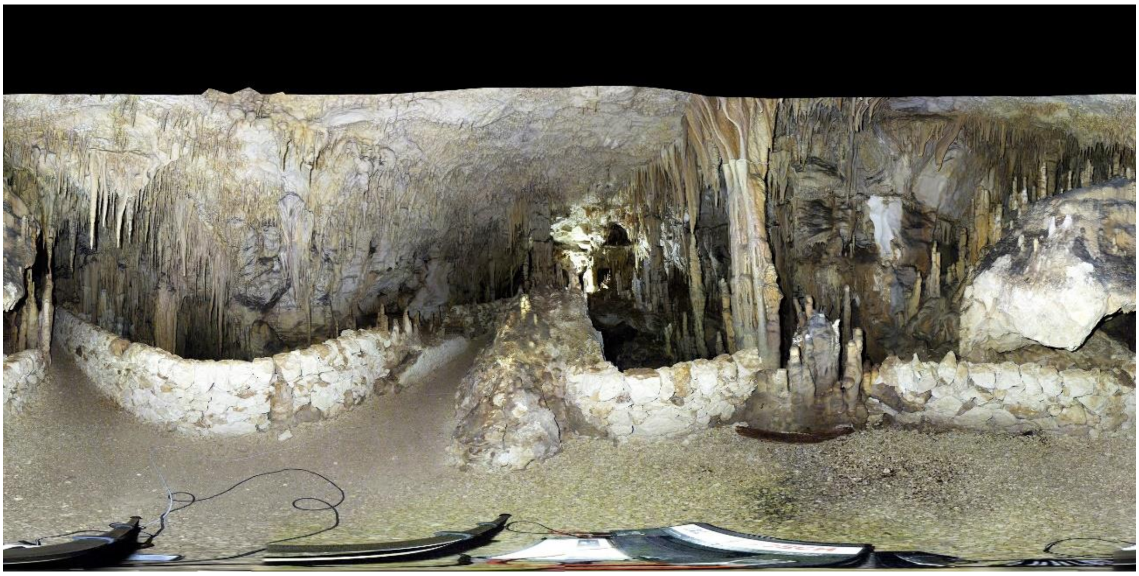


Figure 222 *- Panorama taken using horizontal acquisition with X axis rotation.*

Microsoft ICE vs Autopano

During the stitching phase, I tested the two software to check which one was rendering the best result. They are both powerful, but for different reasons. **Table 26** compares their features.

Microsoft ICE	Autopano
Renders panorama in **partial spherical format**, so the final image is not 360x180 and is not fully compatible with Panotour virtual reality tour that has as requirement images for panoramas with ratio equal to 2:1. To use the rendered pictures inside Panotour, it is necessary to manually edit the resolution of the pictures forcing the width to be the double of the height. In this case, the portion of the image that must be included is in the top of the image.	Renders panoramas in **full spherical format**, in 360x180 format, allowing a full compatible use inside Panotour for virtual reality tours using HMD. No manual editing is needed to fix the spherical format of the rendered panorama.
It provides a very **powerful plugin for auto completion of the image** solving the problem of the black hole in the roof and in the sky or floor.	**It doesn't include any auto completion plugin**, so whenever the stitched image presents black holes, it is necessary to process them manually or with a third-party software like InPaint to rebuilt artificially the missing pieces of the panorama.
It doesn't allow the user to manually move the tiles of the stitched image, so when the stitching presented errors it usually was not possible to manually fix them.	**It allows full control on the tiles of the stitched image to fix any error manually**.
It offers **different acquisition options** specifying the order of the captured images to give a little control on the disposition of the images.	**It recognizes automatically the photos also independently of the order they were captured**.
It is free for commercial purpose.	**It is not free.**
It can process only **Rectilinear lenses** (normal camera lenses) and must have the same focal length for all the pictures. It is not very good for perspective correction lenses (they are not provided in the software for the stitching).	It can process both **Rectilinear lenses** and **Fisheye lenses**, allowing also different focal lengths in a single panorama. It also performs **Perspective correction lenses** very well.
It doesn't offer anything else.	It also offers **Tonemapping functionality, Exposure fusion, HDR output, it can merge HDR from LDR images, it supports HDR input.**
It is **supported only in Microsoft Windows** operative systems from Vista to Windows 10.	It is **multiplatform**, working in Windows 2000 or later, Mac OS X 10.5 or later with Intel processor, Linux 2.6 or later.

Table 26 *- Microsoft ICE vs Autopano pros and cons.*

Considering all these aspects, for the final generation of the images, *Autopano* was chosen to stitch the pictures, using the third-party software *InPaint* to rebuild the missing information in the black holes, or GIMP in the case of the sky that is easier to artificially rebuild.

Figure 223 shows the settings used inside *Autopano* to generate the final panoramas.

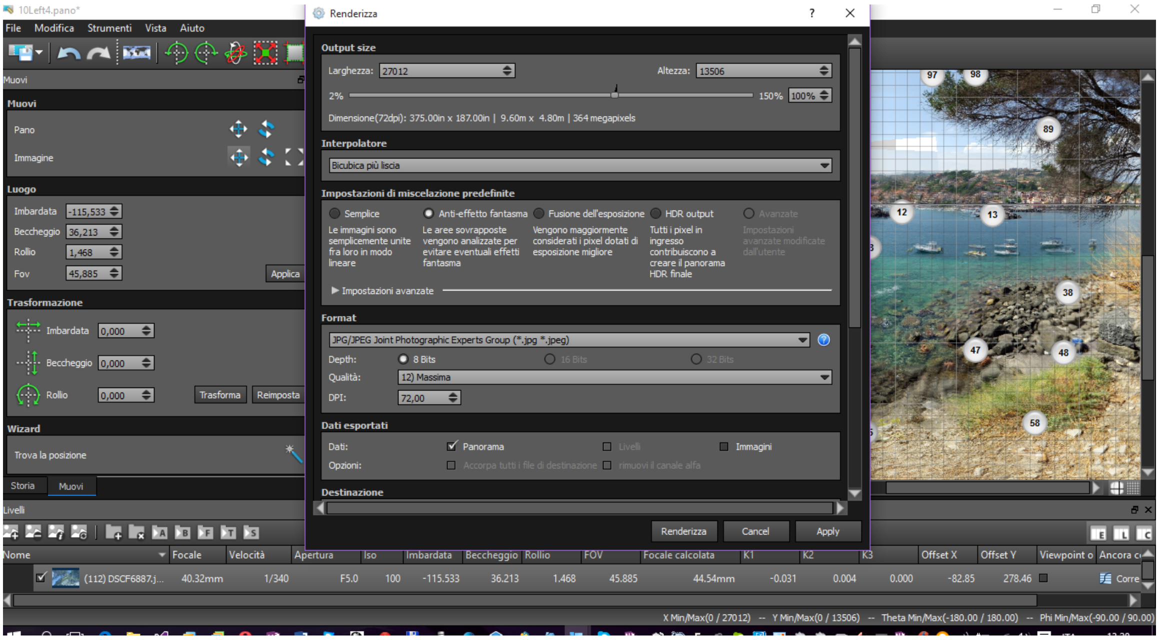


Figure 223 - Autopano settings to generate Lachea's panorama.

I made all the panoramas using the previous explained method to achieve the best results. Furthermore, I skipped the InPaint phase using GIMP to smooth the missing parts using the closest colours in the areas with the GIMP tools. I also discovered that there is a limit for the anti-ghost processing that does not depend on the number of total photos of the scene (I thought it was 250 as max), but on the number of **overlapping photos that make the anti-ghost too hard to complete processing with the low memory used by the program**. In particular, if the free memory is around 1.7 GB, the max number of overlapping images per line to apply the anti-ghost filter must be under 50 images.

Figure 224 shows an example of a rendered panorama in Microsoft ICE instead of Autopano, to show the different result.

Figure 224 *- Testing panorama stitching using Microsoft ICE for Lachea's island.*

As explained, no control on the light or exposure or uniform color is available for post-production, so a third-party software would be needed. The sky was autocompleted.

Figure 225 shows the same scene but rendered using Autopano. The sky was added in post-production, as you can see from the line in the centre that was left to let the reader understand where the black hole was.

Figure 225 *- Testing panorama stitching using Autopano for Lachea's island.*

This picture presents a much uniform color thanks to Autopano filters. This is another reason for choosing this software above Microsoft ICE.

Stitching problems and first temporary solutions

The main problem of the stitching is that it is based on the edges of the pictures and on the exposure and other information of each photo. So, using left eyes images and right eye images would produce consistent errors of vertical and horizontal parallax. The following figures show an example of left and right panorama with vertical parallax error. **Figure 226** is left eye, while **Figure 227** is right eye.

Figure 226 *- Example of stitching and vertical parallax error (left eye panorama of Lachea's island).*

Figure 227 *- Example of stitching and vertical parallax error (right eye panorama of Lachea's island).*

Figure 228 and **Figure 229** show the overlapping result of the two left and right eye images.

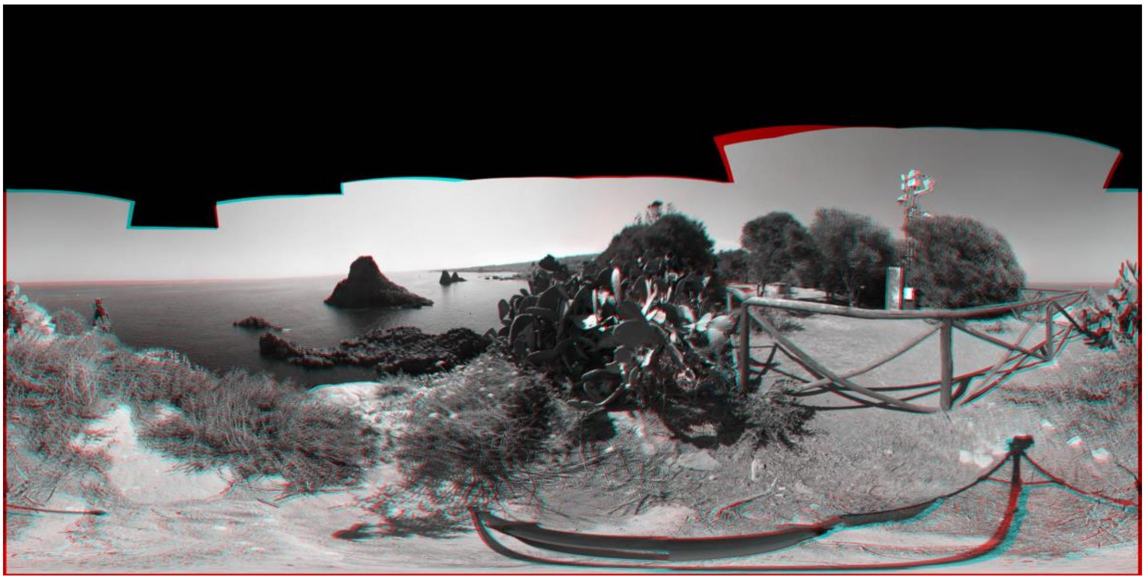


Figure 228 *- Vertical parallax error overlapping left and right eye (anaglyph view).*

Figure 229 *- Vertical parallax error overlapping left and right eye (stereo view).*

Figure 230 shows a detail of the vertical parallax effect.

Figure 230 *- Detail of the parallax effect (anaglyph view).*

In the picture, the right camera has a different vertical position from the left camera position. This problem was present in many details of the photo, making the 3D effect disturbing and not natural. Moreover, the horizontal parallax is not correct at all.

The trick of the double stitching with Autopano

To solve this problem, considering that the project in *Autopano* can be saved and that the name of a 3D picture has the same name of its corresponding left and right pictures in the "L" folder and "R" folder, the trick was to **save the stitching project for the left images, than close the program, rename the "L" folder to "LL" and the "R" folder to "L" and reload the project:** in this way, ***Autopano* will load the right images positioning them in the exact position of their corresponding left images**, without leaving any detail or particular of the scene. When the process is complete, the result is astonishing and free from almost any artificial parallax error.

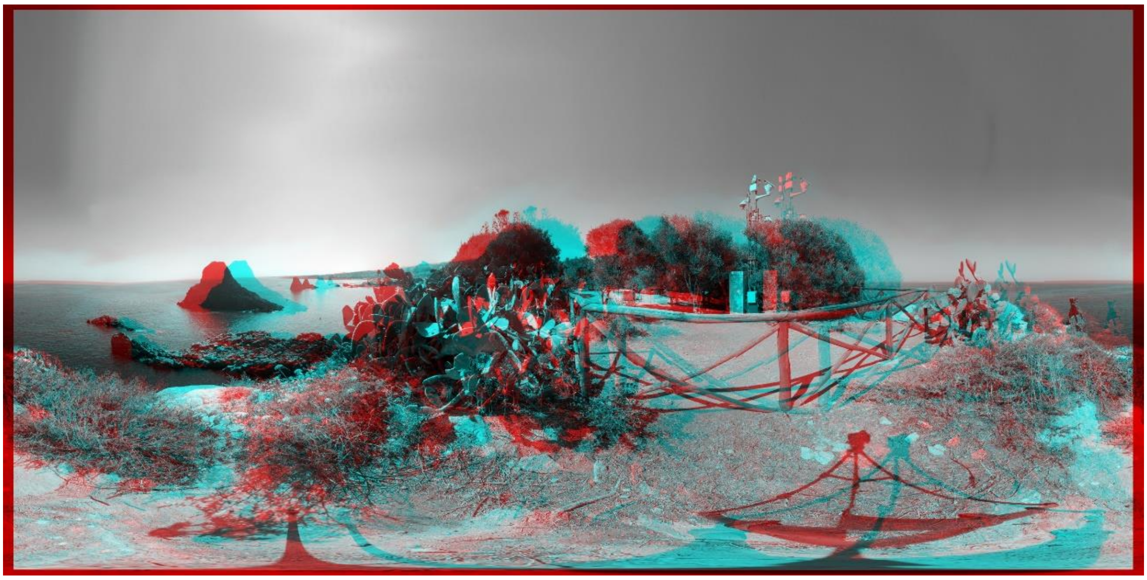


Figure 231 *- Corrected vertical parallax error and result of 3D panorama of Lachea (anaglyph view).*

Figure 232 *- Corrected vertical parallax error and result of 3D panorama of Lachea (stereo view).*

Figure 231 and **Figure 232** show the final result of the 3D panorama, with the correct natural horizontal parallax and no vertical parallax.

Figure 233 shows me having fun taking 400 photos of the Lachea island's first hotspot. This required a manual rotation of the camera, and a very long patience. The result was then tried in Virtual Reality and used for the proposed investigation's user studies.

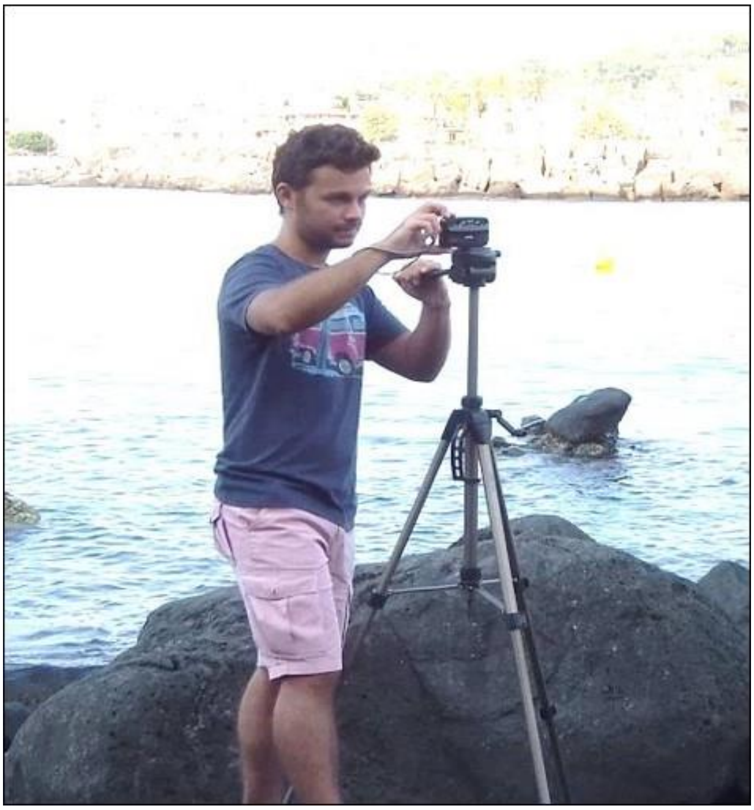


Figure 233 – Alessio Regalbuto, taking the first 3D shots of what will become the 3D panorama of Lachea Island, which was used for the proposed Mobile VR user studies of this thesis.