



HAL
open science

A framework to characterize and influence spatial perception in heterogeneous virtual reality displays

Jose Dorado

► **To cite this version:**

Jose Dorado. A framework to characterize and influence spatial perception in heterogeneous virtual reality displays. Signal and Image Processing. HESAM Université; Universidad de los Andes (Bogotá), 2022. English. NNT : 2022HESAE010 . tel-03622834

HAL Id: tel-03622834

<https://theses.hal.science/tel-03622834v1>

Submitted on 29 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ÉCOLE DOCTORALE SCIENCES ET MÉTIERS DE L'INGÉNIEUR
[UBFC, HESAM, Institut Image]

THESIS

presented by : **José DORADO**
to obtain the degree in : **Docteur d'HESAM Université**

developed at : **École Nationale Supérieure d'Arts et Métiers**
Speciality : **Computer science - Signal processing**

**A FRAMEWORK TO CHARACTERIZE AND
INFLUENCE SPATIAL PERCEPTION IN
HETEROGENEOUS VIRTUAL REALITY DISPLAYS**

THESIS directed by :
[PhDs MERIENNE Frédéric, HERNANDEZ Tiberio]

in co-supervision by :
[PhDs CHARDONNET Jean-Rémy, FIGUEROA Pablo]

Jury

Fernando DE LA ROSA	Full Professor, Faculty of Engineering, University of Los Andes	President
Frédéric LE-MOUEL	Full Professor, INRIA CITI Lab, University of Lyon / INSA Lyon	Reporter
Lola BAUTISTA	Full Professor, Faculty of Engineering, University of Santander	Reporter
Frédéric MERIENNE	Full Professor, Institut Image, Arts et Métiers, UBFC, HESAM	Examiner
Jean-Rémy CHARDONNET	Full Professor, Institut Image, Arts et Métiers, UBFC, HESAM	Examiner
Pablo FIGUEROA	Full Professor, Faculty of Engineering, University of Los Andes	Examiner
Tiberio HERNANDEZ	Full Professor, Faculty of Engineering, University of Los Andes	Examiner

**T
H
È
S
E**

Dedicated to my family for all their unconditional support and to my friends, who assisted me
throughout the process.

Acknowledgements

Thanks to my advisers, to the Imagine research group and the Colivri staff at Uniandes-Bogota, to the Institut Image at Chalon-sur-Saone-France, and my family for all their support.

ACKNOWLEDGEMENTS

Résumé

Nous proposons un cadre de référence pour caractériser et influencer la perception spatiale dans des écrans de réalité virtuelle (RV). Il est difficile, en réalité virtuelle, d'obtenir une perception spatiale précise en raison du phénomène de compression de la profondeur, qui donne lieu à une sous-estimation systématique des distances. La solution de ce problème est complexe car la technologie actuelle ne permet pas de simuler, dans des conditions raisonnables de temps et de complexité, la perfection de l'œil humain. Le problème est d'autant plus compliqué qu'il existe sur le marché une grande variété de systèmes d'affichage, ce qui rend difficile la conception d'expériences partagées nécessitant des performances spatiales précises. Il est vrai que différentes techniques ont été proposées pour améliorer la perception à distance, mais une solution mise en place pour un écran peut ne pas fonctionner pour un autre. En ce sens, nous proposons un cadre de référence pour caractériser et influencer la perception spatiale en réalité virtuelle. Ce cadre de référence prend en considération la variété des facteurs d'affichage qui affectent les performances pour construire un "prédicteur de performances linéaire", une fonction de régression linéaire qui décrit le phénomène de compression de la profondeur à l'écran. Ce prédicteur peut ensuite être utilisé pour appliquer une technique de projection alternative qui exerce une influence positive sur la perception spatiale du sujet en jouant avec les lois de la perspective. Par conséquent, le cadre de référence peut être utilisé pour améliorer la perception des distances en réalité virtuelle et la conception d'expériences partagées plus naturellement entre des écrans hétérogènes.

Mots-clés : perception spatiale, réalité virtuelle, perception de la distance, perspective linéaire

RESUME

Abstract

We propose a framework to characterize and influence spatial perception in heterogeneous virtual reality displays. Accurate spatial perception in VR is challenging due to the phenomenon of depth compression, where distances are systematically underestimated. Solving this issue is complex due to the impossibility with the current technology of simulating, under a reasonable amount of time and hardware complexity, the perfectness of the human eye. The issue becomes more challenging due to the explosion of system designs and displays available in the market, making difficult the design of shared experiences that requires an accurate spatial performance. Although some techniques have been proposed to improve the perception of distances, a solution implemented for a display is likely not suitable for others. In this sense, we propose a framework to characterizing and influencing spatial perception in VR. The framework takes into consideration the variety of display factors that cause differences in performance to build a "linear performance predictor", a linear regression function that describes the depth compression phenomenon on the display. Then, this predictor can be used to apply an alternative projection technique that influences positively subject's spatial perception by playing with the laws of perspective. Thus, the framework can be used to improve the perception of distances in VR and the design of more naturally shared experiences between heterogeneous displays.

Keywords : virtual reality, spatial perception, distance perception, linear perspective).

ABSTRACT

Contents

Acknowledgements	3
Résumé	5
Abstract	7
Liste des tableaux	15
Liste des figures	20
1 Introduction	21
1.1 Context	21
1.2 Problem statement	25
1.3 Research question	29
1.4 Research objectives	29
1.5 Scope	30
2 Literature review	31
2.1 Overview on spatial perception	31
2.1.1 Depth cues	31
2.1.2 Depth illusions and cue conflicts	33
2.1.3 Egocentric distance perception	34

CONTENTS

2.1.4	Measurement methods	35
2.1.5	Action and perception	36
2.2	Display issues	37
2.2.1	Vergence-accommodation mismatch	38
2.2.1.1	Zone of comfortable fusion	40
2.2.1.2	Depth compression and focal distance	42
2.2.1.3	Stereoscopic parallax condition	44
2.2.2	Incorrect peripheral stimulation	46
2.3	Display issues and influence space	50
2.4	Spatial perception in VR	52
2.4.1	Distance perception and VR displays	52
2.4.2	Influence of adaptation	54
2.4.3	Methods that improve spatial perception in VR	55
3	A framework to characterize and influence spatial perception in heterogeneous virtual reality displays	57
3.1	Overview	57
3.2	Preliminary studies	59
3.2.1	Size-Constancy test	59
3.2.2	Comparing Size-Constancy between HMD and CAVE	62
3.2.2.1	Hypotheses	63
3.2.2.2	Method	64
3.2.2.3	Results	65
3.2.2.4	Discussion	66
3.2.3	Blind-Walking Test for HMDs	68
3.2.4	Comparing Blind-Walking between HTC Vive vs Samsung Gear VR	69

CONTENTS

3.2.4.1	Hypotheses	69
3.2.4.2	Method	70
3.2.4.3	Discussion	72
3.3	Summary	73
4	Part I. Characterizing spatial performance	77
4.1	Determinant factors	77
4.1.1	Focal distance	77
4.1.2	Stereoscopic parallax condition	79
4.1.3	Limited FOV and lens distortions	81
4.1.4	Linear depth compression	83
4.1.5	Summary of determinant factors and assumptions	85
4.2	Second study	85
4.2.1	Size-constancy test for a HMD	85
4.2.1.1	Hypotheses	86
4.2.1.2	Method	86
4.2.1.3	Discussion	89
4.2.2	Blind-walking test	90
4.2.2.1	Hypotheses	91
4.2.2.2	Method	91
4.2.2.3	Discussion	93
4.3	Predicting spatial performance	95
4.3.1	Stimuli selection	95
4.3.2	Linear performance predictor	96
4.4	Summary	98
5	Part II. Influencing spatial perception by composition	101

CONTENTS

5.1	Overview	101
5.2	Transformation in perspective projections	102
5.2.1	Lowering the horizon	102
5.2.2	FOV Minification	104
5.2.3	Implementation	107
5.3	Validation of alternative projection techniques	108
5.3.1	Influence of lowering the horizon on size-constancy	109
5.3.1.1	Hypotheses	110
5.3.1.2	Method	110
5.3.1.3	Discussion	111
5.3.2	Influence of FOV minification on blind-walking	112
5.3.2.1	Hypotheses	112
5.3.2.2	Method	113
5.3.2.3	Discussion	115
5.4	Limitations	115
5.5	Improving spatial performance	116
5.5.1	Optimized lowering horizon technique	117
5.5.2	Optimized FOV minification technique	119
5.6	Summary	119
6	Conclusions	121
7	Un cadre de référence pour caractériser et influencer la perception spatiale sur des écrans hétérogènes de réalité virtuelle	125
7.1	Résumé	125
7.2	Introduction	126
7.2.1	Le contexte	126

CONTENTS

7.2.2	Énoncé du problème	128
7.2.3	Question de recherche	129
7.2.4	Objectifs de la recherche	129
7.3	Revue de la littérature	130
7.3.1	Vue d'ensemble de la perception spatiale	130
7.3.1.1	Signaux de la perception de la profondeur	130
7.3.1.2	Illusions de profondeur et conflits de signaux	130
7.3.1.3	Méthodes de mesure	131
7.3.1.4	Action and perception	132
7.3.2	Problèmes d'affichage	133
7.3.2.1	Conflit vergence-accomodation (conflit V-A)	133
7.3.2.2	Stimulation périphérique incorrecte (SPI)	136
7.3.3	Methods that improve spatial perception in VR	137
7.4	Un cadre de référence pour caractériser et influencer la perception spatiale sur des écrans hétérogènes de réalité virtuelle	138
7.4.1	Vue d'ensemble	138
7.4.2	Sélection des méthodes d'évaluation	139
7.4.2.1	Test de constance de la taille	140
7.4.2.2	Comparaison de la constance de la taille entre HMD et CAVE	141
7.4.2.3	Test de marche en aveugle	144
7.4.2.4	Comparando el HTC Vive vs Gear VR	145
7.4.3	Résumé	148
7.5	Partie I - Caractérisation de la compression de la profondeur dans les écrans RV	148
7.5.1	Vue d'ensemble	148
7.5.2	Facteurs déterminants	149
7.5.2.1	Distance focale	149

CONTENTS

7.5.2.2	Condition de parallaxe stéréoscopique	150
7.5.2.3	Stimulation périphérique incorrecte (SPI)	152
7.5.3	Prediction de la performance spatiale	153
7.5.3.1	Choix des stimuli	153
7.5.3.2	Prédicteur de performance linéaire	154
7.5.4	Résumé	156
7.6	Partie II. Influencer la perception spatiale par composition	157
7.6.1	Vue d'ensemble	157
7.6.2	Techniques alternatives de projection	159
7.6.2.1	Influence de l'abaissement de l'horizon sur la constance de la taille	159
7.6.2.2	Influence de la minification du champ de vision pour la marche en aveugle	161
7.6.3	Amélioration de la performance spatiale	163
7.6.3.1	Optimized lowering horizon technique	164
7.6.3.2	Technique de minification du champ de vision optimisée	166
7.6.4	Résumé	167
7.7	Conclusion	168
	Bibliographie	171

List of Tables

1.1	Report of display factors and perceptual issues in HMDs from Patterson et al. (2006)	22
2.1	Spatial perception in a variety of studies using different VR displays and measurement tasks.	53
2.2	Comparative studies about distance perception using different HMDs displays and blind walking tasks.	54

LIST OF TABLES

List of Figures

1.1	Trends in VR displays development.	24
1.2	Influence of VR in the perception-action-feedback loop.	26
1.3	Areas that influence spatial perception in VR.	27
1.4	Illusion of depth by composition	30
2.1	Power to discriminate depth of different cues according to the distance.	32
2.2	Ames Room illusion	34
2.3	Methods to assess the perception of distances	36
2.4	The physiological cues to depth perception.	39
2.5	The vergence accommodation mismatch	40
2.6	Zone of comfortable fusion and depth of field	41
2.7	Functional depth range and depth compression	43
2.8	Functional depth range	44
2.9	Conflict size as a function of the focal distance	47
2.10	Lens distortions in HMDs and angular resolution	48
2.11	Motion pursuit law	49
2.12	Influence of the HTC Vive lenses on motion parallax	50
2.13	Display issues and influence space.	51
3.1	The proposed framework as an extension of the action-perception loop in VR.	58

LIST OF FIGURES

3.2	Candidate assessment methods and influence space	60
3.3	Size-constancy test	61
3.4	Adapted size-constancy test	63
3.5	Results of the adapted size-constancy test (CAVE vs HMD)	66
3.6	Blind-walking-tests using the Virtuix Omni treadmill	70
3.7	Results of the blind-walking tests using the Virtuix Omni treadmill	72
4.1	The main determinant factors associated with the display issues	78
4.2	DOF of a modern HMD compared with the human natural DOF	80
4.3	A comparison of the V-A conflict size between a HMD and the CAVE	81
4.4	Influence of the Oculus Rift and HTC Vive lenses on motion parallax	82
4.5	Depth from motion parallax as a function of the lens	84
4.6	Improved size-constancy test for HMDs.	87
4.7	Result of the adapted size-constancy-tests for the HMD	89
4.8	Blind walking test	92
4.9	Results of the blind walking test with adaptation	94
4.10	The selection of the stimulus as a function of the focal distance f	96
4.11	Stimuli selection according to the display issue.	97
4.12	Subject performance as a function of the V-A conflict size.	98
4.13	Linear performance predictor	99
5.1	Alternative projection techniques	102
5.2	Angular declination hypothesis	103
5.3	The oblique perspective projection	104
5.4	The influence of FOV minification in the perception of distance	105
5.5	The influence of a non-linear projection technique in the perception of distance	106
5.6	Implementation of the "oblique perspective view"-space technique	109

LIST OF FIGURES

5.7	Influence of the lowering horizon technique on the perception of size constancy	111
5.8	Implementation of the "minificated perspective view"-space technique	113
5.9	Influence of the FOV minification technique on the perception of distance	114
5.10	Distortion induced by applying an oblique perspective projection	116
5.11	Incorrect computation of shadow	117
5.12	Effect of reducing the horizon line as a function of the focal distance.	118
5.13	Effect of scaling the FOV as a function of the variation on the subtended visual angle.	120
7.1	Tendances dans la création d'écrans pour la réalité virtuelle..	127
7.2	Zones qui influent sur la perception spatiale en réalité virtuelle.	128
7.3	Capacité à discriminer la profondeur de différents signaux en fonction de la distance.	130
7.4	Illusion de la chambre d'Ames	131
7.5	Méthodes pour évaluer la perception des distances	132
7.6	Le conflit entre accommodation et vergence	134
7.7	Zone de fusion confortable et profondeur de champ	135
7.8	Plage de profondeur fonctionnelle et compression de profondeur.	136
7.9	Taille du conflit en fonction de la distance focale	137
7.10	Le cadre de référence proposé vu comme une extension du cycle action-perception en réalité virtuelle.	139
7.11	Test de constance de la taille	141
7.12	Adapted size-constancy test	142
7.13	Résultats du test de constance de la taille adapté (CAVE vs HMD)	144
7.14	Tests de marche en aveugle avec le tapis roulant Virtuix Omni	146
7.15	Résultats des test de marche en aveugle avec le tapis roulant Virtuix Omni.	147
7.16	Les principaux facteurs déterminants associés aux problèmes de visualisation	149
7.17	PDC d'un HMD moderne par rapport à la PDC naturelle humaine.	151

LIST OF FIGURES

7.18	Une comparaison de la taille du conflit V-A entre un HMD et le CAVE	152
7.19	Compression de la profondeur en fonction de la résolution angulaire induite par la lentille	153
7.20	Le choix du stimulus en fonction de la distance focale f	154
7.21	Le choix des stimuli selon le problème d'affichage.	155
7.22	Performance du sujet en fonction de la taille du conflit V-A.	156
7.23	Prédicteur de performance linéaire	157
7.24	Techniques alternatives de projection	158
7.25	Mise en oeuvre de la technique de perspective de vision oblique (PPO)	159
7.26	Influence de la technique OPP/PPO sur la perception de la constance de la taille . . .	160
7.27	Test de marche en aveugle	162
7.28	Mise en oeuvre de la technique de projection de perspective minifiée (PPM)	163
7.29	Influence de la technique de minification du champ de vision sur la perception des distances	164
7.30	Effet de la réduction de la ligne d'horizon en fonction de la distance focale.	165
7.31	- Effet de la mise à échelle du champ de vision (FOV) en fonction de la variation de l'angle visuel sous-tendu.	167

Chapter 1

Introduction

1.1 Context

In the decade of the 90's, the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) initiated a series of studies on the ability to accurately estimate distances in virtual environments as an essential component of navigating large-scale spaces. Wright (1995) tested helicopter pilots in a virtual environment (VE) finding that distance estimates varied from 41 to 72 percent of the actual distances. These underestimations were obtained despite using one of the most advanced displays of the time, a "helmet-mounted" display with a horizontal field of view (FOV) of 125 degrees and a resolution of 1.5 arc minutes. Other researchers at the institute noticed that the same underestimation occurred on stationary observers on the ground, an underestimation that affected the performance of several spatial tasks (Witmer and Kline (1998), Lampton et al. (1995), Kline and Witmer (1996)). They requested subjects to perform judgments of distances in a variety of VR setups, which had different FOVs, texture resolutions, graphic qualities, and movement techniques, reporting a systematic underestimation of distances. As a result, they found that introducing more realism or more natural methods of movement through the VE do not eliminate the underestimation effects. Researchers at ARI concluded that the interaction of multiple technical factors and perceptual issues could be the cause of the underestimation effects. They did not realize that making spatial perception veridical it would continue to be one of the most challenging problems in VR.

Patterson et al. (2006), a group of researchers at ARI, developed a comprehensive review of different display factors and the perceptual issues that they induce (Table 1.1), where some of the reported technical limitations and perceptual issues, continue as open research areas. Following Patterson et al.

1.1. CONTEXT

Display factor	Perceptual issue
Restricted field of view	Optic flow behavior affects peripheral vision analysis
Poor visual acuity	Incorrect distance scaling, non perceptual constancy
Accommodation-vergence mismatch	Changes in pupil size affects depth of field
Mismatched binocular input	Erroneous depth perception due to interocular differences
System latency	Imprecise visual suppression makes visual perception unstable

Table 1.1: Report of display factors and perceptual issues in HMDs from Patterson et al. (2006)

(2006), other researchers interested on on a better understanding of the phenomenon tried to isolate the contribution of the different display factors, finding questions instead of answers. Neither the limited FOV (Knapp and Loomis (2004), Creem-Regehr et al. (2005), Willemsen et al. (2009)) nor the stereo viewing conditions (Willemsen et al. (2008), Creem-Regehr et al. (2005)) nor the lack of realistic graphics (Thompson et al. (2004), Kunz et al. (2009)) nor the ergonomics (Willemsen et al. (2004), Willemsen et al. (2009)) contribute significantly to the underestimation effects. Thus, these researchers suggested that the causes could be associated with a complex interaction between display factors and perceptual issues, particularly with the way the visual system assimilates the visual geometry.

As the technology improved, especially with the renewed interest in VR in the last decade, the phenomenon of underestimation of distances has been reduced. This amelioration could be a consequence of the fact that many of the technical limitations reported by Patterson et al. (2006) have been addressed and improved in the last years, thanks to the advances in OLED display technology, computer graphics, and optics. For example, modern consumer-oriented HMDs provide a greater FOV, visual acuity, and less latency, compared with second-generation HMDs. While distances were underestimated between a 40% and 70% of the actual distances using the previous generation of HMDs (see Renner et al. (2013), El Jamiy and Marsh (2019) for a complete review), the underestimation today is between 80% and 90% using current generation HMDs. Despite this improvement, we are still far from having a perfect solution for most of issues reported by Patterson et al. (2006).

Compared with the human eye, HMDs are far to be perfect systems. Popular devices like the Oculus Rift and the HTC Vive only provide a 130° horizontal FOV, which is smaller than the human FOV (approx 200°). Also, display resolution in the central vision is poor. The maximum visual acuity is estimated somewhere around 20/60, very far from 20/20 or normal vision (Kreylos (2017)).

While there are promising projects such as the Pimax project, the first HMD that promises to bring a wide FOV of 200° horizontal, 120° vertical, and 8K resolution (550 PPI, 4K per eye) (Apress (2019)), other limitations reported by Patterson et al. (2006), like the vergence-accommodation mismatch are still very challenging issues. Also, the use of lenses in HMDs induces other undesired visual artifacts, such as screen door effects, mura, and chromatic aberrations, which distort the visual stimuli. These limitations affect the sensory fidelity of HMDs.

Aware of these limitations, a group of researchers developed a different kind of display usually denoted as large immersive projection-based displays (LIPDs). By using very large screens and active/passive stereo glasses, these systems are less susceptible to the optical distortions and other visual artifacts typically found in HMDs. The precursor of this kind of system is the CAVE (Cruz-Neira et al. (1993)), a cubic shape room designed as an approximation of a sphere, where the side walls are made up typically of rear-projection screens, whereas, the floor and ceiling use front- and rear projection screens respectively. Other systems use different projection approaches, like spherical domes Jo et al. (2006) or cylindrical-shape displays Doi et al. (2000), but they follow the same principles of the CAVE. By using these large projection screens, LIPDs have less limited ranges than HMDs, potentially providing an almost natural FOV, higher visual acuity, and less influence of optical distortions. Since its conception, the CAVE was designed with two objectives in mind: (1) to overcome the limitations of HMDs, and (2) to foster collaboration, making multiple people interact naturally, visualizing the same imagery in the same space. Despite all these advantages, LIPDs are very expensive and hardly portable system, which make them more suitable for industrial applications. Figure 1.1 presents two examples of the main trends in display development for VR.

In terms of sensory fidelity, LIPDs like the CAVE were always a step further than HMDs because they were only limited by the available space and costs. To obtain a natural FOV, dynamic range, and visual acuity is only necessary to add more projectors and/or increasing the room size, using multi-tiled-projection-wall techniques (Li et al. (2004)). An evolution of the CAVE, called CAVE2, replaced the projectors with an array of 72 LCD stereo screens inside a cylindrical room. This system required three times the space of the original CAVE to provide a resolution 13 times greater, a huge dynamic range, and a maximum visual acuity of 20/20 (Febretti et al. (2013)). To our knowledge, the most advanced CAVE in the world is the C6, located at the Iowa State University's Virtual Reality Applications Center. C6 is a system composed of twenty-four projectors powered by 48 computers,



Figure 1.1: Trends in VR displays development.

Left: Head Mounted Display (HMD). Right: Large immersive projection display (LIPD) known as the CAVE.

giving a resolution greater than 100 million pixels on a 6-sided room of $3.6m^2$. The system is used for very complex tasks, such as simulating the international space station (IowaC6 (2014)). This sensory fidelity seems too much, but tasks that demand very precise dexterity, such as the maintenance of the international space station, require an accurate spatial perception. Indeed, there are very few studies about spatial perception in LIPDs, but the evidence has shown that perception of distances tends to be better on this kind of system (Plumert et al. (2005), Naceri et al. (2010), Murgia et al. (2009), Marsh et al. (2014), Kenyon et al. (2008), Luo et al. (2007)). Besides a better sensory fidelity, LIPDs like the CAVE have another fundamental advantage compared with HMDs: users can visualize the physical and virtual environment simultaneously (including their own body), allowing them to use references from the physical world. Although the enthusiasm for LIPDs like the CAVE has decreased in the last decades due to their elevated costs and the recent advances in HMDs, projects like the Immersis display, a projector that can literally "beam" a virtual environment inside any room of any shape (Wired (2015)), could potentially renovate the interest in this kind of technology.

A good sensory fidelity is important to interact and collaborate more naturally. VR allows the design of shared experiences where people can interact remotely using different VR displays. For example, VR is used nowadays in interactive design scenarios where an engineer and an operator collaborate remotely, using a CAVE and an HMD respectively, to evaluate some aspects of a prototype, such as the accessibility and visibility of different components. Then, imagine a scenario where the

1.2. PROBLEM STATEMENT

accessibility of a component was designed using the length/size of the operator's body as a reference. The different spatial perceptions between both displays could cause that the criterium of accessibility of the component differs between the engineer and the operator. The operator could adapt his/her perception to this situation, but this would affect their performance when using the prototype in the physical world, and in the worst scenario, risking his/her life. Aware of these differences, a developer could implement some clever solutions, such as scaling the virtual prototype to improve the spatial performance; however, these solutions would be optimized only for a particular user/display. Another developer could argue that is a matter of adaptation, users can adapt their spatial perception to the VE; however, adaptation is problematic. Studies have shown that using VR during prolonged periods, can alter our spatial perception and the performance of different cognitive tasks in the physical world (Szpak et al. (2019), Wright (2014), Mittelstaedt et al. (2019)). Hence, VR is still far from being practical for several human tasks in terms of spatial performance, difficulting the design of shared experiences.

In this section, we saw a historic review of the phenomenon of underestimation of distance in VR, its evolution across the different trends in VR display development and how it difficulties the design of shared experiences, where people can interact and collaborate more naturally. HMDs are more susceptible to induce a greater error due to their limitations regarding the restricted FOV, visual acuity, and optical distortions. On the other hand, LIPDs like the CAVE are less susceptible to these distortions and the evidence has shown that distance perception tends to be more accurate in this kind of system. The selection of a system requires a trade-off between sensory fidelity and portability. HMDs are very portable systems but they have an important cost in sensory fidelity. In contrast, LIPDs provide higher sensory fidelity but can be used only in a fixed space. Despite that both systems are based on the same stereoscopic principle, they are very heterogeneous, not only in terms of sensory fidelity but also, as we will see further, in the way we study and assess spatial perception.

1.2 Problem statement

Spatial perception in VR differs from the physical world. Reports describe the sensation of depth compression, where distances are generally underestimated (see Renner et al. (2013), El Jamiy and Marsh (2019) for a complete review). Since the first reports of the phenomenon were published almost

1.2. PROBLEM STATEMENT

30 years ago, this issue has become a puzzling and challenging problem. Bühlhoff and van Veen (2001) developed a pertinent description of this challenge. They describe VR as an encapsulated environment that intercepts the sensory-motor feedback loop (Figure 1.2). An artificial environment is created inside the physical environment by using human-computer interfaces, where the devices interfacing the organism inherently suffer from distortions and limited ranges. The limitations of the technology influence not only the way we perceive the environment, but also our actions and responses to different stimuli, which in turn impacts the performance of several spatial tasks. The problem becomes more complex due to the broad variety of displays available in the market today. Thus, currently there are not formal approaches for creating shared experiences, where people can interact and collaborate naturally inside VEs using a variety of displays. In order to achieve this ideal, we need to overcome several obstacles related to how we study and influence spatial perception in VR.

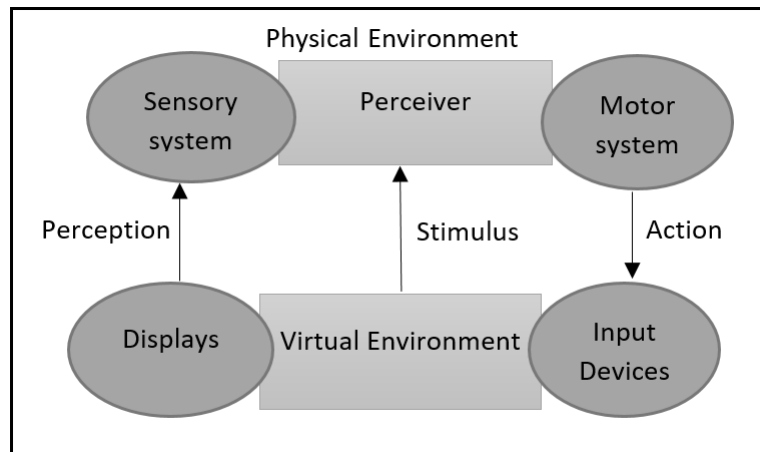


Figure 1.2: Influence of VR in the perception-action-feedback loop.

VR is an encapsulated environment that intercepts the sensory-motor feedback loop. An artificial environment is created inside the physical environment by using human-computer interfaces, where the devices interfacing the organism inherently suffer of distortions and limited ranges. Adapted from Bühlhoff and van Veen (2001).

The first obstacle is that studying spatial perception in VR is not straightforward. Renner et al. (2013) categorize the phenomenology in three areas: *technical* (display technology and their limitations), *compositional* (techniques that creates the illusion of depth, e.g. stereoscopy, linear perspective and light simulation), and *human* (calibration and adaptation) (see Figure 1.3). The *measurement method* is another important element to consider, because some assessments are not suitable for all

1.2. PROBLEM STATEMENT

kinds of displays. Thus, studying spatial perception in VR has become technology-specific, where methods typically used for HMDs are not suitable for LIPDs. Despite the fact that several techniques have been proposed to improve the perception of distances in VR, when a developer designs an application that requires accurate spatial perception for HMDs, the solutions are not suitable for other kinds of systems. To our knowledge, there is a lack of approaches to study and influence spatial perception in heterogeneous displays, which takes into consideration the different technical factors that induce differences in performance. In this sense, we propose a framework for studying and influencing spatial perception in heterogeneous VR displays that allows the design of shared experiences.

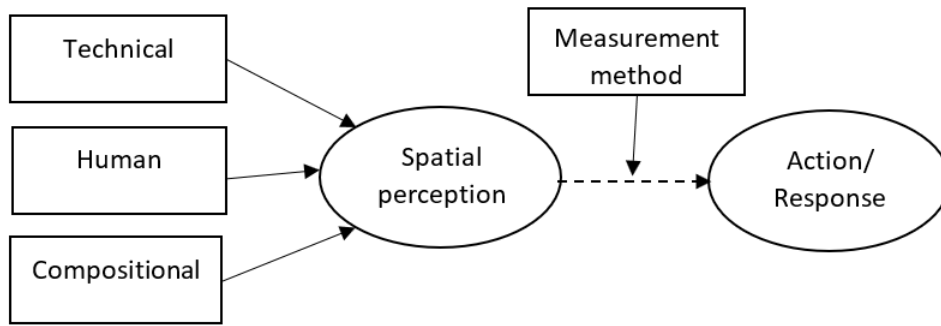


Figure 1.3: Areas that influence spatial perception in VR.

Adapted from Renner et al. (2013).

The second obstacle is that *spatial perception differs between heterogeneous VR displays*. Modern VR displays based on traditional stereoscopy suffer from the inevitable issue that depth cannot be inferred physically based on the light paths of the objects and surfaces, but it must be inferred artificially based on two projections for each eye on a flat-screen (Hoffman et al. (2008), Jones et al. (2001)). This limitation induces natural problems on depth perception, such as the vergence-accommodation mismatch, an issue that has received important attention in the last decade. Other limitations are the restricted FOV and the nature of light stimulation and resolution induced on the peripheral vision. Thus, some displays induce more distortions in spatial perception than others. For example, the phenomenon of depth compression seems to be stronger in HMDs compared with LIPDs, which indicates that depth cues are more distorted in these kinds of displays. This could be a direct consequence of how the CAVE was originally devised to overcome the limitations of HMDs and stereoscopic widescreens, providing an almost natural field of view (FOV), higher visual acuity, and less influence of optical distortions (Cruz-Neira et al. (1993)). These distortions not only differ between hetero-

1.2. PROBLEM STATEMENT

geneous displays, even inside HMDs could exist divergences in spatial perception depending on the vendor. Although the differences in immersion are an objective feature of the displays, few studies compare spatial perception using different kinds of systems. We consider these comparative studies fundamental to understand the contribution of different display factors in the phenomenon and the design of shared experiences.

A third obstacle is the *lack of transversal assessment methods*. Several methods have been proposed to assess spatial perception in VR. Most of the previous work have focused on the perception of egocentric distances in HMDs. The most popular method requires subjects to estimate the distance of a previously visualized target by walking under non-visual conditions (usually denoted as blind-walking). However, these assessment methods are not suitable for LIPDs due to their spatial restrictions. Thus, several alternative methods based on percept-couplings have been proposed. For example, methods that assess the perception of distances indirectly using the perception of size, or combining the perception of distances with the perception of directions. Although LIPDs' spatial restriction seems a limitation, we consider that some of the methods proposed for these kinds of displays are more informative than classic blind-methods, because most of the human spatial tasks are based on percept-couplings. For example, the coupling distance-size for grasping objects or the coupling distance-direction for pointing, two fundamental tasks whose performance is influenced negatively by the phenomenon of depth compression (Barrera Machuca and Stuerzlinger (2019), Chessa et al. (2019)).

An important and sometimes forgotten obstacle is the human factor. It consists of fitting the display to the observer's visual system (*calibration*) and studying how the visual system "fits" to the technical limitations of the display (*adaptation*). Several methods have been proposed to calibrate VR displays, including manual, semi-automatic and automatic methods (see Grubert et al. (2018) for a complete review), in which previous work has been done mostly in the field of augmented reality. Calibration of non-see-through VR displays is not straightforward and there is some evidence that even when the display is calibrated, underestimation effects are reduced but not substantially (Kellner et al. (2012), Swan et al. (2007), Livingston et al. (2009)). Based on this evidence, we would rather focus on adaptation than calibration. Researchers know a century ago that human beings can recalibrate their motor actions to unnatural/distorted visual stimuli so that they can adjust their spatial perception. For example, in the famous inverted glasses experiment (Sachse et al. (2017)) or the funky goggles experiment (Channel (2013)), where perspective is shifted. VR is another kind of unusual visual

1.3. RESEARCH QUESTION

stimulus, in which studies have shown that adaptation occurs rapidly. Unfortunately, there is no agreement between researchers if adaptation is a solution to the depth compression phenomenon in VR.

A last obstacle is that *researchers have focused mostly on solving the technical factors rather than taking advantage of the compositional factors*. Although light field displays (Huang et al. (2015)), focal surface displays (Matsuda et al. (2017)) and eye-tracking techniques (Fu et al. (2016)) are promising technologies that can solve partially the depth compression issue, researchers are still working to reduce their computational demand and hardware complexity. In contrast, modern computer graphics are based on the techniques of perspective that were discovered during the Italian Renaissance, involving several techniques that create the illusion of depth on a two-dimensional surface. Linear perspective is one of these methods and forms the basis of the pin-hole camera model. Artists and architects have played with the rules of perspective for centuries to create illusions of depth that do not follow the rules of the physical world. For example in Holbein's *The Ambassadors painting*, which uses an oblique perspective to create the illusion of a skull leaving the picture; and the *Santa Maria presso San Satiro* church in Milan, which use a "forced perspective" to create the illusion of a choir with a depth greater than its actual dimensions. The power of composition on spatial perception is so strong that can override the influence of other cues, such as accommodation and vergence, and can constitute a more simple and effective solution to the problem.

1.3 Research question

Can we improve spatial perception in VR, taking into consideration the different display factors that influence the performance, so we can design more natural shared experiences between heterogeneous displays?.

1.4 Research objectives

- To study and characterize the influence of the different display factors on spatial perception.
- To find/design transversal assessment methods that are suitable to compare spatial perception between heterogeneous displays.



Figure 1.4: Illusion of depth by composition

Top: Holbein's The Ambassadors painting: a oblique perspective projection creates the illusion of a skulk leaving the picture (National (2016)). Bottom: Santa Maria presso San Satiro church in Milan: a "forced perspective" creates the illusion of a choir with a depth greater than its actual dimensions (National (2018)).

- To analyse the influence of adaptation on spatial perception.
- To determine how composition can be used to improve spatial perception.

1.5 Scope

Since studying spatial perception in VR is complex, this research is focused mostly on the perception of egocentric distances, which is the most fundamental and studied percept. Although the framework was motivated to compare heterogeneous displays, we focused mostly on HMDs because these kinds of displays are the most problematic. However, the findings of the framework can be applied perfectly to LIPDs. As a framework, it consists of a particular set of rules, ideas, or beliefs used in order to deal with problems or to decide what to do. Thus, it is not a perfect solution and it could be considered as a first approach to deal with the phenomenon of underestimation of distances in VR. Our idea is to continue improving the framework, as we get more insights into the phenomenon.

Chapter 2

Literature review

2.1 Overview on spatial perception

2.1.1 Depth cues

Cutting and Vishton (1995) developed a model about the sensitivity of different depth cues according to the distance between the perceiver and a stimulus (see Figure 2.1). They introduced the notion that humans divide their spatial environment into three regions: (1) Personal Space, focused on the manipulation of objects and fine motor tasks; (2) Action Space, focused on tasks associated with self-motion; and (3) Vista Space, focused on tasks associated with navigation, such as orientation and path planning. The power to discriminate depth of each cue varies according to the distance so that some cues become more prominent for some spaces than others. As a general rule, the cues that provide absolute information about depth (i.e. binocular disparity and vergence-accommodation) tend to be more functional at shortest distances than cues that provide more relative information (e.g., linear perspective or motion parallax). Other cues like occlusion and relative size are highly reliable and they work equally well at near and far distances. However, they are highly dependent on additional information (e.g., objects of familiar size close to the target object). Thus, Cutting and Vishton (1995) empathize that the nature of the visual system is adaptive and dynamic, exploiting the cues that are most functional in the region/space of interest and integrating the information from other spaces when they are available. Cutting and Vishton (1995)'s spatial classification of depth cues is an important element of the proposed framework and it will be referenced recurrently in this document.

Several models have been proposed about how the visual system integrates all the depth cues to create the sensation that we denominate distance perception, where the information provided by

2.1. OVERVIEW ON SPATIAL PERCEPTION

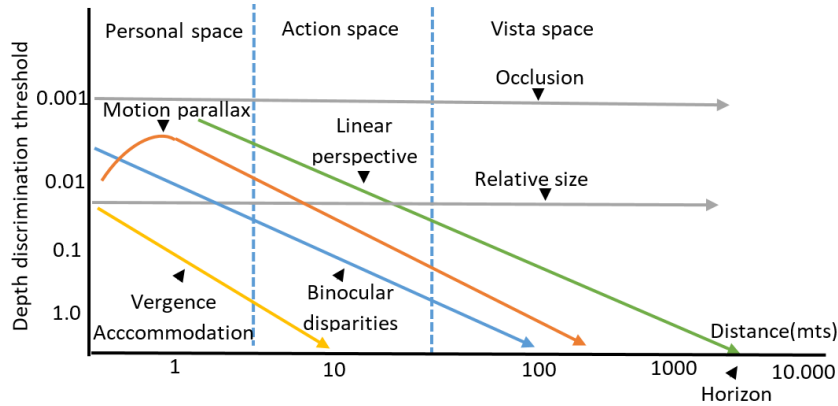


Figure 2.1: Power to discriminate depth of different cues according to the distance.

Adapted from Cutting and Vishton (1995)

different depth cues is integrated to yield a single depth estimate for each region of the visual stimulus. The dominant view, inspired by works in computer vision, is the *Bayesian Integration model* (Jacobs (1999), Ghahramani et al. (1997), Knill and Saunders (2003)). The model states that perceived depth corresponds to the weighted sum of the depth values signaled by different cues. Thus, the relative importance of different cues is governed by their weights, where the most reliable cues are weighted with the higher values. There are two mathematical models derived from this view. The *weak fusion model* (Landy et al. (1995)) assumes that the processing of a given cue is independent of the processing of the other cues. In contrast, the *strong fusion model* states that the processing of all cues is not independent, but based on a feedback loop with a recurrent behavior. In this research, we consider the weak fusion model as the most plausible.

According to Ghahramani et al. (1997), the weak fusion model assumes that each cue depth estimate is independent with a noise associated, and all we know about them is its uncertainty or variance. Thus, to estimate the overall depth, the visual system combine them linearly to minimize the variance by applying a simple Bayes' rule for producing the minimum-variance estimate (MVE):

$$D = \sum_i w_i d_i \quad (2.1)$$

where d_i is the relative depth value from cue i , $w_i = -\sigma_i^{-2} / \sum_j \sigma_j^{-2}$, the inverse of its variance normalized, and D is the overall estimated depth. This equation simply means that the visual system

2.1. OVERVIEW ON SPATIAL PERCEPTION

linearly combines the depth cues, weighted by their inverse variances. Because the weights are proportional to the normalized inverse variances, more weight is assigned to less variable (more reliable) cues. Indeed, the model is in agreement with Cutting and Vishton (1995) depth cues classification, where oculomotor cues tend to be less reliable than pictorial cues, due to the fact that their depth estimates can vary significantly over time. This could be a consequence of the saccadic movements, an oculomotor mechanism in which the eyes are constantly scanning the environment (Ibbotson and Krekelberg (2011)). In contrast, the pictorial cues are as stable as the environment itself, so they become more reliable over time. However, we should never ignore the importance of any cue regardless of the kind of spatial task involved. The visual system exploits the cues that are most reliable in the region/space of interest and integrates the cues from other spaces, when they are available. Thus, every cue is relevant and contributes to the perception of depth regardless of whether the task correspond to the personal- or action space. As consequence, any cue whose depth estimates deviate from the natural values, will induce a bias on the overall perception of depth and distance. We will refer to the weak fusion model as an important assumption of this research.

2.1.2 Depth illusions and cue conflicts

Although the Bayesian integration model can theoretically predict the perception of distances, the integration of depth cues in the human visual system is prone to many biases. An example of this phenomenon, is the famous Ames room illusion, an optical illusion where people apparently can change size (Figure 2.2). Two conditions cause this illusion: a restricted viewing condition, where when looking through an advantageous viewing hole with one eye, binocular cues are removed and the room is perceived as normal; and the particular shape of the room, which induces a cue conflict, between the oculomotor cues and the pictorial cues, where the last succeed (Dorward and Day (1997), King et al. (1976)). Thus, Ames' room is the most generally accepted evidence of the theories of indirect perception and the importance of experience in the perception phenomena. Some studies suggest that the illusion is a consequence of a break-in shape-constancy, a bias in human perception towards symmetry and regular shape patterns, such as rectangular rooms (Dorward and Day (1997), King et al. (1976)). Indeed, people from other cultures who have not been brought up in rectangular environments do not have difficulty in perceiving the actual shape of the room (Seckel, 2006).

This phenomenon can be explained by Cutting and Vishton (1995) model. Because vergence and

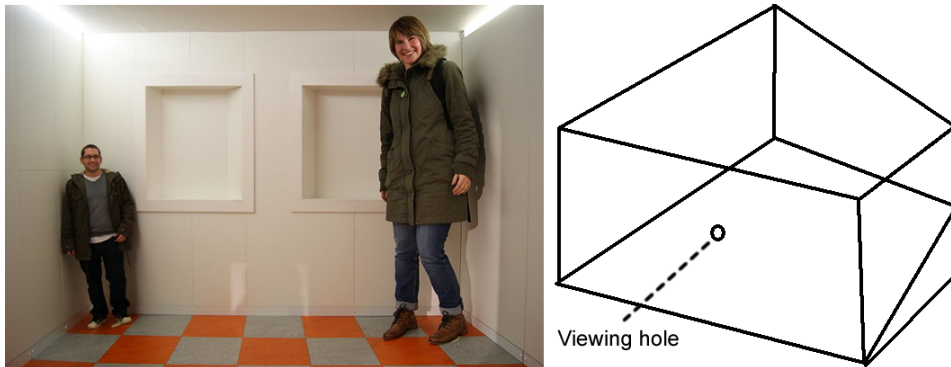


Figure 2.2: Ames Room illusion

Left: Ames room illusion, where people apparently change their size. Right: The actual shape of the room. Photo by Ian Stannard/flickr (Stannard (2010)).

accommodation are not strong enough cues for depth perception, they cannot break the stronger perspective/pictorial cues that are imprinted in our minds by our experience of habiting buildings (Gehring and Engel (1986), Glennerster et al. (2006)). Similarly, the phenomenon has a perfectly explanation from the Bayesian integration model. Since the weight of the pictorial cues is much greater than the weight of vergence-accommodation in the overall perception of depth, a bias in the perception of depth is induced. Indeed, the way in which perspective influences our perception of depth is fundamental to create the illusion of depth in VR. The vanishing points and the horizon line are cues that do not have a physical equivalent, but they are imprinted in our minds since we were born and they are so strong, that can make us perceive depth from an image generated in a flat surface. In this research, we took advantage of the particular bias related to perspective to influence the perception of distances in VR.

2.1.3 Egocentric distance perception

Although depth and distance perception have been used indistinctly in the literature, distance perception refers to a more complex cognitive process that involves not only picking information from the visual stimulus; but also based on the sensorimotor acquisition of information from the body or from the environment during locomotion (Montello (1997), Gärling and Golledge (1989)). The question "how can we perceive the distance of an object given two retinal images?", is one of the oldest questions in cognition research. The literature on this topic is voluminous and dense, with a

2.1. OVERVIEW ON SPATIAL PERCEPTION

diversity of theoretical approaches and empirical findings, that make it impossible to give a precise answer. Even at a functional level, we are still far from fully understanding some fundamental issues, such as the mapping between physical and visual space, the connection between perceived distance and action, and the extent to which the signals in the visual stimulus determine the sensation that we denominate distance. To make things more complicated, distance perception can come with different flavors. First, distance can be perceived using different sensory modalities, including visual, acoustic, haptic and proprioceptive (Loomis et al. (2013), Kolarik et al. (2016), Lederman and Klatzky (2009), Angelaki et al. (2009)). Second, human beings can use different frames of reference (egocentric or allocentric) (Klatzky (1998), Goodale and Haffenden (1998)), and spatial encodings (viewer-centered, object-centered or landmark-centered) (Galati et al. (2010), Sun and Wang (2010)), where the evidence has shown that different areas of the brain can be involved. Finally, action can influence the judgment of distances, so that the observer's perception can be biased depending on the kind of task performed by the observer (Proffitt (2006), Witt et al. (2016)). These characteristics show the complexity of studying this topic; for this reason, the proposed framework does not pretend to be exhaustive and it is more a frame of reference for approaching the perception of distances in VR. In this sense, we will focus mostly on visual perception of egocentric distances, and we will refer to proprioception as an alternative method to estimate distances in VR when vision is not reliable.

2.1.4 Measurement methods

Most work in spatial perception has been focused on the perception of egocentric distances. Several methods have been proposed to assess how people perceive distances. (1) verbal estimates, the most straightforward method but also the less accurate (Andre and Rogers (2006)); (2) perceptual matching, where subjects are asked to reproduce a virtual depth span based on a previously seen physical target (Sinai et al. (1999), Witt et al. (2007)), or where distance perception is assessed indirectly using percept couplings (e.g. taking advantage of its relationship with the perception of size) (Viguier et al. (2001), Haber and Levin (2001)); (3) visually directed action (also known as blind-methods), where subjects are asked to estimate a distance performing an equivalent action physically and usually blinded. From this category, we can highlight blind-walking, where subjects must walk to a previously seen target under non-visual-conditions (see Loomis et al. (2003) for a review), and blind-triangulation, where subjects must estimate the distance of a previously seen target by walking some steps in a direction

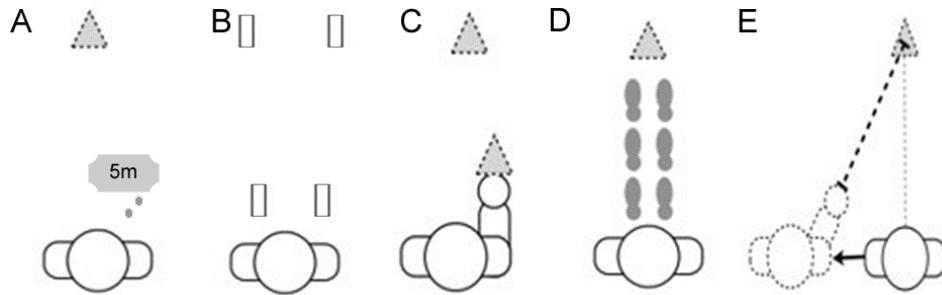


Figure 2.3: Methods to assess the perception of distances

The most popular assessment methods. A: Verbal estimates. Perceptual matching methods: B: Virtual depth span. C: Distance-size coupling. Blind-methods: D: Blind-walking. E: Blind-triangulation.

perpendicular to the target (Fukushima et al. (1997)). The evidence has shown that, under natural conditions, humans beings are highly accurate performing blind-walking tasks until distances of 20m (see Loomis et al. (2003) for a complete review), which explains the popularity of this method. Figure 2.3 depicts the most popular assessment methods found in the literature.

2.1.5 Action and perception

There are two diametrically different ways of perceiving distances. Researchers agree about the existence of different visual pathways in the cerebral cortex, the ventral stream (or vision-for-perception) is believed to serve spatial characteristics of the object (i.e. distance, shape, and size), whereas the dorsal stream (or vision-for-action) has been primarily associated with visually guided motor actions on moment-to-moment analysis. Goodale and Milner (1992) found that different regions of the brain were activated during an experiment that requested subjects to adjust the aperture of the hand either to judge the size of an object or to grasp an object. Similarly, the evidence suggests that there exist two different perceptions of distances, a ventral dedicated to serving spatial judgments of objects (i.e. near, far, short and long), and a dorsal specialized in transforming moment-to-moment information about the location and disposition of objects in efferent motor tasks. Loomis et al. (1992), who were fascinated by the accuracy of the blind methods for assessing the perception of distances, argued that different spatial representations existed, whether the response involved motoric responses (i.e blind-methods) or perceptual reports (i.e. verbal or perceptual-matching methods). Studies have shown that egocentric distances beyond 3m are naturally compressed when the assessment method involves

a perceptual report (Foley (1985), Gogel and Da Silva (1987)). However, under motoric responses, the perceiver corrects (consciously or unconsciously) this distortion when producing the walking responses, implying that other elements may play an important role.

Nowadays, it has become clear that distance perception is influenced by the particular action being performed, or more correctly by the perceiver's ability to perform intended actions, where optical and oculomotor information are scaled to action-specific aspects of the perceiver (Proffitt et al. (2003), Witt et al. (2004), Witt and Proffitt (2008)). Researchers denominated this physiological potential to perform intended actions as the *effort*. For example, Proffitt et al. (2003) demonstrated that perception of egocentric distances are expanded when an observer is wearing a heavy backpack. Also, the researchers showed that manipulating the presence or absence of optic flow while people walked on a treadmill influenced their perception of distances. This discovery was not new, Rieser et al. (1995) performed an experiment in which participants walked on treadmills placed on trailers being pulled across a field by a tractor. This procedure decoupled the rate of induced optic flow from the rate that the participants were walking. Following this procedure, participants performed a blind-walking test. Participants whose treadmill-walking speed was greater than the tractor's speed overestimate the distances, and conversely, those who walked at a slower speed than the tractor underestimate the distances. Together, these studies clearly show that walking effort and optic flow are dynamically calibrated within the visual-motor system. In this research, we put special emphasis to the differences that exist between perceptual-matching methods whose responses depend on perceptual reports, and the blind methods whose responses depend on motoric responses.

2.2 Display issues

Stereoscopic displays suffer from the inevitable issue that an object cannot be recreated based on the light beams reflected on its surface and falling on the retina, but it must be recreated artificially using two images on a flat screen. Assuming that modern computer graphics simulated light in a physically accurate manner (which is probably true), light is in some way intercepted, projected on two flat screens, transformed into pixels, and magnified using an optic lenses to make objects look larger and further away. Unfortunately, this is the best we can do. There is no way with current technology to simulate, in a reasonable amount of time and in an optimal hardware complexity, the induced light on the retina from every object and surface in the scene, and in response for every oculomotor action. Indeed,

the evolution of the technology makes evident the difficulties of emulating the perfectness of the human visual system. Despite the advances in autostereoscopic displays, volumetric displays and holographic displays, only displays based on traditional stereoscopy have been successful, less problematic, and affordable to consumers (Hong et al., 2011). The dependency on the stereoscopic paradigm causes natural problems on depth perception, where we can highlight two challenging problems: the *vergence-accommodation mismatch issue* and the *incorrect peripheral light stimulation*.

2.2.1 Vergence-accommodation mismatch

Under natural conditions, vergence and accommodation work in a perfect geometrical harmony that depend on a close relationship between motor signals (proprioceptive and efferent) and the visual stimulus itself (Figure 2.4). When focusing on an object of interest, light beams from points located at an equivalent distance of the fixation point fall in the same point on the retina, so that they are perceived sharp, an area that is known as the horopter (Hershenson (1999), Schreiber et al. (2008)). Objects that lie along the horopter are perceived as single unified objects when viewed with both eyes on the retina (i.e. the cube), and this area is denoted as the Panum's area of fusion, whose size varies with pupil size. When a point fall in this area, depth can be inferred by comparing the horizontal disparity between both retinal images (binocular disparity) and the motor signals indicating the orientation of the eyes (vergence) (Viguier et al. (2001), Tresilian et al. (1999), Mon-Williams et al. (2000) Viguier et al. (2001)). Conversely, light beams from objects outside the Panum's area do not fall in the same point on the retina and the points are perceived with blur, an area denoted as crossed/uncrossed disparity zone (i.e. pyramid and sphere). Depth cannot be inferred easily in this situation, but some spatial information can be obtained via "blur image analysis" and the motor signals about the lens' shape (accommodation) (Mather (1997), Watt et al. (2005), Held et al. (2010)). According to Held et al. (2012), since all these cues have analogous geometries, they are complementary and enough to provide veridical and quantitative information about depth.

Unfortunately, the harmony of these cues can be broken in VR, as show in Figure 2.5. When looking at a virtual object, the eyes accommodate on the screen plane but converge based on the apparent location of the object, causing a conflict on the habituated relationship between these cues. The Panum's area and the horopter are adjusted according to the screen distance rather than the actual position of the object, making objects that under natural conditions should be out focus, in focus. The

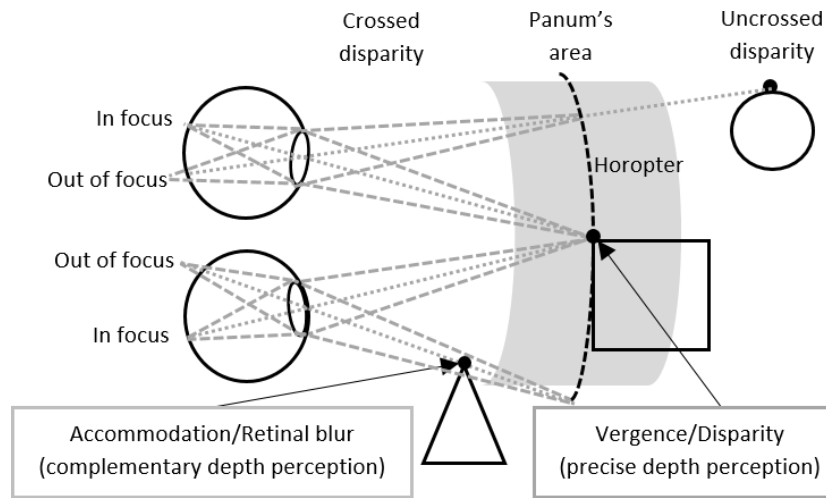


Figure 2.4: The physiological cues to depth perception.

Depth can be inferred from disparity/vergence and accommodation/retinal blur. Adapted from Wartell (2002).

visual system enters in a conflict such as "the objects seem closer based on the accommodative signals, but seem further based on vergence". According to (Lipton (1982)), a pioneer on the stereoscopic display industry, the vergence-accommodation mismatch, is the the only significant issue in which stereoscopic displays differs from the way we see objects in the real world. For him, the phenomenon is a consequence of conflicting information within the vergence-accommodation feedback loop: "different sets of muscles control these functions, and the neurological pathways for these two processes are also separate, the conflict is actually based on a departure from the habituated relationship of two sets of neural pathways and the muscles they control" (Lipton, 1997). Indeed, there is actually a dual parallel feedback loop. First, vergence and accommodation are visually driven cues, where retinal blur is the actual visual cue driving the oculomotor response of accommodation (or the adjustment of the eye's lens to focus and minimizing the blur), and retinal disparity is the visual cue that drives vergence (or the adjustment of the eye's orientation to reduce the image disparity and regain single vision of the object) (Hung et al., 1996). However, there is also a feedback loop between vergence and accommodation, where one becomes a secondary cue influencing the other (Hung et al. (1996), Suryakumar et al. (2007)).

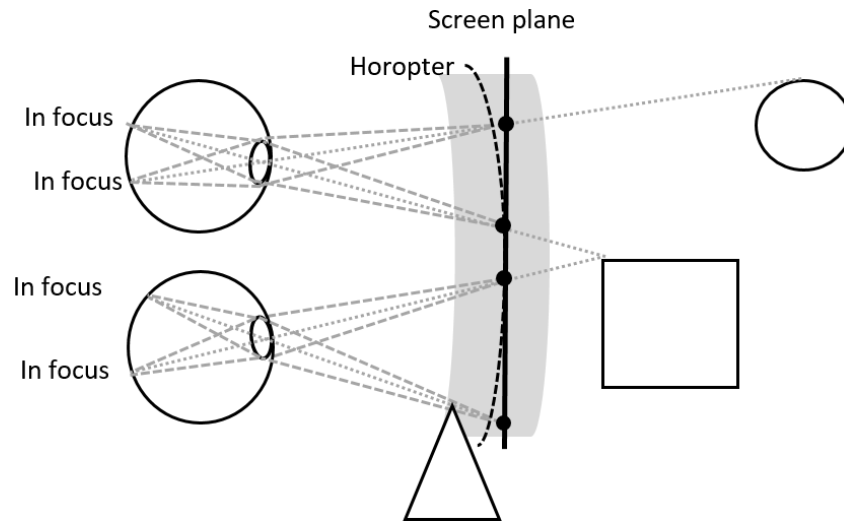


Figure 2.5: The vergence accommodation mismatch

The eyes converge at the apparent location of stimulus, while the eyes are incorrectly accommodate at the screen, making objects that should be out of focus in focus. Adapted from Wartell (2002).

2.2.1.1 Zone of comfortable fusion

Under stereoscopic viewing, accommodation is subject to conflicting demands whose severity depends upon the associated vergence response. An object that has negative or positive disparity will evoke a vergence response whose purpose is to reduce the disparity. Thus, the larger is the disparity, the larger will be the vergence response. As vergence is induced, an accommodation response is elicited, attempting to adjust the pupil size towards the point of convergence. However, if accommodation becomes excessive and moves away from the screen, then the observer will have difficulties focusing on the object and he/she will suffer discomfort. Stereoscopic display developers have studied the limits in which vergence and accommodation could differ from each other without causing discomfort. This limit is known as the *zone of comfortable fusion* (Hoffman et al. (2008), Reichelt et al. (2010), Shibata et al. (2011), Tam et al. (2011)), and it is estimated to be intrinsically related the human depth of field (DOF), the zone in which images disparities are fused without effort and whose size is estimate to be around $\pm 0.3D$ (Campbell (1957)), where D stand for diopters . A diopter is a unit of measurement of the optical power, which is equal to the reciprocal of the focal length measured in meters ($1D = 1m^{-1}$, it is thus a unit of reciprocal length). In a comprehensive study, Shibata et al. (2011) estimate that the zone of comfort was around $0.5D$. They used a very complex apparatus of lenses, polarizers and

prisms that simulate almost any kind of stereoscopic display, showing the great tolerance of the human visual system to the divergences between vergence and accommodation (Figure 2.6). The orange zone represents the zone of comfortable fusion, and the green zone represents the DOF (normal-light-day DOF of ± 0.3 diopters). Thus, the figure describes the intrinsic relationship between the display’s focal distance and the zone of comfort.

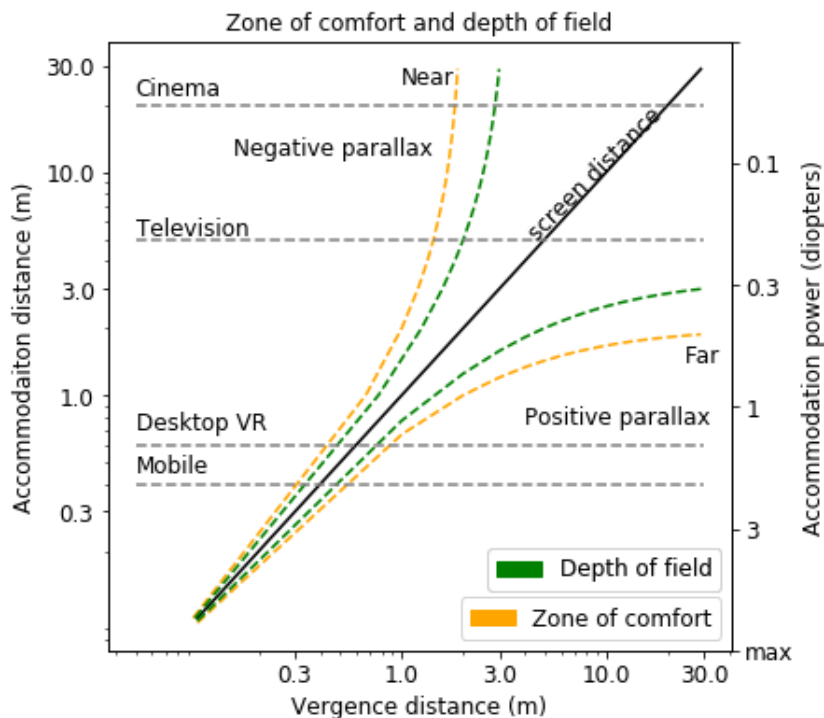


Figure 2.6: Zone of comfortable fusion and depth of field

The zone of comfortable fusion define the limits in which vergence and accommodation could differ each other without causing discomfort. It depends on the display’s focal distance and the human DOF. Adapted from Shibata et al. (2011).

It is evident from Figure 2.6 that the range of simulated distances of objects that we can comfortably present in a VR setting is very dependent on viewing distance. The focal distance is the driver that defines the zone of comfort and the *functional depth range* available for the display or the limits in which vergence and accommodation can be used to perceive depth. For example, in stereoscopic cinema, the range extends from 1.6 m to infinity. This range is ideal, as 3D cinema users are expected to be at more than 2.0m of the screen, and they need to focus either on objects close or far away. In contrast, in mobile 3D devices, the range extends from only 0.28m to 0.44 m, which corresponds to the

distance between the human hands porting the device and the eyes. Thus, display developers face the challenge of restricting the disparity between the left and right images within the limits of this zone, so that the conflict between vergence and accommodation is minimized. Consequently, the functional depth range is formed by the limits of the DOF (Equation 2.2), where F_d is the focal distance and the T_{dof} is the size of the DOF in diopters.

$$F_r = \frac{1}{1/F_d \pm T_{dof}} \quad (2.2)$$

2.2.1.2 Depth compression and focal distance

A consequence of the limited functional depth range of stereo displays is that the 3D scene has to be squeezed in-depth more drastically in the display with the smaller focal distance (if the scene depth is larger than the available depth range) (Hoffman et al. (2008)). For example, if we have a virtual scene of a huge auditorium and we render such scene in a mobile 3D device, everything in the scene will be squeezed in the range from 0.28m to 0.44 m, so that the whole auditorium will look flat. Thus, depth compression is a consequence of the fixed focal distance of stereoscopic displays, which limits the *functional depth range*, the minimum and maximum spatial range in which the eyes could be accommodated to perceive depth while maintaining a sharp image of the screen and without inducing a severe conflict and discomfort. Figure 2.7 shows the compression induced when the perceiver is focused at the screen plane. The image shows the asymmetric compression effect, which is a consequence of the expanding nature of the depth of field. Objects tend to be less compressed when they are located in front of the screen than objects that are located behind the screen. As consequence, the underestimation of distances becomes stronger as the distance increase. HMDs are very susceptible to this phenomenon due to the use of lenses with fixed focal distances, which restricts the zone of comfort to a limited range, a range more suitable for objects located far away than closer to the perceiver. In contrast, LIPDs are more 'flexible', providing different focal distances depending on the location of the perceiver in relation to the projection wall.

Figure 2.8 presents the zone of comfortable fusion, the DOF, and the functional depth range of an HMD compared with a CAVE system, based on the experiment by Shibata et al. (2011). The black line presents the natural condition where vergence and accommodation responses are equal. The limits

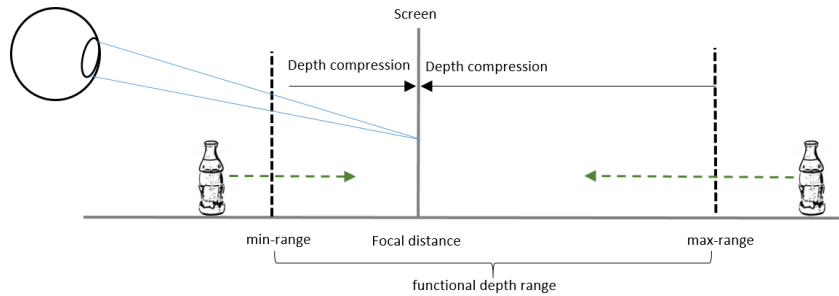


Figure 2.7: Functional depth range and depth compression

The focal distance determines the functional depth range in which the whole scene is compressed.

of the DOF given the screen distance determine the functional depth range, depicted with the red line. HMDs have lenses with a fixed focal distance that creates a virtual image at infinity. Following the estimation of Shibata et al. (2011), the minimum focal distance that falls inside the comfortable zone is 2m ($\frac{1}{\frac{1}{\infty+0.5}} = 2m$), resulting in a functional range from 1.25m to 5m ($\frac{1}{\frac{1}{2\pm 0.3}}$). This range means that our eyes can accommodate and maintain a sharp image of the screen between a range of 1.25m to 5m while being able to converge the eyes from 1.25m to infinity!, a range that demonstrates the amazing tolerance of the human visual system to the divergences between accommodation and vergence. Indeed, first and second generation HMDs used this strategy, subjects were forced to focus far away even when the stimulus was closer, an issue that caused discomfort and fatigue. Thanks to the advances in optics and computer graphics, modern HMD developers have managed to achieve even smaller focal distances than the minimum 2m calculated by Shibata et al. (2011), allowing subjects to focus near and far without inducing discomfort. However, the conflict persists, as we are still forcing the perceiver to unnaturally adapt to conflicting cues.

The same analysis could be performed on LIPDs, like the CAVE, where the focal distance varies depending on the location of the perceiver in relation to the projection wall. For example, if we assume a minimum focal distance of 0.2m and a maximum focal distance of 3m, we will have a functional range from 0.18m to 0.22m, when the observer is at 0.2m from the screen ($\frac{1}{\frac{1}{0.2\pm 0.3}}$), and a functional range from 1.5m to 30m, when the perceiver is at 3m from the screen ($\frac{1}{\frac{1}{3\pm 0.3}}$). Since the DOF expands significantly with the focal distance, there is a functional depth range approximately seven times greater in the CAVE compared with the HMD, by just increasing the focal distance by 1m!. Similarly, the CAVE is a system so flexible, that we can emulate the functional range of a mobile 3D device when the observer is located at the minimum focal distance of 0.2m. Although it could cause a

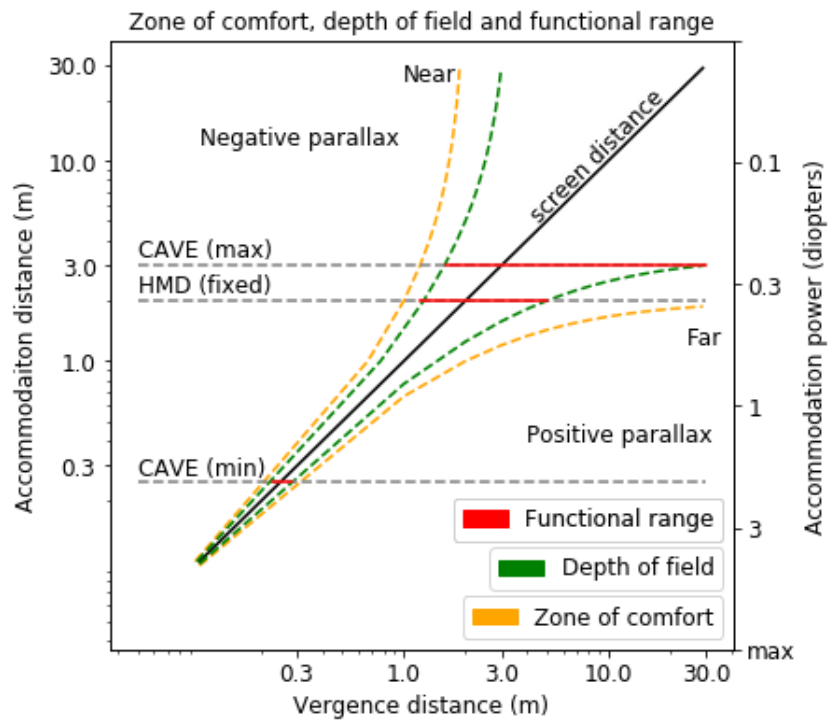


Figure 2.8: Functional depth range

The functional range determines the limits in which the eyes could be accommodate to perceive depth, while maintaining a sharp image of the screen and without inducing a severe conflict and discomfort.

loss of visual acuity, this can be compensated by having a high resolution like the one found in Iowa’s C6. Additional to the ‘flexible functional range’, the CAVE has another strategic advantage compared with the HMD, and it is the possibility of seeing the virtual- and physical environment simultaneously. By using physical cues either from the ground, the boundaries of the system, or the subject’s own body, the perceiver can compensate for the distortions induced by the vergence-accommodation mismatch, even when the virtual object is too close and outside the limits of the zone of comfort. Hence, all these advantages make LIPDs more adaptable for tasks either in the personal or the action space.

2.2.1.3 Stereoscopic parallax condition

The location of the virtual object in relation with the screen is usually referred as the stereoscopic parallax condition, and it is an indicator of the amount and characteristics of the conflict induced between vergence and accommodation. Objects at *zero parallax* means that the object is located

exactly at the screen’s position, so there is not disagreement between vergence and accommodation. In contrast, objects at *positive parallax* and *negative parallax* means that they are located behind or in front of the screen respectively, so a disagreement between these cues have been induced. Previous studies have shown a direct relationship between the amount of underestimation or overestimation in the perception of distances, and the location of the stimulus in relation with the screen. These studies have been done mostly in LIPDs, where is technically possible to accommodate the eyes at different target distances (Naceri et al. (2010), Murgia et al. (2009), Marsh et al. (2014), Bruder et al. (2016), Kenyon et al. (2008), Luo et al. (2007)). Their results have shown not only that the greater the conflict between vergence and accommodation, the greater the perceptual biases, but also the existence of asymmetric effects, with slight overestimation for negative parallax (object in front of the screen plane), and underestimation for positive parallax (object behind the screen plane). The evidence also suggests that at zero parallax, spatial perception tends to be accurate. Regarding HMDs, the evidence suggests that underestimation of distances is linearly related to the amount conflict induced between vergence and accommodation. Vienne et al. (2020) compared the influence of the focal distance between an LIPD and an HMD, finding that while both systems showed similar depth perception for the short distances, the display that induced the smaller conflict size show the best accuracy for the larger distances. Thus, they argued that the underestimation of distances was linearly correlated with the amount of conflict size induced between vergence and accommodation.

According to Vienne et al. (2018) and Vienne et al. (2020), the conflict size can be calculated by subtracting the inverse of the accommodation from the inverse of the vergence:

$$C_{va} = \frac{1}{V_d} - \frac{1}{A_d} \quad (2.3)$$

where C_{va} denotes the size of the vergence accommodation conflict (V-A conflict size), V_d denotes the vergence distance, and A_d denotes the accommodation distance. A positive conflict arises when the vergence distance is smaller than the accommodation distance (negative parallax), and a negative conflict when the vergence distance is greater than the accommodation distance (positive parallax). The sign is important, because at positive parallax the distances are inverted, as they are in function of the focal distance rather than in the function of the perceiver. According to Vienne et al. (2018) and Vienne et al. (2020), the bias in depth perception should be proportional to A-V conflict size, but should also exhibit asymmetry between positive and negative conflicts, with a greater bias for negative

conflicts. Figure 2.9 depicts how the conflict size changes as a function of different focal distances. The equation implies there are two ways of reducing the conflict between vergence and accommodation: (1) by making accommodation and vergence equal, which is only possible when the stimulus is located at zero parallax, and (2) by increasing the display focal distance. Augmenting the distance between the eyes and the screen will force the eyes to accommodate at a large distance, reducing its optical need for focus (in diopters) and evoking a similar vergence response. As the focus demands are reduced, the DOF is expanded, increasing consequently the comfort zone. Indeed, the effect of increasing the focal distance can be seen in the zone of comfortable fusion, as the focal distance increases, the zone of tolerance to the divergences between accommodation and vergence increases. Bruder et al. (2016) validate this hypothesis in a 10m length CAVE system, finding that a distance of 6-7m of the screen was enough to give an almost perfect performance, when using a blind-triangulated pointing method. In this research, we paid special attention to the V-A conflict size, as a method to characterize the depth compression phenomenon induced by a display as a function of the focal distance.

2.2.2 Incorrect peripheral stimulation

In section 2.1.5, we described how action influence perception and how the optic flow is naturally calibrated with motoric tasks, such as human walking. HMDs induce several distortions that impact negatively the optic flow, especially in the peripheral vision. The first factor is the limited FOV provided by the display. Although previous studies suggested that the limited FOV was not a contributing factor on the phenomenon of underestimation of distances. New evidence suggests that the limited FOV, and the nature of light stimulation in the peripheral vision, influence the perception of distances in VR. Jones et al. (2013) demonstrated that by adding a physical light frame in the periphery of the display, it is possible to induce an increase of approximately 13% in the perceived distances. Another option was using a portion of the imagery to create a computer-generated peripheral frame and then adjust artificially the amount of brightness (Li et al. (2018)). The researchers found that there is a threshold between 5%-15% of luminance increments, where the perception of distances becomes more accurate. Also, in another experiment, they change the lighted frame with a pixelated region, obtaining a direct relationship between induced resolution in pixels/degree and distance perception. These results confirm that the limited FOV in VR displays causes an adjustment in the apparent effort (action influencing perception), as the induced optic flow does not correspond

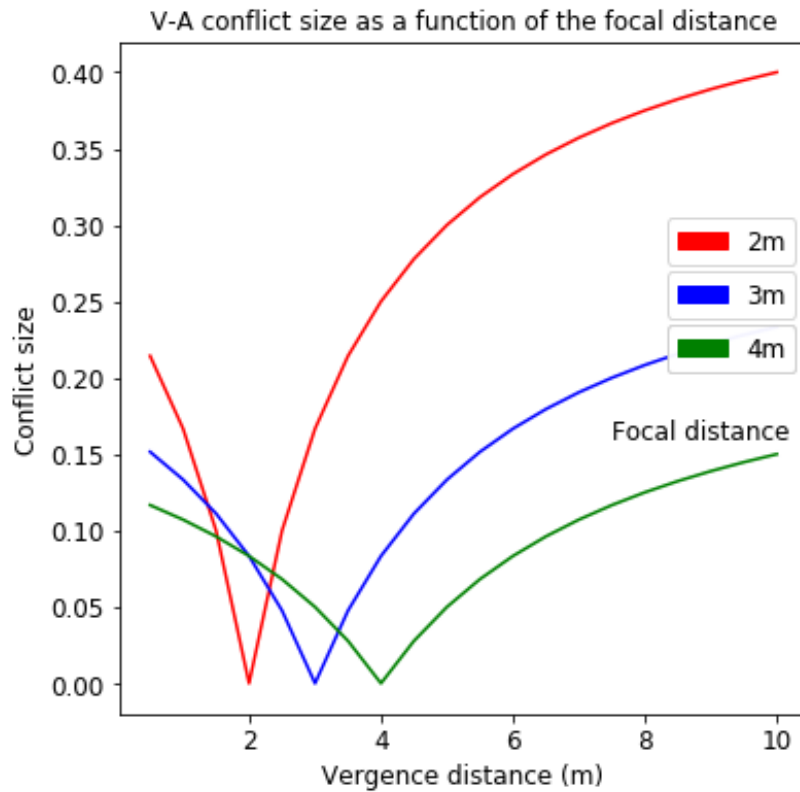


Figure 2.9: Conflict size as a function of the focal distance

The V-A conflict size exhibit asymmetry between positive and negative conflicts, with greater bias for negative conflicts (positive parallax).

with expected one under natural conditions.

The distortions induced by the lenses are another factor that impacts the optic flow. HMDs' lenses suffer from distortions such as god rays, chromatic aberrations, and screen door effects, which impact the natural perception of the optic flow during the execution of motor tasks. They also distort the visual acuity, as the angular resolution does not vary uniformly from pixel located at the center of the screen to the pixels located at the edges. Using a photogrammetric procedure, Kreylos (2017) calculated this variation for two popular headsets (an HTC Vive CV1 with Fresnel lenses and an Oculus Rift CV1 with Hybrid lenses). Figure 2.10 shows the variation of the angular resolution through the display's lenses, analyzed horizontally and vertically. The calculated sampling factor indicates that the angular resolution is relatively constant for pixels located around 50% of the center of the screen and then scales vertically for pixels in the edges. The sampling factor shows that the HTC Vive has

2.2. DISPLAY ISSUES

a better angular resolution in the central vision (a ratio 1 arcminute/subpixel) compared with the Oculus Rift (a ratio 1.5 arcminute/pixel). In contrast, the Oculus Rift has better resolution in the peripheral vision, (a ratio of 2.5 arcminue/subpixel, compared with 4 arcminute/subpixel of the HTC Vive.

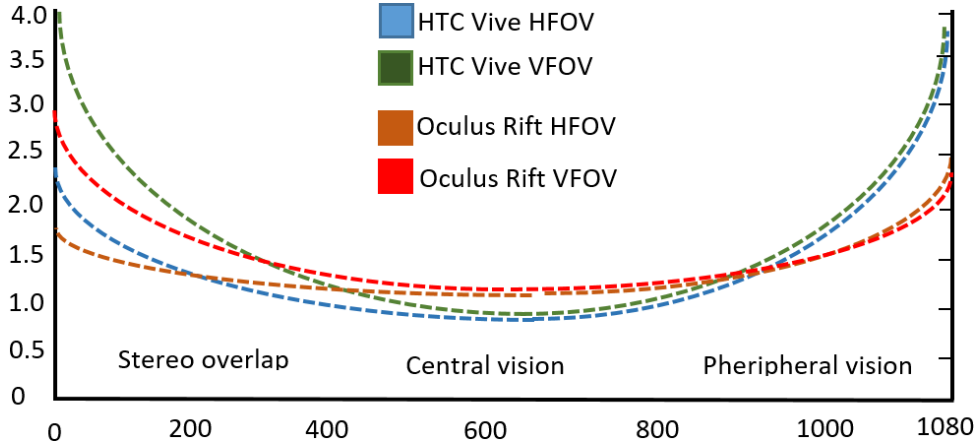


Figure 2.10: Lens distortions in HMDs and angular resolution

The variation of the angular resolution according to the pixel position in the HTC Vive and the Oculus Rift. Top: Horizontal resolution. Bottom: Vertical resolution. Adapted from Kreylos (2017).

Although analysing the optic flow could be interesting for the proposed framework, researcher usually study this cue in isolation by inducing abstract flow patterns (e.g. a radial cylindrical flow), patterns that are easily measured and controlled (Fujimoto and Ashida (2020)). Since these kind of stimuli do not correspond to the found under natural conditions, we will focus more in motion parallax, as this cue is intrinsically related with the optic flow. Motion parallax refers to the phenomenon where objects moving at a constant speed across the visual field will appear to move faster if they are closer to the observer (or camera) than if they are at a greater distance. Studies have shown that the interaction of motion parallax and binocular disparity provides absolute information about the distance and size of objects either in the personal- or action space (Bradshaw and Rogers (1996), Mansour et al. (2019), Kellnhofer et al. (2016)). According to Nawrot and Stroyan (2009), the perception of depth from motion parallax is given by the *motion pursuit law*:

$$\frac{d}{f} = \frac{d\theta}{d(\alpha - \theta)} \quad (2.4)$$

2.2. DISPLAY ISSUES

where f is the distance of the fixation point, d is the distance a point relative to the fixation point, θ is the horizontal/angular disparity between both points, and α is the eye's convergence angle. Figure 2.11 depicts the motion pursuit law graphically. Observer fixation on F produces the convergence angle α that changes over time at the rate $d\alpha/dt$, as the observer moves and the eyes rotate accordingly. This motion also causes the retinal image of D , being displaced from the retinal image of F by the angle θ , which changes at the rate $d\theta/dt$. Thus, the perception of distance of the point F can be estimated by analyzing the rate of change between θ and α over time.

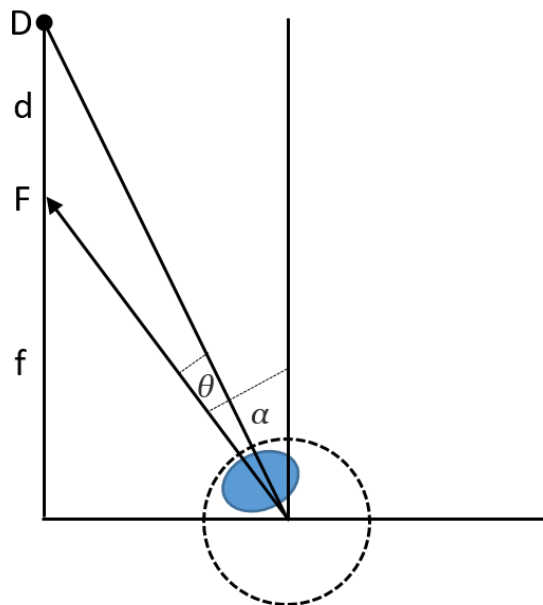


Figure 2.11: Motion pursuit law

The perception of distance of the point F , can be estimated by analysing the rate of change of the horizontal/angular disparity θ between F and D , and the convergence angle α over time. Adapted from Nawrot and Stroyan (2009).

Under natural conditions the motion pursuit law holds. However, the influence of the lenses in HMDs causes that when objects approaches at the peripheral vision, they move at a faster rate towards the periphery of the visual field, creating the illusion of being closer than they are, and influencing the perception of distances of the stimulus over time. This effect is reinforced by the limited FOV, as the objects disappear from the visual field sooner than under natural conditions. Figure 2.12 depicts how motion parallax is distorted in HMDs. The variation causes that point D moves according to the motion pursuit law in the central vision but starts to move at a faster rate in the visual field as it

approaches to the peripheral vision, increasing the rate of change of the angular disparity θ over time and decreasing the perceive distance of point f . In this sense, we will take into consideration how the perception of depth from motion parallax is influenced by the limited FOV and the distortions induced by the lenses.

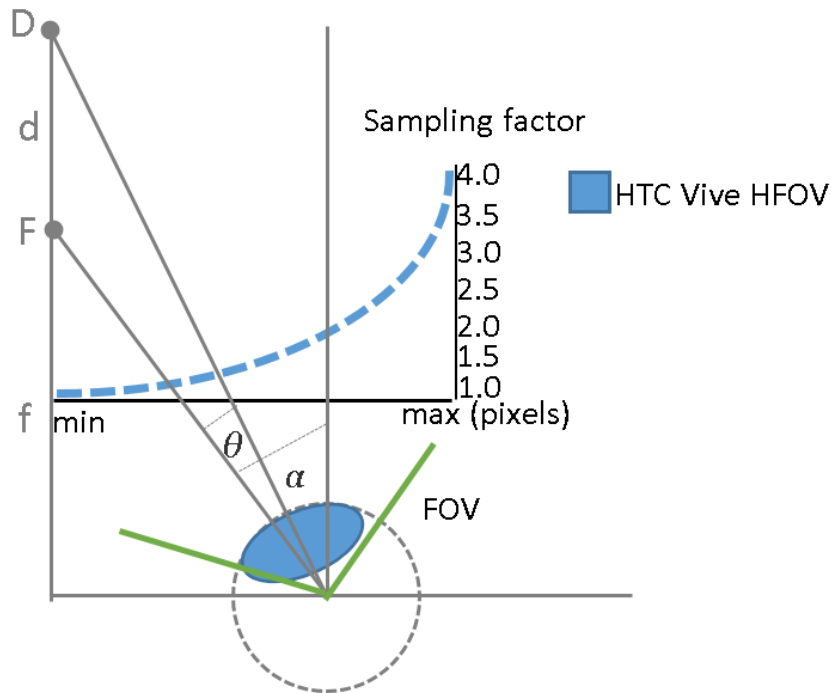


Figure 2.12: Influence of the HTC Vive lenses on motion parallax

Because the angular resolution varies not uniformly from the central vision to the peripheral vision, point D move at a faster rate in the visual field as it approaches to the peripheral vision, increasing the rate of change of the angular disparity θ over time and decreasing the perceive distance of point f .

2.3 Display issues and influence space

Figure 2.13 depicts the main issues described before and their influence space, based on the classification of depth cues by Cutting and Vishton (1995) . The vergence-accommodation mismatch issue is directly associated with the vergence and accommodation cues, so its influence is greater on the personal space, where the power of these cues to discriminate depth is stronger. Thus, this issue affects the performance of spatial tasks in the near space, which involves mostly tasks associated

2.3. DISPLAY ISSUES AND INFLUENCE SPACE

with subject's interaction with objects (i.e. reaching or grasping, perceptual matching tasks). Studies have shown that, when tested in isolation, the useful range seems to be less than 3m for these tasks (Linton (2020)). Regarding the incorrect peripheral stimulation, this issue is more associated with the perception of motion parallax and the optic flow, cues not only associated with the perception of depth but also with the action-based distance task. Hence, this issue involves tasks in the personal space (e.g. matching or aligning) and tasks in the action space (e.g. aligning, pointing, throwing, blind-walking). Depending on the task, the useful range of these cues could vary ranging from 3m to 20m (Bradshaw et al. (2000), Loomis et al. (1992)). However, we must emphasize again that even though cues like vergence and accommodation have the smallest discrimination threshold and their performance decay quickly, their importance cannot be underestimated because these cues contribute to the overall perception of depth (Held et al. (2012), Viguier et al. (2001)). Similarly, motion parallax has the highest discrimination threshold at shorter distances becoming a very suitable depth cue for task in the personal space (Bradshaw and Rogers (1996), Mansour et al. (2019), Kellnhofer et al. (2016)). Hence, the influence space of the display issues is dynamic, they always influence the perception of depth, regardless of the useful range of the different depth cues but also depending on the kind of spatial task.

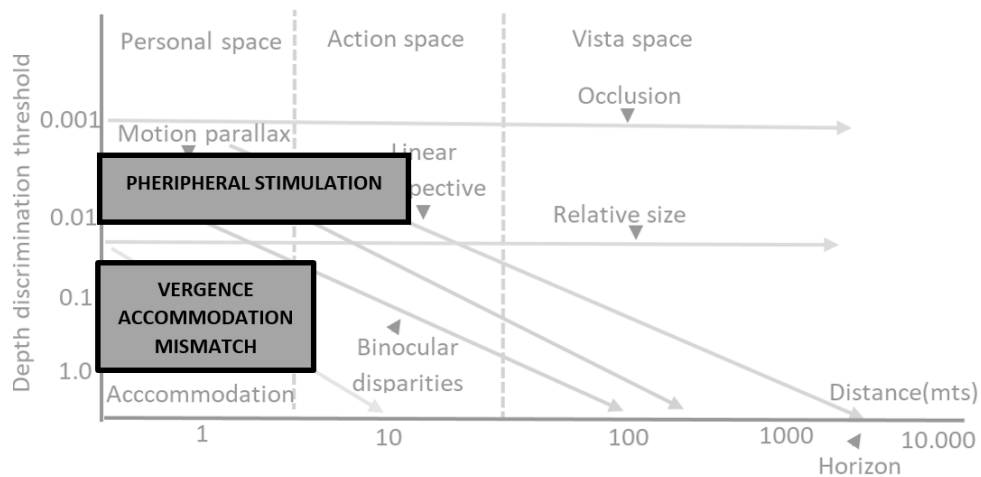


Figure 2.13: Display issues and influence space.

Main display issues and their influence on the personal and action space. Adapted from Cutting and Vishton (1995).

2.4 Spatial perception in VR

2.4.1 Distance perception and VR displays

Table 2.1 presents a summary of studies using different displays and measurement methods found in the literature. The most common methods used by the researchers to assess the perception of distances in HMDs are undoubtedly the blind-methods (Kelly et al. (2017), Creem-Regehr et al. (2015), Peer and Ponto (2017), Messing and Durgin (2005), Kellner et al. (2012), Willemsen et al. (2008), Sahm et al. (2005)). This preference is an obvious consequence of the possibility of modern headsets of integrating positional tracking systems, allowing subjects to walk naturally in VR. In contrast, there are very few studies that used perceptual matching methods for assessing the perception of distances in HMDs. This little interest could be caused by the need of these kinds of methods to support subjects' distance estimates based on a physical reference, which would require subjects to raise or lower the headset. Conversely, the impossibility of walking greater distances in LIPDs has motivated researchers to propose alternative methods for assessing the perception of distances different from blind-walking. A popular alternative is perceptual matching methods using the relationship between the perception of distance and size (Kenyon et al. (2008), Luo et al. (2007), Murgia et al. (2009), Naceri et al. (2010)). Perceptual matching methods are ideal for LIPDs, as the subject is able to see the virtual- and physical environment simultaneously. Another popular method is blind-triangulation, an alternative to blind-walking that does not require subjects to walk larger distances (Bruder et al. (2016), Peer and Ponto (2017), Willemsen et al. (2008)).

Since an objective of this research is the design of traversal methods to compare the perception of distances between heterogeneous displays, we selected three assessment methods as good study cases. From these tests, we selected the assessment method proposed by Kenyon et al. (2008) as an interesting study case because it is based on a strong perceptual phenomenon called *size-constancy*. Because the method was originally devised for CAVE systems, we adapted it for HMDs. The second assessment method was the classic blind-walking tests. Although we stated before that blind-walking is not a method suitable for LIPDs, we explored the possibility of assessing the perception of distances using a modern consumer-oriented treadmill, as this technology has improved significantly in recent years.

2.4. SPATIAL PERCEPTION IN VR

Display	Task	Percepts	Ref
HTC Vive	Perceptual matching Blind walking	Distance Distance-Size	Kelly et al. (2017)
Oculus Rift	Blind walking	Distance	Creem-Regehr et al. (2015)
HTC Vive Oculus Rift	Blind-triangulation Blind throwing	Distance Distance-Direction	Peer and Ponto (2017)
VRS V8 HMD	Blind walking	Distance	Messing and Durgin (2005)
nVisor ST60 HMD	Blind walking	Distance	Kellner et al. (2012)
NVision HMD	Blind-triangulation	Distance-Direction	Willemsen et al. (2008)
nVisor SX HMD	Blind walking Blind-throwing	Distance	Sahm et al. (2005)
Immersive wall Cybermind HMD	Perceptual matching	Distance-Size	Naceri et al. (2010)
CAVE	Perceptual matching	Distance-Size	Murgia et al. (2009)
CAVE	Perceptual matching	Distance-Size	Kenyon et al. (2008)
CAVE	Perceptual matching	Distance-Size	Luo et al. (2007)
CAVE	Perceptual matching	Distance-Size	Marsh et al. (2014)
CAVE	Blind-triangulation	Distance-Direction	Bruder et al. (2016)

Table 2.1: Spatial perception in a variety of studies using different VR displays and measurement tasks.

Finally, we took into consideration the method proposed by Bruder et al. (2016), which was performed in a CAVE system and is based on blind-triangulation. Because this method does not require subjects to translate larger distances, is ideal to assess the perception of distances either in HMDs or LIPDs.

There exists a direct relationship between the phenomenon of underestimation of distances and the evolution of the technology. Table 2.2 shows a review of studies that measured the perception of egocentric distances using different HMDs and based on blind-walking tasks. Before the advent of modern consumer-oriented devices, distances were generally underestimated by a mean of 75% of the actual distances (see Renner et al. (2013) for a complete review). With the arrival of displays with greater capabilities, the underestimation effects have become weaker, with distances underestimated around 80% for virtual outdoor spaces and 90% for virtual indoor spaces. It is surprising that despite the huge advances in display optics, tracking, and rendering capabilities, there has been only a slight improvement, especially for outdoor spaces. These results suggest that the phenomenon is still relevant and the display issues distort spatial perception.

There are a few studies about spatial perception in LIPDs, either using immersive walls Plumert et al.

2.4. SPATIAL PERCEPTION IN VR

Environment	Distance	Display	Ave. Under.	Study
Outdoor	5m	Oculus Rift CV1	83%	Buck et al. (2018)
		Oculus Rift DK2	65%	
		HTC Vive	86%	
Indoor	5m	nVisor SX111	92%	Kelly et al. (2018)
Indoor	5m	HTC Vive	88%	Kelly et al. (2017)
Outdoor	5m	nVisor SX111	70-73%	Siegel and Kelly (2017)
Indoor	5m	Oculus Rift DK2	90%	Li et al. (2016)
Outdoor	4.5m	Oculus Rift DK2	82%	Creem-Regehr et al. (2015)
		Nvis SX60	65%	
Indoor	4.5m	Oculus Rift DK2	89%	Creem-Regehr et al. (2015)
		Nvis SX60	77%	
Indoor	5m	Oculus Rift DK1	83%	Andrus et al. (2014)
		Nvis SX60	52%	
		NvisSX111	47%	

Table 2.2: Comparative studies about distance perception using different HMDs displays and blind walking tasks.

(2005), Naceri et al. (2010)) or CAVE systems (Murgia et al. (2009), Marsh et al. (2014), Bruder et al. (2016), Kenyon et al. (2008), Luo et al. (2007)). We can extract three conclusions from these studies. (1) The evidence suggests that underestimation effects are smaller than in HMDs, (2) the physical space between the user and the projection screen seems to be the most important factor for distance perception, and (3) asymmetric effects are found with slight overestimation if the object is located before the projection screen and underestimation if the object is located after the projection screen. Thus, the distortion effects in LIPDs seem to be smaller than HMDs and they are a function of the physical space.

2.4.2 Influence of adaptation

A frequent question is whether continually interacting with the VE induces an adaptation process in which spatial perception is improved. There are at least two hypotheses of the influence of VR in spatial perception: the *recalibration hypothesis* states that adaptation is task-dependent, and walking through the VE with continuous visual feedback modifies the action-perception loop but only for tasks related to walking (Waller and Richardson (2008)). This theory is supported by the empirical studies about the influence of action in perception (see section 2.1.5). In contrast, the *rescaling theory* states that walking with continuous visual feedback causes a re-scaling of the perceived space, where the perception of distances should be transferable across different categories of spatial tasks (e.g.

grasping, throwing, or pointing) (Richardson and Waller (2007), Kelly et al. (2013)).

In this research, we consider the recalibration hypothesis as the most plausible. We believe that adaptation is task-specific, Interacting with the VE induces a different perception of distance that could vary depending on the task and the nature of the apparent effort induced (i.e optic flow vs walking speed). Indeed, the recalibration hypothesis has been tested in a CAVE system using a treadmill. Subjects walking at a constant speed were induced with different visual speeds, influencing their perception of distances and showing that recalibration is highly dependent on the motion parallax (Mohler et al. (2007)). Another study showed that motor recalibration occurs rapidly and transfers asymmetrically across scale, which means that its main effects occur just in the first interactions, and interacting with distant objects improves near distances but not vice-versa (Kelly et al. (2014)). Hence, we consider important to analyze the influence of adaptation, especially in assessment methods that depend on motoric responses.

2.4.3 Methods that improve spatial perception in VR

An objective of this framework is to explore methods to influence the perception of distances using composition. Composition refers to the capability of the system of representing 3D objects on a flat surface using different techniques that creates the illusion of depth, where linear perspective is one of these methods. Thus, some methods have been proposed to improve the perception of distances in VR that are more focused on the compositional factors than on the technical factors. These methods are based on alternative projection techniques that do not follow the standard linear perspective projection used in computer graphics (see Salomon (2007) for an overview). Although there are several methods, we selected from the literature two methods that seemed promising: *fov minification* and *lowering the horizon*.

- *FOV minification* is a method that compresses the visual imagery by artificially increases the geometric FOV of the display. The method have shown that can improve distance perception and reduce the underestimation effects (Kuhl et al. (2006), Bolte et al. (2010), Zhang et al. (2012), Li et al. (2014), Steinicke et al. (2009a)). By applying a scaling of 0.7%-0.82% in the visual imagery, their results showed an average increase in perceived distances between the 13%-20%. The initial implementation of FOV minification was straightforward to apply in older HMDs because developers had complete control of the rendering pipeline. However, modern HMDs do

not allow to change the display's rendering settings easily, so we used an alternative approach to adapt its implementation to the current technology.

- *Lowering the horizon* is another method that has shown good results improving distance perception in VR (Messing and Durgin (2005), Williams et al. (2009), Kuhl et al. (2009)). By artificially applying an angular declination in the horizon line of 5%-11%, their results have shown an increase in the perceived distances directly related to the amount of angular declination. This method is based on the *angular declination hypothesis* $d = h/\tan(\alpha)$ (Ooi et al. (2001), Gajewski et al. (2014), Todorovic and Toskovic (2012), Bunch (2014)), a theory that states that distance perception d is mostly influenced by the observer's eye height h and the angular declination of the target regarding the horizon line α . Ooi et al. (2001) validate this hypothesis in the physical reality by requesting subjects to estimate the perceived distance of a target using base-up prism goggles. This device refracts the induced light, "moving" everything upwards, changing the apparent location of the horizon line. Similarly, we proposed a different approach to adapt the implementation of this method to the current technology.

Chapter 3

A framework to characterize and influence spatial perception in heterogeneous virtual reality displays

3.1 Overview

We propose a framework to characterize and influence the perception of distances in heterogeneous displays, which takes into consideration the different display factors that distort spatial perception. The proposed framework can be seen as an extension of the action-perception loop in VR (Figure 3.1). A typical feedback loop is composed of two main entities, the perceiver and the virtual environment (VE). A virtual stimulus is presented to the perceiver using a display. The perceiver creates a spatial representation of the stimulus based on his/her spatial perception, and performs an action accordingly using his/her motor system. The VE provides visual feedback according to the perceiver's actions, altering the spatial representation of the stimulus itself. In optimal conditions, the perceiver's spatial representation and the virtual representation are approximated, as the depth cues from the VE are simulated based on the rules of the physical world. However, the display distorts some of these cues, affecting the perceiver's perception of distances. To determine how their spatial perception is affected, an assessment method is required. Some assessment methods measure the perception of distances indirectly (perceptual matching), while others measured it directly but can be affected by adaptation (blind-methods), and their main difference is related to the use of continuous visual feedback. Thus, the framework is based on the idea that if we are able to predict how spatial perception is distorted by the different display issues (using an assessment method), we can alter the simulation itself to influence

3.1. OVERVIEW

positively the perception of depth. To achieve this objective, the framework requires two steps. First, the induced depth compression is measured and predicted using an assessment method, whose stimuli is strategically selected based on the associated display factors. Then, a compositional technique is used to influence the perception of distances positively based on the predicted depth compression. Thus, the framework can be used to characterize and improve the perception of distances in VR.

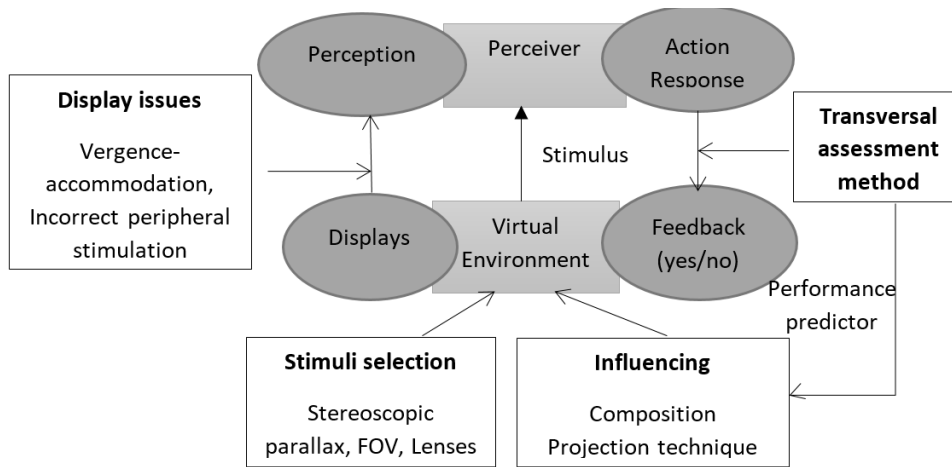


Figure 3.1: The proposed framework as an extension of the action-perception loop in VR.

An objective of the framework is that the selected assessment methods must be transversal. In section 3.2 some preliminary studies were developed to design good assessment methods for the proposed framework. A natural method to compare the influence of the display issues is by directly comparing displays with technical differences in critical factors. Thus, we performed two comparative studies between heterogeneous displays. In chapter 4, we discuss the first part of the framework, a formal approach to characterize the depth compression phenomenon on a display based on the nature of the spatial task and the display factors. In chapter 5, describes how the linear performance predictor can be used to influence subjects' spatial perception using a *compositional technique*. For example, we can use the performance predictor to calculate the optimal angular declination to apply at the horizon line in the *lowering horizon* technique. Similarly, the same process can be used to calculate the ideal scaling factor in the *FOV minification* technique. Thus, the framework can be used either to characterize the depth compression phenomenon in the display and to influence subjects' spatial perception.

3.2 Preliminary studies

As a first approach, we selected two spatial perception tests as good candidate methods to validate the proposed framework. These tests were selected based on two criteria: (1) They should be easily adaptable to other displays without altering significantly their nature or main objective; and (2) they must represent the spectrum of human spatial tasks, either in the personal- or action space, where distance perception is relatively accurate. Figure 3.2 depicts the two selected methods and their applicable space based on (Cutting and Vishton (1995)) spatial classification. In the personal space, we selected the size-constancy test (Kenyon et al. (2007)), a test that belongs to the category of perceptual matching tasks, where the subject is requested to perform distance estimates based on a physical reference. As a test based on perceptual matching, its application space is better at shorter distances, not too far from the personal space. Regarding the action space, we selected the classic blind-walking test (Loomis et al. (1992)). This test corresponds to the blind-methods category, where the subject is requested to estimate a distance by performing a motoric action under non-visual conditions. Alike the methods based on perceptual matching, the blind methods are more suitable for assessing spatial perception at larger distances. Some of these methods were initially intended to be used for a particular display. For example, the size-constancy test was devised to assess spatial perception in a CAVE system, due to its advantage for seeing the physical- and virtual environments simultaneously. Hence, the challenge consisted in adapting and extending their applicability to different kinds of displays.

Once the assessment methods were adapted to the needs of the current framework, we were interested in comparing the performance between different displays. If the method is pertinent, subjects using a display that has a technical advantage will perform better than subjects using a more limited display. Thus, we studied the influence of the accommodation-vergence mismatch and the incorrect peripheral stimulation using comparative studies between displays. To achieve this objective, we had to determine which assessment methods and what kind of displays were more useful to study according to each display factor.

3.2.1 Size-Constancy test

Kenyon et al. (2007) and Luo et al. (2007) proposed a method to assess spatial perception in VR

3.2. PRELIMINARY STUDIES

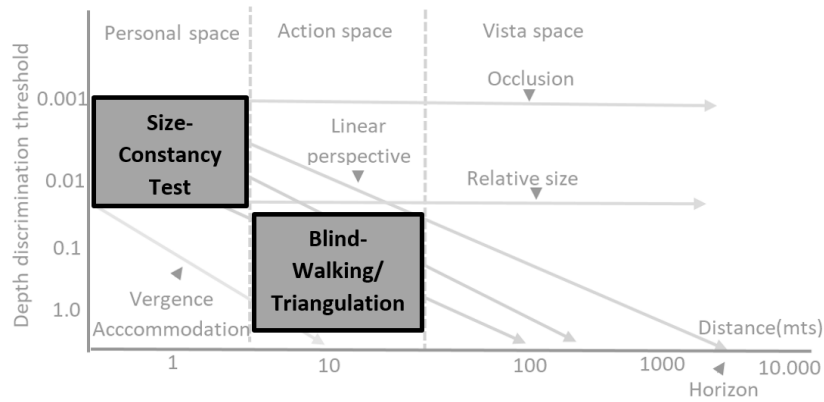


Figure 3.2: Candidate assessment methods and influence space

Two distance perception tests were selected as good candidates methods to validate the proposed framework.

taking advantage of the coupling between the perception of size and distance. Size-constancy is a perceptual phenomenon where an object is perceived of the same size regardless of its distance and its visual angle subtended by the object in the retina. This phenomenon is imprinted by experience so that when we approach or walk away from a familiar object, we do not perceive a change in its size despite it becomes greater/smaller on the retina. This is a consequence of the context, where cues like linear perspective and the relative size of objects around influence the perceived distance of the object. Thus, it is possible to assess the perception of distances, using the relative size of a familiar object as reference. Taking advantage of this phenomenon, the researchers studied the perception of size-constancy in a CAVE system, requesting participants to estimate the relative size of a familiar object located over a virtual table at different distances and using a physical reference (Figure 3.3). The reference object is placed at one side and at the same height of the virtual table. Then, some virtual replicas of the reference object are presented with exaggerated dimensions, where the table acts as the context to estimate its distance. Subjects scaled the virtual replicas until their size perceptually matched the size of the physical reference. These replicas were located at different target distances strategically selected to match the zero, positive and negative stereoscopic parallax conditions. Under natural conditions, it is not expected to find significant differences in size perception between target distances. However, the performance could be different in VR due to the amount of V-A conflict induced by the display, which in turn induces an apparent shortening of the virtual table by depth compression. Thus, the test is a good method to assess the perception of distance indirectly using the

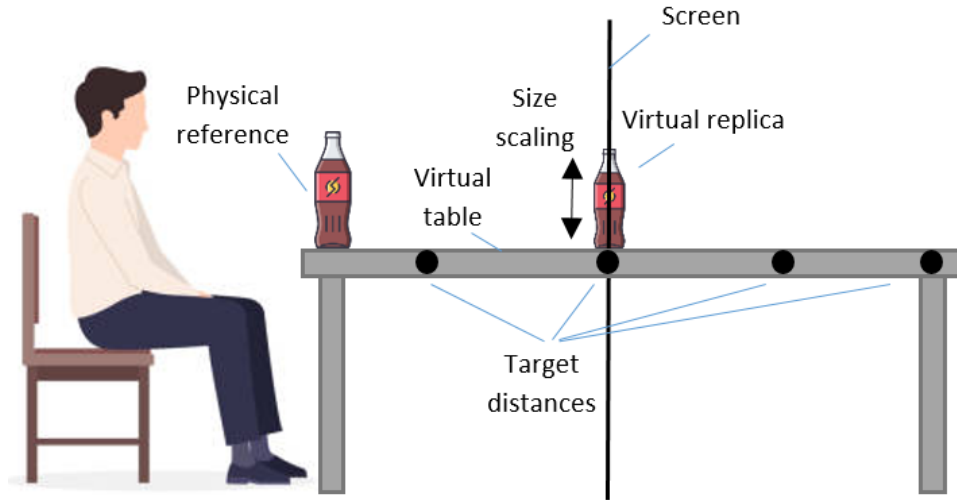


Figure 3.3: Size-constancy test

Participants must estimate the relative size of a virtual object located at different distances using a physical reference.

perception of size.

In their CAVE system, the researchers found that, when good perspective cues are provided, participants tend to slightly overestimate the size of the replicas but they develop size-constancy. There were no significant differences in the perception of size between virtual replicas compared with the physical reference. This result implies that overall, subjects perceived the length of the virtual table a little bit smaller than its actual length. The performance was evaluated using a basic measure called *Size-Ratio*, which represents the estimated size of the virtual object compared to the correct size of the physical reference object (Equation 3.1). A ratio of 1 means that the subject perceives the size (and the distance) of the object perfectly, a *ratio* < 1 means that its size is underestimated (or its distance is overestimated), and a *ratio* > 1 means that size is overestimated (or its distance is underestimated). Indeed, the performance can be predicted based on the size-distance-invariance hypothesis (SDIH), $s = d \tan(\alpha)$, where s is the perceived size, d is the distance to the object and α its visual angle (Kilpatrick and Ittelson (1953)). In their last study, Luo et al. (2007) found that underestimation of distances and the corresponding overestimation of size tend to increase as the V-A conflict increases, either in positive- or negative parallax. Also, the researchers found that the perception of size was more accurate when the virtual object was positioned at the screen plane (zero

parallax where V-A conflict is zero), which confirms the influence of the focal distance factor and the vergence-accommodation mismatch issue. Consequently, these characteristics make the size-constancy test an interesting assessment method to study.

$$SizeRatio = EstimatedSize/CorrectSize \quad (3.1)$$

To determine if this test was suitable for the proposed framework, we must fulfill two requirements: (1) It can be adapted to other kinds of displays, and (2) the influence of the vergence-accommodation mismatch issue on the perception of distances can be characterized. As a perceptual matching method, the size-constancy test is more suitable for LIPDs than HMDs, because the perceiver is able to see the physical reference and its virtual replicas simultaneously. To make this test suitable for HMDs, we had to provide a method to allow subjects to use the physical object as a reference, without the need of raising and lowering the headset. Thus, we proposed to use the sense of proprioception rather than the sense of sight to perform this task. The objective was to allow subjects to touch the physical object, get an overall idea of its dimensions and use their own body as a reference to perform their estimates. Because proprioception is not influenced by the display, a conflict between the motor/efferent cues and the visual stimulus will be induced. We hypothesize that in the HMD, the perception of size (and distances) will be impacted more negatively than in the CAVE, due to the greater susceptibility of this display to induce depth compression. To validate these ideas, we adapted the size-constancy test and performed a comparative study between a CAVE and an HMD. The next section describes the adaptation of the test and the performed study.

3.2.2 Comparing Size-Constancy between HMD and CAVE

We performed a study to analyze the suitability of the size-constancy test. We compared subjects' performance between a current consumer-oriented HMD (HTC Vive CV1) and a LIPD (a four-wall CAVE system, 3 m X 3 m horizontally and 2.67 m in height, with rear projection on three walls and front projection on the floor). Alike the original tests, we forced subjects to use their sense of proprioception to estimate the size of the physical reference, instead of using their sight. Hence, to enact the proprioceptive cues, we changed the virtual table for a physical one, and we replicated it virtually, allowing participants to touch its surface, sense its borders and get an overall idea of its dimensions. The physical table in which the reference object was located had a length of 50cm. We

3.2. PRELIMINARY STUDIES

calibrated its position and extended it virtually 3 times its length, creating the illusion of a long table. Similarly, the reference object was calibrated, located in a horizontal position, and adhered at the center of the table at a fixed distance from the border, so that subjects can manipulate both and use their body as a reference (Figure 3.4 - right). Proprioception was enacted at the moment participants established a relationship between the apparent length of the table, the proprioceptive size of the object, and the length of their arms. This relationship was later used to estimate the size and location of the "invisible" reference object in relation to their own body. To maintain both conditions similar, in the CAVE we did not allow subjects to see the physical reference and they must also use their sense of proprioception (Fig. 3.4 - left). The object was covered with a styrofoam sheet before participants began the test, so they could not see its actual size. Similarly, in the HMD condition, the reference object was put in place only after the participant wore the headset and sat on the table.

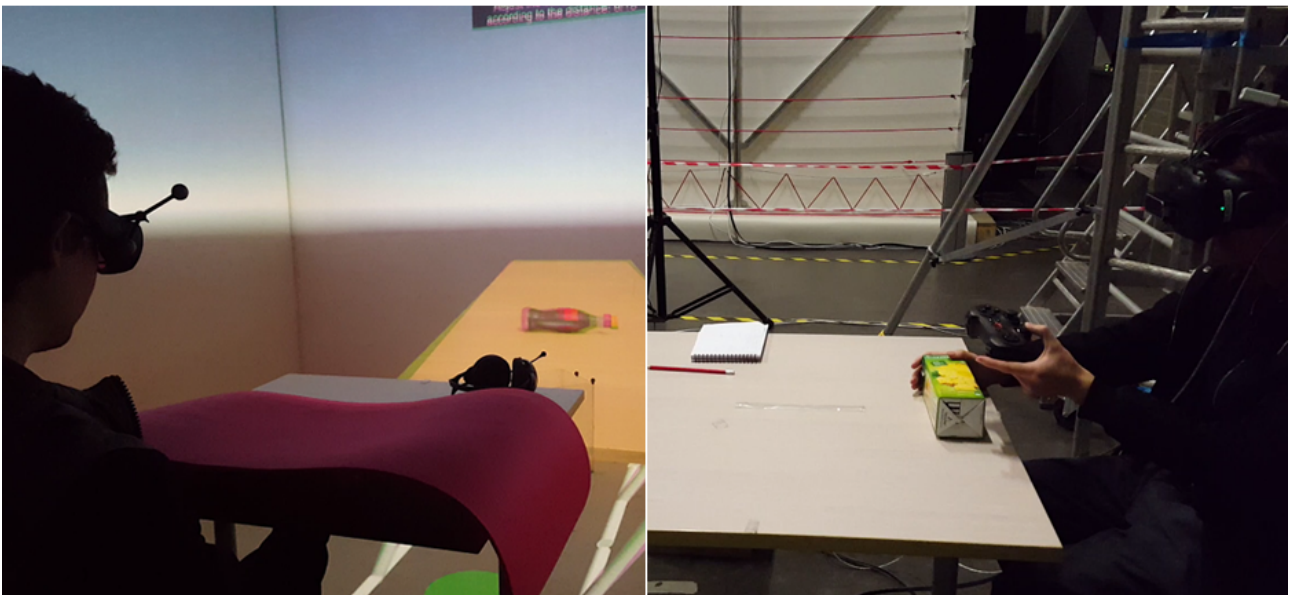


Figure 3.4: Adapted size-constancy test

Left: Cave condition. Right: HMD condition.

3.2.2.1 Hypotheses

Our hypotheses were as follows:

- H1: SizeRatio responses will reflect a greater size-constancy in the CAVE compared with the HMD.

3.2. PRELIMINARY STUDIES

- H2: SizeRatio responses will be influenced by the size of the induced V-A conflict in both displays.

3.2.2.2 Method

We designed a within-subjects experiment where subjects performed the test in both environments with counterbalance order between subjects. To prevent bias with the first environment, we alternated the physical object between a soft drink bottle and a juice box. The virtual replicas were created to match the dimensions of the physical ones. Thus, we requested subjects to estimate the perceived size of each replica located at 0.5m, 1.5m, and 2.5m of the physical reference. These distances represent the zero, positive, and negative parallax conditions, so that subjects were located at approximately 1.5m of the screen in our CAVE. Although we did not measure the focal distance in the HMD, we estimated that the focal distance of HTC Vive’s CV1 lenses was inside this range. Once the distance stimuli were set, subjects adjusted the perceived scale of the replicas six times at each target distance with aleatory order between trials. For each trial, the virtual replica was presented with an exaggerated dimension equivalent to 25% or 400% of its actual size. In short, this gave us a configuration of 6 adjustments x 3 target distances x 2 objects x 2 VEs conditions.

Participants

Eight subjects (all males, $M = 22.85 \pm 1.06$ years old) participated in our experiment. All participants signed a letter of consent reporting normal vision condition and good health at the moment of the experiment, without any previous history of relevant diseases.

Materials

Subjects performed the test in the four-walls CAVE and the HMD. In the CAVE, subjects’ heads were tracked using an ART tracking system and they used a Flystick as the input interface. In the HMD condition, we used an HTC Vive CV1 as display and their heads were tracked using a Lighthouse tracking system. Similarly, subjects adjusted the scale of the replicas using the Vive controllers as the input interface. Both environments were created with similar rendering conditions using the Unity3D game engine. We applied forward lighting for all materials with support for dynamic shadows. A realistic texture with a wood pattern was provided for the table, which had similar characteristics to the physical one, and we did the same to the virtual replicas.

Procedure

In the HMD condition, the procedure was as follows: First, subjects were requested to sit comfortably on a chair in front of the physical table. Then, they were presented with the physical object, either the soft drink or the juice box. After calibrating the subjects' IPDs and wearing the headset, the reference object was oriented perpendicular to the table, and fixed at a distance of 10 cm from the border, so that subjects could reach it, manipulate it with their hands and get familiar with its size. Then, we requested participants to visualize and touch the virtual table for some seconds to get a spatial impression of its dimension, and also to touch the reference object, to get familiar with its size. This familiarization task was performed for one minute. In the CAVE condition, the object and the subject's hands were covered with a Styrofoam sheet, so they could visualize the table but not the reference object. In both conditions, when they got an idea of the object's size, we presented subjects with the first virtual replica with exaggerated dimensions (either 25% or 400% of its physical size) and we requested subjects to scale its size according to the perceived distance, using the physical object and the table as references. Once subjects felt comfortable with the current size, they pressed a button, which aleatorily changed the target distance and prepared the replica for the next trial. Once all trials were performed, they repeated the experiment in the second display but with a different reference object.

3.2.2.3 Results

Results of the adapted size-constancy test are shown in Figure 3.5. Using a paired sample t-test, we found that the perception of size was significantly different between the CAVE and the HMD at 0.5m ($M = 1.28 \pm 0.10$, $M = 1.15 \pm 0.10$, $t(7) = 5.680$, $p = 0.01$), at 1.5m ($M = 1.26 \pm 0.06$, $M = 1.08 \pm 0.11$, $t(7) = 3.046$, $p = 0.023$) and at 2.5m ($M = 1.18 \pm 0.16$, $M = 0.95 \pm 0.16$, $t(7) = 3.42$, $p = 0.014$). Contrary to our expectations, the perception of size was more accurate in the HMD condition (yellow error bars) than in the CAVE condition (blue error bars). The overestimation of size in the CAVE condition was even greater compared with the results of the study by Kenyon et al. (2007) and Luo et al. (2007) (green error bars). However, the results suggest that H2 holds, there is a greater tendency to size-constancy in the CAVE compared with the HMD, which implies that these displays are more susceptible to the depth compression phenomenon. Regarding H2, our results suggest that this hypothesis partially holds: at negative parallax (0.5m), SizeRatio estimations were greater compared

3.2. PRELIMINARY STUDIES

with the estimations at zero parallax (1.5m). However, Size-Ratio estimations at zero parallax were greater when compared with the estimations at positive parallax (2.5m), a result not consistent with Kenyon et al. (2007) and Luo et al. (2007). We believe that this could be a consequence of using different sensory modalities, i.e, sight vs proprioception.

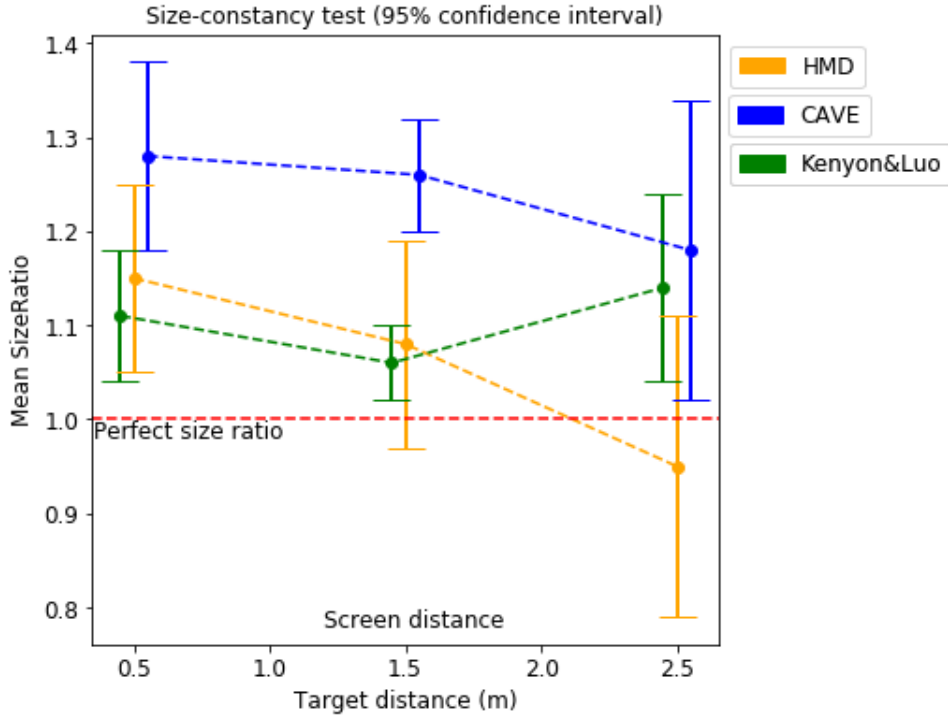


Figure 3.5: Results of the adapted size-constancy test (CAVE vs HMD)

Subjects tend to overestimate the size of the object in both conditions, but there is a greater tendency to size-constancy in the CAVE than in the HMD.

3.2.2.4 Discussion

We validated the suitability of the size-constancy test for the proposed framework. Even though subjects' performance was not completely as expected, some results suggest that this test is a good candidate assessment method. Subjects size estimations reflect partially the influence of the vergence-accommodation mismatch issue and the focal distance factor. As Kenyon et al. (2007) and Luo et al. (2007), our results indicate overestimation of size and underestimation of distances at negative parallax (at 0.5m). This can be confirmed following the SDIH ($s = d \tan(\alpha)$), a decrease in perceived distance d can only be compensated by an increase in perceived size s (the visual angle is assumed

3.2. PRELIMINARY STUDIES

to be constant for size-constancy). Also, the subject's performance improves at zero-parallax (at 1.5m) when the virtual replica is located at the screen plane and the V-A conflict approaches zero. Unfortunately, at the largest distance our results were the opposite to the expected, showing a decrease in size overestimation, a result not consistent with Kenyon et al. (2007) and Luo et al. (2007) and the evidence found in the literature. However, we believe that this could be a consequence of the use of proprioception instead of sight to perform the distance estimates. Proprioception, as a cue to estimate distances, is less precise for distances greater than the personal space. This limitation is evident in subjects' performance at that distance, where there is an apparent improvement in the perceived SizeRatio, but with a significant increase in variability in both display conditions. Thus, we considered important to limit the target distances inside a range where participants can discriminate the differences in size with relative accuracy. Hence, due to these results, particularly at short distances, we are optimistic about the suitability of the test.

An unexpected result was that the overestimation of size was greater in the CAVE condition than in the HMD condition, which is contrary to the evidence about the greater susceptibility of HMDs to the vergence-accommodation mismatch issue. An explanation of this unexpected difference could be the distortions induced due to the simultaneous visualization of the physical table and the virtual table in the CAVE condition. The illusion of fusing a physical object with the virtual one in a convincing way was particularly harder to achieve, they not only look different but also the illusion was easily broken, as subjects move their heads to get another point of view. An alternative solution could be the use of a glass table so that its surface can be fused easily as visible light passes straight through it without being absorbed or reflected. Despite this greater overestimation of size, the variation in the perception of distance was smaller than in the HMD, which confirms that these displays are less susceptible to depth compression. Another problem is that by relying on proprioception, we are sacrificing the most important advantage of the CAVE and it is the possibility of seeing the virtual and physical references simultaneously. This decision was made to emulate the conditions found in the HMD, but under natural conditions, we do not interact with invisible objects, and seeing our hands or our body is fundamental for tasks in the personal space. We consider that providing these visual cues is important for size-constancy, because they not only allow subjects to use the same sensory modality but also provide them with a greater sense of embodiment. Indeed, studies agreed that the stronger the sense of embodiment, the stronger the use of the body size as a fundamental metric to scale the

perceived size of objects (Ogawa et al. (2019), Banakou et al. (2013), Patterson (2019)). According to the authors, visualizing our hands is fundamental because it also influences the perceived size of objects.

3.2.3 Blind-Walking Test for HMDs

Blind walking is a popular method to assess the perception of distances in VR. In this test, a stimulus is located on the floor and at some distance from the subject. Then, he/she is requested to visualize the stimulus and estimate its distance from some seconds. Once the subject is confident about his/her estimation, he/she must walk toward the location of the stimulus under non-visual conditions. Under natural conditions, several studies have shown the absence of systematic error when walking to targets up to 20m away (see Loomis et al. (2003) for a review). Typical studies assess subjects' perception of distances before and/or after adaptation, where walking with continuous feedback some minutes before performing the test has shown an improvement in the perception of distances. As it was discussed before, when distance perception is assessed by motoric tasks, cues such as the optic flow and motion parallax induce a motor calibration with the attempt of adjusting subjects' perception of distances on moment-to-moment analysis (action-influencing perception). According to the recalibration hypothesis (Rieser et al. (1995), Kelly et al. (2013)), walking with continuous feedback in the VE modifies the action-perception feedback loop, which in turn rescales the perception of distances. This characteristic makes the test ideal to study the perception of distances in the action space and analyse the influence of the incorrect peripheral stimulation issue in HMDs

Although the blind walking test seemed an excellent candidate method for the proposed framework, assessing spatial perception in VR using this method is problematic. First, the test is only suitable for HMDs and not for LIPDs. Second, the greater the target distance, the greater the differences between displays, but targets up to 20m away are difficult to achieve in conventional VE setups, and even more difficult due to the spatial restrictions in our lab. Thus, to overcome these limitations, we explored the idea of performing blind walking tests using a modern treadmill, where subjects can walk larger distances in situ. Very few studies have studied the perception of distances in VR using treadmills (Li et al. (2021), Bossard et al. (2020), Santillán and Barraza (2019), Witmer and Sadowski Jr (1998)). By using a variety of displays and different experimental protocols, their results showed underestimation of distances similar to the ones found in VR but with great variability between subjects. We hypothesized

that subjects using VR displays with greater FOV and better optics would perform better than more limited displays. Hence, we designed an experiment to compare the performance between two vendors' HMDs with different capabilities.

3.2.4 Comparing Blind-Walking between HTC Vive vs Samsung Gear VR

How the size of the FOV and the lenses optics influence the perception of distances in HMDs? We performed a study to compare the perception of distances between a HTC Vive CV1 (110° horizontal FOV, 52 mm fresnel lenses) and Samsung Gear VR (96° horizontal FOV, 42 mm biconvex lenses). We requested subjects to estimate the distance of five targets located at 5m, 7.5m, 10m, 12.5m, and 15m using blind-walking tests. We requested subjects to walk using a Virtuix Omni Treadmill, a modern consumer-oriented treadmill that allows omnidirectional walking. This device mimics locomotion by sliding the feet on a slightly concave surface. Although this interaction is quite different from natural walking, some preliminary tests showed that subjects learned and adapted quickly to the locomotion mechanism. To facilitate this process, we performed an adaptation phase, where we requested subjects to walk freely for five minutes, before starting the tests. Figure 3.6 depicts the two conditions studied for the proposed blind walking test. We hypothesized that subjects' perception of distances would be better in the HTC Vive condition compared with the GearVR condition because its display provides a greater FOV and optics. Even though GearVR's biconvex lenses do not suffer from god rays and screen door effects (typical distortions of fresnel lenses), they have a small optical power, which potentially can reduce the FOV and the angular resolution (depending on the eye relief). In this sense, these characteristics make the HTC Vive better than the GearVR for spatial perception.

3.2.4.1 Hypotheses

Our hypotheses were as follows:

- H1: Subject perception of distances will be better in the HTC Vive condition compared with the GearVR condition.
- H2: Significant differences are expected as the distance increases between both displays.

3.2. PRELIMINARY STUDIES



Figure 3.6: Blind-walking-tests using the Virtuix Omni treadmill

Subjects performed the blind walking tests using a Virtuix Omni Tradmill. Left: HTC Vive condition. Right: Gear VR condition

3.2.4.2 Method

We designed a within-subjects experiment where subjects performed the test in both displays with counterbalance order between subjects. We requested subjects to estimate the relative distance of a virtual stimulus located at 4m, 8m, and 1.2m of the subject. Subjects performed the test in the GearVR and the HTC Vive, where half of them start with the first one, and the other half with the last one. A total of three trials at each target distance was requested with aleatory order between trials.

Participants

Sixteen subjects (11 males, 5 females, $M = 23.1 \pm 2.1$ years old) participated in our experiment. As in the previous study, all participants signed a letter of consent informing normal vision and good health.

Materials

Participants performed the tests using the Virtuix Omni treadmill. We set the default walking

3.2. PRELIMINARY STUDIES

speed provided by the platform, and we calibrated the walking interaction as "forward-only", where motion depends on the direction of the subject's waist, as it was recommended by beginner users. Alike the HTC Vive, the GearVR does not provide positional head tracking and an input controller; hence, we implemented these functions, using the Unity3D Networking System. To achieve this integration, we implemented the Vive Controller as a remote object and we implemented remote positional tracking by attaching a tracker at the headset, using the Light House Tracking system. Thus, subjects using the GearVR could move their heads, getting visual feedback; and use the Vive controller as an input device, just like if they were using the HTC Vive.

Procedure

The procedure was based on a typical blind-walking test. Participants were presented with the treadmill, set in the platform, and instructed about how the walking interaction worked. Then, after setting the headset, they were requested to walk for some minutes in the VE, until they felt comfortable walking. Then, subjects were requested to stand at a particular point and look ahead. A virtual stimulus was presented at some particular distance, and subjects were requested to look at it and estimate its distance. Once they felt confident with their estimation, subjects pressed the trigger button, which darkens the screen. After once they could not see anymore, they were asked to walk toward the virtual stimulus and pressed the trigger button once they believe they have reached the target. The position was captured and the error was measured. Once subjects' performance was measured, they were assisted to return to the starting point, the screen was activated again, and they were presented with the next virtual stimulus located at a different distance. Finally, after performing all the trials, the subject rested for three minutes, and they were requested to repeat the test with the other display.

Results

Results of the blind-walking test are shown in Figure 3.7. A paired sample t-test was conducted to compare the perception of distances among the three target distances. Overall all subjects underestimated the distances for all trials and regardless of the display condition. There was a significant difference in the underestimation of distances at 8m between the Vive condition ($M = 0.60m, SD = 0.29$) and the GearVR condition ($M = 1.05, SD = 0.54$), $t(47) = -2.06, p < 0.01$; while no significant difference was found at 4m between the Vive condition ($M = 0.41, SD = 0.27$) and the GearVR condition ($M = 0.35, SD = 0.19$), $t(47) = -2.6, p = 0.42$; and at 12m between the Vive condition

3.2. PRELIMINARY STUDIES

($M = 1.28, SD = 0.39$) and the GearVR condition ($M = 1.49, SD = 0.85$), $t(47) = -1.01, p = 0.32$. Thus, results suggest that H1 partially holds, subjects performance was significantly better in the Vive condition compared with the GearVR condition, but only at 8m. However, we cannot confirm that the subjects using Vive performed better than subjects using the GearVR, due to the great variability in subject responses. In contrast, results indicate that H2 does not hold; except at 8m, distance seems no to be a factor that causes differences in performance between displays.

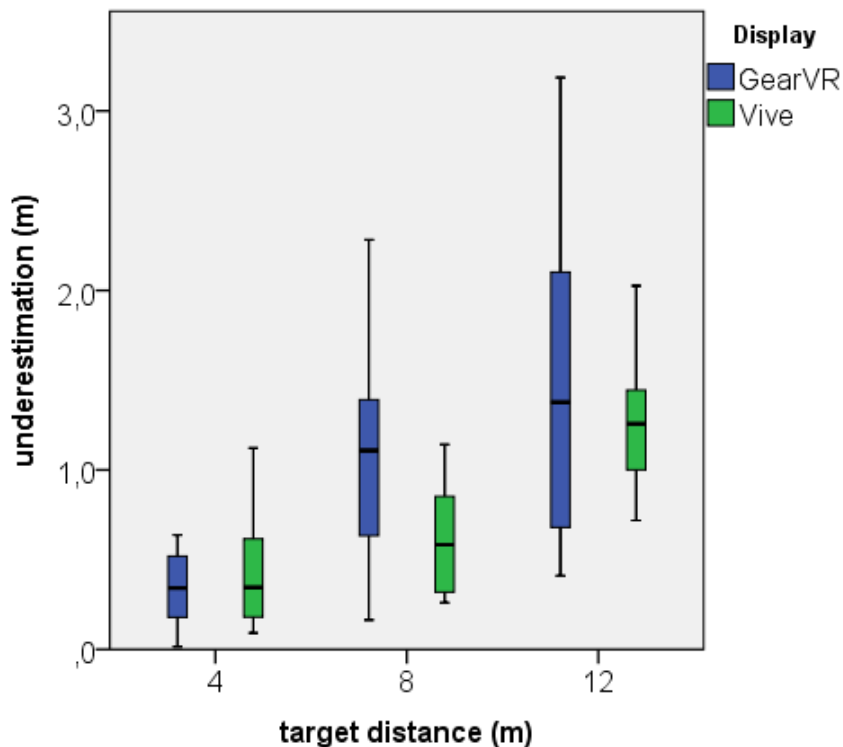


Figure 3.7: Results of the blind-walking tests using the Virtuix Omni treadmill

Distance perception was significantly different between the GearVR condition and the HTC Vive condition at 8m.

3.2.4.3 Discussion

We performed a study about the influence of the incorrect peripheral stimulation issue in the perception of distances during blind-walking tasks. Despite results were not the expected, it is particularly interesting the significant differences in performance at 8m between the HTC Vive and the GearVR condition. However, subjects' performance at 4m and 12m, does not show differences between displays.

3.3. SUMMARY

It is possible that at shorter distances, differences are not significant, as very few steps of similar length are required. In contrast, we expected that differences to increase as the target distance increased, but no differences were found at the largest distance. A cause of this phenomenon could be the great variability between subject responses, which implies that most subjects were not confident about their distance estimates. This variability is not consistent with the studies about blind walking, where subjects can walk naturally and can estimate distances with relative confidence. Even though we allow subjects to acclimatize to the treadmill for some minutes, we notice that not all subjects could adapt easily to the platform. Unfortunately, there is no standardized treadmill acclimatization protocol and insufficient familiarization potentially confounds analyses. Indeed, the locomotion technique of the Virtuix Omni corresponds more to a proprioceptive/efferent task, as subjects must slide their foot to perform a step. Another cause is fatigue, many subjects manifested that walking in the platform was tiring, which could explain the great variability at 12m. Finally, the great variability in the GearVR condition could be a consequence of the positional tracking implemented in the GearVR condition, where correct sensor fusion was not achieved, as the tracking sensor was not perfectly synchronized with the GearVR sensors.

Based on the limitations found, we cannot consider the current test as a good transversal assessment method for the proposed framework. Despite that we found some significant differences in underestimation of distances between displays, the treadmill is an important factor that could introduce many biases in subject responses. In addition to subject's adaptability to the novel locomotion, studies have shown that perception of distances in the action space relies more on the visual cues (motion parallax, optic flow) than in the proprioceptive/efferent cues (Witmer and Kline (1998) Santillán and Barraza (2019), Li et al. (2021)). Thus, the unnatural movement of sliding one's feet on the Virtuix Omni treadmill might have contributed little to the perception of distance and even worse, have distracted subjects from attending to changes in the visual stimulus due to the physical effort. Hence, we considered important to perform another study allowing subjects to walk naturally.

3.3 Summary

We presented an overview of the proposed framework and we studied the suitability of the size-constancy tests and the blind-walking tests as transversal assessment methods. The proposed framework is based on the idea that if we are able to characterize the phenomenon of depth compression

3.3. SUMMARY

induced by a display (using an assessment method), we can influence their perception of distances positively using composition (by playing with the rules of perspective). The preliminary studies showed that the selected assessment methods can measure the perception of distances in VR with relative accuracy and they can characterize the influence of the displays issues on spatial perception. We can extract from these studies the following conclusions:

- The size-constancy test is ideal to study the influence of the vergence-accommodation mismatch issue on the perception of distances. The underestimation of distances and the corresponding overestimation of size increase as a function of the size of the V-A conflict, showing a greater effect for objects at positive parallax and evidencing depth compression. Also, the results showed that LIPDs like the CAVE are less susceptible to the phenomenon of depth compression. Although our results were not fully consistent with Kenyon et al. (2007) and Luo et al. (2007), there was a greater tendency to size-constancy in this our CAVE system, which indicate that subjects perceived the size of objects more or less uniformly between target distances.
- The blind-walking test is ideal to analyse the influence of the incorrect peripheral stimulation issue for distances in the action space. The greater the distortions induced in the peripheral vision by the display, the greater the underestimation of distances. Despite the issues identified in our study related to the treadmill, we believed that significant differences will be found if we allow subjects to walk naturally, but this restriction limits the suitability of the test only for HMDs. Finally, even if it is not possible to perform blind-walking tests in CAVE systems, it is reasonable to think that these displays provide a more accurate perception of distances as they cause less distortions in peripheral vision.
- Developing transversal assessment methods is challenging. Comparing very heterogeneous displays, such as HMDs and LIPDs, was not an straightforward task due to the particularities of each assessment method and the restrictions of each technology. We introduced for the first time the size-constancy for HMDs, but the validity of comparing it with the CAVE condition is debatable, as subjects does not see the physical reference and the table in the same conditions. Similarly, using a treadmill as a substitute for natural walking was not a good idea for the blind-walking tests. An alternative method could be blind-triangulated pointing (Fukushima et al. (1997)), a test similar to blind-walking that does not require a large locomotion space.

3.3. SUMMARY

Thus, we would like to emphasize that the proposed framework is not restricted to these two assessment methods, as other can be explored depending on the restrictions of the display and the nature of the spatial task.

3.3. SUMMARY

Chapter 4

Part I. Characterizing spatial performance

In chapter 2, we presented two spatial perception tests that were selected as transversal assessment methods, the size-constancy test, and the blind-walking test. The proposed framework is based on the idea that if we are able to predict the underestimation of distances in VR, we can influence positively spatial perception. To predict subjects' performance, we analysed how the different display factors associated with the display issues influence the perception of distances. The accommodation-mismatch issue and the incorrect peripheral stimulation issue are the main causes of the incorrect perception of distances in VR, but the degree of underestimation or overestimation varies depending on some associated factors. In terms of the vergence-accommodation issue, the interaction between the *focal distance* and the *stereoscopic parallax condition* of the target stimulus, determine the degree of underestimation or overestimation. Regarding the peripheral stimulation issue, the interaction between the *size of the FOV* and the induced lenses distortions determine the degree of error on the perception of distances. Figure 4.1 depicts the main determinant factors that influence spatial perception in VR. Based on these four factors, we developed a model to characterize the induced depth compression in a display, based on some assumptions from the studies and the evidence found in the literature.

4.1 Determinant factors

4.1.1 Focal distance

In the previous chapter, we discussed how the phenomenon of depth compression was directly associated with the display's focal distance, which limits the functional depth range where vergence

4.1. DETERMINANT FACTORS

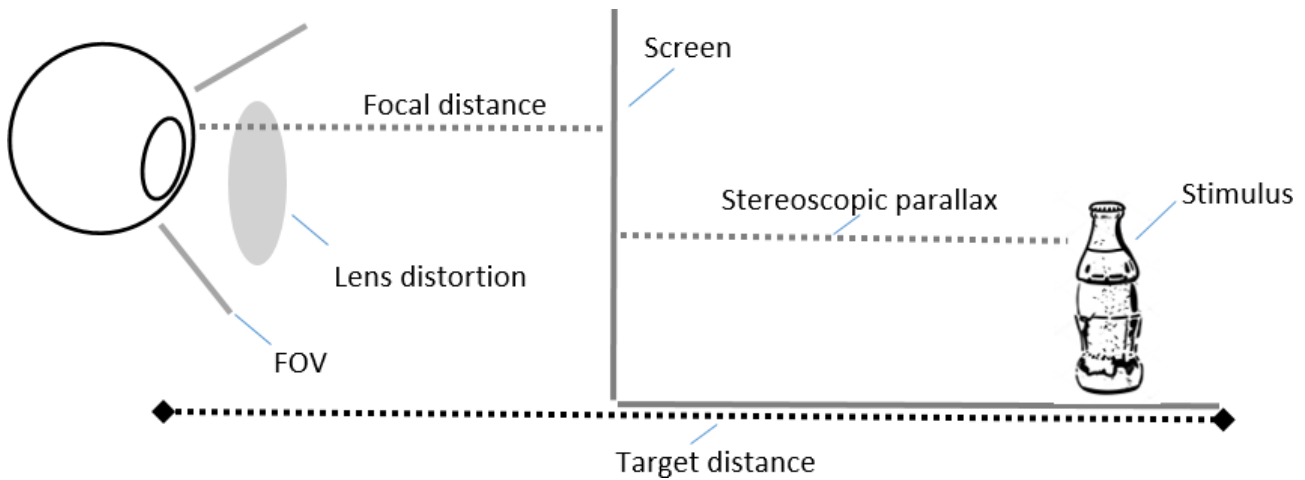


Figure 4.1: The main determinant factors associated with the display issues

The interaction between the size of the FOV, the focal distance, the stereoscopic parallax condition and the distortions of the lenses determine the degree of underestimation or overestimation in the perception of distances.

and accommodation can be used to perceive depth. This range depends on the available depth of field (DOF) and the zone of comfortable fusion induced by the display (Shibata et al. (2011)'s zone of comfort). If the focal distance determines the amount of depth compression, our previous studies using the size-constancy tests raise an important question: *Why spatial performance was very different between the CAVE and the HMD in the size-constancy test, in the condition where their focal distance was similar?* The answer is how HMDs are designed. As it was mentioned before, HMDs have lenses with a fixed focal distance that creates a virtual image at infinity, altering the way the DOF naturally expands and making these displays more susceptible to the depth compression phenomenon. In contrast, LIPDs do not require the use of lenses, so there is a more natural relationship between the focal distance and the induced DOF. Thus, HMDs can alter significantly these ranges as the display is very close to the perceiver's eyes. Indeed, the evidence suggests that the DOF increases inversely with pupil diameter (Hollingsworth et al. (2008)), which implies that HMDs induce a greater DOF compared with the human natural DOF.

Figure 4.2 shows a comparison of the induced DOF of a modern HMD compared with the human natural DOF. For example, the HTC Vive Pro have a focal distance of around 0.75m (Reddit.com (2017)), which implies that the minimum DOF which falls in this range is $DOF = 1/0.75m = 1.33D$. Because the DOF expands with the focal distance, focusing at 0.75m in the HMD is equivalent to

focusing further under natural conditions (we used the estimation of 2m by Shibata et al. (2011) as a reference). Our eyes are effectively focused at 0.75m, but the induced retinal blur (the cue associated with accommodation) corresponds to the one induced when focusing at an object at approximately 2m. Thus, focusing at an object at shorter distances in a HMD (in the personal space), causes a DOF similar to focusing at larger distances (in the action space), increasing the amount of conflict between vergence and accommodation. In contrast, LIPDs like the CAVE do not require the use of lenses, so we can assume that they induce a DOF similar to the found under natural conditions. Because there are not severe optical distortions between the perceiver’s eyes and the screen, the eyes are not only effectively focused at the screen but the induced retinal blur is less distorted. Thus, we will introduce the first assumption of the proposed framework: **The amount of depth compression is directly related with e focal distance and the induced DOF.** Hence, heterogeneous displays with the same focal distance can not have the same functional depth range, and the amount of depth compression will depend on the induced DOF.

4.1.2 Stereoscopic parallax condition

The second factor necessary to predict subject’s spatial performance is the stereoscopic parallax condition, the location of the stimulus in relation with the screen, which also determines the amount of divergence between vergence and accommodation. Our previous studies using the size-constancy tests and the evidence from the literature have shown that greater the size of the V-A conflict, greater the error in the perception of distances. However, an important question was originated from previous studies: *Why there is a greater tendency to size-constancy in the CAVE compared with the HMD, if the location of the stimulus was the same ?*. The answer is again the way HMDs are designed and the nature of the DOF induced by these displays. Due to the incongruence between the focal distance and the induced DOF explained before, HMDs induce a more severe V-A conflict compared with LIPDs.

According to the hypothesis by Vienne et al. (2018), the underestimation of distances should be proportional to the V-A conflict size, which can be calculated by subtracting the inverse of the accommodation from the inverse of the vergence: $C_{va} = \frac{1}{V_d} - \frac{1}{A_d}$. There are two important elements about this equation: (1) It could give rise to positive conflicts ($V_d < A_d$) and a negative conflicts ($V_d > A_d$), as they are in function of the focal distance rather than in the function of the perceiver; and (2) there are only two ways to reduce the conflict, at zero parallax (where vergence and accommodation

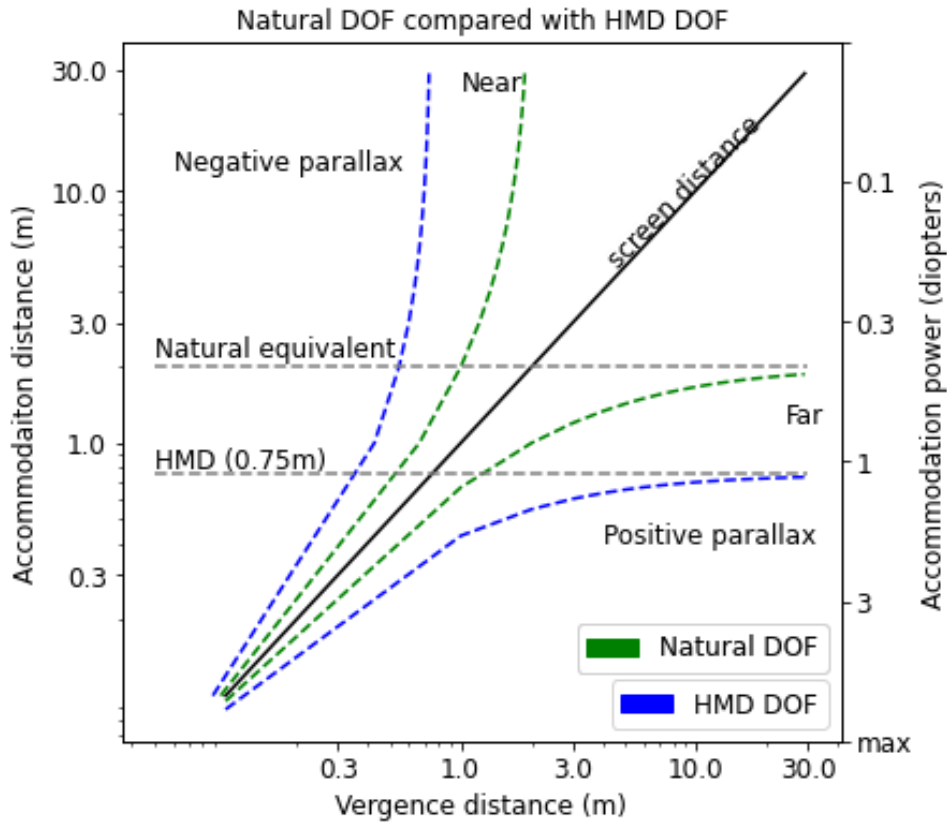


Figure 4.2: DOF of a modern HMD compared with the human natural DOF

The use of lenses in HMD causes the DOF expands more drastically than under natural conditions, so that focusing at 0.75m is approximately equivalent to focusing at 2m under natural conditions.

are equals), and by increasing the vergence distance (which in turns decreases the accommodation demands). However, Vienne et al. (2018)’s equation did not consider the influence of the DOF induced by the display on the perception of distances. If we introduce this element, so that HMDs focal distance is scaled to the corresponding natural DOF equivalent, two different patterns emerge. Figure 4.3 show a comparison of the V-A conflict size between a CAVE focused at 2m and a HMD focused at 0.75m. The chart depicts that while in the CAVE there is positive- and negative conflicts, in the HMD the conflicts are always negative (like if all objects were located at positive parallax). Also, the conflict is more severe in the HMD compared with the CAVE, a direct consequence of the greater difference between vergence and accommodation. These different patterns could explain why there are some reports of overestimation of distances in LIPDs at negative parallax (depending on the task), while distances are always underestimated in HMDs regardless of the location of the stimulus in relation

with the screen. Hence, we introduce the second assumption of the proposed framework: **the degree of underestimation in the perception of distances should be proportional to the V-A conflict size.**

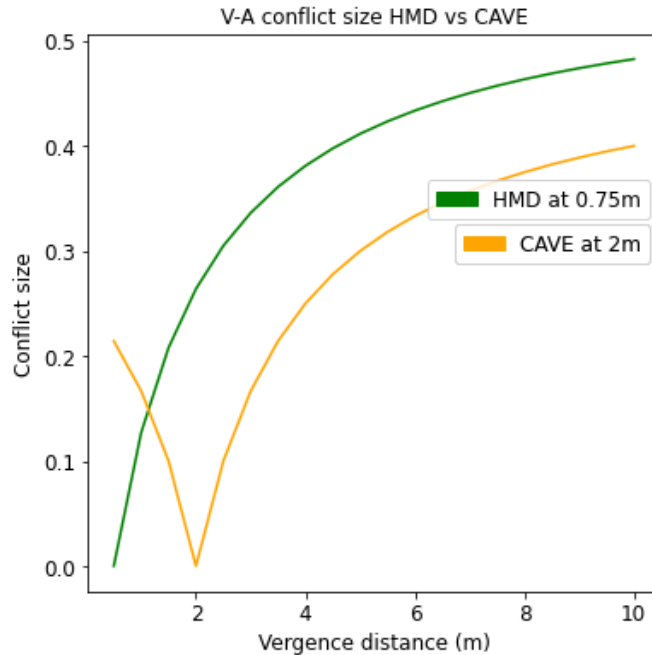


Figure 4.3: A comparison of the V-A conflict size between a HMD and the CAVE

While in the CAVE there are positive- and negative conflicts, in the HMD the conflicts are always negative, a direct consequence of the disparity between the focal distance and the induced DOF in this kind of display.

4.1.3 Limited FOV and lens distortions

A third factor that influences the perception of distances is the incorrect peripheral stimulation issue. Alike the previous factors, these factors involve motion and more specifically adaptation. Motion parallax is an important cue for depth perception, that influences the perception of distances and the motor recalibration process. In the previous chapter, we introduced that the perception of depth from motion parallax can be calculated with the motion pursuit law $\frac{d}{f} = \frac{d\theta}{d(\alpha-\theta)}$, where F is the fixation point, D is a point at a distance d of F , θ the horizontal/angular disparity between both points, and α the convergence angle (Nawrot and Stroyan (2009)) (Figure 4.4). In this equation, θ and α change over time, as the perceiver moves forward/backward of F . Under natural conditions, when the observer moves at a constant speed depth can be predicted based on the constant rate of

4.1. DETERMINANT FACTORS

change between the horizontal/angular disparity θ and the convergence angle α . However, due to the influence of the limited FOV and lens distortions, the motion pursuit law does not follow these rules anymore. Because the angular resolution does not vary uniformly in the peripheral vision, the point D moves at a faster rate in the visual field when it approaches to the limits of the FOV, changing the rate of change of the angular disparity θ and decreasing the perceived depth of the focal point f . Thus, the kind of lenses used by a display can impact severely the perception of depth from motion parallax.

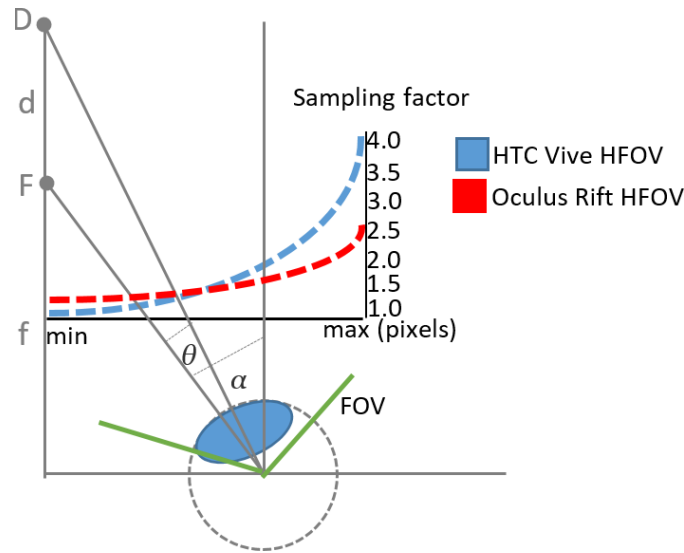


Figure 4.4: Influence of the Oculus Rift and HTC Vive lenses on motion parallax

The variation on angular resolution in the peripheral vision influences negatively the perception of depth from motion parallax.

Figure 4.4 shows a comparison of the influence of the HTC Vive (fresnel lenses) and the Oculus Rift (hybrid lenses) on the motion pursuit law. The effect is stronger in the HTC Vive, as its sampling factor is approximately 4 arcminute/subpixel in the peripheral vision compared with the approximately 2 arcminute/subpixel for the Oculus Rift. Point D not only moves faster in the HTC Vive, but also leaves the visual field faster compared with Oculus Rift. Hence, there could exist significant differences in perception of distances between HMDs, although both displays provide the same FOV. We can calculate influence of the lenses by approximating linearly the sampling factor in terms of the FOV size: $r = \frac{min_{ar} - max_{ar}}{fov/2}$, where min_{ar} and max_{ar} represents the min and max angular resolution respectively; and fov is the size of the FOV provided by the display. For example, the rate of change

4.1. DETERMINANT FACTORS

of HTC Vive with a FOV of 110° is $r = \frac{4.0-1.0}{fov/2} = 0.054$, which means that for each degree, a subpixel will move 0.054 degrees faster from the center of vision to the periphery. Similarly, for the Oculus Rift, the rate of change is $r = \frac{2.5-1.5}{fov/2} = 0.018$, for each degree, a subpixel will move 0.018 degrees faster from the center of vision to the periphery. By assuming that the distance between the fixation point F and the point D is constant, we can estimate the induced depth compression as a function of angular resolution using the motion pursuit law (Equation 2.4).

Figure 4.5 depicts the induced depth compression in response to the distortions of the lens for the HTC Vive and the Oculus Rift. At shorter distances, the underestimation is stronger because the distance stimulus is closer and the horizontal/angular disparity increases, moving faster in the visual field toward the peripheral vision. Conversely, at larger distances, the underestimation moves towards an equilibrium, as the horizontal/angular disparity with objects around is smaller and these objects move more slowly in the visual field. The graph also shows the difference between the HTC Vive and the Oculus Rift, where the best angular resolution of the Oculus Rift in terms of peripheral vision induces a smaller underestimation of distances. Similarly, distances are perceived as compressed but there is a fundamental difference compared with the previous display factors. Because the underestimation of distances tends to a equilibrium, the degree of underestimation should not increase as the target distance increases. Thus, we adhere to Kelly et al. (2014) hypothesis that adaptation not only induces a motor recalibration but also causes a rescaling of the perceived space. Distances are perceived compressed but more homogeneously compared to the nature of the depth compression induced by vergence-accommodation mismatch issue. Hence, the visual space is scaled homogeneously rather than compressed in depth. In this sense, our third assumption is straightforward: **Under adaptation, the underestimation of distances should be more homogeneous and inversely proportional to the angular resolution induced by the display in the peripheral vision.**

4.1.4 Linear depth compression

A question from the previous studies was if the depth compression phenomenon has a linear behaviour. Unfortunately, we did not take enough samples and the variability in subjects' responses did not allow us to get any conclusions. To answer this question, we refer to the Bayesian Integration model, a model that states that depth cues are combined in a linearly optimal form to produce a minimum variance estimate (MVE):

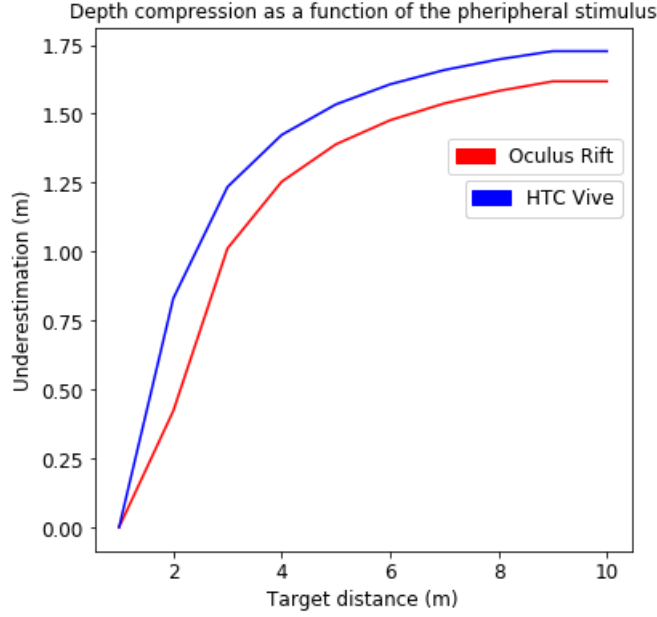


Figure 4.5: Depth from motion parallax as a function of the lens

Estimated depth compression as a function of the rate of change of the horizontal/angular resolution induced by the lens. The best angular resolution of the Oculus Rift induces a smaller underestimation of distances compared with the HTC Vive.

$$d = \sum_i w_i d_i \tag{4.1}$$

where d_i is the relative depth value from cue i , w_i , the inverse of its variance normalized, and d is the overall estimated depth (Jacobs (1999), Ghahramani et al. (1997), Knill and Saunders (2003)). A simplification of Equation 4.1 is given by:

$$d = w_{accommodation} d_{accommodation} + w_{vergence} d_{vergence} + w_{parallax} d_{parallax} + w_{pictorial} d_{pictorial} \tag{4.2}$$

where the depth cues are linearly integrated from the most "noisiest" to the most reliable (pictorial cues). Since every cue is relevant and contributes to the overall perception of depth, any cue whose depth estimates move away from the natural values, will induce a bias on the overall perception of depth. However, following Bayes's MVE rule, the overall perception of depth will maintain a linear behaviour. As the "noise" of the cue increases, its weight is reduced. Hence, the final assumption of

this framework is that **the underestimation of distances has a linear behaviour and can be predicted linearly.**

4.1.5 Summary of determinant factors and assumptions

We review the most important determinant factors associated with the display issues and their influence on the perception of distances in VR. The first assumption of the proposed framework was that the amount of depth compression depends on the display's focal distance and its induced DOF. Thus, HMDs are more susceptible to the depth compression phenomenon due to the incongruence between the focal distance and the DOF induced in this kind of displays. The second assumption was that the degree of underestimation/overestimation in the perception of distances should be proportional to the V-A conflict size. This statement simply means that the size of the V-A conflict can predict the amount of underestimation induced by a display. The third assumption is that the underestimation of distances should be more homogeneous and inversely proportional to the angular resolution induced by the display in the peripheral vision. This sentence just describes the influence of the distortions induced by the lenses in the perception depth from motion parallax. Finally, the last assumption is that the underestimation of distances has a linear behaviour and can be predicted linearly. Thus, these four assumptions form the core of the proposed framework.

4.2 Second study

In order to validate the proposed assumptions, we designed a second study to characterize the phenomenon of depth compression in VR displays and get more insights into the influence of the different display factors on spatial perception. We repeated the same preliminary studies using the size-constancy test and the blind-walking test, but this time only focused on HMDs, as these displays are more susceptible to depth compression phenomenon. We improved some of the methodological issues found in the preliminary studies to reduce the subject's responses variability, and characterize better the influence of the different display factors.

4.2.1 Size-constancy test for a HMD

In the previous chapter, we discussed a first adaptation of the size-constancy test to compare the performance between heterogeneous displays. We used a reductionistic approach, where we tried to

4.2. SECOND STUDY

emulate the conditions found in the HMD in the CAVE so that the display with "more capabilities" was somewhat downgraded to the "less capable" display. However, we considered that a holistic approach was better because it takes advantage of all the capabilities that the display can offer. The CAVE is ideal for this test because subjects can see the physical reference object and they count with the additional cues from their own body. Thus, to maximize the capabilities of the HMD, we provided a mechanism to allow subjects to get visual feedback of the physical reference (based on the perceived proprioceptive size) and visualize their hands, to support their distance/size estimates.

A set of different modifications were performed to the previous experiment. First, we reduced the target distances up to 1.2 m to have a good depth discrimination threshold. Although we did not measure the focal distance of the HTC Vive's Pro lenses, some reports suggest that its focal distance is between 0.70cm and 0.80cm (Reddit.com (2017)). Second, in addition to the possibility of touching the reference object, we provided a virtual representation of it. Third, we provided visual feedback of the subject's hands using a leap-motion controller. Figure 4.6 depicts the task performed at the three target distances with the additional reference cues. We hypothesized that these modifications would improve the perception of size/distance in the HMD and reduce the variability in subjects' responses. Also, if our first and second assumptions were correct, we expected to get depth compression (due to the influence of the fixed focal distance factor), and the underestimation of distances and the corresponding overestimation of size would increase proportionally to the size of the V-A conflict induced.

4.2.1.1 Hypotheses

Our hypotheses were as follows:

- H1: SizeRatios will be influenced by the amount of V-A conflict.
- H2: Participant's performance will show evidence of size-constancy.

4.2.1.2 Method

A similar experiment was conducted to investigate the effects of visual feedback, embodiment, and motion parallax. As in the previous study, we requested subjects to estimate the relative size of each replica but this time located at 0.4m, 0.8m, and 1.2m of the physical reference. A total of 5 trials at

4.2. SECOND STUDY

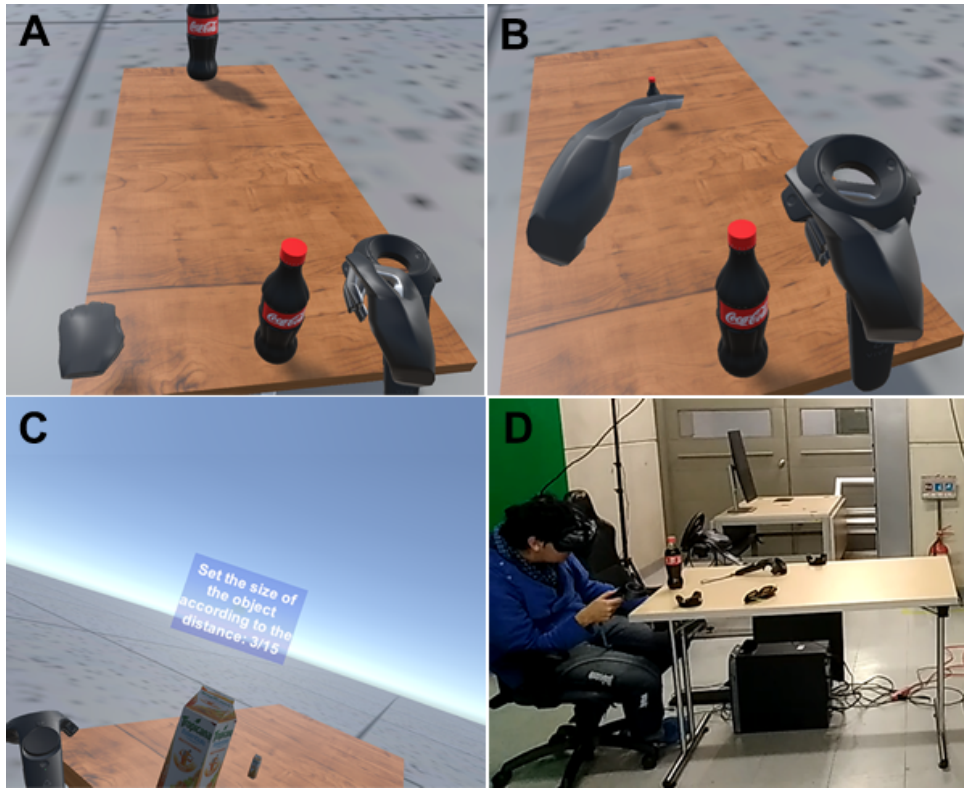


Figure 4.6: Improved size-constancy test for HMDs.

The task performed at the three target distances with the influence of visual feedback, sense of embodiment and motion parallax. A. 1.2m, B. 0.8m, C. 0.4m, D. Participant performing the task.

each target distance was requested with aleatory order between trials.

Participants

Ten subjects (8 males, 2 females, $M = 21.5 \pm 1.12$ years old) participated in our experiment. As in the previous study, all participants signed a letter of consent informing normal vision and good health.

Materials

Participants performed the tests using only the HMD. To provide visual feedback of the hands, a leap-motion controller was attached to the headset and we configured it with a 3D model of human hands.

Procedure

The procedure was also similar to the previous study. Participants sat comfortably on the chair in front of the table and they were presented with the physical object. When setting the headset, the

4.2. SECOND STUDY

reference object was put in the middle, oriented perpendicular to the table, and fixed at a distance of 10 cm from the border. The virtual representation of the object was also calibrated to match the location of the physical object and its dimensions. After perceptually calibrating and providing visual feedback of the reference object, we requested subjects to see their hands, look at the object and the table, touch them and get familiar with their size. This familiarization task was also performed for one minute. Once the familiarization was finished, we performed the test, encouraging subjects to move their head to get a different point of view, inducing motion parallax.

Results

Results of the second size-constancy test are shown in Figure 4.7. A paired sample t-test was conducted to compare the Size-Ratio among the three target distances. There was a significant difference for Size-Ratio between 0.8m ($M = 0.99, SD = 0.12$) and 1.2m ($M = 1.09, SD = 0.11$), $t(19) = -3.66, p = 0.02$; between 0.4m ($M = 1.13, SD = 0.078$) and 0.8m ($M = 0.99, SD = 0.12$), $t(19) = 4.79, p = 0.00$; while no significant difference was found for Size-Ratio between between 0.4m ($M = 1.13, SD = 0.07$) and 1.2m ($M = 1.09, SD = 0.07$), $t(19) = 1.66, p < 0.11$; Although overall, participants overestimated the size at the three target distances, there is an improvement compared with the previous test, not only the variability was reduced but also there is evidence that the use of additional references cues improves the perception of size/distance. The results also showed that the best performance was found at the estimated focal distance of the display (at 0.8m), while in the shortest and longest estimation, the distributions are skewed closer or further to the perfect SizeRatio respectively. These results indicate that H1 holds; the perception of size and distance is influenced by the amount of V-A conflict induced.

Regarding H2, to demonstrate the existence of size-constancy, we must provide support to the null hypothesis "perception of size is equal between target distances". Thus, we performed a two-sided t-test (TOST), an equivalence hypothesis testing procedure complementary to the paired sample t-test (Schuirmann (1987)). Results showed that size perception was statistically equivalent and not different between 0.4m and 0.8m, ($t(19) = 2.16, p = 0.021$), and not equivalent and statistically different between 0.4m and 1.2m ($t(19) = -1.66, p = 0.943$), and between 0.8m and 1.2m ($t(19) = 1.52, p = 0.072$). These result suggest that H2 partially holds. There exist evidence of size-constancy but only at shorter distances, size-perception is equivalent between 0.4m and 0.8m and significantly different at 1.2m.

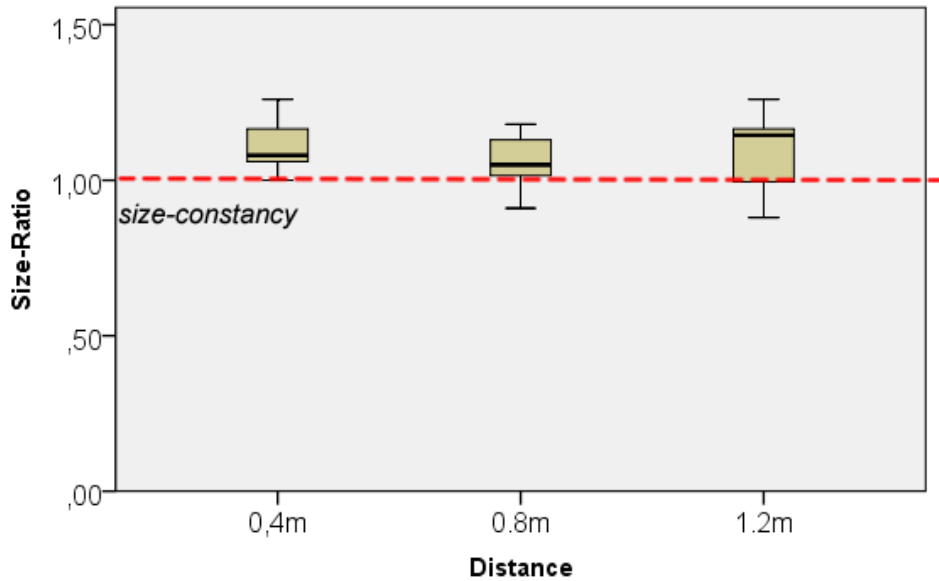


Figure 4.7: Result of the adapted size-constancy-tests for the HMD

Influence of visual feedback, embodiment and motion parallax on the perception of size-constancy in HMD. The use of these additional reference cues improved the perception of size and distance.

4.2.1.3 Discussion

Our results showed that, by enhancing the HMD with the inclusion of additional reference cues, a positive effect on the perception of distance and size is induced. The variability was reduced and subjects' performance became quite similar to the found in CAVE systems but with a greater error, a result that suggests that the degree of underestimation depends on the induced DOF (first assumption). Also, the performance is consistent with the expected results taking into account the influence of the vergence-accommodation mismatch issue and the focal distance factor (second assumption). At negative parallax (0.4m), the V-A conflict causes an overestimation of size and underestimation of distances ($V_d < A_d$). The best performance was found at 0.8m, the estimated focal distance of the display, where the V-A conflict is closer to zero ($V_d = A_d$). Finally, at positive parallax (1.2m), the underestimation of distances was stronger than at negative parallax, as the V-A conflict change its sign ($V_d > A_d$). As consequence, there is a strong tendency to overestimate the size and underestimate the distance of the replica than at positive parallax. This difference is reflected in the asymmetry found in subjects' response distribution. In the shortest and longest estimation, the distributions are skewed closer or further to the perfect size-constancy ratio respectively. Kenyon et al. (2007) and Luo

et al. (2007) found that even though the size of the replicas were overestimated, the perception of size-constancy was maintained; however, this was not the case for our HMD condition. The greater tendency to underestimate the distance of the object at positive parallax affected severely the perception of size constancy, a performance consistent with the evidence about the greater susceptibility of HMDs to the vergence-accommodation mismatch issue. In this sense, these results confirmed how spatial perception is influenced by the V-A conflict size, the amount of underestimation seems to grow proportional to the amount of conflict induced.

4.2.2 Blind-walking test

The preliminary study demonstrated that natural locomotion is fundamental for assessing the perception of distances using blind-walking tests. We repeated the blind-walking test allowing subjects to walk naturally using an HMD, to describe better the influence of the peripheral stimulation issue on depth perception. Typical blind-walking studies in VR assess the perception of distances with and without adaptation, so we were also interested in studying the influence of continuous feedback on depth perception, especially the influence of the optic flow and motion parallax. We hypothesized that without any interaction with the VE, subjects' perception of distances is influenced by the depth compression phenomenon, and they walk naturally closer than the actual distances, according to the information provided by the visual/oculomotor cues. In contrast, with adaptation, subjects' perception of distances is recalibrated, the scale of the perceived space is changed and motor actions are adjusted, due to the greater influence of the motion parallax and the optic-flow cues. As result, the peripheral stimulation issue should induce a more homogeneous underestimation of distance regardless of the location of the visual stimulus. Thus, if our third assumption is correct, the underestimation of distances not only should be smaller in the post-adaptation condition but also, the motor re-calibration should cause a scaled instead of a compressed perception of distance. Following Kelly et al. (2014)'s hypothesis, the underestimation should not vary significantly between target distances. Also, according to Kelly et al. (2014), adaptation occurs fast and transfers asymmetrically according to the distance. Thus, we believe that a few minutes of walking in VE with continuous feedback are enough to induce a motor recalibration.

4.2.2.1 Hypotheses

Our hypotheses were as follows:

- H1: Subject perception of distances will be significantly better after adaptation.
- H2: Underestimation of distances will have a more steady/homogeneous tendency after post-adaptation compared with the pre-adaptation condition

4.2.2.2 Method

We designed a within-subjects experiment where subjects performed a classic blind walking test, first without adaptation and then, after an adaptation step. We requested subjects to estimate the distance of a virtual stimulus located at 3m, 5m, and 7m of the subject without interacting in the VE and under non-visual conditions. Then, subjects performed the adaptation step, where they were requested to walk and collect a set of objects which appears randomly in an area of 7x4m. Subjects performed this task for 5 minutes, walking directly toward the object and getting continuous visual feedback. After the adaptation step was finished, we requested subjects to repeat the blind walking test to analyze the influence of adaptation. To prevent learning effects between the pre- and post adaptation conditions, we alternated the VE between a plaza and a neighborhood; and also changed the target distances at 3.5m, 5.5m, and 7.5m. A rope was set aside and across the walking area to assist subjects walking blind. This was necessary because we noticed that some subjects have difficulties walking straight in VR, so the rope provided them with a sense of direction and assist them to return to the starting position. A total of 3 trials at each target distance was requested with aleatory order between trials. Figure 4.8 shows the experimental setup.

Participants

Sixteen subjects (13 males, 3 females, $M = 22.6 \pm 4.1$ years old) participated in our experiment. As in the previous study, all participants signed a letter of consent informing normal vision and good health.

Materials

Participants performed the tests using an HTC Vive headset and wearing a backpack PC, allowing freedom of movement. The tracking space was an area of 8x4m. This was the maximum area that we

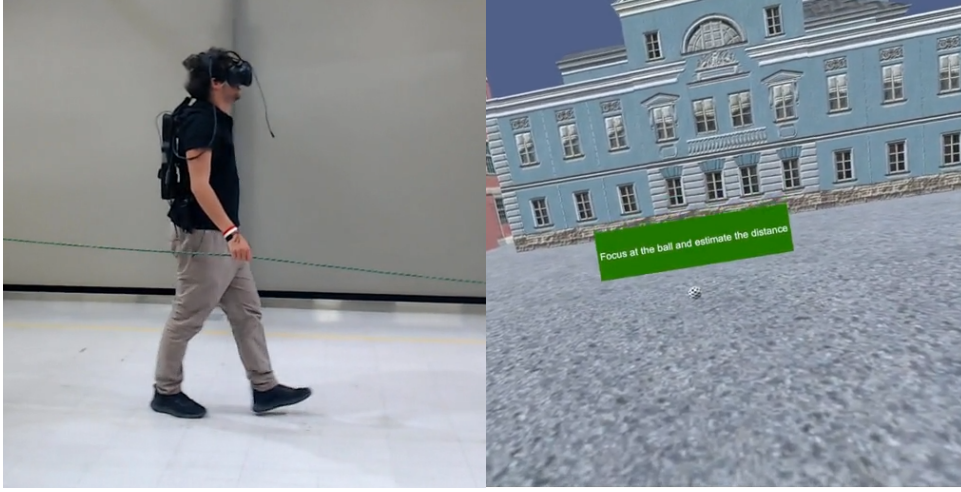


Figure 4.8: Blind walking test

Left: subject performing the blind-walking test. Right: Virtual stimulus, subject must estimate the distance to the soccer ball by walking toward them under non-visual-conditions.

could achieve in our Light House Tracking System via the "sync-cable" configuration and the space limitations in our lab. Two scenes were designed for each blind-walking test using the Unity3D game engine: a plaza and neighborhood, both rendered realistically using deferred lighting with forward shadows.

Procedure

The procedure was based on a typical blind-walking test. First, subjects were requested to walk blind for some minutes in physical reality to get familiarized with the procedure. Then, subjects were located at the starting position, assisted to wear the headset and the backpack, and they are presented with the VE. Then, we provided subjects with the Vive controller and asked them to walk blind for some minutes to get comfortable with it, by turning on/off the visual imagery. Once subjects felt comfortable, they were requested to perform the "pre-adaptation" blind walking test, estimating the distances of the first set of stimuli by walking toward them under nonvisual conditions. Immediately, subjects performed the adaptation step, where they were requested to collect a set of soccer balls that appeared randomly around the tracking space in a consecutive manner. This task was performed for 5 minutes, allowing subjects to adapt to the VE and inducing a motor recalibration. After finishing the adaptation phase, subjects were requested to perform the "post-adaptation" blind-walking test but this time, using the second set of target distances and inside the second VE. Between each step,

subjects were requested to rest for 2 minutes.

Results

Results of the blind-walking test are presented in Figure 4.9. A paired sample t-test was conducted to compare the perception of distances among the three target distances. Overall subjects underestimated the distances for all trials and regardless of the pre/post adaptation condition. There was a significant difference in the underestimation of distances at min target distance (3-3.5m) between the pre-adaptation condition ($M = 0.72m, SD = 0.28$) and the post-adaptation condition ($M = 0.91, SD = 0.22$), $t(47) = 2.4, p < 0.03$; while no significant difference was found at the middle target distance (5-5.5m) between the pre-adaptation condition ($M = 0.95, SD = 0.34$) and the post-adaptation condition ($M = 1.01, SD = 0.33$), $t(47) = 0.59, p = 0.55$; and at the maximum distance (7-7.5m) between the pre-adaptation condition ($M = 1.13, SD = 0.39$) and the post-adaptation condition ($M = 1.01, SD = 0.45$), $t(47) = -1.05, p = 0.30$. These results suggest that H1 does not hold, subjects' perception of distances after adaptation was not significantly better compared to the adaptation condition. However, there is a steady tendency for the post-adaptation condition, suggesting that H2 holds. The underestimation of distances was on average around 0.95m regardless of the target distances, indicating a possible motor recalibration. In contrast, the underestimation of distances for the pre-adaptation condition shows an incremental tendency, which may imply that the error could be significantly greater than the post-adaptation condition for distances greater than the studied here.

4.2.2.3 Discussion

We studied the influence of adaptation on blind walking tests and the impact of the incorrect peripheral stimulation issue. Subject performance in the pre-adaptation condition was consistent with the performance found in the studies about distance perception in VR based on blind-walking tests. Without adaptation, as the distance of the stimulus increases, the underestimation tends to increase because the space is perceived compressed in depth (due to the influence of the visual/oculomotor cues and the vergence-accommodation mismatch issue). Similarly, subject performance after the post-adaptation step is consistent with the studies about adaptation in VR using continuous visual feedback. The underestimation of distances maintains a more steady/homogeneous tendency, indicating that a motor recalibration has been induced and the perception of distances has been adjusted/scaled (assumption 3). Although we expected that subjects' performance after adaptation was significantly

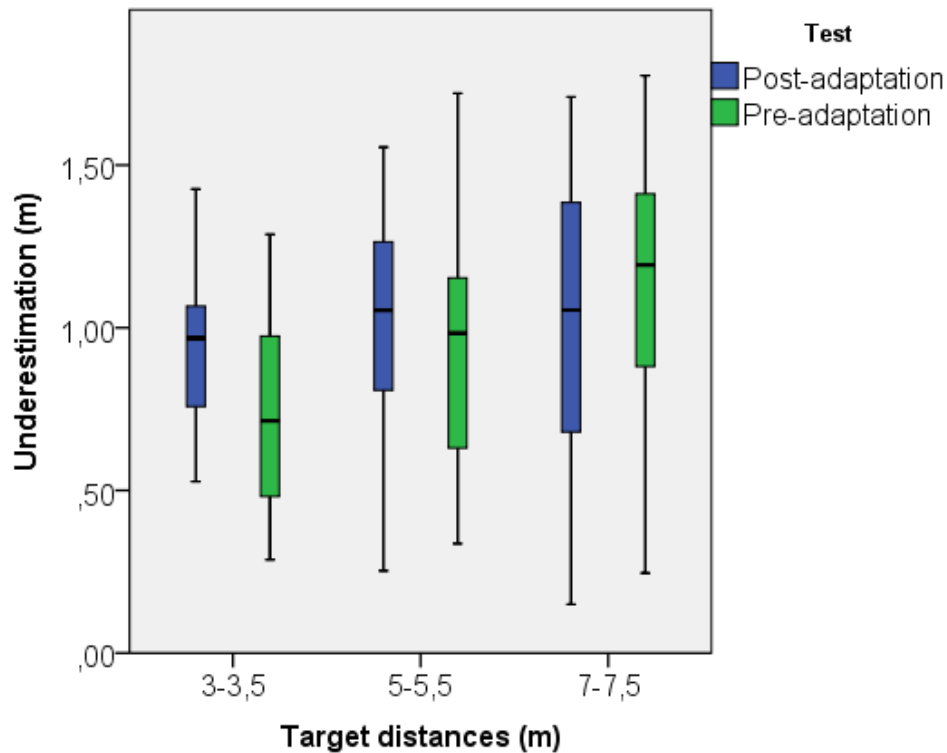


Figure 4.9: Results of the blind walking test with adaptation

Without adaptation, underestimation of distances tend to increase as the target distance increases. After adaptation, underestimation of distances is reduced and shows a more homogeneous/steady tendency, implying a motor recalibration.

greater compared with the pre-adaptation condition, that was not the case. We believe that this could be a consequence of some limitations in our study. Despite these results, the different trends between both conditions suggest that Kelly et al. (2014) hypothesis is true, adaptation not only induces a motor recalibration but also causes a rescaling of the perceived space. Thus, the space is compressed and scaled, in the pre-adaptation and post-adaptation conditions respectively.

Even though that our results were consistent with the evidence found in the literature, there were some limitations in the present study. The main problem was the limitations of space in our lab, which limited the differences between both conditions as very few steps were required to walk to the target distances. The rope that was introduced to solve the issue of maintaining a sense of direction, could introduce some noise, as some subjects use them as an additional cue to perform their estimates (the problem of proprioception as a cue for distance perception again). These limitations

could explain the variability in subjects' responses, which is high considering the short target distances used in this study. Despite these limitations, the results suggest that without adaptation, subjects' performance follows the pattern of increased underestimation of distances, which is characteristic of the depth compression phenomenon. The results also showed evidence that adaptation, induces a motor recalibration that adjusts the perception of distances. Also, the results confirmed that distances are still underestimated after adaptation, which implies that factors such as the limited FOV and the lenses distortions, associated with the incorrect peripheral stimulation issue, influence negatively the motor recalibration process, but there is not the characteristic depth compression found in the vergence-accommodation mismatch issue, as the perception of distance behaves uniformly between stimulus.

4.3 Predicting spatial performance

4.3.1 Stimuli selection

Based on the results from the previous studies, we developed the first step to generating a prediction model. To predict subject's spatial performance, we must characterize the depth compression phenomenon induced by a display in a smarter way. We defined a set assumptions that describe the influence of the different display factors on spatial perception. To characterize better the phenomenon of underestimation of distances in a particular display, we can perform an assessment method, whose stimuli are strategically selected to represent the influence of the different displays factors. For example, to analyze the influence of the focal distance and the stereoscopic parallax condition, both associated with the vergence-accommodation mismatch issue, the ideal stimulus configuration is a set of target distances linearly correlated with the focal distance, where positive values refer to stimulus located at positive parallax and negative values refer to stimulus located at negative parallax (Figure 4.10). These stimuli selection depend also on the kind of spatial task performed and the cues involved. For example, in the size-constancy tests, a small range is enough, as the useful range of these cues for perceptual matching task is less than 3m (Linton (2020)). Regardless of the spatial task, we hypothesize that a stimuli composed of five target distances is enough to characterize the depth compression phenomenon on the display. Figure 4.11 shows the stimuli selection as a function of the influence of each display issue in spatial perception. For the vergence-accommodation mismatch issue,

4.3. PREDICTING SPATIAL PERFORMANCE

five stimulus equally spaced at zero, positive and negative parallax are enough to characterize the induced depth compression. Conversely, for the incorrect peripheral stimulation issue, five stimulus non linearly spaced are enough to characterize the induced depth compression. In this case, the target distances are doubly spaced to represent the inflexion points in the induced compression curve.

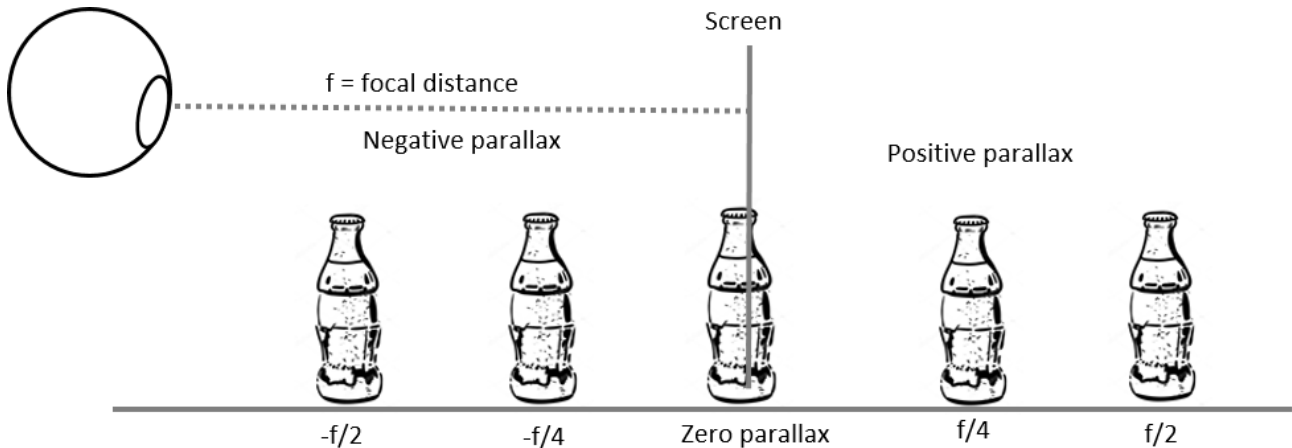


Figure 4.10: The selection of the stimulus as a function of the focal distance f .

4.3.2 Linear performance predictor

The second step to predict spatial performance is to calculate subjects' linear performance predictor (LPP) using a transversal assessment method. The LPP can be calculated using a linear regression model (e.g. using ordinary least squares) that correlates the degree of underestimation or overestimation with the target distances. Figure 4.12 shows a scatter plot of a subject's performance using the HMD and the CAVE, during the size-constancy test. The chart is presented in terms not only of the target distances but also in terms of the stereoscopic parallax condition. An assumption of the framework is that the amount of underestimation is proportional to the size of the V-A conflict. Thus, we hypothesize that the underestimation slope must be proportional to the size of the divergence between vergence and accommodation. The chart depicts how the CAVE induces a smaller V-A conflict compared with the HMD, which in turn causes a small underestimation of distances at positive parallax and a slight underestimation of distances in negative parallax. In contrast, the HMD induces a greater V-A conflict, causing a more severe underestimation of the distances. These results reflect the greater susceptibility of HMDs to the depth compression phenomenon.

Calculating the LPP of a display is straightforward. We can estimate subjects performance by

4.3. PREDICTING SPATIAL PERFORMANCE

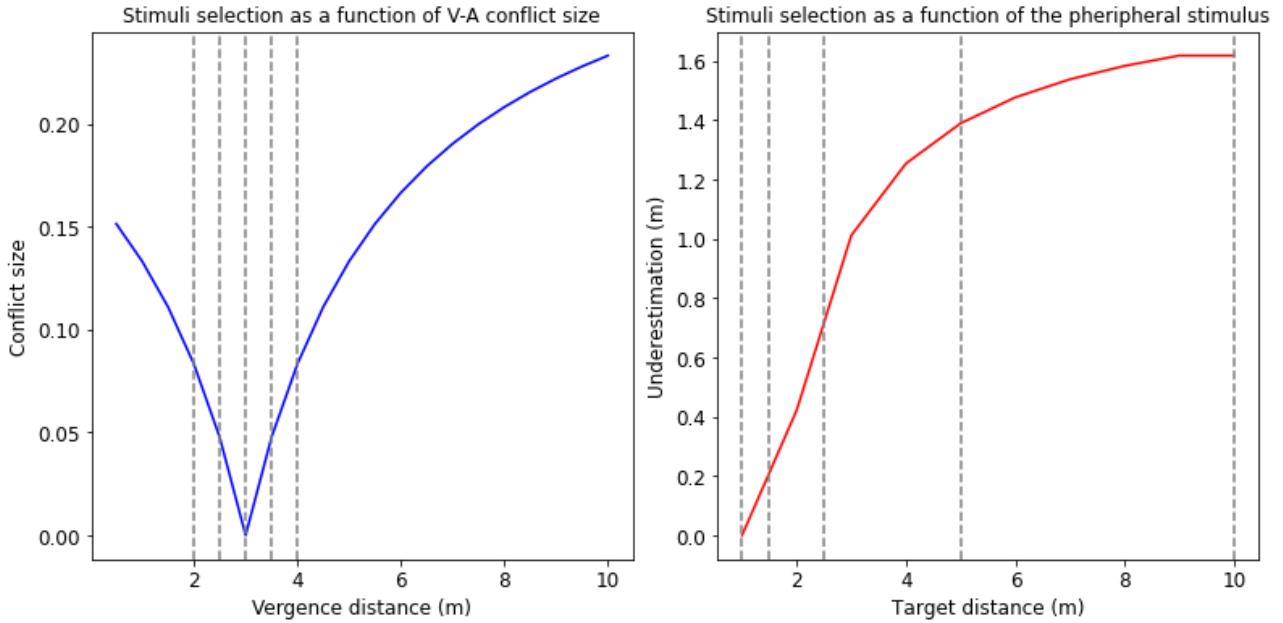


Figure 4.11: Stimuli selection according to the display issue.

A stimuli composed of five target distances is enough to characterize the depth compression phenomenon on the display. The selection depends on the influence of the display factors on the perception of distances.

means of a linear regression function that estimate the underestimation of distances induced by the display, which could be defined as:

$$d' = (1 - c)d \quad (4.3)$$

where d' is the expected underestimated distance, d is the target distance stimulus, and c is the overall induced compression slope by the display. Figure 4.13 compares subject's performance, between the HMD and the perfect performance. In the ideal condition, distances are not underestimated, so the slope is equal to one. In contrast, the linear regression slope based on subject performance for the display is $1.07 - 0.29d$, which means that the induced compression in this case is $c = 0.29d$. Thus, the LPP for the HMD is $d' = (1 - 0.29)d$, for each meter, distances are underestimated a 29%.

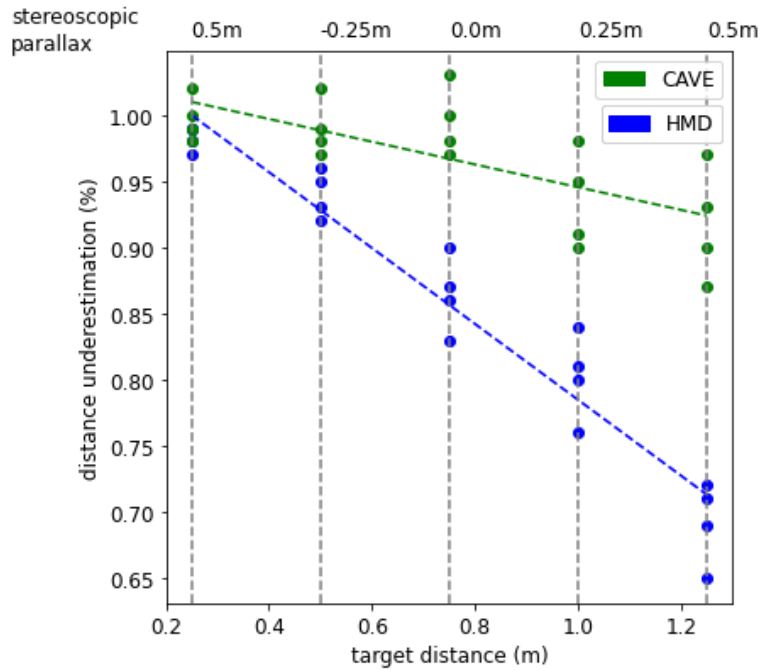


Figure 4.12: Subject performance as a function of the V-A conflict size.

The CAVE induces a smaller conflict between vergence and accommodation compared with the HMD, which in turn causes an smaller underestimation over distances.

4.4 Summary

In this chapter, we presented the first part of the proposed framework, which is focused on characterizing the influence of the different display issues in spatial perception. The first objective of the framework was to characterize the influence of the different display issues on spatial perception. The results of the previous studies confirmed the influence of the V-A conflict and the incorrect peripheral stimulation issue in spatial perception. The size-constancy test showed results consistent with the evidence about the influence of the focal distance and the stereoscopic parallax condition on spatial perception. Distances were compressed due to the HMD's fixed focal distance (and the induce DOF) (assumption 1) and the degree of underestimation increased proportionally to the size of the V-A conflict (assumption 2). Similarly, the blind walking tests demonstrated the influence of the limited FOV and the distortions induced by the lenses on the perception of distances. Distances were also underestimated, but more homogeneously after adaptation, which suggests that a motor calibration was induced and the perception of distances was adjusted/scaled (assumption 3). Regarding if the

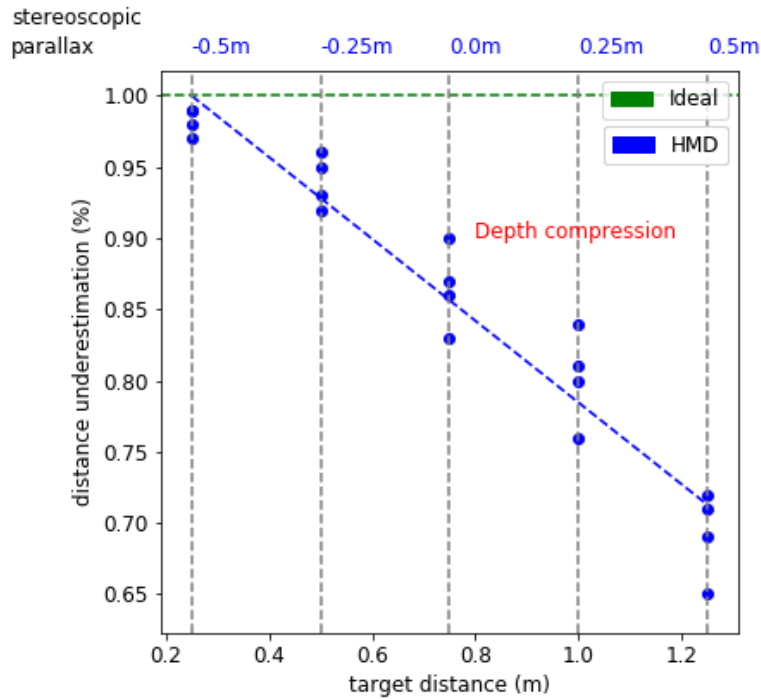


Figure 4.13: Linear performance predictor

Subject performance as a function of the distance stimuli and the stereoscopic parallax condition. The linear regression slope characterizes the amount of depth compression induced by the display.

underestimation behaves in all cases linearly (assumption 4), we need to perform more studies. The results from the blind-walking tests suggest this behaviour, but in the size-constancy test, it was not clear and we need to consider the inversion of sign in the Vienne et al. (2018)'s V-A conflict hypothesis.

Finally, we described how the stimuli must be strategically selected to assess the influence of the display issues in spatial perception based on how the different display factors influence the perception of depth. Once the subject's spatial perception is measured, a linear performance predictor (LPP) can be calculated to predict the induced depth compression in the display. Although the results were not totally the expected, we were optimistic about the validity of the assumptions, which provided us more insights about the phenomenon and expands the possibility of generating a more accurate model. In the next chapter, we discuss the second part of the framework, about how composition can be used to influence spatial perception, and how the slope of the LPP can be used to improve the perception of distances in VR.

4.4. SUMMARY

Chapter 5

Part II. Influencing spatial perception by composition

5.1 Overview

In the previous chapter, we discussed how the different display factors influence the perception of distances in VR, and how the depth compression phenomenon can be predicted using an assessment method and calculating a linear performance predictor (LPP). A objective of the framework is that if we are able to predict the spatial performance, we can influence positively the perception of distances using composition. Composition refers to all the techniques that creates the illusion of depth that are originated by human experience and imprinted in our minds. From all the depth cues discussed in this document, we have paid little attention to the pictorial cues, particularly linear perspective. To influence subject performance, we can take advantage of bias on the human beings toward symmetry and regular shape patterns associated with perspective (such as those described in the Ames' room illusion). Thus, we took advantage of some techniques that plays with the rules of perspective and have shown good results influencing the perception of distances in VR: *lowering the horizon* and *FOV minification*. Figure 5.1 show two screenshots of each technique, where the default perspective projection is altered. Lowering the horizon uses an oblique perspective projection that applies a subtle declination in the horizon line, moving the vanishing points downwards and increasing the perception of distances. FOV minification applies a different geometric FOV than the FOV defined for the display. The technique causes that a larger portion of the VE is "compressed" in the viewport, changing the rate at which the vanishing point converges at the horizon and making object look smaller and further. In the next sections, we will discuss how we integrated these techniques in the proposed framework

and why we implement these techniques using a different approach.

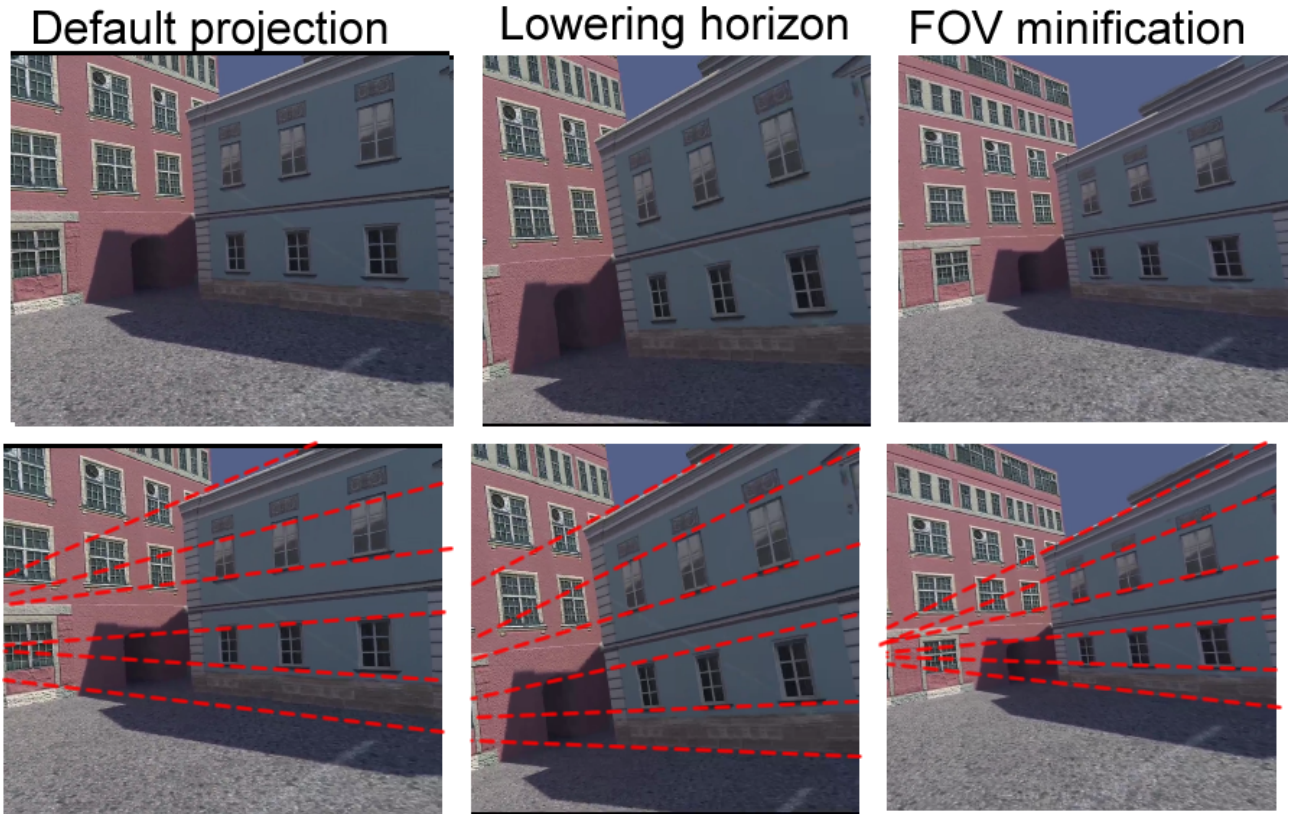


Figure 5.1: Alternative projection techniques

A. Default perspective projection. B. Lowering the horizon. C. FOV minification.

5.2 Transformation in perspective projections

5.2.1 Lowering the horizon

During the Italian Renaissance, when artists developed the technique of perspective, they recognize the importance of the horizon line and the vanishing points in spatial perception. Today, their importance has been quantified: the *angular declination hypothesis* states that the human beings can estimate a distance d based on the eye height h and the angular declination below the horizon α , a relation that can be expressed in the trigonometric relationship $d = h/\tan(\alpha)$ (Figure 5.2) (Gajewski et al. (2014)). Ooi et al. (2001) validated this hypothesis in the physical reality by requesting subjects to estimate the perceived distance of a target using a base-up prism goggles. This device induces a

refraction in the induce light that "moves" everything upwards, changing the apparent location of the horizon line. Ooi demonstrated that perception of distances was influenced by the apparent location of the horizon under the prism-based stimulus and also after adaptation.

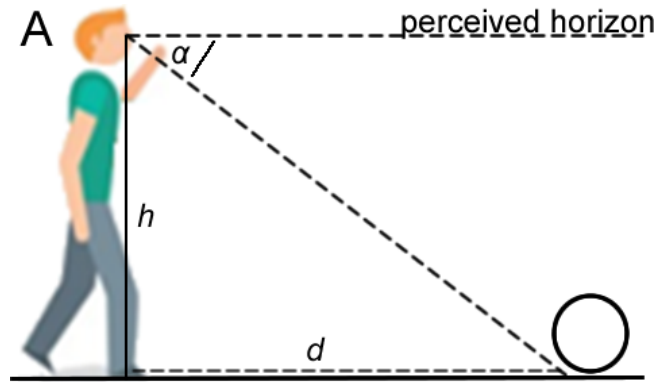


Figure 5.2: Angular declination hypothesis

Distance d can be estimated based on the eye height h and the angular declination below the horizon α .

Messing and Durgin (2005) replicated this test in a VE. Alike Ooi study, they preferred to directly changed the apparent location of the horizon by altering the virtual geometry. Although this approach is straightforward to implement, it has the inconvenient that the virtual geometry must be altered, which could be a restriction in some applications. A better approach is using an oblique perspective projection (OPP) as is presented in Figure 5.3, where a point P is projected in a projection plane with an angular inclination. Since all points in the scene are shifted down, the apparent horizon location also moves proportional to the angle of inclination. The immediate effect is the apparent reduction of the angular declination α and the apparent increase in the perceived distance of the point.

This oblique perspective projection can be performed in two ways. The easiest method is by performing this transformation in projection-space on the calculation of the viewport transform, by simulating an angular declination in the FOV. This was the approach used by Williams et al. (2009) obtaining similar results with previous studies. However, this approach has the inconvenient that the orientation of the geometric FOV enters in disagreement with the orientation of display FOV (or the parameters in which the display was calibrated), introducing some undesired distortions. A better but more complex approach is to perform this transformation in view-space before the projection is calculated. The method will convert a vertex in world-space into an "oblique distorted view"-space,

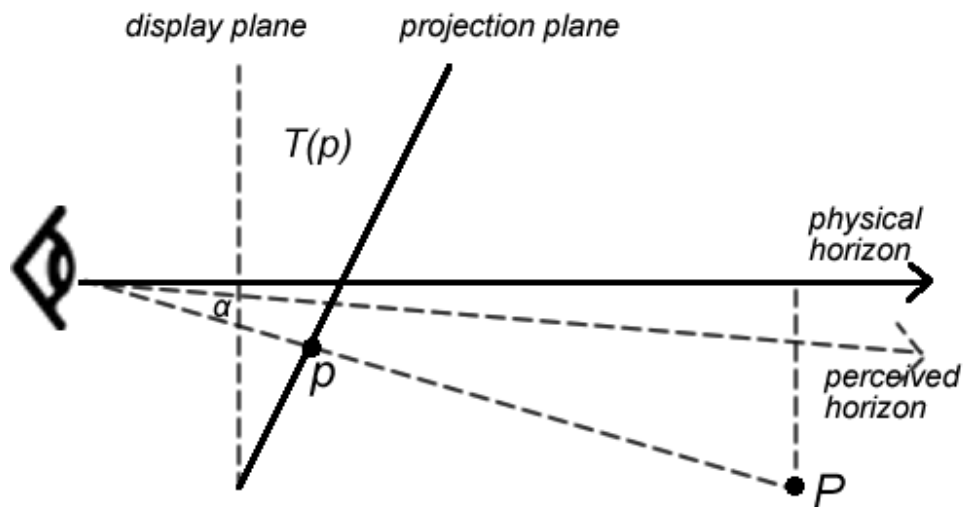


Figure 5.3: The oblique perspective projection

The apparent reduction of the angular declination α induces an apparent increase in the perceived distance of the point P .

which distorts the view of the point simulating an angular declination of the horizon line. Similar approaches have been proposed to optimize lens distortion correction in HMDs (Pohl et al. (2013), Kehler (2016)), where the vertex in view-space are further transformed into an inverted lens coordinate system. This approach has the advantage that we do not need to alter the virtual geometry directly, but it has the disadvantage that we have to deal with shadows and other light-related issues. This approach was the selected in this research, and we hypothesized that this technique would allow us to influence positively the perception of distances.

5.2.2 FOV Minification

VR displays are usually calibrated so that the geometric FOV matches the display's FOV. The viewport is mapped from virtual space onto the real space constituting a "correct" perspective. Kuhl et al. (2006) proposed to apply a different geometric FOV than the provided by the display, to influence spatial perception in VR. Their idea was originated from the old issue of scaling the FOV in videogames to match screens with different sizes and resolutions. Scaling the geometric FOV will result, either in minification or magnification, of the visual imagery. If the geometric FOV is smaller than the display FOV, the image will appear magnified (each pixel will fill a larger subtended angle in real space versus virtual space). Conversely, if the geometric FOV is larger than the display's FOV, the visual imagery

will be minified (each pixel will fill a smaller subtended angle in real space versus virtual space). Thus, minification causes that a larger portion of the VE to be compressed in the viewport, making objects look further and smaller. Consecutive studies following the approach of Kuhl et al. (2006) showed that distance perception is improved significantly using this technique (Bolte et al. (2010), Zhang et al. (2012), Li et al. (2014), Steinicke et al. (2009b)).

Figure 5.4 presents the effect of minification in the perceived distance of a point P . Minification does not change the perceived location of the horizon, but it reduces the angular separation between the projected point p and the center of projection. The movement of the point in the projected image upwards the center of projection (represented as the point p') causes a reduction of the angular declination α , increasing the perceived distance of the point (represented as P').

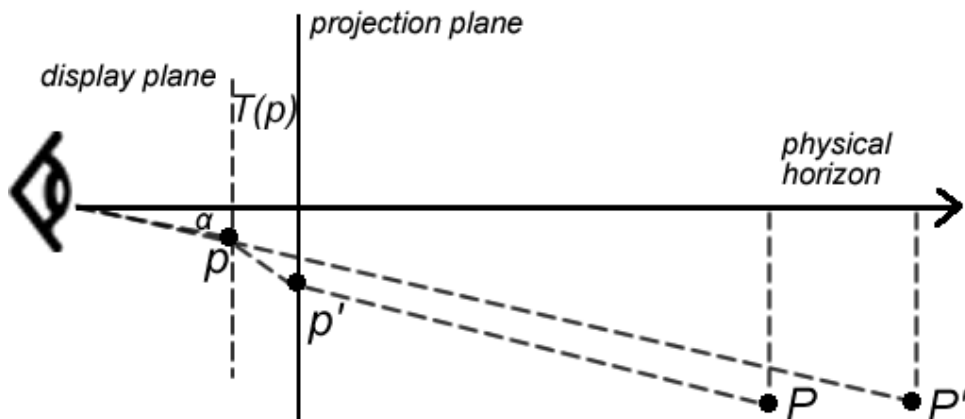


Figure 5.4: The influence of FOV minification in the perception of distance

The movement of the point P in the projected image towards the center of projection causes a reduction of the angular declination α , increasing the perceived distance of the point P .

Expanding the geometric FOV is a straightforward task but has to be done cautiously. Stretching the image arbitrarily to fill the screen will cause significant distortion if this is done without taking into account the aspect ratio in which the display was originally calibrated. In modern videogames, the vertical FOV is fixed, while the horizontal FOV is expandable. This adjustment is ideal because most of screens used for gaming nowadays are widescreen. In contrast, HMD displays sacrifices horizontal FOV to favors stereo overlap. Thus, a method focused on the vertical FOV is preferable. However, because we are using a geometric FOV different from the one recommended by the display, distortions

are inevitable and they become an important issue to deal with. For example, some distortions are induced when the observer moves or turns his/her head. This is because there is no direct mapping between the location of the point P in virtual space and the movement of the projected point p' . The distortion becomes stronger as the point moves towards the peripheral vision. To deal with this undesired effect, Bolte et al. (2010) proposed to apply head rotation gains to reduce the differences between real space and virtual space by taking advantage of the imperfections of the human visual-vestibular system.

An alternative approach is to use a non-linear projection technique as is depicted in Figure 5.5. This approach was used in two studies using, either an exaggerated pincushion distortion (Kuhl et al. (2006)) or a fish-eye lens distortion (Orlosky et al. (2014)), getting mixed results. Non-linear projection techniques can reduce the disparities between real space and virtual space. However, using non-standard projection techniques in HMDs is challenging because most of the commercial headsets do not allow to change the display calibration parameters easily. Thus, a small amount of induced distortion is sufficient to become a distracting and cybersickness-inducing factor. Novel opensource frameworks, such as OpenVR and OSVR give new opportunities to researchers to work with the interesting but sometimes forgotten area of nonlinear projections, but they are outside of the scope of this research.

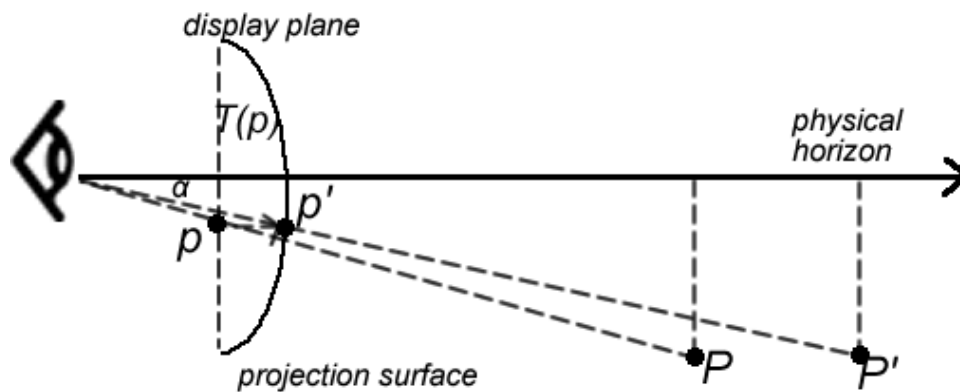


Figure 5.5: The influence of a non-linear projection technique in the perception of distance

The point P is projected in a curved surface, decreasing in a non linear manner the angular separation between the projected points and increasing the perception of distance.

As in the *lowering horizon* technique, a better but more complex approach is to perform the minification in view-space. The method will convert a vertex in world-space into a "minificated perspective

view"-space, which distorts the "projection" of the point due to a non-homogeneous scaling factor. The term non-homogeneous refers to the fact that the scaling factor varies according to the angle and distance of the point. An advantage of this method is that we can apply different minification factors for the horizontal FOV and vertical FOV without suffering severe distortions. A disadvantage is that we have to deal with light-related issues. We hypothesized that these technique would allow us to influence positively the perception of distances in VR.

5.2.3 Implementation

These techniques require to apply a transformation that changes the direction of the vertex in the view-space before the projection is computed. Equations 5.1 and 5.2 describes the transformation chain in VR using LaValle (2016) notation, where the transformation matrices $T_{minification}$ and $T_{lowhorizon}$ are added with the purpose of inducing an angular declination or a minification in the view-space before the projection is computed.

$$T = T_{viewport}T_{canonical}T_{lowhorizon}T_{eye}T_{rbody} \quad (5.1)$$

$$T = T_{viewport}T_{canonical}T_{minification}T_{eye}T_{rbody} \quad (5.2)$$

In the standard rendering pipeline, each transformation is implemented using a chain of matrices that transform the position of the vertex in world-space, to the view-space, and finally to the homogeneous clip-space. In the Unity3D render pipeline, these transformation matrices are implemented in a predefined set of functions, denoted as the *TransformObjectToWorld* (T_{rbody}), *TransformWorldToView* (T_{eye}) and *TransformViewToHClip* ($T_{canonical}$), which receive as parameter the position of vertex in the current space and return its transformed position in the target space. Thus, we introduced two other functions to the chain that implements the lowering horizon and the FOV minification technique: a *TransformViewToCustomHorizonVS* ($T_{lowhorizon}$) and a *TransformViewToCustomFovVS* ($T_{minification}$). The purpose of these functions is to change the orientation of the vertex in the view-space before the homogeneous clip-space is calculated, by applying an angular declination in the horizon line or scaling the FOV respectively. To change the orientation of the vertex, we define a directional vector, which is used to determine the new intended direction of the vertex in view-space

5.3. VALIDATION OF ALTERNATIVE PROJECTION TECHNIQUES

using equation 5.3, where \vec{V} is the vertex in view-space, \vec{D} is the directional vector, δ is the angular rotation of \vec{V} towards \vec{D} , and \vec{V}' is the new vertex position.

$$\vec{V}' = \vec{V} \cos(\delta) + \vec{D} \times \vec{V} \sin(\delta) \quad (5.3)$$

Implementing the lowering horizon technique using this equation is straightforward. The position of the vertex in view-space is rotated towards the vertical axis ($\vec{D} = (0, 1, 0)$) by an angular declination δ . Thus, it only suffices to apply a negative angular declination, so that all objects in the scene are rotated downwards in view-space and move down from the center of the screen in clip-space, simulating an angular declination in the horizon. On the other hand, implementing the FOV minification technique requires a little more effort. We calculated first the amount of uniform scaling induced in the FOV, which is defined by the expression $m = \frac{\tan(vfov/2)}{\tan(vfov*g/2)}$, where m is the amount of uniform scaling, $vfov$ is the vertical display FOV, and g is the geometric scaling factor applied to the FOV. Then, we calculate the "induced angular compression", by calculating the offset between the displays FOV and the minified FOV ($\delta = vfov(1 - m)$). Then, this angular offset can be used in equation 5.3 to induce a rotation of the vertex in the forward axis ($\vec{D} = (0, 0, 1)$), so that all objects in the scene are rotated forward in view-space and converge to the center of the screen in clip-space.

5.3 Validation of alternative projection techniques

We developed two studies to validate if the implemented techniques were suitable for the proposed framework. We believed that the lowering horizon technique was more suitable for task that are based on perceptual matching, such as the size-constancy test, as this technique does not influence the perception of size compared with FOV minification. In contrast, we believed that FOV minification was more suitable for task based on motoric responses, such as the blind-methods, as this technique increases the geometric FOV influencing the perception of motion parallax. Thus, we performed some other tests based on the transversal assessment methods, analysing if the implemented techniques can improve the perception of distances in VR.

5.3.1 Influence of lowering the horizon on size-constancy

We replicated the last size-constancy test, asking subjects to estimate what it would be the relative size of a virtual replica located at different distances and using a physical reference. As in the previous test, we focused only on a HMD, but this time we analysed the influence of the "oblique perspective view"-space technique on spatial perception. We took advantage of the shape-constancy bias in human beings of perceiving tables always straight and parallel to the ground to solve the issue. By applying a subtle declination in the horizon line and affecting the table's vanishing points (Figure 5.6), the visual system enters in a cue conflict. As in Ames' room illusion, we hypothesized that the visual system would ignore the information from the oculomotor depth cues in favor of a "stable table"; a table that is straight and parallel to the ground. The apparent change in the location of the horizon will cause a reduction in the angular declination α , and an increase in the perceived distance of the replica, which in turn would induce a decrease in its perceived size. The subtle declination has the effect of the table vanishing points moving down from the center of the screen, creating the illusion of a table larger than its actual dimensions and influencing the perception of distance and size. We hypothesized that by applying our adapted lowering horizon technique, we would improve the perception of distance and induce size-constancy.

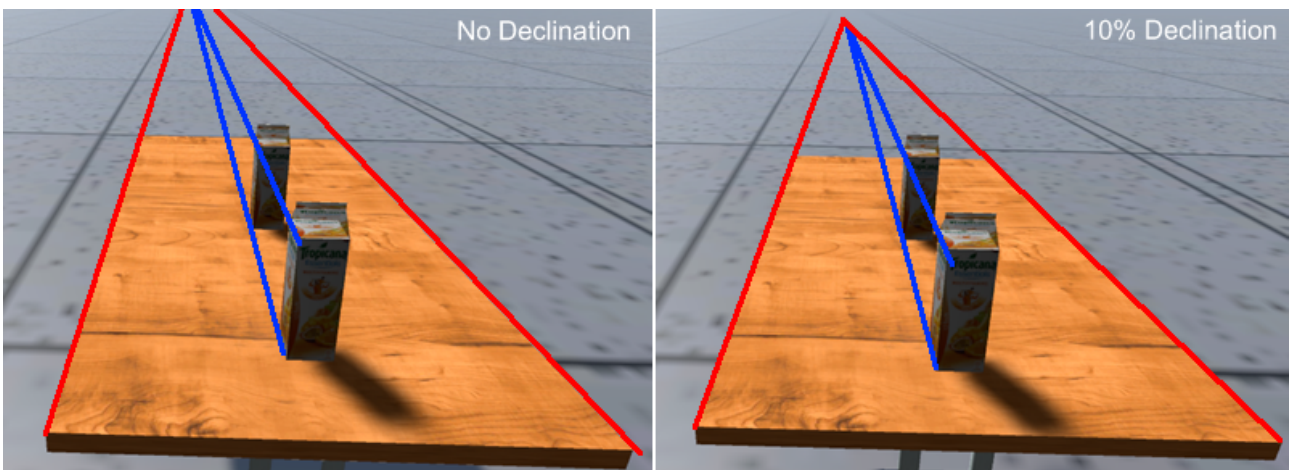


Figure 5.6: Implementation of the "oblique perspective view"-space technique

Left: Default projection. Right: Oblique perspective projection with a declination of 5%.

We developed another study to analyse if by applying the proposed technique, the perception of size constancy was improved. We repeated the same test with the same experimental design but this

time we applied a declination of five degrees. We hypothesized that the use of the "oblique-perspective-view"-space technique would improve the perception of size-constancy in the HMD.

5.3.1.1 Hypotheses

Our hypotheses were as follows:

- H1: SizeRatio responses will be improved compared with the default projection condition.
- H2: SizeRatio responses will reflect size-constancy.

5.3.1.2 Method

The experiment was conducted using the same apparatus and procedure of the last size-constancy test, with the same target distances (0.4m, 0.8m, and 1.2m), physical objects and number of trials.

Participants

Ten participants (8 males, 2 females, $M = 23.5 \pm 1.25$ years old) participated in our experiment. As in the previous study, all participants signed a letter of consent reporting normal health and vision conditions.

Results

The results of the lowering horizon technique compared with the default projection technique are depicted in Fig. 5.7. A paired sample t-test was conducted to compare the Size-Ratio among the three target distances. There was a significant difference for Size-Ratio between the lowering horizon condition at 1.2m ($M = 0.99, SD = 0.12$) and the default projection condition at 1.2m ($M = 1.09, SD = 0.11$), $t(19) = 3.6, p < 0.02$; between the lowering horizon condition at 0.8m ($M = 0.90, SD = 0.08$) and the default projection condition at 0.8m ($M = 0.99, SD = 0.12$), $t(19) = 2.6, p = 0.01$; while not significant differences for Size-Ratio were found between the lowering horizon technique at 0.4m ($M = 1.09, SD = 0.11$) and the default projection technique at 0.4m ($M = 1.13, SD = 0.07$), $t(19) = 1.6, p = 0.11$. The results showed H1 partially holds, the lowering horizon techniques reduce the underestimation of distances and improve the perception of size compared with the default projection technique, particularly for objects at positive parallax.

We also repeated the TOST equivalence test to demonstrate the existence of size-constancy. We got

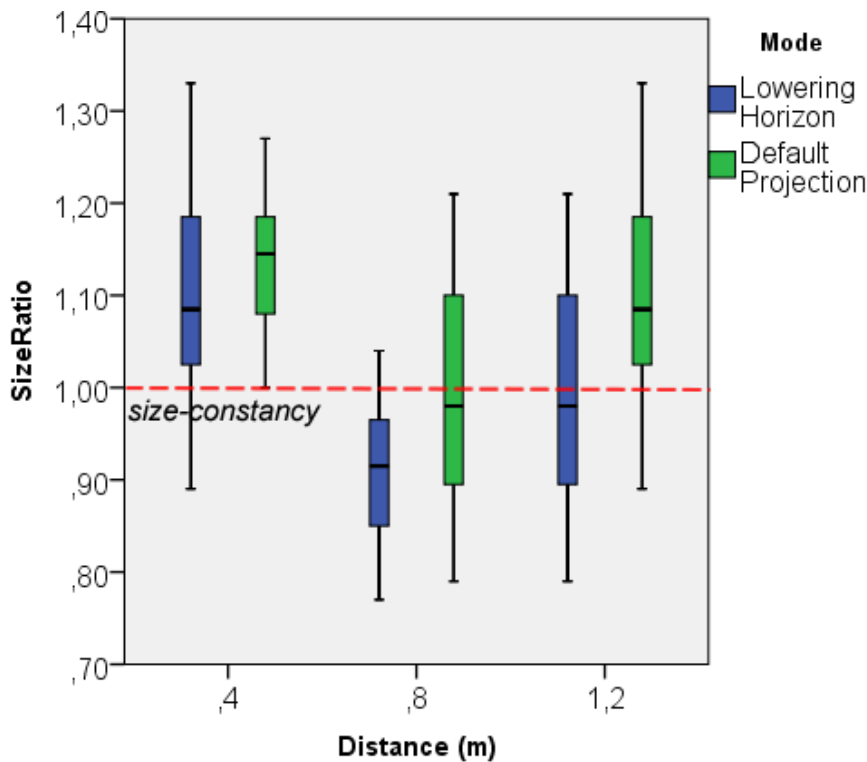


Figure 5.7: Influence of the lowering horizon technique on the perception of size constancy

The "oblique perspective view"-space technique improves the perception of size and distance compared with the default projection technique, particularly at larger distances.

opposite result to the previous study, where size-perception is statistically equivalent and not different between 0.8m and 1.2m $t(19) = 2.50, p = 0.010$, and not equivalent and statistically different between 0.4m and 0.8m $t(19) = 13.68, p = 0.845$, and between 0.4m and 1.2m $t(19) = 7.03, p = 0.076$. These results suggest H2 partially holds, the proposed technique improves the perception of size-constancy for distances at positive parallax but has a negative effect in distances at negative parallax. Interestingly, we obtained similar effects from the previous test at the focal distance (0.8m), but this time with overestimation of distances. Responses are nearly accurate indicating that at the focal distance, the perception of size tends to be accurate.

5.3.1.3 Discussion

The proposed "oblique-perspective-view"-space technique improved partially the perception of size-constancy in the HMD. Although the perception of size seems to improve, the technique also induced

disparate results with overestimation of distances a zero parallax and underestimation of distances at positive parallax. Contrary to the previous study, there is a greater tendency to size-constancy for objects at positive parallax, between 0.8m and 1.2m meters. These results are in agreement with previous studies which suggest that artificially lowering the horizon can positively influence spatial perception in VR. The influence of the pictorial cues is so strong that overrides the influence of the oculomotor cues, vergence and accommodation. Thus, the technique creates the illusion of a table larger than its physical dimensions, influencing the perception of distance and size of the objects. However, the selection of the angle of declination is critical, as excessive declination may influence negatively the perception of short distances. Despite the undesired overestimation effect, it is particularly interesting that the perception of size tended again to be quite accurate at 0.8m, confirming again the strong influence of the focal distance.

5.3.2 Influence of FOV minification on blind-walking

We replicated the last blind-walking test, where subjects estimated the distance of different objects walking naturally under non-visual conditions. As in the previous test, we focused only on a HMD, but this time we analyse the influence of the "minificated perspective view"-space technique on spatial perception after adaptation. The FOV minification have the effect that the vanishing points converge at smaller rate as they approach to the horizon line, influencing the perception of motion parallax and increasing the perceived FOV (Figure 5.8). Objects that were not visible with the default FOV become visible with the scaled FOV and the whole scene is expanded in depth. We hypothesized that by applying our adapted FOV minification technique, we would improve the perception of distance and induce size-constancy. We developed another study to analyse if by applying the proposed technique, the perception of distances during blind walking task was improved. We repeated the same test with the same experimental design but this time we applied a scaling of the FOV of 10%.

5.3.2.1 Hypotheses

Our hypotheses were as follows:

- H1: Perception of distances will be improved compared with the default projection condition.
- H2: Perception of distances will have a steady tendency after post-adaptation condition.

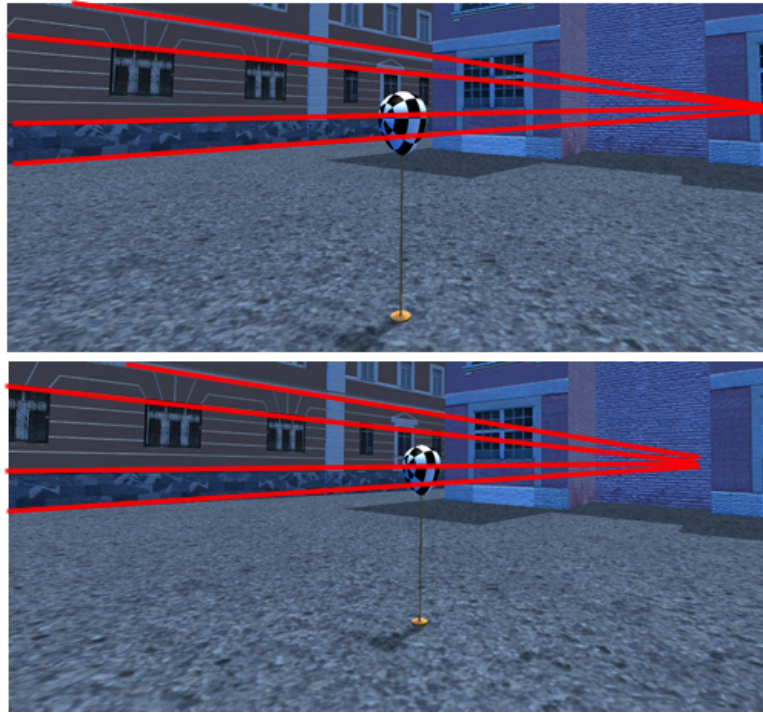


Figure 5.8: Implementation of the "minificated perspective view"-space technique

Left: Default projection. Right: FOV scaled by a 10% factor.

5.3.2.2 Method

The experiment was conducted using the same apparatus and procedure than the last blind-walking experiment, with the same target distances (3.5m, 5.5m, and 7.5m), visual stimulus and number of trials.

Participants

Sixteen subjects (13 males, 3 females, $M = 22.6 \pm 4.1$ years old) participated in our experiment. As in the previous study, all participants signed a letter of consent informing normal vision and good health.

Results

Results of the FOV minification technique compared with the default perspective projection technique are depicted in Figure 5.9. A paired sample t-test was conducted to compare the subjects performance among the three target distances. There was a significant difference for the three target distances, between the FOV minification condition at 3.5m ($M = 0.40, SD = 0.22$) and the

5.3. VALIDATION OF ALTERNATIVE PROJECTION TECHNIQUES

default projection condition at 1.2m ($M = 0.91, SD = 0.22$), $t(47) = -9.1, p = 0.0$; between the FOV minification condition at 5.5m ($M = 0.47, SD = 0.28$) and the default projection condition at 5.5m ($M = 1.01, SD = 0.33$), $t(47) = -6.1, p = 0.0$; and between the FOV minification condition at 7.5m ($M = 0.57, SD = 0.22$) and the default projection condition at 7.5m ($M = 1.01, SD = 0.45$), $t(47) = -4.2, p = 0.0$. Results showed that H1 holds, the FOV minification technique influences positively the perception of distances, with a superior performance compared with the default projection technique. Regarding H2, results suggest that this hypothesis does not hold, there is an apparent steady tendency for 3.5m and 5.5m, but underestimation increases at 7.5m, which suggest that perception of distances is not recalibrated uniformly after adaptation.

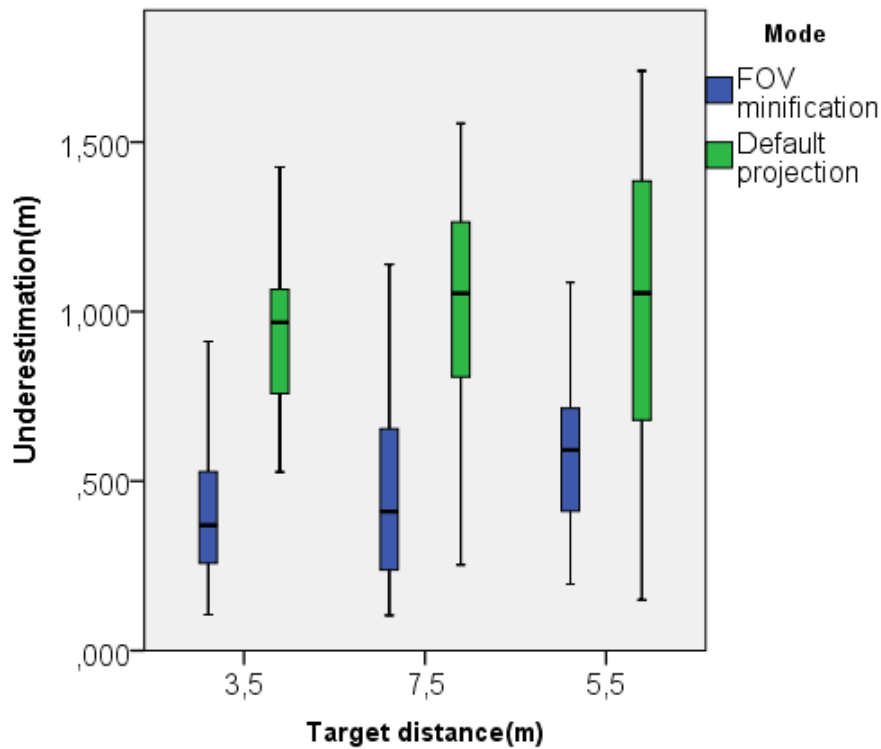


Figure 5.9: Influence of the FOV minification technique on the perception of distance

The "minificated perspective view"-space technique improves significantly the perception of distance compared with the default projection technique.

5.3.2.3 Discussion

The proposed "minificated-perspective-view"-space technique improved the perception of distances in HMD compared with the default projection technique. These results are in agreement with previous studies which suggest that artificially increasing the FOV can positively influence spatial perception in VR. Although the perception of distance seems to improve, the technique seems to influence the adaptation process, where distances seems to be recalibrated in a non-uniform manner. This could be a consequence that FOV minification alters in a non-linear way the behaviour of the motion-pursuit-law, which is the basis of the estimations of depth from motion parallax. However, scaling the FOV by a 10%, reduces the underestimation of distances almost a 50%. These results confirms again the strong influence of the pictorial cues in the overall perception of depth. Thus, the technique have two direct effects: the visual stimulus is perceived further as the whole scene is scaled, but the perceived optic flow and motion parallax cue is increased, inducing a motor recalibration where each step is "traversing more distances with less effort". However, as in the lowering horizon technique, the selection of the scaling factor is critical.

5.4 Limitations

Changing the direction of the vertices before the projection-space is computed has the issue that the table appears normal when it is observed parallel to the line of sight, but it looks distorted when the observer rotates his/her head. This is consequence of the fact that there is not longer a direct mapping between the coordinates of the object in virtual space and the orientation of the screen in the physical space. Tracked head-mounted displays typically use some type of sensor to measure the orientation of the HMD. If the sensor is perfectly aligned with the optical axis of the display, the displayed imagery its perfectly aligned with head orientation. However, since we are changing the direction of the vertex before the projection is computed, this direct mapping is lost. The distortion become evident when subjects tilt their head to the left or the right (Figure 5.10).

There was another issue associated with the computation of shadows. Because we do not deal with the computation of light, when the participant moves his/her head to get another point of view, the direction of vertex is altered according to the proposed projection technique (Figure 5.11). However, their shadows become synchronized incorrectly with respect to the motion of the head.

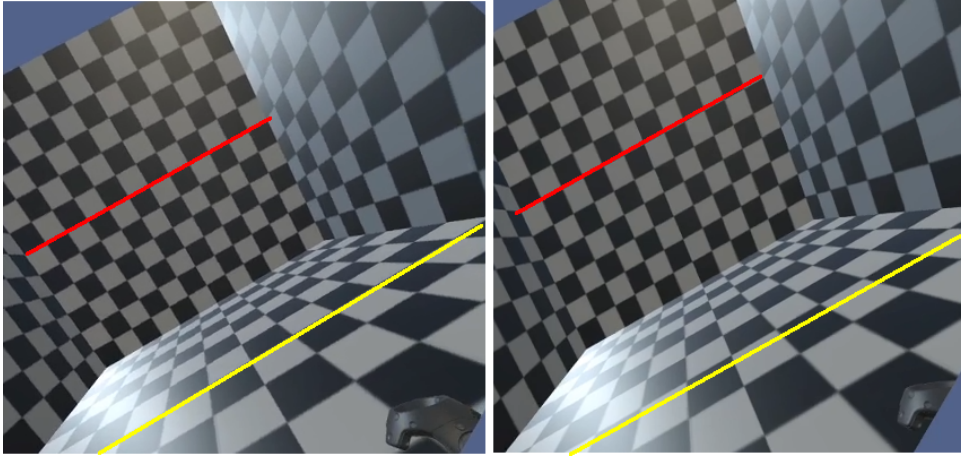


Figure 5.10: Distortion induced by applying an oblique perspective projection

The red line is located at the center of optical axis and the yellow line is parallel to the red line. Left: Default perspective projection, where vertices are aligned with the optical axis. Right: Oblique perspective projection, where vertices become misaligned.

Modern game engines such as Unity3D and Unreal3D implemented their own versions of high level shader programming languages, with a very large of pre-implemented macros, to make transparent and friendly for developers to deal with light, shadows and other complex stuffs. Unfortunately, these implementations are optimized to the default perspective projection technique, and they do not provide enough flexibility to implement some alternative projection techniques like the exposed here.

These issues can be solved thanks to the recent release of the universal render pipeline (URP), a framework that provides VR developers the flexibility to control each aspect of the rendering pipeline. We are currently porting and improving our implemented techniques to this framework.

5.5 Improving spatial performance

In the previous section, we presented and validate the suitability of the using composition to influence the perception of distance in VR. By influencing the depth cues associated with linear perspective, it is possible to influence positively the perception of distances in VR. In the previous chapter, we describe the calculation of a linear performance predictor (LPP), as a mean to characterize the induced depth compression by the display. In this section, we will put everything together by describing how we can use the slope of the LPP to apply an optimal angular declination in the



Figure 5.11: Incorrect computation of shadow

Because we do not deal with the computation of light, when the participant moves his/her head to get another point of view, the direction of vertex is altered and their shadows become synchronized incorrectly.

lowering horizon technique, or an optimal scaling factor in the FOV minification technique, to improve the perception of distances in VR.

5.5.1 Optimized lowering horizon technique

According to the angular declination hypothesis $d = h/\tan(\alpha)$, the perceive distance d can be estimated as function of the subject's eye height h and the angular declination under the horizon line α . Based on this hypothesis, we can use the LPP to determine the amount of angular declination that must be applied in the horizon line technique to improve the perception of distances. The perceived horizon is based on a image at the focal distance, so we can use as reference the measured subject performance at that distance, and estimate the angular offset between the perceived and actual focal distances. Figure 5.12 shows the expected effect when decreasing the angular declination of the horizon line as a function of the focal distance. Lowering the horizon line by an angular offset β , increases the perception of distances of points either at zero parallax (at the focal distance), negative and positive parallax, with a greater influence at positive parallax. To calculate the angular offset β , the process is straightforward. First, we calculate the LPP using any assessment method, taking in consideration the

adequate stimuli selection. Then, we take subject performance at the focal distance f' and calculate the corresponding declination angle in relation with the horizon line. Thus, the angular declination at the underestimated focal distance is $\alpha' = \arctan(h/f')$. Finally, to raise the perception of distances, we can calculate the angular offset β , between the focal distance f and the predicted subject performance at the focal distance f' , which is the angle required to reduce the apparent location of the horizon line. Hence, the angular offset β is given by:

$$\beta = \alpha - \alpha' = \arctan(h/f) - \arctan(h/f') \tag{5.4}$$

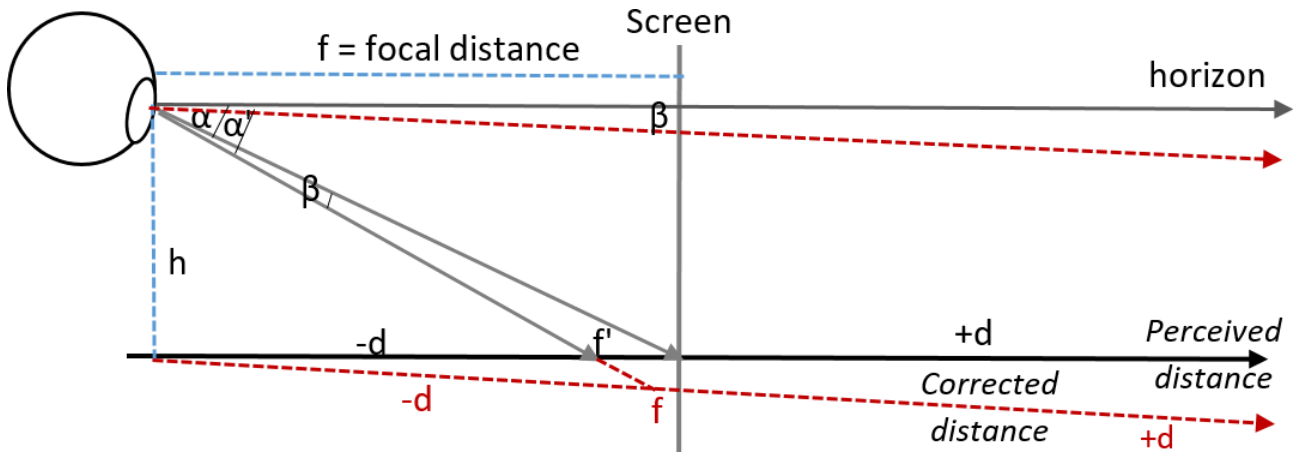


Figure 5.12: Effect of reducing the horizon line as a function of the focal distance.

After reducing the apparent location of the horizon line by a the angular offset β , the perception of distances is increased either in positive, negative and zero parallax, with a greater effect at positive parallax.

We can use an example to describe the process: after a perceptual matching test, subject's linear performance predictor was $d' = (1 - 0.20)d$, which means that for each meter, distances are underestimated a 20%. Assuming a focal distance of 0.75m (like in the HTC Vive's), the predicted perceived focal distance would be $f' = (0.8)(0.75m) = 0.6m$. If subject eye's height in relation with the ground is 1.6m , the angular declination required to increase the perception of distances a 20% would be $\arctan(1.6m/0.75m) - \arctan(1.6m/0.6m) = -4.5^\circ$. Thus, to increase the perception of distances a 20% in the display, the horizon line must be decreased 4.5 degrees.

5.5.2 Optimized FOV minification technique

The same analogy could be performed to the FOV minification technique. This technique scales the visual geometry with the purpose of artificially increasing the FOV. Steinicke et al. (2009a) defines minification/magnification as the ratio $m = \frac{\tan(vfov/2)}{\tan(vfov*g/2)}$, where m is the amount of uniform scaling that is required to map the viewport (rendered with a certain GFOV) to the display, $vfov$ is the vertical display FOV, and g is the geometric scaling factor applied to the FOV. If ($m = 1$), a person will perceive a spatially accurate image, as defined by the spatial dimensions of the virtual environment. When the geometric FOV is increased ($g > 1$), the resulting image is minified ($m < 1$), whereas a decreased geometric FOV ($g < 1$) results in a magnified image ($m > 1$). The amount of induced minification changes the perception of size, scaling the subtended angle of objects' projected image on the retina (Figure 5.13). Thus, following the size-distance-invariance hypothesis $s = d * \tan(\theta)$, where s is the perceived size of the object, θ is the subtended visual angle, and d is the object's distance; we can calculate the variance in the subtended angle θ' , using the LPP, so that $\theta' = 2 * \text{atan}(s/d')$, and then calculate the ratio θ/θ' . Because this ratio is equivalent to $\frac{\tan(vfov/2)}{\tan(vfov*g/2)}$, the necessary geometric scaling factor g required to improve the perception of distances is given by:

$$g = \frac{\theta}{\theta'} = \frac{\text{atan}(s/d)}{\text{atan}(s/d')} \quad (5.5)$$

As in the lowering horizon technique, we can use an example to describe calculation of the geometric factor g : after a blind assessment method, subject's linear performance predictor was $d' = (1 - 0.30)d$, which means that for each meter, distances are underestimated a 30%. Assuming that the size of the visual stimulus is $s = 1m$ and it is located at the focal distance of $0.75m$, which implies that the predicted perceived focal distance would be $f' = (0.7)(0.75m) = 0.52m$; then the geometric factor necessary to increase the perception of distances a 30%, it would be $g = \frac{\text{atan}(1m/0.75m)}{\text{atan}(1m/0.52m)} = 0.84$. Thus, to increase the perception of distances a 30% in the display, the amount of minification that must be induced in a display with a FOV of 110° is $m = \frac{\tan(110^\circ/2)}{\tan(110^\circ*0.84/2)} = 1.1$, the FOV must be scaled a 10%.

5.6 Summary

In this chapter, we presented the last part of the proposed framework, which is focused on influencing spatial perception using composition. We presented our implementation of the lowering horizon

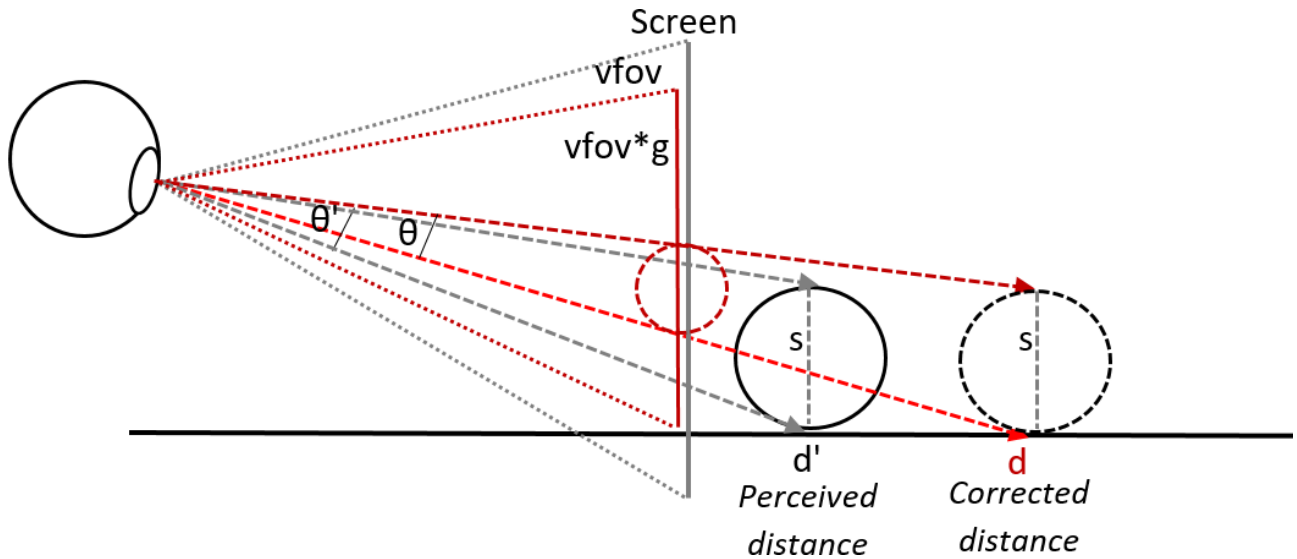


Figure 5.13: Effect of scaling the FOV as a function of the variation on the subtended visual angle.

After scaling the vertical FOV by the geometric factor g , the subtended angle of the object θ is reduced, according to its projected image and increasing its perception of distance.

technique and the FOV minification technique, and we validate the suitability of these techniques to improve spatial perception in VR. Our results suggest that both techniques can influence positively the perception of distances, but the selection of the required angular declination or the scaling factor for each technique is critical, as a value outside the optimal parameters could give rise to unexpected results. Thus, we presented two preliminary models about how the slope of the LPP can be used to determine the optimal angular declination for the lowering horizon technique and the optimal scaling factor for the FOV minification technique. We were interested in validating these optimized techniques in a final study, but due to the COVID 19 emergency, we cannot continue performing subjects' tests in our controlled environment, and we were forced to use a different approach based on remote testing. This restriction implied that we had to alter our experimental conditions or use different assessment methods. The results from this study are presented in Appendix A. We choose to not include these results in the document, because the changes performed in the studies may become confusing to the reader and we experience several methodological and technical issues. However, it shows the flexibility of the framework. We expect to continue this work and validate the suitability of the framework in the future, when the conditions become more favorable.

Chapter 6

Conclusions

This thesis presents a framework to characterize and influence spatial perception between heterogeneous VR displays, which has shown some promising results according to the research objectives proposed. The first objective was study and characterize the influence of the different display factors on spatial perception, where we described the influence of two of the most challenging issues in VR, the vergence-accommodation mismatch issue and the incorrect peripheral stimulation issue. The second objective was to find transversal assessment methods suitable to compare spatial perception between heterogeneous displays, where we adapted some study cases of assessment methods found in the literature. Our third objective was to analyse the influence of adaptation on spatial perception, where we study the effects of interacting with the VE on spatial perception and the influence of the different display issues. Finally, the last objective was to determine how composition can influence spatial perception, where we implemented some alternative projection techniques found in the literature, that can be optimized to improve the perception of distances. Although the results were not always aligned with our expectations, we are optimistic about the suitability of the framework. Some of the limitations and problems found in the different studies can be easily solved, while others likely required a different approach.

A fundamental contribution of this thesis is a formal approach to characterize the depth compression phenomenon in VR displays, by considering the influence of display factors, such as the focal distance, the stereoscopic parallax condition, the lens distortions and the limited FOV. Although these display factors have been studied for decades, there was not an approach that tried to integrate all these factors in a unique framework. Our approach was more holistic, where we considered the in-

fluence of all depth cues and display factors as a whole. However, it could be sensate to explore a more reductionistic, isolating the contribution of different depth cues and display factors. Although we consider the holistic approach more appropriate as it reflects how spatial perception works under natural conditions, a better solution could be a mix of both perspectives. Another conclusion of this thesis is that studying and influencing spatial perception in VR is as challenging as studying the same in the physical world. The proposed models were not 100% accurate as human perception is incredibly complex, so more studies are required about how the depth cues interact, how we perceived distances and how action influences perception. Thus, more future work is required to refine and improve the equations and models proposed in this document.

Another contribution of this thesis was the adaptation for the first time of the size-constancy test for HMDs, where the results of the studies demonstrated that it is possible to design transversal assessment methods. These studies not only demonstrated that the phenomenon of underestimation of distances is still relevant, but also that it is directly related to the different display factors, as it was reported by several researchers. We demonstrated again using the size-constancy test, that the perception of distances is influenced by the size of the V-A conflict and the stereoscopic parallax condition of the stimulus. Similarly, we validated using blind-walking tests, that the perception of distances is influenced by the size of the FOV and the distortions induced by the lens. These assessment methods were presented as study cases, but they are not the only spatial perception tests found in the literature. We hope this thesis can work as an inspiration, so that other researchers can adapt and design more transversal assessment methods that allow us to not only study spatial perception in VR, but also compare the performance between devices.

The linear performance predictor is a good method to characterize the depth compression phenomenon in the display but is not perfect, as it depends on subject confidence during their distance estimates. As more variability is introduced in the model, less accuracy in predicting subject performance occurs. Thus, a great part of this thesis was focused in reducing subjects' response variability, by introducing additional reference cues. Also, the framework starts from a set of assumptions, such as the idea that depth cues are integrated linearly and independently (following the Bayesian integration model), but it is possible that this does not occur in VR, or it depends on the kind of spatial task involved. Although the Bayesian integration model is the most accepted, other models and variations of the same model have been proposed (Landy et al. (2011)). Thus, more studies and knowledge about

the phenomenon are required.

A final contribution of this thesis is that we modernize some of the techniques based on composition that were proposed during the second generation of HMDs to influence the perception of distances in VR. Because the display calibration parameters cannot be modified easily using modern HMDs, we developed a object-level approach, introducing what we denominate "alternative-perspective"-view space transformations in the render pipeline before the projection is computed. Our studies showed that by using these alternative projection techniques, the perception of distances can be influenced positively, either in the personal or action space. Thus, we proposed a preliminary model to optimize these techniques based on the linear performance predictor, but we need to validate these in a further study and solve some of the limitations of the techniques themselves.

Finally, as a framework, this is a continuous work. We need to perform more studies to validate the assumptions of the framework, improve the alternative projection techniques, explore different methods to influence spatial perception, and study other kind of spatial tasks. We believe that this thesis, along with future improvements in technology, will help overcome the problem of underestimation of distances in VR. However, we believe that this thesis can have other applications than the ones exposed here. There exists situations in which an accurate spatial perception is not convenient. For example, experts golf players perceive the size of the hole much greater than amateur players (Memmert et al. (2009)), the best goalkeepers overestimate their own size (Masters et al. (2010)), drivers always underestimate distances when traveling at speeds between 40km/h and 60km/h (Baumberger et al. (2005)). Thus, the potential of using VR to characterize and influence these phenomena, where we can manipulate subject's spatial perception in situations that is not possible in the physical reality, becomes relevant.

Chapter 7

Un cadre de référence pour caractériser et influencer la perception spatiale sur des écrans hétérogènes de réalité virtuelle

7.1 Résumé

Nous proposons un cadre de référence pour caractériser et influencer la perception spatiale dans des écrans de réalité virtuelle (RV). Il est difficile, en réalité virtuelle, d'obtenir une perception spatiale précise en raison du phénomène de compression de la profondeur, qui donne lieu à une sous-estimation systématique des distances. La solution de ce problème est complexe car la technologie actuelle ne permet pas de simuler, dans des conditions raisonnables de temps et de complexité, la perfection de l'œil humain. Le problème est d'autant plus compliqué qu'il existe sur le marché une grande variété de systèmes d'affichage, ce qui rend difficile la conception d'expériences partagées nécessitant des performances spatiales précises. Il est vrai que différentes techniques ont été proposées pour améliorer la perception à distance, mais une solution mise en place pour un écran peut ne pas fonctionner pour un autre. En ce sens, nous proposons un cadre de référence pour caractériser et influencer la perception spatiale en réalité virtuelle. Ce cadre de référence prend en considération la variété des facteurs d'affichage qui affectent les performances pour construire un "prédicteur de performances linéaire", une fonction de régression linéaire qui décrit le phénomène de compression de la profondeur à l'écran. Ce prédicteur peut ensuite être utilisé pour appliquer une technique de projection alternative qui exerce une influence positive sur la perception spatiale du sujet en jouant avec les lois de la perspective. Par conséquent, le cadre de référence peut être utilisé pour améliorer la perception des distances en réalité

virtuelle et la conception d'expériences partagées plus naturellement entre des écrans hétérogènes.

Mots-clés : perception spatiale, réalité virtuelle, perception de la distance, perspective linéaire

7.2 Introduction

7.2.1 Le contexte

À la fin du vingtième siècle, l'Institut de recherche de l'Armée des États-Unis d'Amérique pour les sciences sociales et du comportement (ARI) a réalisé des études sur la perception des distances dans des environnements de réalité virtuelle (RV) comme un élément essentiel de la navigation dans des espaces de grande taille. Les chercheurs ont demandé aux usagers d'estimer des distances dans des environnements divers, et ont trouvé que de 41 à 72 pour cent des distances réelles étaient sous-estimées systématiquement (Witmer and Kline (1998), Lampton et al. (1995), Kline and Witmer (1996)). Ils ont trouvé également que l'introduction d'un plus grand réalisme ou de méthodes plus naturelles de mouvement n'éliminait pas les effets de sous-estimation. Patterson et al. (2006), un groupe de chercheurs de l'ARI, ont fait une revue exhaustive des différents facteurs d'affichage et des problèmes perceptifs qu'ils induisent, précisant que quelques unes des limitations techniques restent des thèmes de recherche ouverts. Par comparaison avec l'œil humain, les écrans de RV sont loin d'être des systèmes parfaits. Les casques de RV actuels (HMDs) ne fournissent qu'un champ de vision horizontal de 130°, plus petit que le champ de vision humain (environ 200°). En outre, la résolution à la vision centrale est pauvre. L'acuité visuelle maximale est estimée aux alentours de 20/60, très loin de 20/20 ou vision normale (Kreylos (2017)). Bien qu'il y ait des projets prometteurs tels que le projet Pimax, premier HMD avec un champ de vision large (200°) et résolution 8K (Apress (2019)), d'autres limitations connues, à savoir, par exemple, le conflit entre vergence et accommodation, restent des problèmes complexes. En plus, l'utilisation de lentilles induit d'autres artefacts visuels non désirés, comme les aberrations chromatiques qui déforment le stimulus visuel. Ces limitations affectent la fidélité sensorielle des HMD.

Un groupe de chercheurs, conscients de ces limitations, a conçu un type d'écran différent, un écran immersif fondé sur la projection (connu par le sigle LIPD). Il s'agit de systèmes qui utilisent des écrans très grands et des lunettes stéréo actives/passives qui les rendent moins susceptibles aux distorsions

7.2. INTRODUCTION

optiques et autres artefacts visuels fréquents dans les HMD. Le précurseur de ce type de système est le CAVE (Cruz-Neira et al. (1993)), une salle de forme cubique conçue comme l'approximation d'une sphère dont les parois latérales sont formées typiquement par des écrans de rétroprojection. Grâce à l'utilisation de grands écrans, les LIPD fournissent un champ de vision presque naturel, une plus grande acuité visuelle et une moindre influence des distorsions optiques. Malgré tous ces avantages, les LIPD sont des systèmes très coûteux et peu portables, ce qui les rend plus adaptés à des applications industrielles. La figure 7.1 présente deux exemples des principales tendances dans la création d'écrans pour la réalité virtuelle.



Figure 7.1: Tendances dans la création d'écrans pour la réalité virtuelle..

À gauche: écran monté sur la tête (HMD). À droite : grand écran de projection immersive (LIPD) connu comme CAVE.

Dans cette section nous décrivons les problèmes de perception spatiale en réalité virtuelle et le problème de la sous-estimation des distances. Les HMD sont plus sensibles à ce phénomène en raison de leurs limitations par rapport au champ de vision, à l'acuité visuelle et aux distorsions optiques. D'autre part, les LIPD tels que le CAVE sont moins susceptibles à de telles distorsions. Choisir un système consiste à choisir entre la fidélité sensorielle et la portabilité. Les HMD sont des systèmes très portables mais ils ont un coût en termes de fidélité sensorielle. En revanche, les LIPD offrent une meilleure fidélité sensorielle, mais ils ne peuvent être utilisés que dans un espace fixe. Bien que les deux systèmes se fondent sur le même principe stéréoscopique, ils sont très hétérogènes, non seulement en termes de fidélité sensorielle mais aussi, comme nous le verrons plus loin, quant à la façon dont nous étudions et évaluons la perception spatiale.

7.2.2 Énoncé du problème

La perception spatiale en réalité virtuelle diffère du monde physique. Les études décrivent une sensation de compression de la profondeur où les distances sont sous-estimées (pour une revue complète, voir Renner et al. (2013), El Jamiy and Marsh (2019)). Depuis que les premières informations sur le phénomène ont été publiées il y a presque 30 ans, ce sujet est devenu un problème déconcertant et stimulant. Renner et al. (2013) ont établi trois catégories pour cette phénoménologie: technique (technologie d’affichage et ses limitations), compositionnelle (techniques qui créent l’illusion de profondeur: par exemple, la stéréoscopie et la perspective linéaire) et humaine (calibration et adaptation) (voir Figure 7.2). Il est important de tenir compte également de la méthode de mesure, puisque certains tests ne conviennent pas à tous les types d’écran.

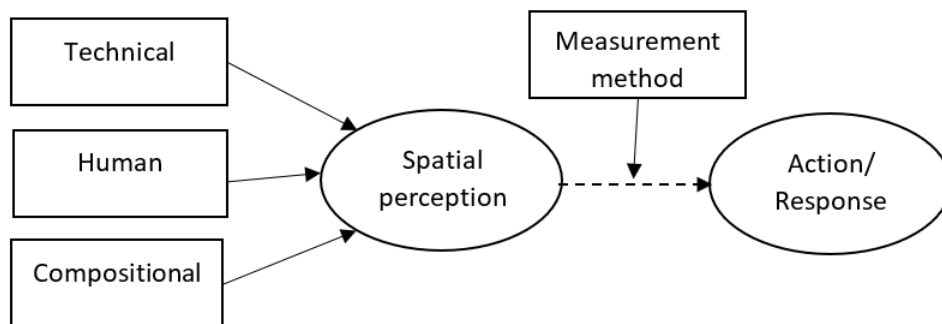


Figure 7.2: Zones qui influent sur la perception spatiale en réalité virtuelle.

Adapté de Renner et al. (2013).

Les facteurs techniques expliquent la différence entre la perception spatiale et le monde physique. Les écrans de RV modernes fondés sur la stéréoscopie sont confrontés à un problème inévitable, celui de l’impossibilité d’inférer la profondeur d’un objet en fonction des trajectoires de lumière des objets, cette profondeur devant plutôt être inférée de façon artificielle en fonction de deux projections pour chaque œil sur un écran plat (Hoffman et al. (2008), Jones et al. (2001)). Ceci pose des problèmes naturels de perception de la profondeur, tels que le conflit entre la vergence et l’accommodation, un problème qui a été étudié considérablement pendant la dernière décennie. De même, le champ de vision réduit et la nature de la lumière et sa résolution induite sur la vision périphérique constituent d’autres limitations. Par conséquent, certains écrans induisent davantage de distorsions dans la perception spatiale que d’autres. Même entre plusieurs HMD, il peut y avoir des différences en fonction du

modèle, ce qui rend difficile la réalisation d'expériences partagées de façon naturelle.

En ce qui concerne les méthodes d'évaluation, des méthodes d'évaluation transversales font défaut. Plusieurs méthodes ont été proposées pour évaluer la perception spatiale en réalité virtuelle. La plupart des travaux se sont concentrés sur la perception de la distance de façon égocentrique dans des HMD. La méthode la plus populaire demande aux participants d'estimer la distance d'une cible en marchant en aveugle. Cependant, cette méthode d'évaluation ne convient pas pour les LIPD en raison de leurs contraintes spatiales. Ainsi, des méthodes alternatives fondées sur des couplages perceptifs ont été proposées. Par exemple, des méthodes qui évaluent la perception des distances de façon indirecte en utilisant la perception de la taille.

Enfin, les chercheurs se sont penchés principalement sur la solution des facteurs techniques, sans considérer les facteurs de composition. Bien que les techniques d'affichage à champ lumineux (Light Field Displays) (Huang et al. (2015)), d'affichage à surface focale (Focal Surface Displays) (Matsuda et al. (2017)) et de suivi oculaire (Fu et al. (2016)) soient des technologies prometteuses qui peuvent résoudre le problème partiellement, les chercheurs ne cessent de travailler pour réduire leur complexité. Par contraste, la création des graphiques modernes sur l'ordinateur se fonde sur les techniques de la perspective découvertes pendant la Renaissance, qui rendent une illusion de profondeur sur une surface bidimensionnelle. Les artistes et les architectes ont joué avec les règles de la perspective au long des siècles pour créer des illusions de profondeur qui ne suivent pas les règles du monde physique. Le pouvoir de la composition dans la perception spatiale est tellement fort qu'il peut annuler l'influence d'autres signaux incorrects, tels que ceux d'accommodation-vergence, et peut constituer une solution plus simple et efficace du problème.

7.2.3 Question de recherche

La perception spatiale en réalité virtuelle peut-elle être caractérisée et améliorée, prenant en considération les différents facteurs d'affichage qui influent sur la performance, afin de concevoir des expériences partagées plus naturelles entre les écrans ?.

7.2.4 Objectifs de la recherche

- Étudier et caractériser l'influence des différents facteurs d'affichage sur la perception spatiale.

- Trouver/concevoir des méthodes d'évaluation qui servent à comparer la perception spatiale entre les écrans.
- Déterminer comment la composition peut-elle être utilisée pour améliorer la perception spatiale.

7.3 Revue de la littérature

7.3.1 Vue d'ensemble de la perception spatiale

7.3.1.1 Signaux de la perception de la profondeur

Cutting and Vishton (1995) ont développé un modèle sur la sensibilité de différents signaux de profondeur en fonction de la distance (voir Figure 7.3). Ils ont introduit la notion selon laquelle les humains divisent leur environnement en trois régions : (1) espace personnel, centré sur la manipulation d'objets et sur des tâches motrices fines ; (2) espace d'action, centré sur des tâches associées au mouvement autonome ; et (3) espace éloigné, visant des tâches associées à la navigation. La capacité à pouvoir discriminer la profondeur de chaque signal varie avec la distance, de sorte que certains signaux deviennent plus importants pour certains espaces que pour d'autres.

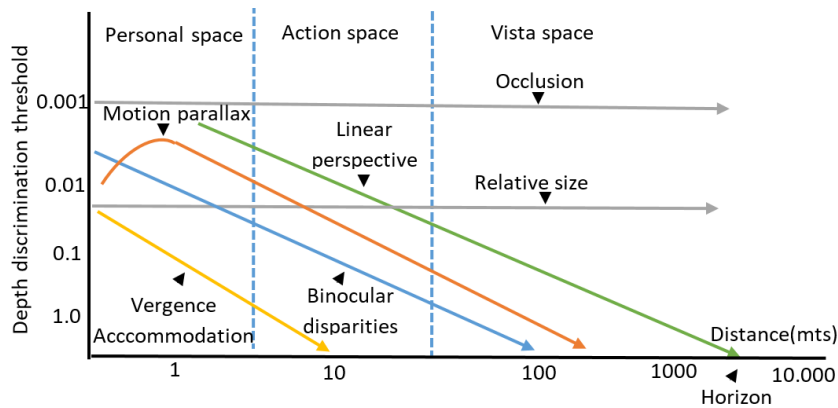


Figure 7.3: Capacité à discriminer la profondeur de différents signaux en fonction de la distance.

Adapté de Cutting and Vishton (1995)

7.3.1.2 Illusions de profondeur et conflits de signaux

L'intégration de signaux de profondeur dans le système visuel humain est susceptible à des biais. Un exemple en est la fameuse illusion optique de la chambre d'Ames, où les personnes peuvent ap-

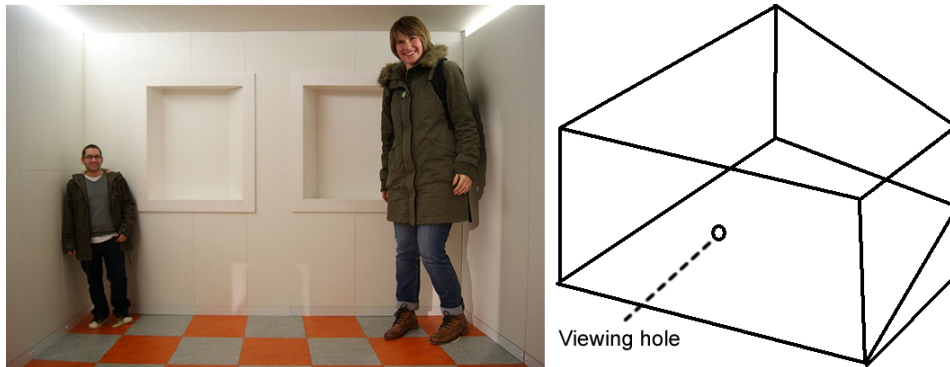


Figure 7.4: Illusion de la chambre d'Ames

À gauche: illusion de la chambre d'Ames, où les personnes changent apparemment de taille. À droite: la forme réelle de la chambre. Photo par Ian Stannard/flickr (Stannard (2010)).

paremment changer de taille (voir figure 7.4). La forme particulière de la chambre induit un conflit de signaux entre les signaux oculomoteurs et les signaux picturaux, dans lequel ces derniers l'emportent (Dorward and Day (1997), King et al. (1976)). Ainsi, cette illusion est une évidence acceptée des théories de la perception indirecte et l'importance de l'expérience dans les phénomènes de la perception. Certaines études suggèrent que l'illusion se produit à cause d'une rupture de la constance de la forme, d'un biais dans la perception humaine vers la symétrie et les schémas de forme réguliers, tels que les chambres rectangulaires (Dorward and Day (1997), King et al. (1976)).

7.3.1.3 Méthodes de mesure

La plupart des travaux en RV se sont concentrés sur la perception de distances égocentriques. Plusieurs méthodes ont été proposées pour évaluer la façon dont les personnes perçoivent les distances: (1) l'estimation verbale, méthode la plus simple mais aussi la moins précise (Andre and Rogers (2006)); (2) l'appariement perceptif, où la perception de la distance est évaluée de façon indirecte au moyen de couplages perceptifs (faisant appel, par exemple, à sa relation avec la perception de la taille) (Viguiet et al. (2001), Haber and Levin (2001)); (3) les méthodes en aveugle, où les participants estiment une distance en réalisant une action généralement sans vision. De cette catégorie, nous pouvons faire ressortir la méthode de la marche en aveugle, où les participants doivent se déplacer vers une cible précédemment observée (Fukushima et al. (1997)). Il a été mis en évidence que, dans des conditions naturelles, les êtres humains sont assez précis quand ils réalisent des tâches de marche en aveugle

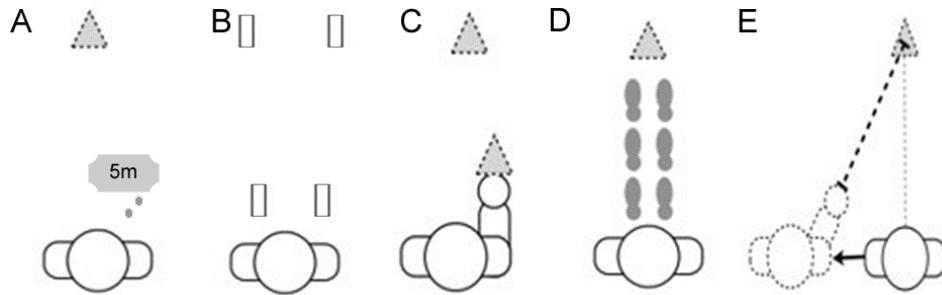


Figure 7.5: Méthodes pour évaluer la perception des distances

Méthode d'évaluation les plus populaires. A : l'estimation verbale. B/C : l'appariement perceptif. D : la marche en aveugle. D : la triangulation aveugle.

(voir, pour une revue complète, Loomis et al. (2003)). La figure 7.5 décrit les méthodes d'évaluation les plus populaires.

7.3.1.4 Action and perception

Il existe deux façons différentes de percevoir l'espace. Deux processus visuels différents ont été mis en évidence dans le cortex cérébral. Le processus ventral (ou vision pour la perception), centré sur les caractéristiques spatiales de l'objet (distance, forme, taille), et le processus dorsal (ou vision pour l'action), centré sur les actions motrices guidées visuellement Goodale and Milner (1992). Il existe également deux perceptions différentes des distances: une perception ventrale, centrée sur les jugements spatiaux des objets (près, loin, court, long), et une perception dorsale, spécialisée dans la transformation de l'information en tâches motrices, à chaque instant. Selon Loomis et al. (1992), les représentations spatiales variaient en fonction de la réponse, selon que celle-ci comprenait des réponses motrices (méthodes en aveugle) ou des rapports perceptifs (méthodes verbales ou d'appariement perceptif). Aujourd'hui, il est clair que la perception des distances est influencée par l'action particulière réalisée, l'information optique (flux optique, parallaxe de mouvement) étant calibrée naturellement en fonction des signaux moteurs du perceuteur (Proffitt et al. (2003), Witt et al. (2004), Witt and Proffitt (2008)).

7.3.2 Problèmes d’affichage

Stereoscopic displays suffer from the inevitable issue that an object cannot be recreated based on the light beams reflected on its surface and falling on the retina, but it must be recreated artificially using two images on a flat screen. Assuming that modern computer graphics simulated light in a physically accurate manner (which is probably true), light is in some way intercepted, projected on two flat screens, transformed into pixels, and magnified using an optic lenses to make objects look larger and further away. Unfortunately, this is the best we can do. There is no way with current technology to simulate, in a reasonable amount of time and in a optimal hardware complexity, the induced light on the retina from every object and surface in the scene, and in response for every oculomotor action. Indeed, the evolution of the technology makes evident the difficulties of emulating the perfectness of the human visual system. Despite the advances in autostereoscopic displays, volumetric displays and holographic displays, only displays based on traditional stereoscopy have been successful, less problematic, and affordable to consumers (Hong et al., 2011). The dependency on the stereoscopic paradigm causes natural problems on depth perception, where we can highlight two challenging problems: the *vergence-accommodation mismatch issue* and the *incorrect peripheral light stimulation*.

7.3.2.1 Conflit vergence-accomodation (conflit V-A)

Dans des conditions naturelles, il y a une parfaite harmonie entre la vergence et l’accommodation; cette harmonie est fonction d’une relation étroite entre les signaux moteurs (proprioceptifs et efférents) et le stimulus visuel proprement dit. La profondeur d’un objet peut être déduite soit en comparant la disparité horizontale entre les images rétiniennes (disparité binoculaire) et les signaux moteurs qui indiquent l’orientation des yeux (vergence) (Viguiet et al. (2001), Tresilian et al. (1999), Mon-Williams et al. (2000) Viguiet et al. (2001)), soit en analysant le flou de l’image dans la rétine et les signaux sur la forme de la lentille/iris (accommodation) (Mather (1997), Watt et al. (2005), Held et al. (2010)). Malheureusement, cette harmonie est brisée en réalité virtuelle, comme on peut le voir sur la figure 7.6. En regardant un objet, les yeux s’accommodent sur le plan de l’écran, mais convergent vers la localisation apparente de l’objet, créant ainsi un conflit entre ces signaux. Par conséquent, les objets qui devraient, dans des conditions naturelles, être flous, sont mis au point. Le système visuel se trouve dans un conflit du type “les objets apparaissent plus proches en fonction des signaux accommodatifs, mais semblent plus proches en fonction de la vergence”.

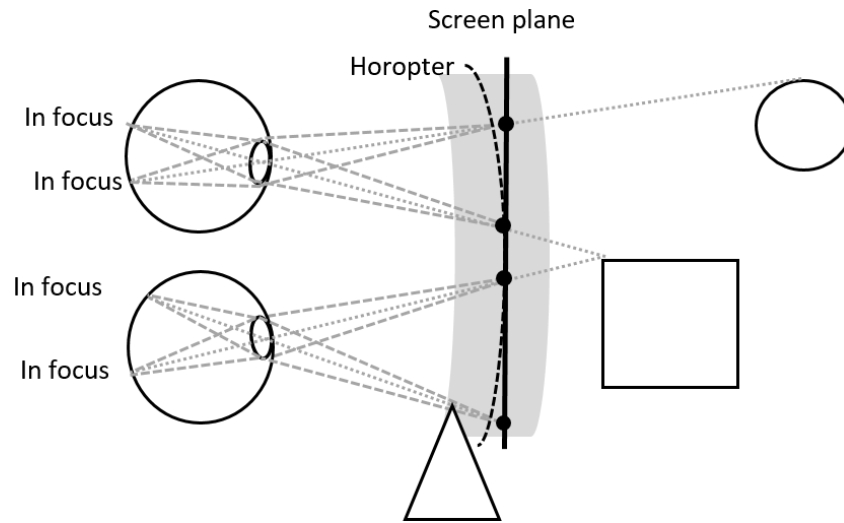


Figure 7.6: Le conflit entre accommodation et vergence

Les yeux convergent sur la localisation apparente du stimulus, mais ils s'accommodent de façon incorrecte sur l'écran, de sorte que les objets qui devraient être flous sont mis au point. Adapté de Wartell (2002).

Zone de fusion confortable

En vision stéréoscopique, l'accommodation doit répondre à des demandes opposées dont la sévérité dépend de la vergence associée. Il existe une limite à laquelle le percepteur peut orienter ses yeux pour se centrer sur un objet (vergence) et ajuster la taille des pupilles pour le mettre au point (accommodation), et en même temps maintenir une image nette de l'écran sans induire un conflit sévère. Les développeurs d'écrans ont étudié les limites entre lesquelles ces signaux peuvent différer les uns des autres, une limite connue comme la zone de fusion confortable (Hoffman et al. (2008), Reichelt et al. (2010), Shibata et al. (2011), Tam et al. (2011)) et qui est intrinsèquement liée à la profondeur de champ humaine (PDC). Dans une étude exhaustive, Shibata et al. (2011)) ont estimé la zone de confort pour différentes technologies 3D. Ils ont conclu que l'éventail de distances que nous pouvons présenter confortablement dépend beaucoup de la distance focale. La distance focale est le conducteur qui définit la zone de confort et la plage de profondeur fonctionnelle disponible pour l'écran ou les limites entre lesquelles la vergence et l'accommodation peuvent être utilisées pour percevoir la profondeur.

Compression de la profondeur et distance focale

En raison des limites de la plage de profondeur fonctionnelle des écrans stéréo, la compression en

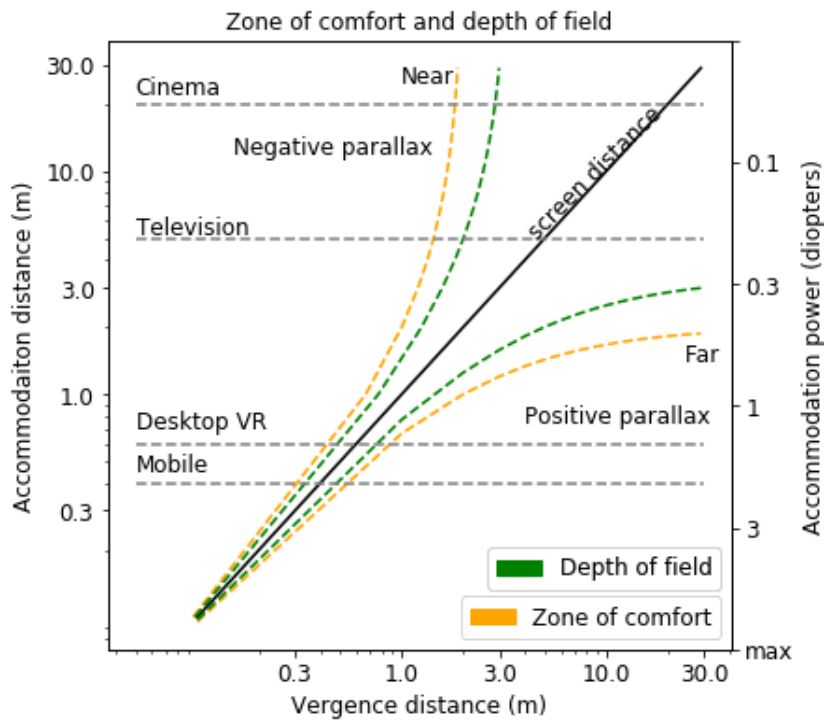


Figure 7.7: Zone de fusion confortable et profondeur de champ

La zone de fusion confortable définit les limites au sein desquelles la vergence et l'accommodation pourraient différer sans provoquer un malaise. Cela dépend de la distance focale de l'écran et de la profondeur de champ humaine. Adapté de Shibata et al. (2011).

profondeur de la scène 3D doit être plus importante sur l'écran ayant la plus petite distance focale (si la profondeur de la scène est plus grande que la plage de profondeur disponible) (Hoffman et al. (2008)). La figure 7.8 montre la compression induite quand le perceuteur fixe les yeux sur le plan de l'écran. L'image montre l'effet de compression asymétrique, qui est une conséquence de la nature expansive du champ de profondeur. Les objets ont tendance à être moins compressés quand ils se trouvent en face de l'écran que ceux qui se trouvent derrière lui. Par conséquent, la sous-estimation est plus forte à mesure de l'augmentation de la distance. Les casques HMD sont très sensibles à ce phénomène car ils utilisent des lentilles avec des distances focales fixes, ce qui restreint la zone de confort à une plage limitée, plus utile pour des objets situés loin du perceuteur que pour des objets situés près de lui. En revanche, les LIPD sont plus "flexibles" car ils permettent plusieurs distances focales en fonction de la localisation du perceuteur par rapport au mur de projection.

Condition de parallaxe stéréoscopique

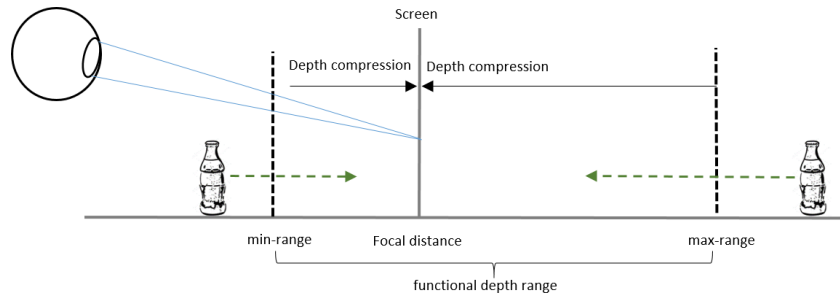


Figure 7.8: Plage de profondeur fonctionnelle et compression de profondeur.

La distance focale détermine la plage de profondeur fonctionnelle dans laquelle toute la scène est compressée.

La localisation de l’objet virtuel par rapport à l’écran est décrite généralement comme la condition de parallaxe stéréoscopique, selon laquelle les objets placés en face de l’écran sont décrits comme situés en parallaxe négative, et les objets placés derrière l’écran sont décrits comme situés en parallaxe positive. Des études ont montré une relation directe entre le degré de sous-estimation et la condition de parallaxe stéréoscopique. Ces études ont été réalisées principalement avec des LIPD, qui permettent d’accommoder les yeux à des distances différentes (Nacéri et al. (2010), Murgia et al. (2009), Marsh et al. (2014), Bruder et al. (2016), Kenyon et al. (2008), Luo et al. (2007)). Selon Vienne et al. (2020), la taille du conflit peut être calculée en soustrayant l’inverse de l’accommodation de l’inverse de la vergence :

$$C_{va} = \frac{1}{V_d} - \frac{1}{A_d} \quad (7.1)$$

où C_{va} désigne la taille du conflit V-A, V_d désigne la distance de la convergence et A_d désigne la distance d’accommodation. La figure 7.1 montre le taille du conflit en fonction de la distance focale.

7.3.2.2 Stimulation périphérique incorrecte (SPI)

Dans la section 7.3.1.4, nous avons décrit comment l’action influe sur la perception et comment le flux optique est calibré naturellement avec des tâches motrices, telles que la marche. Les HMD induisent plusieurs distorsions qui ont un impact négatif sur le flux optique, particulièrement sur la vision périphérique. Le premier facteur est le champ de vision limité, où il a été mis en évidence que la nature de la stimulation induite dans la vision périphérique influe sur la perception des distances (Jones

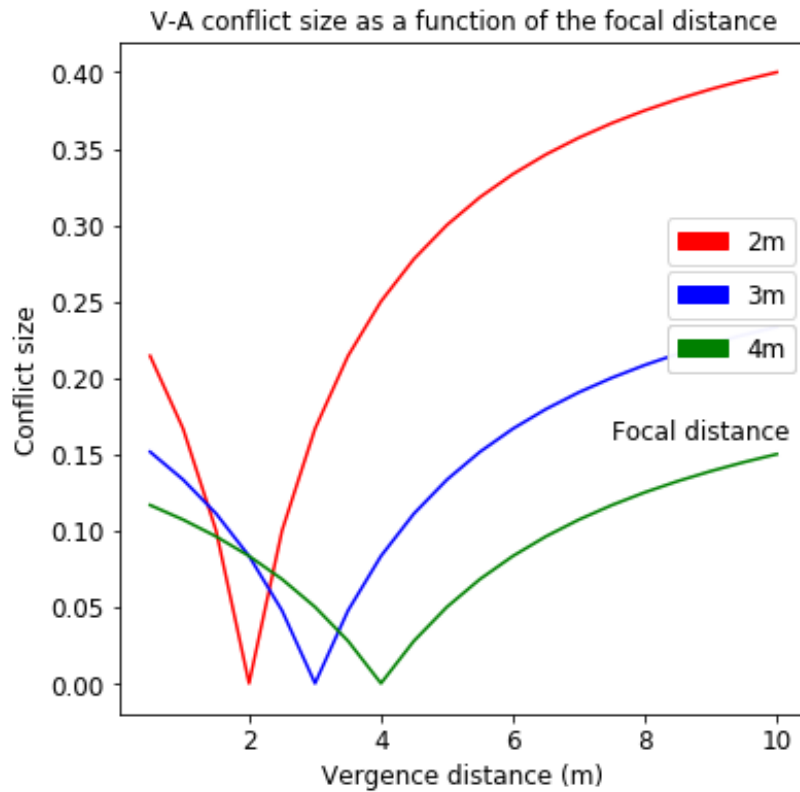


Figure 7.9: Taille du conflit en fonction de la distance focale

Le biais dans la perception des distancias est proportionnel à la taille du conflit V-A.

et al. (2013), Li et al. (2018)). Un autre facteur qui influe sur le flux optique est celui des distorsions induites par les lentilles. Les lentilles déforment l'acuité visuelle, car la résolution angulaire ne subit pas une variation uniforme depuis un pixel situé au centre de l'écran jusqu'aux pixels situés sur les bords (Kreylos (2017)). Ceci affecte la parallaxe de mouvement, car lorsque les objets s'approchent de la vision périphérique, ils bougent à un rythme plus rapide vers la périphérie du champ visuel, influant sur la perception des distances du stimulus dans le temps.

7.3.3 Methods that improve spatial perception in VR

Un objectif de ce cadre de référence est d'explorer des méthodes pour influencer sur la perception des distances en utilisant la composition. La composition consiste à représenter des objets 3D sur une surface plane en utilisant différentes techniques qui créent l'illusion de profondeur, telles que la perspective linéaire. Plusieurs méthodes ont été proposées pour améliorer la perception des distances

7.4. UN CADRE DE RÉFÉRENCE POUR CARACTÉRISER ET INFLUENCER LA PERCEPTION SPATIALE SUR DES ÉCRANS HÉTÉROGÈNES DE RÉALITÉ VIRTUELLE

en RV qui jouent avec les lois de la perspective. Ces méthodes sont basées sur des techniques de projection alternatives qui ne suivent pas la projection de perspective linéaire standard utilisée pour la création de graphiques sur ordinateur (pour une vue d'ensemble, voir Salomon (2007)). Parmi ces méthodes, nous en avons choisi deux qui nous semblaient prometteuses: la *minification du champ de vision* et la *déclinaison de l'horizon*.

- La *minification du champ de vision* est une méthode qui augmente artificiellement le champ de vision géométrique de l'écran. La méthode a démontré qu'elle peut améliorer la perception de la distance et réduire les effets de sous-estimation (Kuhl et al. (2006), Bolte et al. (2010), Zhang et al. (2012), Li et al. (2014), Steinicke et al. (2009a)). Ayant appliqué une échelle de 70%-82% aux images visuelles, ses résultats ont montré une augmentation de 13%-20% des distances perçues.
- La *déclinaison de l'horizon* est une autre méthode qui applique artificiellement une déclinaison angulaire sur la ligne d'horizon (Messing and Durgin (2005), Williams et al. (2009), Kuhl et al. (2009)). La méthode a montré une augmentation dans les distances perçues, liée directement au degré de la déclinaison angulaire. Cette méthode se fonde sur l'hypothèse de la déclinaison angulaire $d = h/\tan(\alpha)$ (Ooi et al. (2001), Gajewski et al. (2014), Todorovic and Toskovic (2012), Bunch (2014)), une théorie selon laquelle la perception d'une distance d est influencée par la hauteur oculaire h de l'observateur et la déclinaison angulaire de la cible par rapport à la ligne d'horizon α .

7.4 Un cadre de référence pour caractériser et influencer la perception spatiale sur des écrans hétérogènes de réalité virtuelle

7.4.1 Vue d'ensemble

Nous proposons un cadre de référence pour caractériser et influencer la perception des distances dans des écrans hétérogènes de réalité virtuelle. Le cadre de référence proposé tient compte des différents facteurs d'affichage qui déforment la perception spatiale et peut être vu comme une extension du cycle de perception de l'action en réalité virtuelle (figure 7.10). Un stimulus virtuel est présenté au perceuteur à l'aide d'un écran. Le perceuteur crée une représentation spatiale du stimulus basée sur sa perception spatiale et réalise une action motrice en conséquence. L'environnement virtuel (VE/EV)

7.4. UN CADRE DE RÉFÉRENCE POUR CARACTÉRISER ET INFLUENCER LA PERCEPTION SPATIALE SUR DES ÉCRANS HÉTÉROGÈNES DE RÉALITÉ VIRTUELLE

fournit un retour visuel en fonction de ses actions, modifiant la représentation spatiale du stimulus en soi. Dans des conditions optimales, la représentation spatiale du perceuteur et la représentation spatiale se rapprochent, puisque les signaux de profondeur se comportent naturellement. Toutefois, l'écran déforme quelques uns de ces signaux, ce qui affecte la perception des distances. Ainsi, pour déterminer comment la perception est affectée, une méthode d'évaluation est nécessaire. Certaines méthodes mesurent la perception des distances indirectement (appariement perceptif), tandis que d'autres en font une évaluation directe (méthodes en aveugle). Le cadre de référence proposé se fonde donc sur l'idée que si nous sommes capables de prédire comment la perception est déformée par les différents facteurs d'affichage// (en utilisant une méthode d'évaluation et en calculant un prédicteur de performance linéaire), nous pouvons modifier la simulation elle-même pour influencer positivement la perception des distances en utilisant une technique compositionnelle. Par conséquent, le cadre de référence peut être utilisé pour caractériser et améliorer la perception des distances dans des écrans de réalité virtuelle.

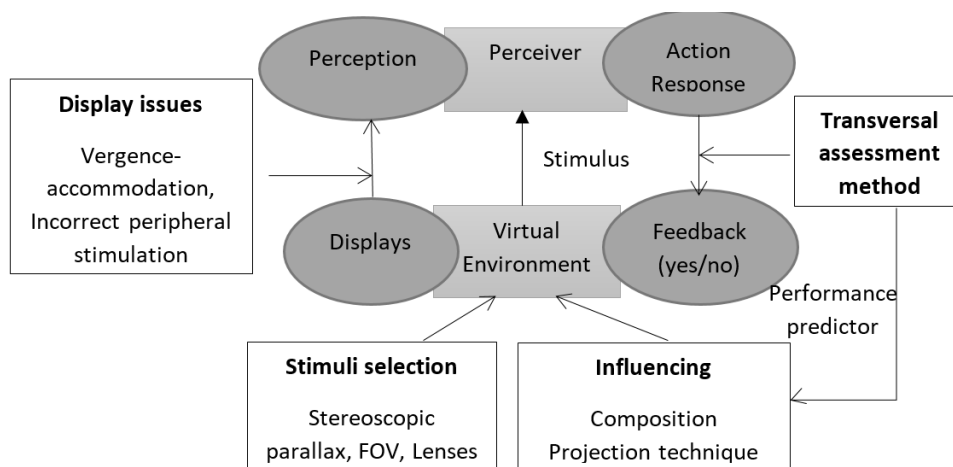


Figure 7.10: Le cadre de référence proposé vu comme une extension du cycle action-perception en réalité virtuelle.

7.4.2 Sélection des méthodes d'évaluation

Dans une première étape, nous avons retenu deux tests de perception spatiale comme méthodes d'étude pour valider le cadre de référence proposé. Le test de constance de la taille (Kenyon et al. (2007)) un test d'appariement perceptif qui demande au sujet de réaliser des estimations basées sur une référence physique. S'agissant d'une méthode d'appariement perceptif, les distances courtes con-

7.4. UN CADRE DE RÉFÉRENCE POUR CARACTÉRISER ET INFLUENCER LA PERCEPTION SPATIALE SUR DES ÉCRANS HÉTÉROGÈNES DE RÉALITÉ VIRTUELLE

stituent l'espace le mieux indiqué pour son application, pas très loin de l'espace personnel. Pour l'espace d'action, nous avons retenu le test classique de marche en aveugle (Loomis et al. (1992)). Ce test correspond à la catégorie des méthodes en aveugle qui demandent au sujet d'estimer une distance en marchant sans vision.

7.4.2.1 Test de constance de la taille

Kenyon et al. (2007) ont proposé une méthode pour évaluer la perception spatiale en réalité virtuelle à partir de la relation entre la perception de la taille et la distance. La constance de la taille est un phénomène perceptif qui fait qu'un objet soit perçu de la même taille indépendamment de la distance à laquelle il se trouve et de la taille de son image projetée sur la rétine. Ceci est dû au contexte, où des signaux tels que la perspective linéaire et la taille des objets environnants influent sur la distance et la taille perçue de l'objet. Il est donc possible d'évaluer la perception des distances d'une manière indirecte, en utilisant la taille relative d'un objet familier comme référence. Sur cette base, les chercheurs ont étudié la perception de constance de la taille dans un système CAVE, en demandant aux participants d'estimer la taille relative d'un objet familier placé sur une table virtuelle à des distances différentes et en utilisant une référence physique (figure 7.11). L'objet de référence est placé sur un côté et à la même hauteur de la table. Ensuite, quelques répliques virtuelles de l'objet sont présentées avec des dimensions exagérées, la table faisant fonction de contexte. Les participants ont ajusté la taille des répliques virtuelles jusqu'à atteindre une coïncidence perceptive avec la taille de la référence physique. La performance a été évaluée en utilisant une mesure de base appelée *Size-Ratio*, qui représente la taille estimée de l'objet virtuel par rapport à la taille réelle de l'objet (équation 7.2). Un rapport de 1 signifie que la taille (et la distance) de l'objet est parfaitement perçue; un rapport < 1 signifie que sa taille est sous-estimée (ou que sa distance est surestimée), et un rapport > 1 signifie que la taille est surestimée (ou que sa distance est sous-estimée).

$$SizeRatio = EstimatedSize/CorrectSize \quad (7.2)$$

S'agissant d'une méthode d'appariement perceptif, ce test fonctionne mieux avec des LIPD qu'avec des HMD, parce que le percepteur peut voir simultanément la référence physique et ses répliques virtuelles. Afin de le rendre apte aux HMD, nous avons proposé une méthode permettant aux usagers d'utiliser l'objet physique comme référence sans qu'ils aient à enlever le casque. Nous avons donc

7.4. UN CADRE DE RÉFÉRENCE POUR CARACTÉRISER ET INFLUENCER LA PERCEPTION SPATIALE SUR DES ÉCRANS HÉTÉROGÈNES DE RÉALITÉ VIRTUELLE

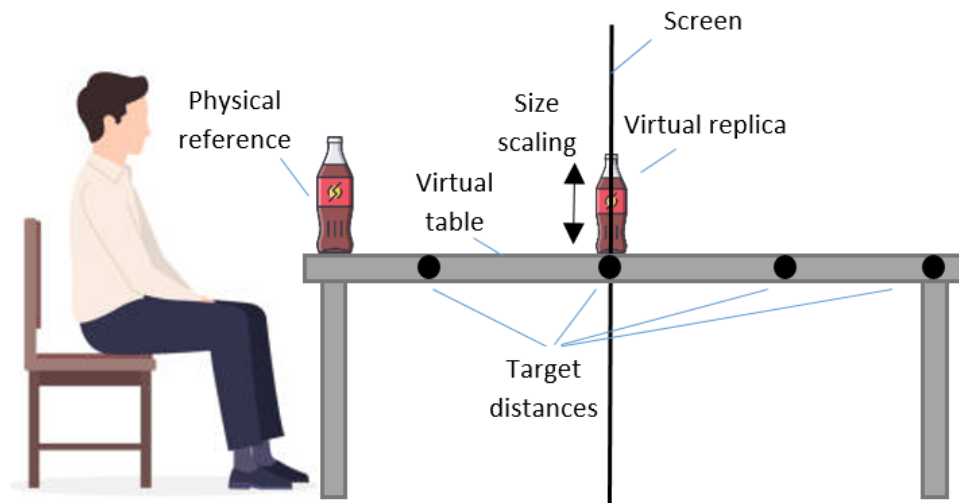


Figure 7.11: Test de constance de la taille

Les participants doivent estimer la taille relative d'un objet virtuel situé à des distances différentes en utilisant une référence physique.

proposé de réaliser cette tâche en utilisant le sens de la proprioception plutôt que le sens de la vue. La proprioception n'étant pas influencée par la visualisation, un conflit est induit entre les signaux moteurs et le stimulus visuel. Nous avons posé comme hypothèse que, pour le HMD, l'effet négatif sur la perception de la taille (et des distances) est plus important que pour le LIPD, puisqu'il s'agit d'un écran plus susceptible d'induire une compression de la profondeur. Ainsi, nous avons adapté le test de constance de la taille et réalisé une étude comparative entre un CAVE et un HMD.

7.4.2.2 Comparaison de la constance de la taille entre HMD et CAVE

Une étude a été réalisée pour comparer la perception de la distance entre un HMD moderne (HTC Vive CV1) et un CAVE. Contrairement au test original, les utilisateurs ont estimé la taille de la référence physique en utilisant leur proprioception, au lieu d'utiliser leur vue. Nous avons donc remplacé la table virtuelle par une table physique et adhérent l'objet de référence au centre de la table, à une distance fixe du bord, de sorte que les participants pouvaient les manipuler et utiliser leur corps comme référence (figure 7.12 - droite). Pour assurer la similarité entre les deux conditions, les participants sur CAVE n'étaient pas autorisés à voir la référence physique et devaient aussi utiliser leur proprioception (figure 7.12 - gauche).

7.4. UN CADRE DE RÉFÉRENCE POUR CARACTÉRISER ET INFLUENCER LA PERCEPTION SPATIALE SUR DES ÉCRANS HÉTÉROGÈNES DE RÉALITÉ VIRTUELLE

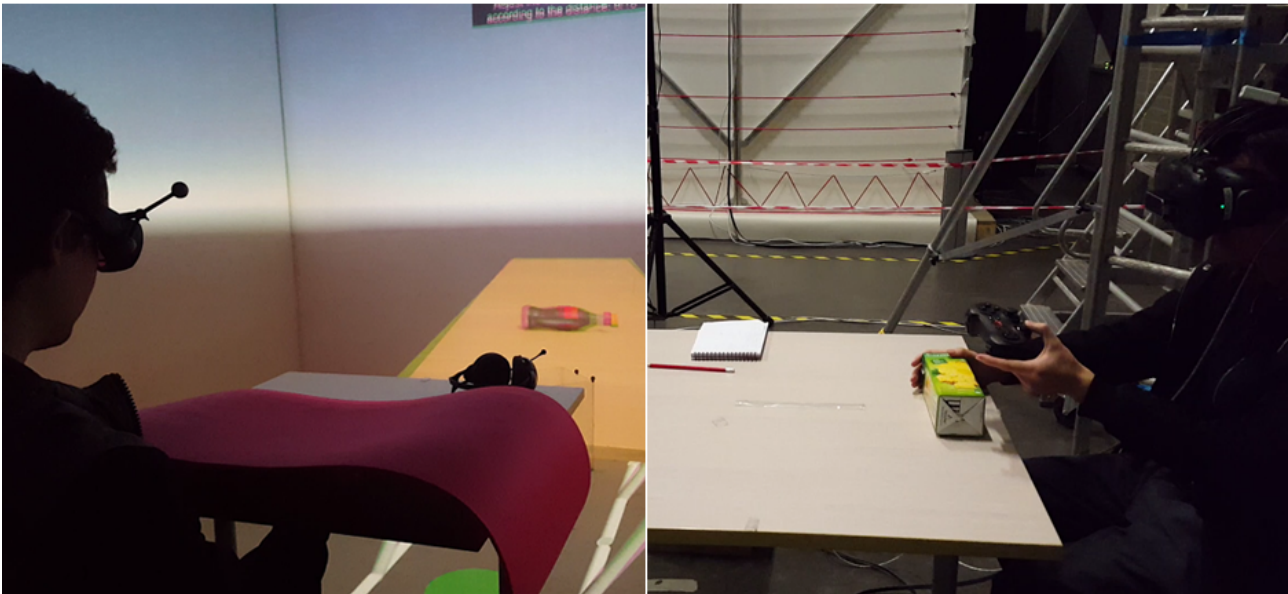


Figure 7.12: Adapted size-constancy test

Left: Cave condition. Right: HMD condition.

Hypothèses

- H1: Les estimations de Size-Ratio montreront une plus grande constance de la taille dans CAVE par rapport au HMD.
- H2: Les estimations de Size-Ratio seront influencées par la taille du conflit V-A induit dans les deux écrans.

Méthode

Nous avons conçu une expérience dans laquelle les participants devaient réaliser le test dans les deux environnements avec un ordre de contreponds. Pour éviter des biais avec le premier environnement, nous avons alterné l'objet physique entre une bouteille de soda et une boîte à jus de fruit. Ainsi, les participantes ont estimé la taille de chaque réplique située à 0,5m, 1,5m et 2,5m de la référence physique. Ces distances représentent les conditions de parallaxe nulle, positive et négative, de sorte que les participants étaient situés à 1,5m environ de l'écran frontal dans notre CAVE. Les participants ont ajusté l'échelle perçue des répliques six fois à chaque distance cible dans un ordre aléatoire entre les essais. Pour chaque essai, la réplique virtuelle a été présentée avec une dimension exagérée équivalente

7.4. UN CADRE DE RÉFÉRENCE POUR CARACTÉRISER ET INFLUENCER LA PERCEPTION SPATIALE SUR DES ÉCRANS HÉTÉROGÈNES DE RÉALITÉ VIRTUELLE

à 25% ou 400% de sa taille réelle.

Résultats

Les résultats du test de constance de la taille sont présentés à la figure 7.13. En utilisant un test-t pour échantillons appariés, nous avons trouvé une différence significative entre la perception de la taille avec le CAVE et avec le HMD à 0,5m ($M = 1.28 \pm 0.10$, $M = 1.15 \pm 0.10$, $t(7) = 5.680$, $p = 0.01$), à 1,5m ($M = 1.26 \pm 0.06$, $M = 1.08 \pm 0.11$, $t(7) = 3.046$, $p = 0.023$) et à 2,5m ($M = 1.18 \pm 0.16$, $M = 0.95 \pm 0.16$, $t(7) = 3.42$, $p = 0.014$). Contrairement à nos attentes, la perception de la taille a été plus précise sous la condition HMD (barres d'erreur jaunes) que sous la condition CAVE (barres d'erreur bleues). Cependant, les résultats suggèrent que H1 se maintient, que la tendance à la constance de la taille est plus grande pour le CAVE que pour le HMD, ce qui implique que ces écrans sont plus sensibles au phénomène de compression de la profondeur. En ce qui concerne H2, nos résultats suggèrent que cette hypothèse se maintient partiellement : en parallaxe négative (0,5m), les estimations de SizeRatio ont été plus grandes par rapport à celles en parallaxe nulle (1,5m). Pourtant, les estimations en parallaxe nulle (1,5m) ont été plus grandes par rapport à celles en parallaxe positive (2,5m), un résultat qui ne coïncide pas avec ceux de Kenyon et al. (2007).

Discussion

Les estimations reflètent partiellement l'influence du conflit V-A et le facteur de la distance focale. De même que chez Kenyon et al., (2007), nos résultats montrent une surestimation de la taille et une sous-estimation des distances en parallaxe négative (à 0,5m), et la meilleure performance en parallaxe nulle (à 1,5m) lorsque la réplique virtuelle est placée sur le plan de l'écran et que le conflit V-A se rapproche de zéro. Malheureusement, à plus grande distance, nos résultats ont été contraires à ceux attendus. Ils ont montré une diminution de la surestimation de la taille, ce qui ne coïncide pas avec Kenyon et al. (2007) ni avec l'évidence trouvée dans la littérature. Cependant, nous pensons que cela pourrait être une conséquence de l'utilisation de la proprioception au lieu de la vue, puisque la proprioception, en tant que signal pour l'estimation des distances, est moins précise pour des distances plus grandes que l'espace personnel. C'est pourquoi, au vu de ces résultats, nous sommes optimistes quant à la pertinence du test.

7.4. UN CADRE DE RÉFÉRENCE POUR CARACTÉRISER ET INFLUENCER LA PERCEPTION SPATIALE SUR DES ÉCRANS HÉTÉROGÈNES DE RÉALITÉ VIRTUELLE

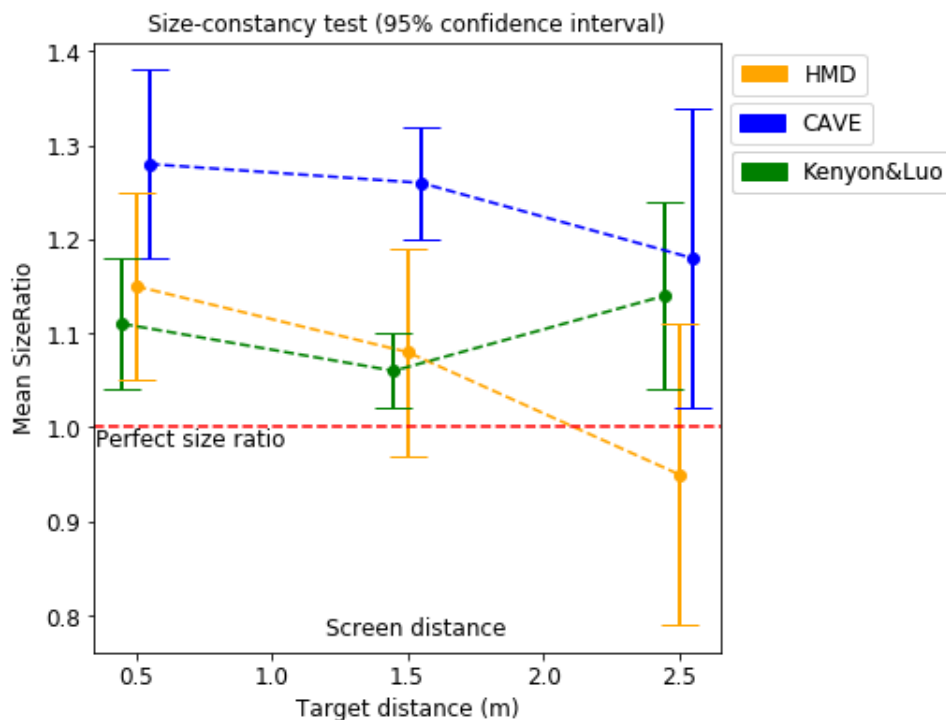


Figure 7.13: Résultats du test de constance de la taille adapté (CAVE vs HMD)

Les participants ont tendance à surestimer la taille de l'objet dans les deux conditions, mais la tendance à la constance de la taille est plus grande sous CAVE que sous HMD.

7.4.2.3 Test de marche en aveugle

Dans ce test, un stimulus est placé au sol et à une certaine distance du sujet. Ensuite, il est demandé au sujet de visualiser le stimulus et d'en estimer la distance. Une fois que le sujet est sûr de son estimation, il doit marcher sans vision vers l'emplacement du stimulus. Dans des conditions naturelles, des études ont montré l'absence d'erreur systématique en marchant vers des cibles situées jusqu'à 20m de distance (pour une revue, voir Loomis et al. (2003)). Comme il a été dit précédemment, lorsque la perception de la distance est évaluée au moyen de tâches motrices, des signaux tels que le flux optique et la parallaxe de mouvement induisent une calibration motrice par laquelle la perception des distances est ajustée à tout moment (c'est l'action qui influe sur la perception). Vu les limitations d'espace dans notre laboratoire, nous avons décidé d'utiliser un tapis roulant moderne, où les participants peuvent parcourir de plus grandes distances in situ. Nous avons fait l'hypothèse que des écrans avec un champ de vision plus grand et une meilleure optique fonctionneraient mieux que des écrans plus limités,

7.4. UN CADRE DE RÉFÉRENCE POUR CARACTÉRISER ET INFLUENCER LA PERCEPTION SPATIALE SUR DES ÉCRANS HÉTÉROGÈNES DE RÉALITÉ VIRTUELLE

puisqu'ils induisent une déformation moindre de la vision périphérique. Nous avons donc conçu une expérience pour comparer la performance entre deux HMD ayant des capacités différentes.

Although the blind walking test seemed an excellent candidate method for the proposed framework, assessing spatial perception in VR using this method is problematic. First, the test is only suitable for HMDs and not for LIPDs. Second, the greater the target distance, the greater the differences between displays, but targets up to 20m away are difficult to achieve in conventional VE setups, and even more difficult due to the spatial restrictions in our lab. Thus, to overcome these limitations, we explored the idea of performing blind walking tests using a modern treadmill, where subjects can walk larger distances in situ. Very few studies have studied the perception of distances in VR using treadmills (Li et al. (2021), Bossard et al. (2020), Santillán and Barraza (2019), Witmer and Sadowski Jr (1998)). By using a variety of displays and different experimental protocols, their results showed underestimation of distances similar to the ones found in VR but with great variability between subjects. We hypothesized that subjects using VR displays with greater FOV and better optics would perform better than more limited displays. Hence, we designed an experiment to compare the performance between two vendors' HMDs with different capabilities.

7.4.2.4 Comparando el HTC Vive vs Gear VR

Nous avons mené une étude pour comparer la perception des distances entre un HTC Vive CV1 (110° champ de vision horizontal, lentilles Fresnel de 52 mm) et un Samsung Gear VR (96° champ de vision horizontal, lentilles biconvexes de 42 mm). Les participants devaient estimer la distance de cinq objets placés à 4m, 8m et 12m en utilisant un tapis roulant Virtueux Omni, un tapis de course moderne conçu pour le consommateur, qui permet de marcher dans toutes les directions. C'est un appareil qui imite la locomotion en faisant glisser les pieds sur une surface légèrement concave. Bien que cette interaction soit différente de la marche naturelle, des tests préliminaires avaient montré que les participants s'adaptent rapidement au mécanisme de locomotion. Pour faciliter le processus, nous avons prévu une phase d'adaptation durant laquelle les participants ont marché librement pendant cinq minutes avant de commencer le test. La figure 7.14 présente les deux conditions étudiées. Nous avons posé comme hypothèse que la perception des distances serait meilleure avec le HTC Vive par rapport au GearVR, puisque son écran fournit un champ de vision plus grand et une meilleure optique.

Hypothèse

7.4. UN CADRE DE RÉFÉRENCE POUR CARACTÉRISER ET INFLUENCER LA PERCEPTION SPATIALE SUR DES ÉCRANS HÉTÉROGÈNES DE RÉALITÉ VIRTUELLE



Figure 7.14: Tests de marche en aveugle avec le tapis roulant Virtuix Omni

Les participants ont fait le test de marche en aveugle en utilisant un Virtuix Omni Treadmill. À gauche: condition HTC Vive. À droite: condition Gear VR

- H1: La perception des distances sera meilleure sous la condition HTC Vive que sous la condition GearVR.

Méthode

Nous avons conçu une expérience dans laquelle les participants devaient réaliser le test dans les deux écrans avec un ordre de contre-poids. Ils devaient estimer la distance relative d'un stimulus virtuel placé à 4m, 8m et 12m. Trois essais au total pour chaque distance cible dans un ordre aléatoire ont été demandés.

Résultats

Les résultats du test de marche en aveugle sont présentés à la figure 7.15. Un test-t pour échantillons appariés a permis de comparer la perception des distances entre les trois distances cible. En général, tous les participants ont sous-estimé les distances, indépendamment de la condition de visualisation. Une différence significative a été observée dans la sous-estimation des distances à 8m entre la condition de Vive ($M = 0.60m, SD = 0.29$) et la condition GearVR ($M = 1.05, SD = 0.54$), $t(47) = -2.06, p <$

7.4. UN CADRE DE RÉFÉRENCE POUR CARACTÉRISER ET INFLUENCER LA PERCEPTION SPATIALE SUR DES ÉCRANS HÉTÉROGÈNES DE RÉALITÉ VIRTUELLE

0.01); en revanche, aucune différence significative n'a été observée à 4m ni à 12m. Par conséquent, les résultats suggèrent que H1 se maintient partiellement, la performance des participants ayant été significativement meilleure sous la condition Vive que sous la condition GearVR, mais uniquement à 8m.

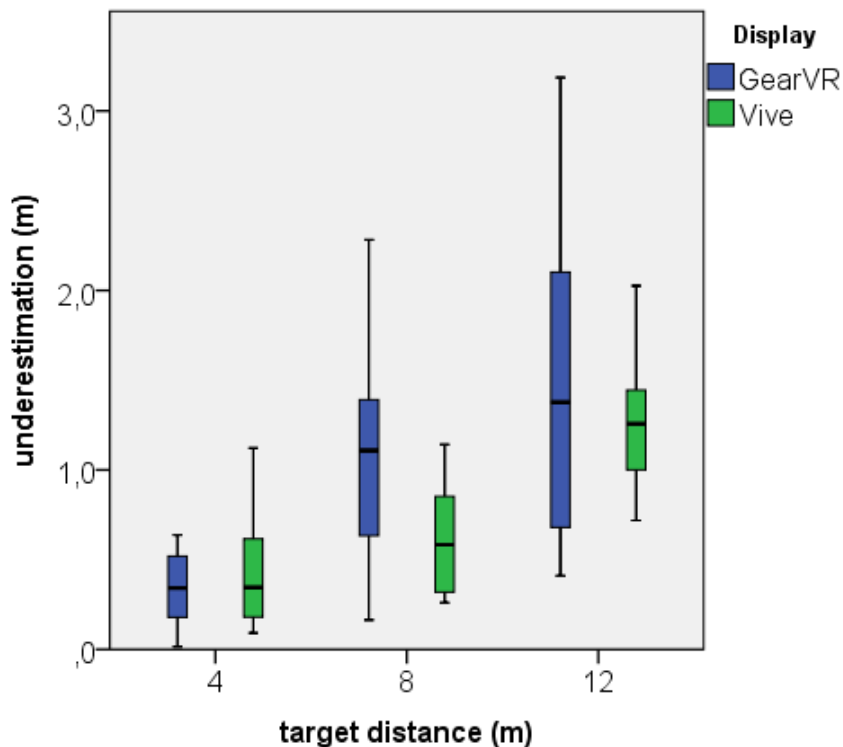


Figure 7.15: Résultats des test de marche en aveugle avec le tapis roulant Virtuix Omni.

Une différence significative a été observée entre la perception de la distance sous la condition GearVR et sous la condition HTC Vive à 8 m.

Discussion

Los résultats sont cohérents par rapport à la littérature, selon laquelle la sous-estimation des distances a tendance à augmenter à mesure de l'augmentation de la distance du stimulus. Bien que les résultats n'aient pas toujours été ceux attendus, les différences significatives dans la performance à 8m entre le HTC Vive et le Gear VR présentent un intérêt tout particulier. Cependant, aucune différence significative entre les écrans n'a été observée dans la performance des participants à 4m et 12m. Même si nous avons permis aux participants de s'habituer au tapis roulant pendant quelques minutes, nous avons remarqué qu'ils ne se sont pas tous adaptés facilement à la plate-forme, certains ayant même

7.5. PARTIE I - CARACTÉRISATION DE LA COMPRESSION DE LA PROFONDEUR DANS LES ÉCRANS RV

déclaré être fatigués, ce qui peut expliquer la grande variabilité dans les réponses. Compte tenu de ces limitations, nous ne pouvons pas considérer ce test comme une bonne méthode d'évaluation pour le cadre de référence proposé. Malgré les quelques différences significatives observées, le tapis roulant est un facteur qui pourrait introduire beaucoup de biais. Pour cette raison, nous considérons qu'il est important de permettre au sujet de marcher naturellement.

7.4.3 Résumé

Le test de constance de la taille est idéal pour étudier l'influence du conflit V-A. La sous-estimation des distances et la surestimation de la taille correspondante augmentent en fonction de la taille du conflit V-A. Les résultats ont montré que les LIPD tels que le CAVE sont moins sensibles au phénomène de compression de la profondeur. La tendance à la constance de la taille a été plus grande dans le CAVE, ce qui montre que les participants ont perçu la taille des objets d'une manière plus ou moins uniforme entre les distances cible. Le test de la marche en aveugle est idéal pour analyser l'influence du problème SPI. Plus la distorsion induite dans la vision périphérique est grande, plus la sous-estimation des distances est importante. Malgré les problèmes liés au tapis roulant, nous pensons que des différences significatives pourraient apparaître si nous permettons aux participants de marcher naturellement.

7.5 Partie I - Caractérisation de la compression de la profondeur dans les écrans RV

7.5.1 Vue d'ensemble

Dans le chapitre précédent nous avons présenté deux tests de perception spatiale que nous avons utilisés comme méthodes d'évaluation, le test de constance de la taille et le test de la marche en aveugle. Le cadre de référence proposé se fonde sur l'idée que, si nous sommes capables de prédire la sous-estimation des distances en réalité virtuelle, nous pourrions avoir une influence positive sur la perception spatiale. Pour prédire la performance des participants, il nous faut analyser comment les différents facteurs d'affichage associés aux problèmes de visualisation influent sur la perception des distances. Le problème du conflit entre vergence et accommodation (conflit V-A) et le problème de la stimulation périphérique incorrecte (problème SPI) sont les causes principales de la perception incor-

7.5. PARTIE I - CARACTÉRISATION DE LA COMPRESSION DE LA PROFONDEUR DANS LES ÉCRANS RV

recte des distances, mais le degré de sous-estimation varie en fonction de quelques facteurs associés. En termes du conflit V-A, l'interaction entre la distance focale et la condition de parallaxe stéréoscopique déterminent le degré de sous-estimation. Quant au problème EPI/SPI, c'est l'interaction entre la taille du champ de vision et les distorsions des lentilles qui va déterminer le degré de sous-estimation. La figure 7.16 présente les principaux facteurs déterminants qui influent sur la perception spatiale en réalité virtuelle. À partir de ces quatre facteurs, nous avons développé un modèle initial pour caractériser la compression de la profondeur induite dans un écran, basés sur quelques suppositions des études préalables et sur l'évidence trouvée dans la littérature.

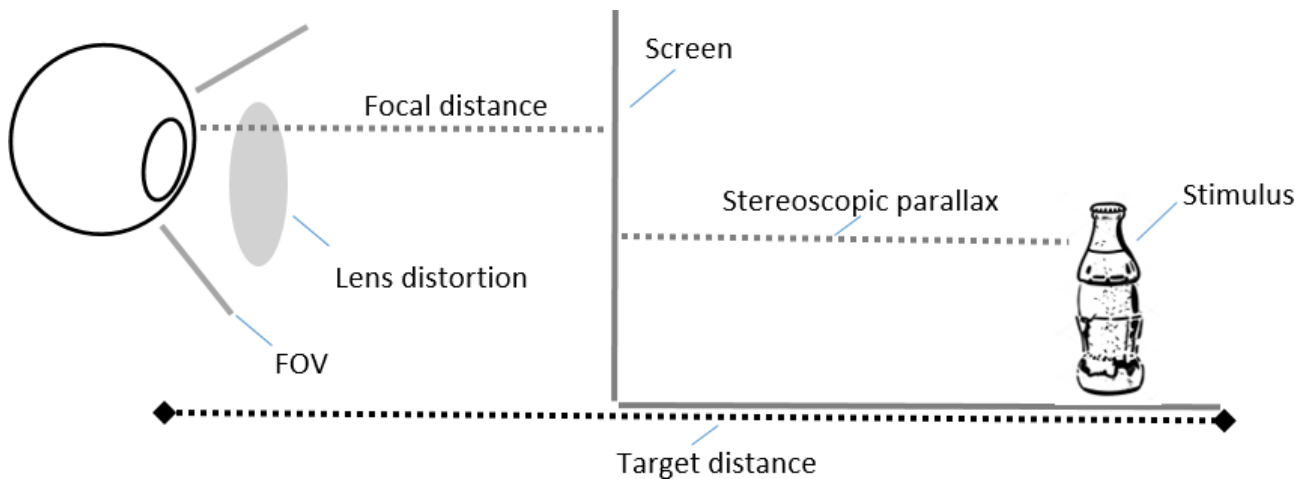


Figure 7.16: Les principaux facteurs déterminants associés aux problèmes de visualisation

L'interaction entre la taille du champ de vision, la distance focale, la condition de parallaxe stéréoscopique et les distorsions des lentilles déterminent le degré de sous-estimation dans la perception des distances.

7.5.2 Facteurs déterminants

7.5.2.1 Distance focale

Dans le chapitre 2 nous avons montré que le phénomène de la compression de la profondeur est associé directement à la distance focale de l'écran, ce qui limite la plage de profondeur fonctionnelle où la vergence et l'accommodation peuvent être utilisées pour percevoir la profondeur. Cette plage dépend de la profondeur de champ disponible (PDC) et de la zone de fusion confortable induite par l'écran (zone de confort de Shibata et al. (2011)). Si la distance focale détermine la compression de la profondeur, une question importante qui a été soulevée par notre étude préalable à partir du test de la

7.5. PARTIE I - CARACTÉRISATION DE LA COMPRESSION DE LA PROFONDEUR DANS LES ÉCRANS RV

constance de la taille est de savoir pourquoi les performances ont-elles été si différentes entre le CAVE et le HMD. La réponse en est dans la façon dont les HMD sont conçus. Comme il a été mentionné précédemment, les HMD ont des lentilles avec une distance focale fixe qui crée une image virtuelle à l'infini, ce qui influe sur la façon dont la DOF/PDC augmente naturellement, rendant ces écrans plus sensibles au phénomène de la compression de la profondeur. Au contraire, les LIPD n'ont pas besoin de lentilles, ce qui permet une relation plus naturelle entre la distance focale et la DOF/PDC induite. Les HMD peuvent modifier considérablement ces plages, car leurs écrans sont placés très près des yeux du perceur.

La figure 7.17 présente une comparaison de la PDC induite dans un HMD moderne par rapport à la PDC naturelle humaine. L'écran HTC Vive Pro a une distance focale d'environ 0,75m (Reddit.com (2017)). Comme la PDC augmente avec la distance focale, la mise au point sur un objet à 0,75m dans le HMD est l'équivalent de la mise au point à des distances plus grandes dans des conditions naturelles. Nos yeux sont, en effet, mis au point à 0,75m, mais le flou rétinien induit (le signal associé à l'accommodation) correspond à celui de la mise au point sur un objet à 2m environ (en utilisant comme référence l'estimation de Shibata et al. (2011)). Ainsi, la mise au point sur un objet à des distances courtes dans un HMD crée une PDC similaire à celle qui se crée lors de la mise au point à de grandes distances, en augmentant la taille du conflit V-A. Nous allons donc introduire la première hypothèse du cadre de référence proposé: **le degré de compression spatiale est en rapport direct avec la distance focale et la PDC induite**. Pour cette raison, des écrans ayant la même distance focale peuvent ne pas avoir la même plage de profondeur fonctionnelle, et le degré de compression sera fonction de la PDC induite.

7.5.2.2 Condition de parallaxe stéréoscopique

Le deuxième facteur nécessaire pour prédire la performance spatiale est la parallaxe stéréoscopique, l'emplacement du stimulus par rapport à l'écran, qui détermine également le degré du conflit V-A. Nos études préalables en utilisant les tests de constance de la taille et l'évidence de la littérature ont montré que plus le conflit V-A est grand, plus l'erreur dans la perception des distances est grande. Cependant, ces études préalables ont soulevé une question importante: pourquoi la tendance à la constance de la taille est-elle plus grande dans le CAVE que dans le HMD, si l'emplacement du stimulus est le même? La réponse en est, encore une fois, dans la façon dont les HMD sont conçus et dans la nature de la

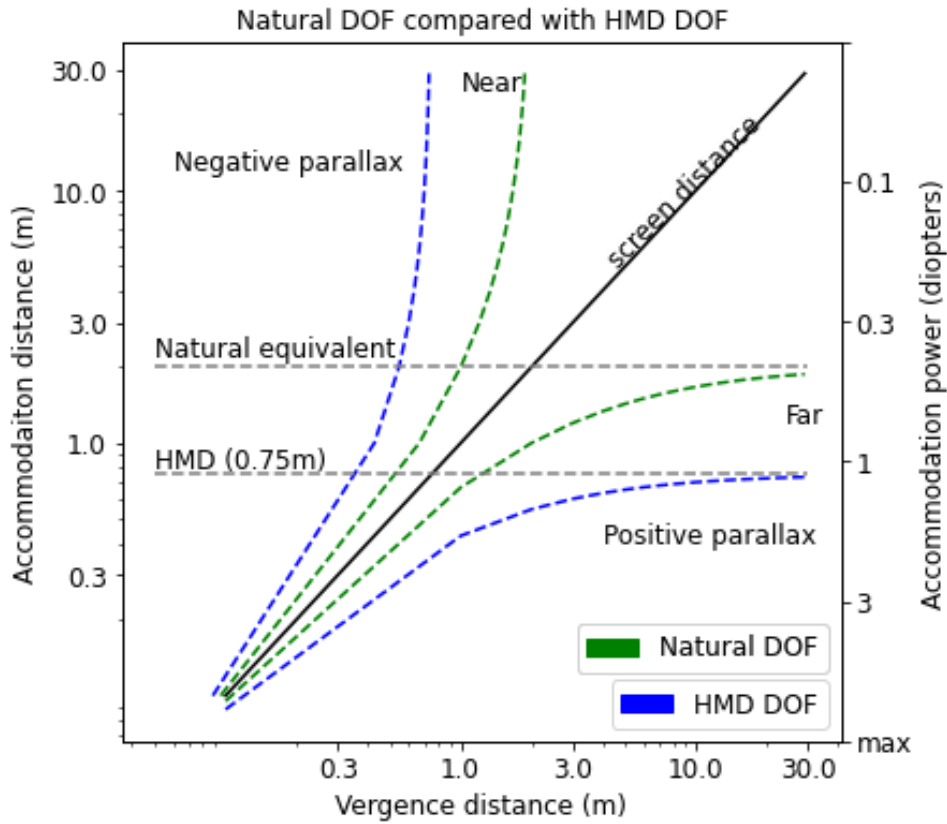


Figure 7.17: PDC d'un HMD moderne par rapport à la PDC naturelle humaine.

L'utilisation de lentilles dans un HMD fait que l'augmentation de la PDC soit plus importante que dans des conditions naturelles, de sorte qu'une mise au point à 0,75m est à peu près équivalente à une mise au point à 2m dans des conditions naturelles.

DOF/PDC induite par ces écrans. Du fait de l'incohérence entre la distance focale et la DOF/PDC induite, expliquée précédemment, les HMD induisent un conflit V-A plus important que les LIPD.

Selon l'hypothèse de Vienne et al. (2018), la sous-estimation des distances doit être proportionnelle à la taille du conflit V-A, qui peut être calculée en soustrayant l'inverse de l'accommodation de l'inverse de la vergence: $C_{va} = \frac{1}{V_d} - \frac{1}{A_d}$. Il y a deux éléments importants à propos de cette équation: (1) elle pourrait donner lieu à des conflits positifs ($V_d < A_d$) et à un conflit négatif ($V_d > A_d$), puisqu'ils sont fonction de la distance focale, plus que du percepteur; et (2) il n'y a que deux façons de réduire le conflit, en parallaxe nulle (où la vergence et l'accommodation sont égales) et en augmentant la distance de convergence (ce qui, à son tour, diminue les demandes en accommodation). Cependant, l'équation de Vienne et al. (2018) n'a pas tenu compte de l'influence de la PDC induite dans la perception des

7.5. PARTIE I - CARACTÉRISATION DE LA COMPRESSION DE LA PROFONDEUR DANS LES ÉCRANS RV

distances. Si nous introduisons cet élément, de façon à ce que la distance focale des HMD s'ajuste à l'équivalent de la PDC naturelle correspondante, deux schémas différents émergent. La figure 7.18 présente une comparaison de la taille du conflit V-A entre un CAVE mis au point à 2m et un HMD mis au point à 0.75m. Selon le graphique, tandis que dans le CAVE il y a des conflits positifs et négatifs, dans le HMD les conflits sont toujours négatifs (comme si tous les objets étaient situés en parallaxe positive). En outre, le conflit est plus important dans le HMD que dans le CAVE, conséquence directe du fait qu'il y a une plus grande différence entre la vergence et l'accommodation. Nous allons donc introduire la deuxième hypothèse du cadre de référence proposé: **le degré de sous-estimation dans la perception des distances doit être proportionnel à la taille du conflit V-A.**

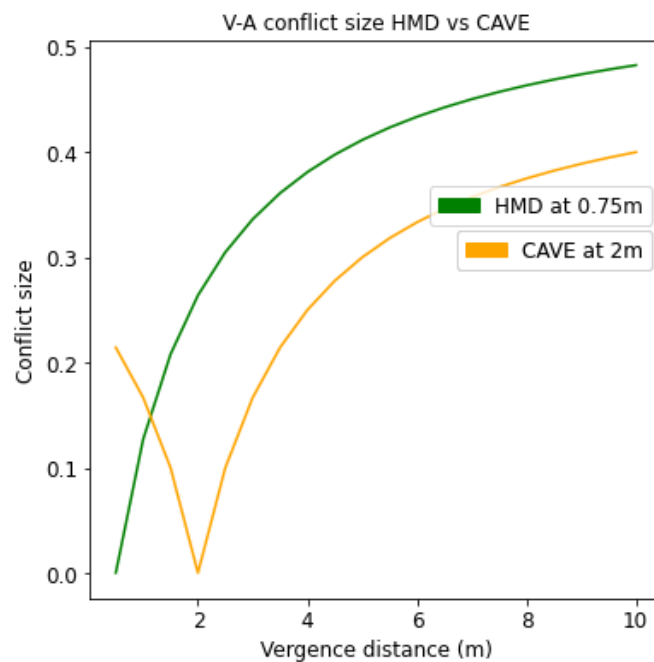


Figure 7.18: Une comparaison de la taille du conflit V-A entre un HMD et le CAVE

Tandis que dans le CAVE il y a des conflits positifs et négatifs, dans le HMD les conflits sont toujours négatifs, ce qui est une conséquence directe de la disparité entre la distance focale et la DOF/PDC induite dans ce type d'écran.

7.5.2.3 Stimulation périphérique incorrecte (SPI)

Un troisième facteur qui affecte la perception des distances est le problème de la SPI. La parallaxe de mouvement est un signal qui influe sur la perception des distances et sur le processus de recalibration

motrice. La figure 7.19 présente une comparaison de l'influence de la SPI sur la compression de la profondeur induite en réponse aux distorsions des lentilles. un HTC Vive (lentilles fresnel) et l'Oculus Rift (lentilles hybrides). L'effet est plus fort dans le HTC Vive, son facteur d'échantillonnage étant de 4 minutes d'arc/sous-pixel environ dans la vision périphérique, par rapport aux 2 minutes d'arc/sous-pixel environ pour l'Oculus Rift. Par conséquent, il pourrait y avoir des différences significatives dans la perception des distances entre les HMD, même si les deux écrans offrent le même champ de vision.

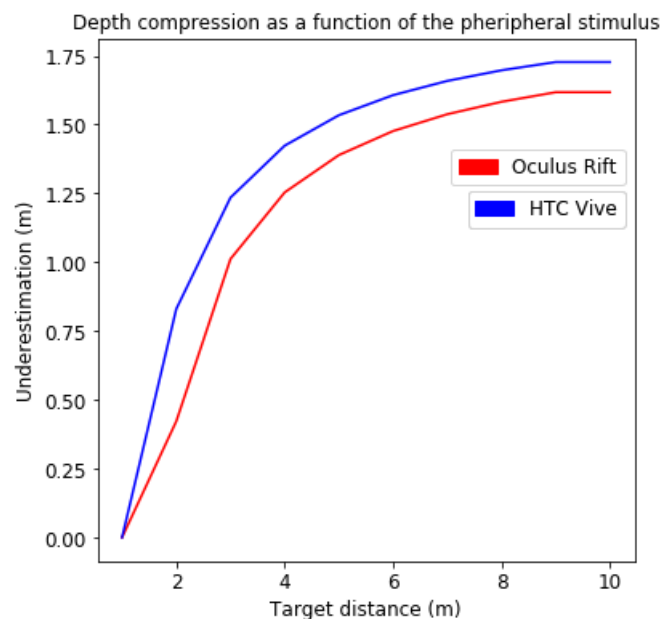


Figure 7.19: Compression de la profondeur en fonction de la résolution angulaire induite par la lentille
La meilleure résolution angulaire de l'Oculus Rift induit une sous-estimation des distances plus petite par rapport au HTC Vive.

7.5.3 Prediction de la performance spatiale

7.5.3.1 Choix des stimuli

Pour prédire la performance spatiale du sujet, nous devons caractériser la compression de la profondeur induite par un écran. Nous définissons un ensemble d'hypothèses qui décrivent l'influence des différents facteurs d'affichage dans la perception spatiale. Pour valider ces hypothèses, nous pouvons réaliser une méthode d'évaluation, pour laquelle les stimuli font l'objet d'un choix stratégique afin qu'ils représentent l'influence des différents facteurs d'affichage. Par exemple, pour analyser l'influence

7.5. PARTIE I - CARACTÉRISATION DE LA COMPRESSION DE LA PROFONDEUR DANS LES ÉCRANS RV

de la distance focale et de la condition de parallaxe stéréoscopique, toutes deux associées au conflit V-A, la configuration idéale du stimulus est un ensemble de distances cible en corrélation linéaire avec la distance focale, les valeurs positives faisant référence au stimulus placé en parallaxe positive et les valeurs négatives faisant référence au stimulus placé en parallaxe négative (figure 7.20). Indépendamment de la tâche spatiale, nous posons comme hypothèse qu'un stimulus formé de cinq distances cible est suffisant pour caractériser le phénomène de compression de la profondeur dans l'écran (figure 7.21 - gauche). Au contraire, pour le problème de l'SPI, cinq stimuli non espacés linéairement sont suffisants. Dans ce cas, les distances cible sont doublement espacées pour représenter les points de basculement dans la courbe de compression induite (figure 7.21 - droite).

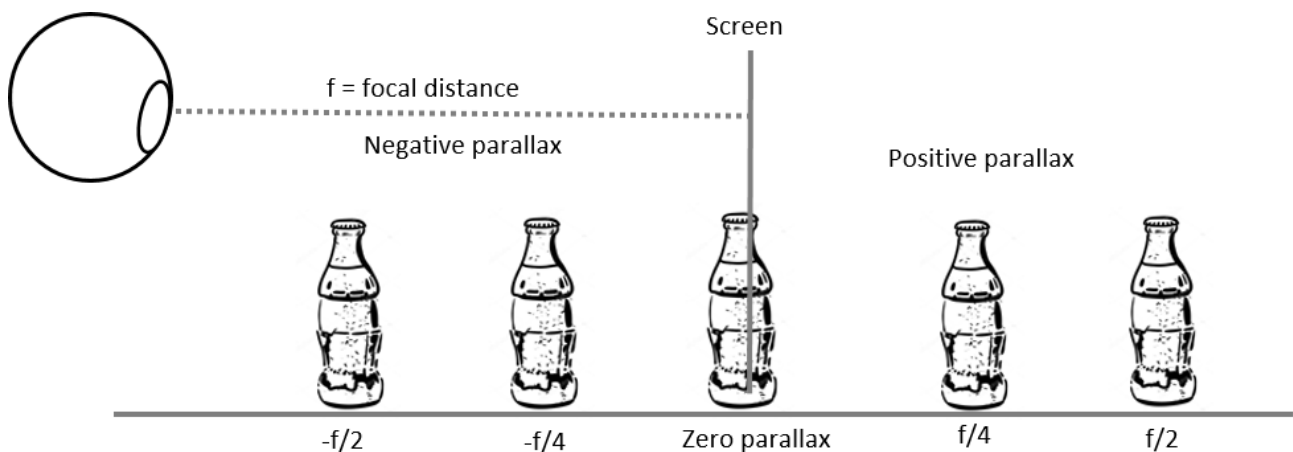


Figure 7.20: Le choix du stimulus en fonction de la distance focale f .

7.5.3.2 Prédicteur de performance linéaire

Le deuxième pas pour prédire la performance spatiale consiste à calculer le prédicteur de performance linéaire (PPL) à l'aide d'une des méthodes d'évaluation. Le PPL peut être calculé en utilisant un modèle de régression linéaire (par exemple, les moindres carrés ordinaires) qui établit la corrélation entre le degré de sous-estimation et les distances cible. La figure 7.22 présente un diagramme de dispersion de la performance d'un sujet en utilisant le HMD et le CAVE lors du test de constance de la taille. Le graphique est construit non seulement en termes des distances cible, mais aussi en termes de la condition de parallaxe stéréoscopique. Une hypothèse du cadre de référence proposé est que le degré de sous-estimation est proportionnel à la taille du conflit V-A. C'est pourquoi nous posons comme hypothèse que la pente de sous-estimation doit être proportionnelle à la taille du conflit V-A.

7.5. PARTIE I - CARACTÉRISATION DE LA COMPRESSION DE LA PROFONDEUR DANS LES ÉCRANS RV

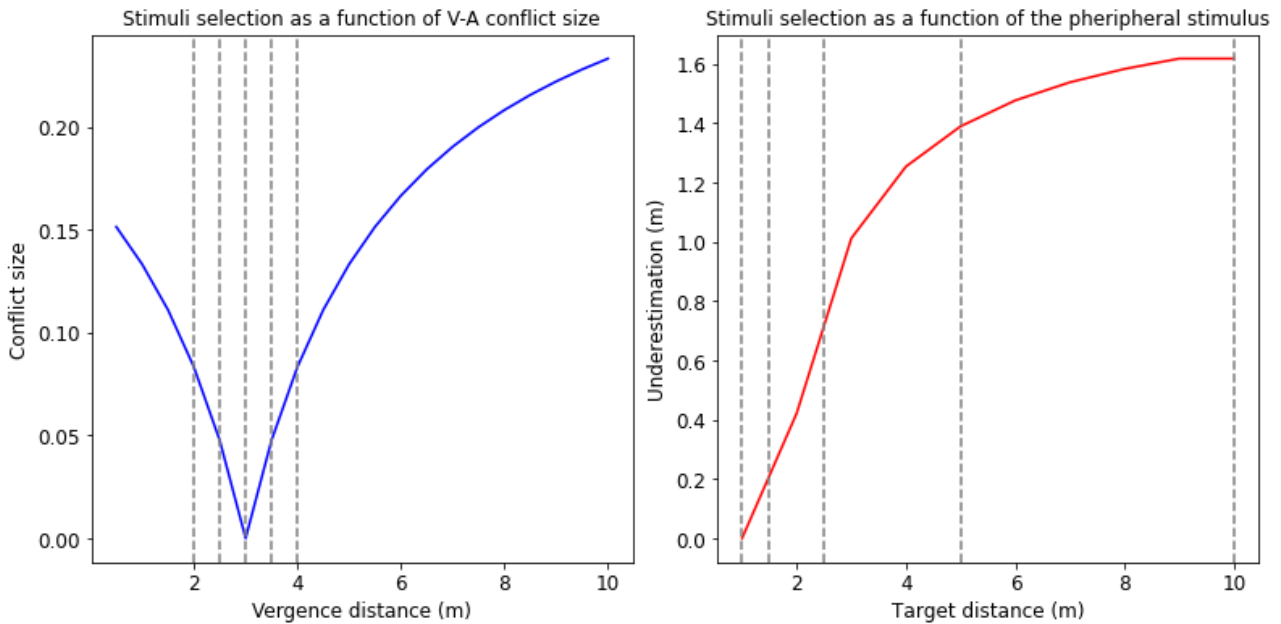


Figure 7.21: Le choix des stimuli selon le problème d’affichage.

Un stimulus formé de cinq distances cible est suffisant pour caractériser le phénomène de compression de la profondeur dans l’écran.

Le graphique montre comment le CAVE induit un conflit V-A plus petit par rapport au HMD, ce qui à son tour entraîne une légère sous-estimation des distances. Par contre, le HMD induit un conflit V-A plus grand, entraînant une sous-estimation plus grande des distances, ce qui reflète la plus grande sensibilité des HMD au phénomène de la compression de la profondeur.

Le calcul du LPP/PPL d’un écran est une opération simple. Nous pouvons estimer la performance des participants au moyen d’une fonction de régression linéaire qui estime la sous-estimation des distances induite par l’écran, qui pourrait se définir ainsi:

$$d' = (1 - c)d \quad (7.3)$$

où d' est la distance sous-estimée attendue, d est la distance du stimulus et c est la pente globale de compression induite par l’écran. La figure 7.23 compare la performance du sujet, entre le HMD et la performance parfaite. Dans la condition idéale, les distances ne sont pas sous-estimées, la pente étant donc égale à un. Au contraire, la pente de régression linéaire basée sur la performance du sujet pour l’écran est $1.07 - 0.29d$, ce qui signifie que la compression induite dans ce cas est $c = 0.29d$. Ainsi, le

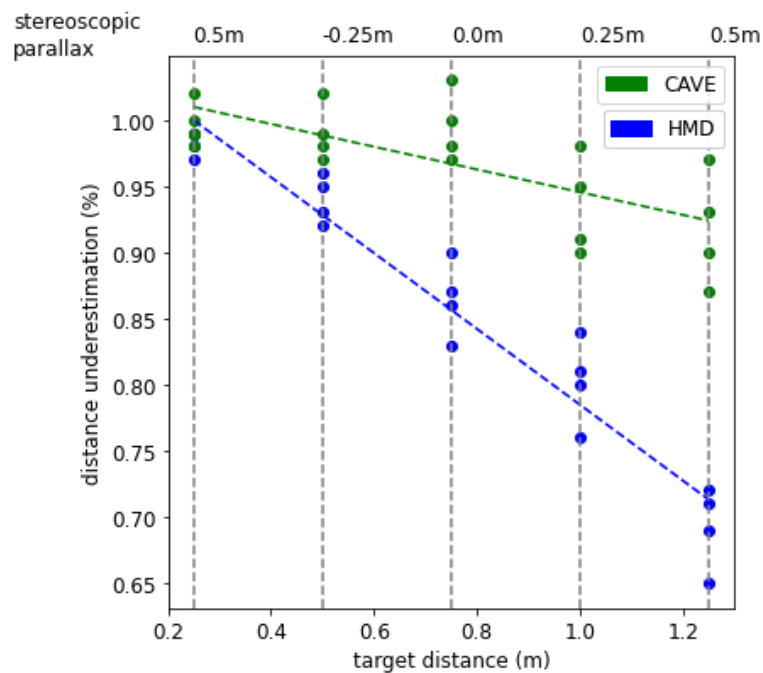


Figure 7.22: Performance du sujet en fonction de la taille du conflit V-A.

Le CAVE induit un conflit V-A plus petit par rapport au HMD, ce qui à son tour entraîne une sous-estimation plus petite tout au long des distances.

PPL pour le HMD est $d' = (1 - 0.29)d$, d pour chaque mètre, et les distances sont sous-estimées de 29%.

7.5.4 Résumé

Dans ce chapitre nous avons proposé une approche formelle pour caractériser le phénomène de compression de la profondeur induite par un écran. Nous avons défini un ensemble d'hypothèses qui décrivent l'influence des différents facteurs de l'écran pour ensuite montrer comment doit être fait le choix stratégique des stimuli afin d'évaluer l'influence de ces facteurs. Une fois que la perception spatiale du sujet est mesurée, un prédicteur de performance linéaire (LPP/PPL) peut être calculé pour prédire la compression de la profondeur induite. Le chapitre suivant est consacré à la présentation de la deuxième partie du cadre de référence proposé, nous discuterons de la façon dont la composition peut être utilisée pour améliorer la perception spatiale et de la façon dont la pente du LPP/PPL peut être utilisée pour améliorer la perception des distances en réalité virtuelle.

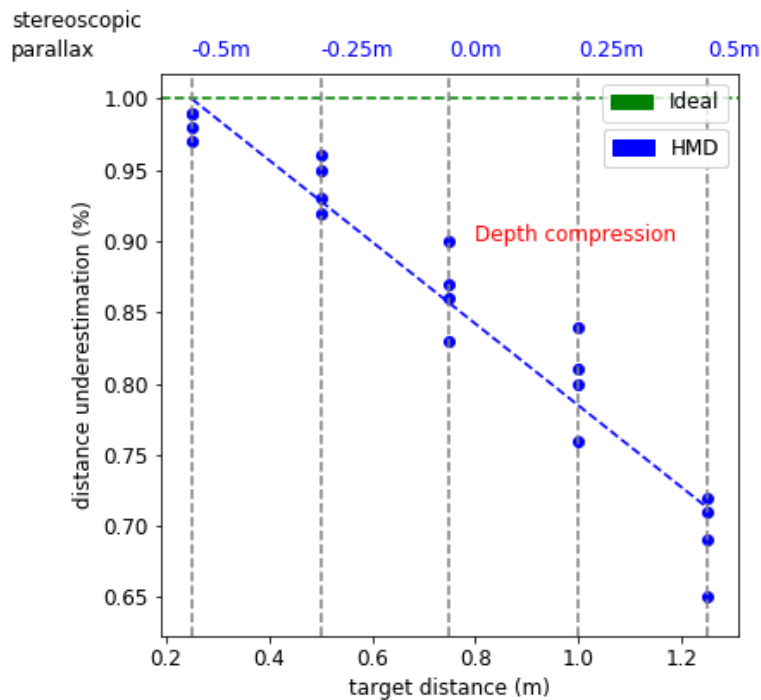


Figure 7.23: Prédicteur de performance linéaire

Performance du sujet en fonction des stimuli de distance et la condition de parallaxe stéréoscopique. La pente de régression linéaire caractérise le degré de la compression de la profondeur induite par l'écran

7.6 Partie II. Influencer la perception spatiale par composition

7.6.1 Vue d'ensemble

Dans le chapitre précédent nous avons expliqué comment les différents facteurs d'affichage influent sur la perception des distances, et comment le phénomène de compression de la profondeur peut-il être prédit en utilisant une méthode d'évaluation et en calculant un prédicteur de performance linéaire (PPL). Un objectif du cadre de référence proposé est que, si nous sommes capables de prédire la performance spatiale, nous pouvons avoir une influence positive sur la perception des distances à l'aide de la composition. Pour influencer la perception du sujet, nous pouvons faire appel au biais dans la perception humaine vers la symétrie et les schémas de forme réguliers associés à la perspective (comme ceux décrits dans l'illusion de la chambre d'Ames). Nous avons donc retenu quelques techniques qui jouent avec les règles de la perspective et qui ont montré de bons résultats en termes de l'influence sur

7.6. PARTIE II. INFLUENCER LA PERCEPTION SPATIALE PAR COMPOSITION

la perception des distances: l'abaissement de l'horizon et la minification du champ de vision. La figure 7.24 présente deux captures d'écran de chaque technique, où la projection de perspective standard est altérée. Pour l'abaissement de l'horizon, une projection de perspective oblique (PPO) est utilisée, qui applique une déclinaison subtile sur la ligne d'horizon, déplaçant les points de fuite vers le bas et augmentant la perception des distances. La minification applique un champ de vision géométrique différent au champ de vision défini pour l'écran. La technique change le taux auquel les points de fuite convergent vers l'horizon, de sorte que les objets apparaissent plus petits et plus lointains. Dans les sections qui suivent nous discuterons de la façon dont nous avons intégré ces techniques au cadre de référence proposé et expliquerons pourquoi nous les avons mises en oeuvre en utilisant une approche différente.

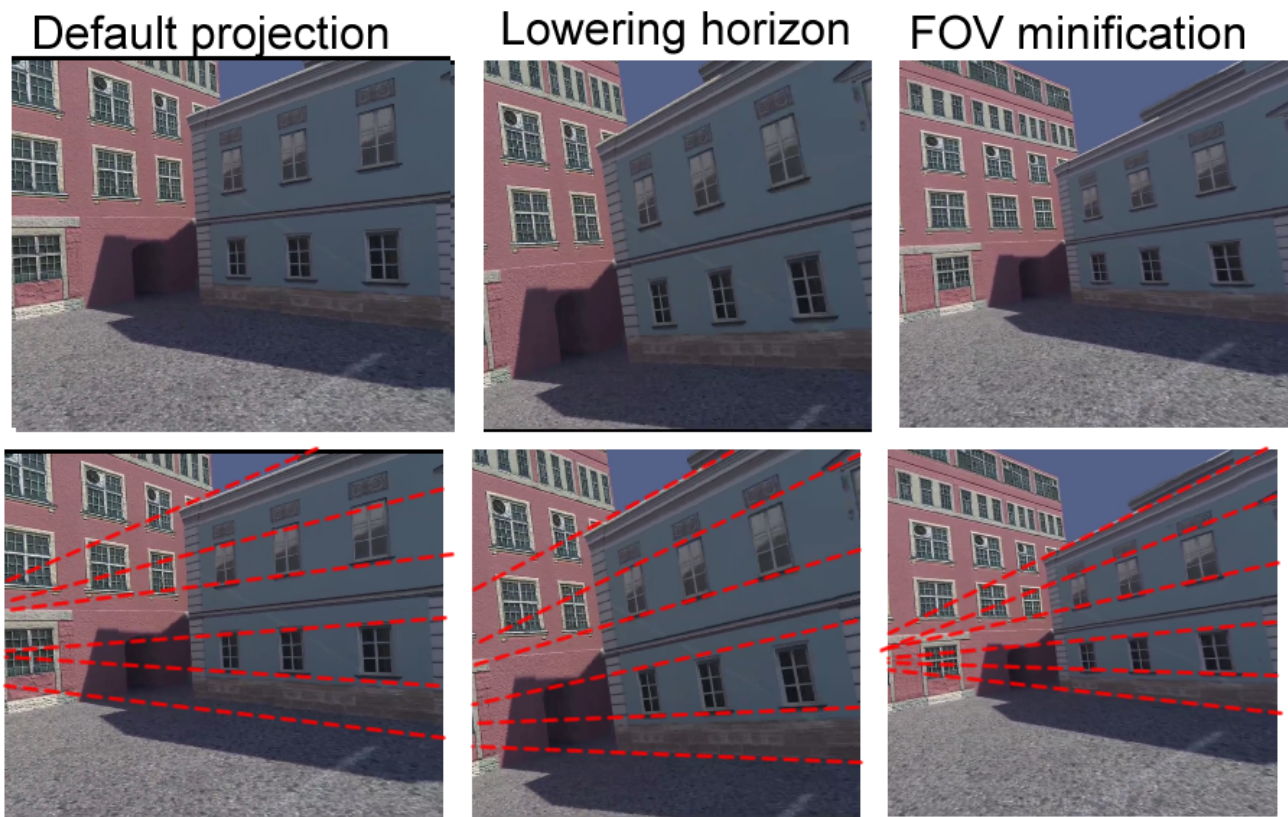


Figure 7.24: Techniques alternatives de projection

A. Projection de perspective par défaut. B. Abaisser l'horizon. C. Minification du champ de vision.

7.6.2 Techniques alternatives de projection

7.6.2.1 Influence de l'abaissement de l'horizon sur la constance de la taille

Nous avons reproduit le dernier test de constance de la taille, mais en nous concentrant cette fois uniquement sur les HMD, et nous avons analysé l'influence de la technique PPO sur la perception spatiale. L'application d'une déclinaison subtile sur la ligne d'horizon et la modification des points de fuite de la table (figure 7.25) fait entrer le système visuel en un conflit de signaux. De même que pour l'illusion de la chambre d'Ames, nous avons fait l'hypothèse que le système visuel ferait peu de cas de l'information contradictoire des signaux de profondeur oculomotrice pour favoriser une "table stable", une table droite et parallèle au sol. La déclinaison subtile a pour effet de faire bouger vers le bas les points de fuite de la table, depuis le centre de l'écran, créant l'illusion d'une table plus grande que ses dimensions réelles et influant sur la perception de la distance et de la taille. Nous avons posé comme hypothèse que l'application de cette technique en simulant une déclinaison angulaire de 5 degrés par rapport à l'horizon entraînerait l'amélioration de la perception des distances et induirait la constance de la taille.

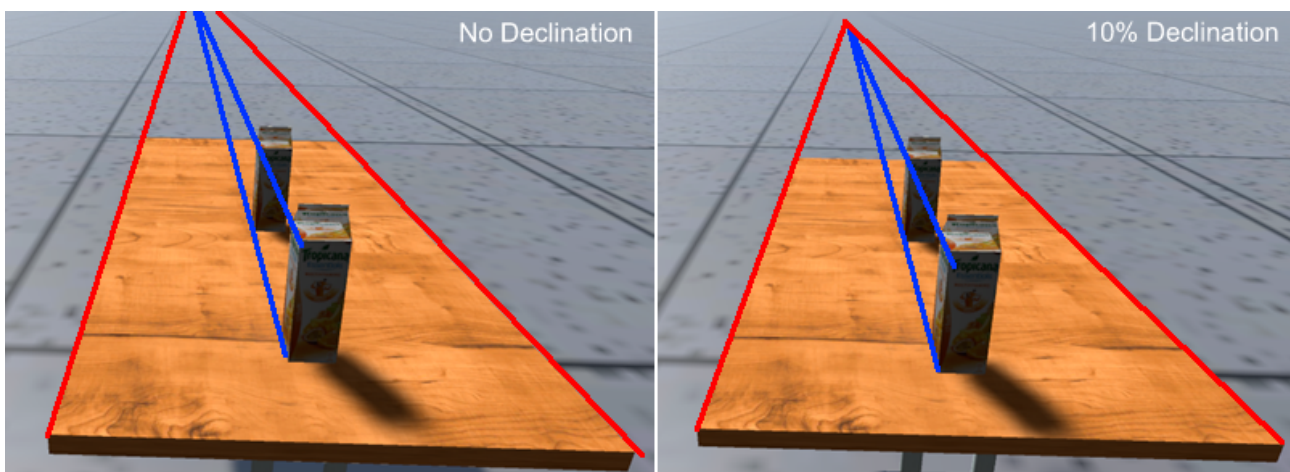


Figure 7.25: Mise en oeuvre de la technique de perspective de vision oblique (PPO)

À gauche: Projection par défaut. À droite: Projection oblique avec une déclinaison de 5%.

Hypothèse

- H1: Une amélioration des taux de SizeRatio par rapport à la projection standard aura lieu.

Méthode

7.6. PARTIE II. INFLUENCER LA PERCEPTION SPATIALE PAR COMPOSITION

L'expérience s'est déroulée avec le même appareil et en suivant la même procédure du dernier test de taille, cette fois-ci avec des distances cible de (0,4m, 0,8m et 1,2m).

Résultats

Les résultats de la technique PPO par rapport à la projection par défaut sont présentés à la figure 7.26. Un test-t pour échantillons appariés a permis de comparer le SizeRatio entre les trois distances cible. Il y a eu une différence significative pour le rapport de taille entre la technique OPP/PPO et la projection par défaut à 1,2m ($M = 0.99, SD = 0.12, M = 1.09, SD = 0.11, t(19) = 3.6, p < 0.02$), et à 0,8m ($M = 0.90, SD = 0.08, M = 0.99, SD = 0.12, t(19) = 2.6, p = 0.01$), tandis qu'aucune différence significative n'a été trouvée à 0,4m. Les résultats ont montré que H1 se maintient partiellement, que la technique proposée réduit la sous-estimation des distances et qu'elle améliore la perception de la taille para rapport à la technique de projection prédéterminée, particulièrement pour des objets en parallaxe positive.

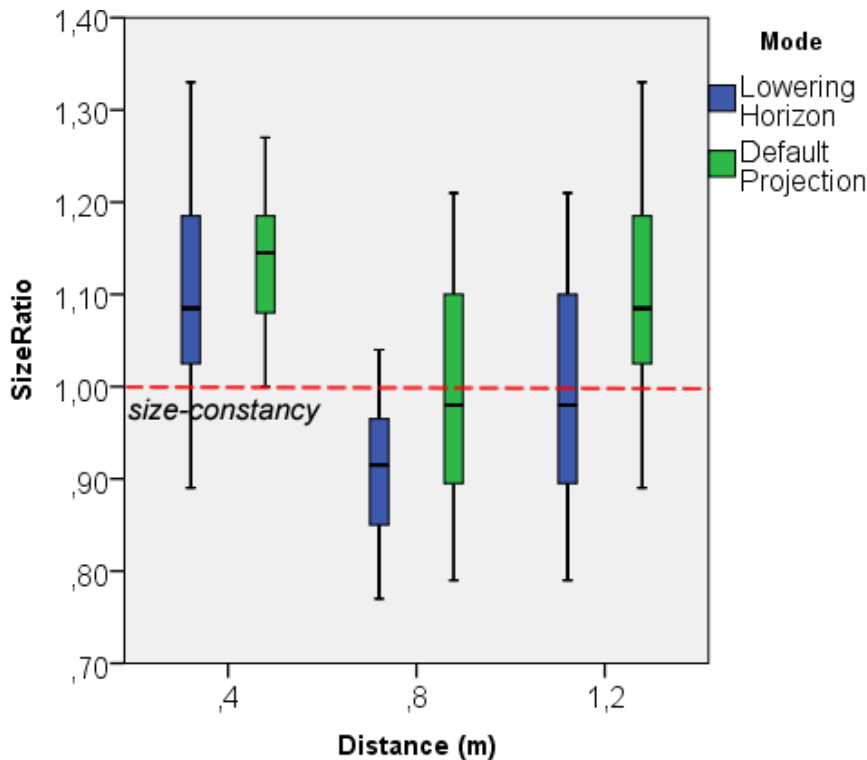


Figure 7.26: Influence de la technique OPP/PPO sur la perception de la constance de la taille

La technique améliore la perception de la taille et de la distance par rapport à la technique de projection standard, particulièrement à des distances plus grandes.

Discussion

La technique proposée a permis une amélioration partielle de la perception de la constance de la taille dans le HMD. Bien que la perception de la taille semble s'améliorer avec cette technique, des résultats inégaux ont aussi été observés, avec une surestimation des distances en parallaxe nulle et une sous-estimation de distances en parallaxe positive. Contrairement à l'étude précédente, la tendance à la constance de la taille est plus grande pour les objets en parallaxe positive, entre 0,8m et 1,2m. Ces résultats coïncident avec des études antérieures qui suggèrent qu'une réduction de l'horizon peut avoir une influence positive sur la perception des distances. L'influence des signaux picturaux est si forte qu'elle annule l'influence des signaux oculomoteurs. Ainsi, la technique crée l'illusion d'une table plus grande que ses dimensions physiques, ce qui influe sur la perception de la distance et de la taille des objets.

7.6.2.2 Influence de la minification du champ de vision pour la marche en aveugle

Nous avons reproduit le dernier test de marche en aveugle, où les participants ont estimé la distance de différents objets en marchant dans des conditions non visuelles (figure 7.27). De même que dans le test précédent, nous nous sommes concentrés uniquement sur un HMD, mais cette fois-ci nous avons permis aux participants de marcher naturellement et nous avons analysé l'influence de la technique de projection de perspective minifiée (PPM) sur la perception spatiale (figure 7.28). La PPM fait converger les points de fuite à une vitesse moindre sur la ligne d'horizon, ce qui influe sur la parallaxe de mouvement et augmente le champ de vision perçu. Les objets qui n'étaient pas visibles avec le champ de vision prédéterminé deviennent visibles avec le champ de vision mis à l'échelle, et toute la scène est agrandie en profondeur. Nous avons posé comme hypothèse que l'application de cette technique améliorerait la perception des distances. Nous avons réalisé le même test avec la même conception expérimentale mais en appliquant cette fois-ci un facteur d'échelle de 10% sur le champ de vision.

Hypothèse

- H1: La perception des distances sera améliorée par rapport à la projection par défaut.

Méthode

7.6. PARTIE II. INFLUENCER LA PERCEPTION SPATIALE PAR COMPOSITION



Figure 7.27: Test de marche en aveugle

À gauche : sujet faisant le test de marche en aveugle. À droite : stimulus virtuel, le sujet doit estimer la distance jusqu'au stimulus.

L'expérience s'est déroulée avec le même appareil et en suivant la même procédure de la dernière expérience de marche en aveugle, cette fois-ci avec des distances cible de (3,5m, 5,5m et 7,5m).

Résultats

Les résultats de la technique PPM par rapport à la projection de perspective prédéterminée sont présentés à la figure 5.7. Un test-t pour échantillons appariés a permis de comparer la performance des participants entre les trois distances cible. Il y a eu une différence significative pour les trois distances cible entre la technique MPP/PPM et la projection standard à 3,5m ($M = 0.40, SD = 0.22, M = 0.91, SD = 0.22, t(47) = -9.1, p = 0.0$); à 5,5m ($M = 0.47, SD = 0.28, M = 1.01, SD = 0.33, t(47) = -6.1, p = 0.0$), et à 7,5 m ($M = 0.57, SD = 0.22, M = 1.01, SD = 0.45, t(47) = -4.2, p = 0.0$). Les résultats ont montré que H1 se maintient partiellement, que la technique de minification du champ de vision a une influence positive sur la perception des distances, avec une performance supérieure par rapport à la technique de projection par défaut.

Discussion

La technique MPP/PPM proposée a amélioré la perception des distances dans HMD par rapport à la technique de projection prédéterminée. Ces résultats coïncident avec des études antérieures qui suggèrent que l'augmentation artificielle du champ de vision peut avoir une influence positive sur la perception spatiale. Bien que la perception de la distance semble s'améliorer, la technique pourrait influencer sur le processus d'adaptation, où les distances semblent se recalibrer d'une manière

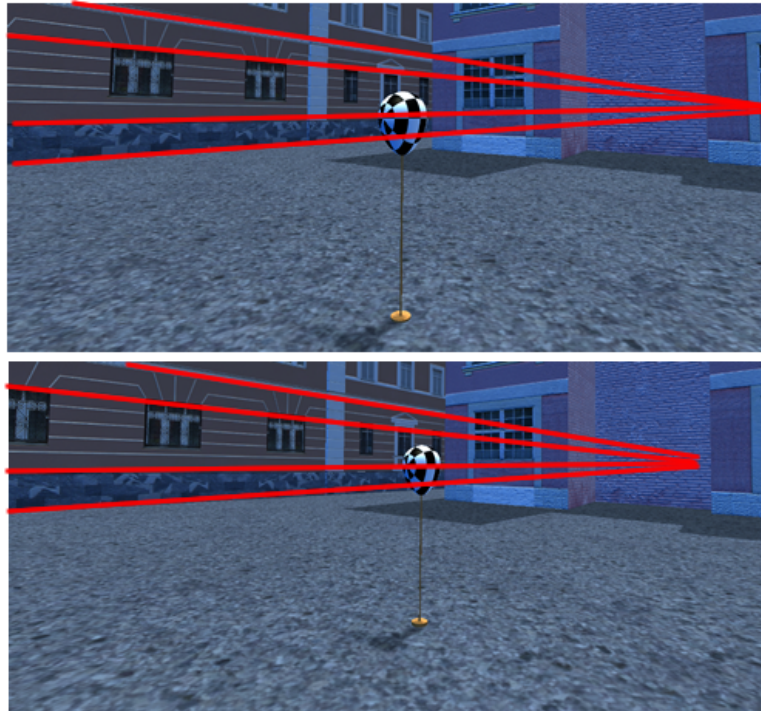


Figure 7.28: Mise en oeuvre de la technique de projection de perspective minifiée (PPM)

À gauche : projection par défaut. À droite : champ de vision mis à l'échelle par un facteur de 10%.

non uniforme. Cependant, l'augmentation du champ de vision selon un facteur d'échelle de 10% réduit de près de 50% la sous-estimation des distances. Ces résultats confirment une fois de plus la forte influence des signaux picturaux sur la perception générale de profondeur.

7.6.3 Amélioration de la performance spatiale

Dans la section précédente, nous avons présenté et validé la pertinence des techniques de composition pour influencer la perception des distances. Dans ce chapitre nous allons tout réunir et donnerons quelques idées sur la façon d'utiliser la pente du prédicteur de performance linéaire (PPL) pour appliquer une déclinaison angulaire optimale dans la technique de déclinaison de l'horizon (PPO), ou un facteur d'échelle optimal dans la technique de minification du champ de vision (PPM), afin de rapprocher la perception des distances en réalité virtuelle des valeurs trouvées dans des conditions naturelles. Ce sont des idées que nous avons explorées pour surmonter quelques limitations au cours des études antérieures. Nous comptons, dans un travail futur, les valider dans le cadre d'études expérimentales afin de créer certaines directrices/modèles sur la façon d'améliorer la perception des distances

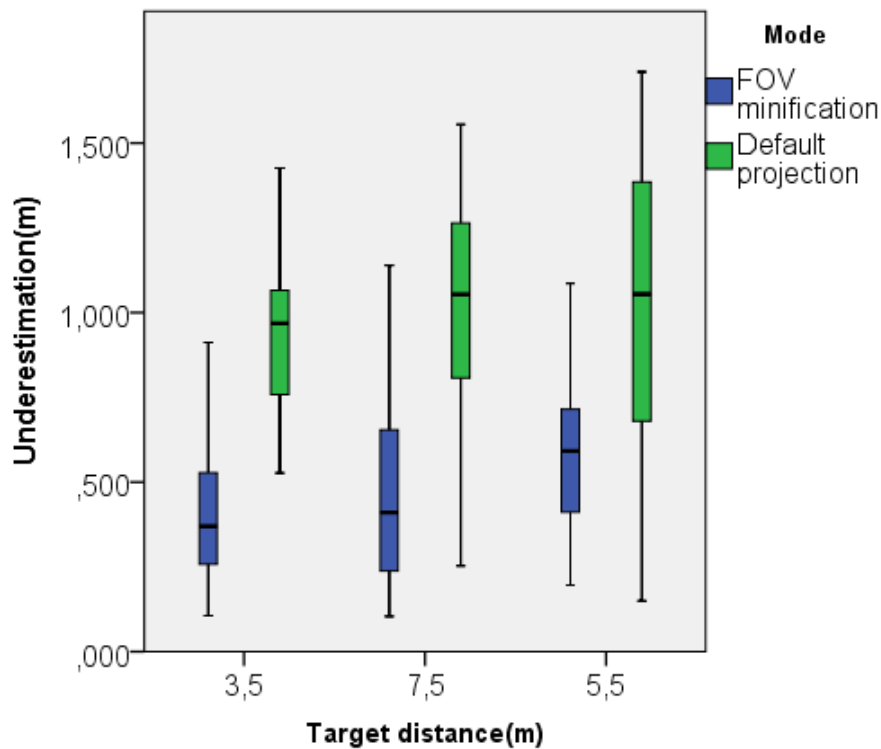


Figure 7.29: Influence de la technique de minification du champ de vision sur la perception des distances

La technique PPM permet une amélioration significative de la perception de la distance par rapport à la technique de projection prédéterminée.

en réalité virtuelle basés sur le PPL.

7.6.3.1 Optimized lowering horizon technique

Conformément à l'hypothèse de déclinaison angulaire $d = h/\tan(\alpha)$, les distances peuvent être estimées en fonction de la hauteur de l'oeil du sujet h et de la déclinaison angulaire en dessous de la ligne d'horizon α . À partir de cette hypothèse, nous pouvons utiliser le PPL pour déterminer le degré de déclinaison angulaire qui doit être appliqué à la ligne d'horizon afin d'améliorer la perception des distances. L'horizon perçu est basé sur une image à la distance focale; pour cette raison, nous pouvons utiliser comme référence la performance du sujet mesurée à cette distance, et estimer le déplacement angulaire entre les distances focales perçues et réelles. La figure 7.30 montre l'effet attendu lorsqu'on diminue la déclinaison angulaire de la ligne d'horizon en fonction de la distance

7.6. PARTIE II. INFLUENCER LA PERCEPTION SPATIALE PAR COMPOSITION

focale. Un abaissement de la ligne d'horizon par un déplacement angulaire β , augmente la perception des distances des points, que ce soit en parallaxe nulle (à distance focale), en parallaxe négative ou en parallaxe positive, l'influence étant plus grande en parallaxe positive. Pour calculer le déplacement angulaire β , le processus est simple. En premier lieu, nous calculons le PPL en utilisant une méthode d'évaluation quelconque, compte tenu de la sélection adéquate des stimuli. Ensuite, à partir de la performance du sujet à la distance focale f' , nous calculons l'angle de déclinaison correspondant par rapport à la ligne d'horizon. Par conséquent, la déclinaison angulaire à la distance focale sous-estimée est $\alpha' = \arctan(h/f')$. Enfin, pour augmenter la perception des distances, nous pouvons calculer le déplacement angulaire β , entre la distance focale f et la performance prévue du sujet à la distance focale f' , qui est l'angle requis pour réduire l'emplacement apparent de la ligne d'horizon. Ainsi, le déplacement angulaire β est obtenu comme suit:

$$\beta = \alpha - \alpha' = \arctan(h/f) - \arctan(h/f') \quad (7.4)$$

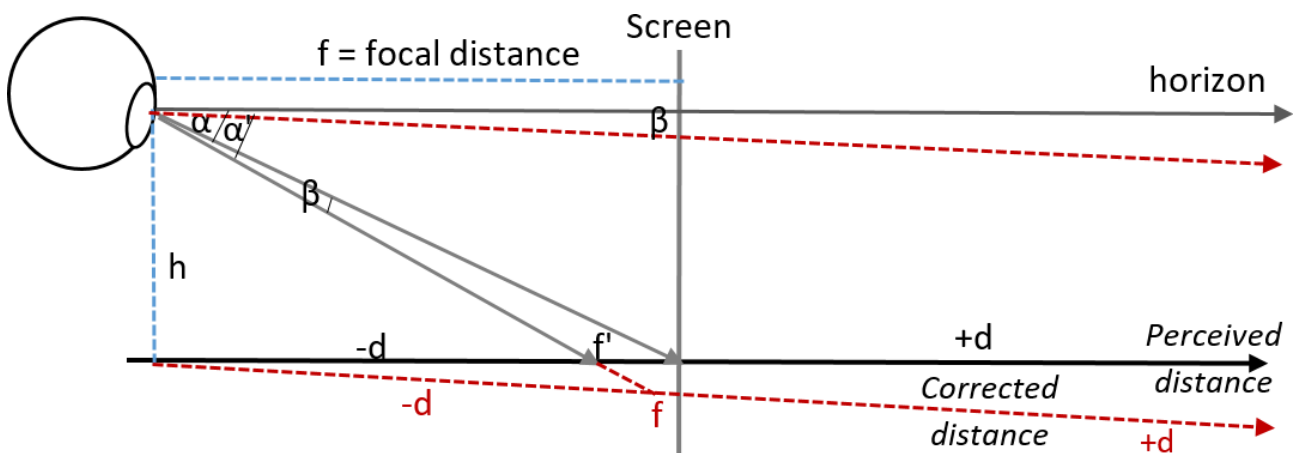


Figure 7.30: Effet de la réduction de la ligne d'horizon en fonction de la distance focale.

Après réduction de l'emplacement apparent de la ligne d'horizon par un déplacement angulaire β , la perception des distances augmente, que ce soit en parallaxe positive, négative ou nulle, l'effet étant plus important en parallaxe positive.

Le processus peut être décrit par un exemple : soit $d' = (1 - 0.20)d$ le prédicteur de performance linéaire du sujet; ceci veut dire que, pour chaque mètre, les distances sont sous-estimées de 20%. Soit une distance focale de 0,75m (comme dans le HTC Vive) ; la distance focale perçue prévue serait $f' = (0.8)(0.75m) = 0.6m$. Si la hauteur de l'oeil du sujet par rapport au sol est de 1,6m, la déclinaison

angulaire requise pour augmenter de 20% la perception des distances serait $\arctan(1.6m/0.75m) - \arctan(1.6m/0.6m) = -4.5^\circ$. Par conséquent, pour augmenter de 20% la perception des distances sur l'écran, la ligne d'horizon doit être réduite de 4,5 degrés.

7.6.3.2 Technique de minification du champ de vision optimisée

La même analogie peut être utilisée pour la technique de minification du champ de vision. Steinicke et al. (2009a) définissent la minification comme le rapport $m = \frac{\tan(vfov/2)}{\tan(vfov*g/2)}$, où m est le degré d'échelle uniforme nécessaire pour représenter la fenêtre graphique (rendue avec un certain GFOV) sur l'écran, $vfov$ étant le FOV (champ de vision) de visualisation verticale et g le facteur d'échelle géométrique appliqué au FOV (champ de vision). Si ($m = 1$), une personne va percevoir une image spatialement précise, selon ce qui est défini par les dimensions spatiales de l'environnement virtuel. Quand le champ de vision géométrique est augmenté ($g > 1$), l'image résultante est minimisée ($m < 1$), tandis qu'un champ de vision géométrique diminué ($g < 1$) donne comme résultat une image élargie ($m > 1$). Le degré de minification induite change la perception de la taille, mettant à échelle l'angle sous-tendu de l'image projetée des objets dans la rétine (figure 7.31). Ainsi, suivant l'hypothèse de constance de la taille $s = d * \tan(\theta)$, où s est la taille perçue de l'objet, θ est l'angle visuel sous-tendu, et d est la distance de l'objet, nous pouvons calculer la variance dans l'angle sous-tendu θ' , en utilisant le PPL, de sorte que $\theta' = 2 * \text{atan}(s/d')$, et calculer le rapport θ/θ' . Parce que cette proportion est équivalente à $\frac{\tan(vfov/2)}{\tan(vfov*g/2)}$, le facteur d'échelle géométrique nécessaire g requis pour améliorer la perception des distances, est obtenu comme suit:

$$g = \frac{\theta}{\theta'} = \frac{\text{atan}(s/d)}{\text{atan}(s/d')} \quad (7.5)$$

De même que pour la technique de l'abaissement de l'horizon, un exemple nous aidera à décrire le calcul du facteur géométrique g : suite à une méthode de marche en aveugle, le prédicteur de performance linéaire du sujet a été $d' = (1 - 0.30)d$, ce qui veut dire que, pour chaque mètre, les distances sont sous-estimées de 30%. Soit un stimulus visuel de taille $s = 1m$, placé à la distance focale de 0.75m, ce qui implique que la distance focale perçue prévue serait $f' = (0.7)(0.75m) = 0.52$; alors, le facteur géométrique nécessaire pour augmenter la perception des distances de 30% serait $g = \frac{\text{atan}(1m/0.75m)}{\text{atan}(1m/0.52m)} = 0.84$. Ainsi, pour augmenter la perception des distances de 30% sur l'écran, le degré de minification qui peut être induit dans un écran avec un champ de vision de 110° étant

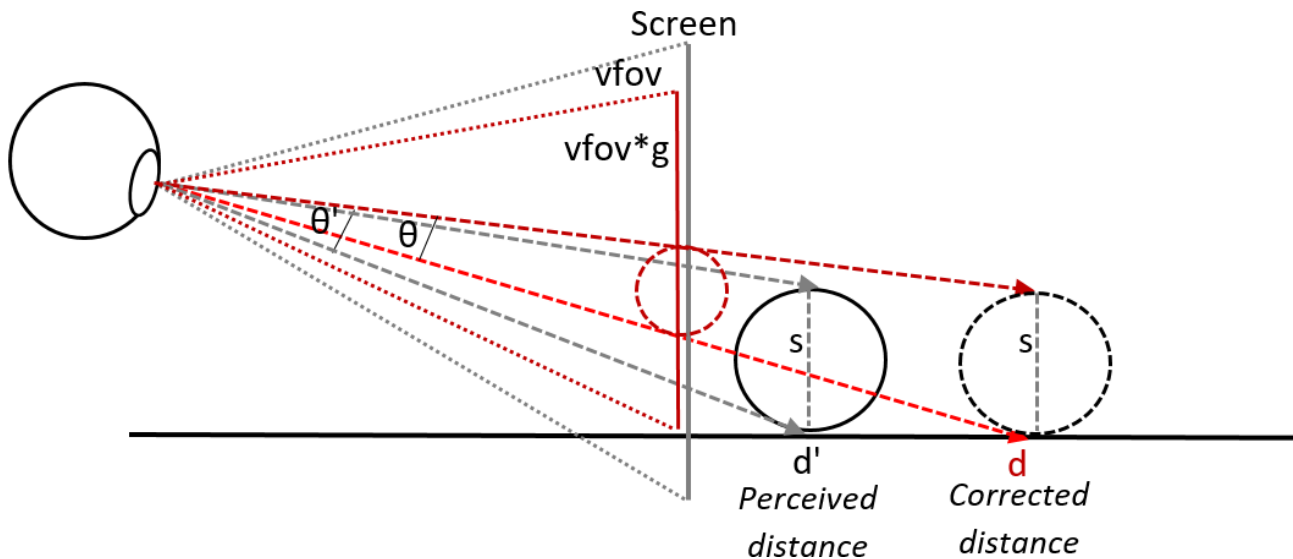


Figure 7.31: - Effet de la mise à échelle du champ de vision (FOV) en fonction de la variation de l'angle visuel sous-tendu.

Suite à la mise à échelle du champ de vision (FOV) vertical par le facteur géométrique g , l'angle sous-tendu de l'objet θ se réduit, selon son image projetée et en augmentant la perception de la distance

$m = \frac{\tan(110^\circ/2)}{\tan(110^\circ * 0.84/2)} = 1.1$, the FOV must be scaled a 10%, le champ de vision doit être mis à l'échelle par un facteur de 10

7.6.4 Résumé

Dans ce chapitre, nous avons présenté la dernière partie du cadre de référence proposé, qui se concentre sur la manière d'influencer la perception spatiale en utilisant la composition. Nous avons présenté notre mise en oeuvre de la technique de réduction de l'horizon et la technique de minification du champ de vision, et nous avons validé la pertinence de ces techniques pour améliorer la perception spatiale en réalité virtuelle. Nos résultats suggèrent que les deux techniques peuvent avoir une influence positive sur la perception des distances, mais le choix de la déclinaison angulaire requise ou du facteur d'échelle requis pour chaque technique est critique, puisqu'une valeur en dehors des paramètres optimaux pourrait conduire à des résultats inattendus. Nous avons donc exploré quelques idées sur la façon d'utiliser la pente du LPP/PPL pour améliorer la perception des distances et comme un moyen de calcul des valeurs optimales pour chaque technique.

7.7 Conclusion

Cette thèse présente un cadre de référence pour caractériser et influencer la perception spatiale entre des écrans hétérogènes de réalité virtuelle. Ce cadre de référence a montré des résultats prometteurs conformément aux objectifs de recherche. Le premier objectif était d'étudier et de caractériser l'influence des différents facteurs d'affichage sur la perception spatiale. À cet égard, nous avons décrit l'influence de deux des problèmes les plus complexes : le problème du conflit V-A et le problème de la stimulation périphérique incorrecte. Le deuxième objectif était de trouver des méthodes d'évaluation pour comparer la perception spatiale entre des écrans hétérogènes. Ici, nous avons adapté quelques méthodes d'évaluation trouvées dans la littérature. Enfin, le dernier objectif était de déterminer comment la composition peut influencer la perception spatiale. Pour cela, nous avons mis en œuvre des techniques de projection alternatives, qui peuvent être optimisées pour améliorer la perception des distances. Si les résultats n'ont pas toujours été en ligne avec nos attentes, nous sommes toutefois optimistes quant à la pertinence du cadre de référence proposé. Parmi les limitations et les problèmes rencontrés, certains peuvent se résoudre facilement, tandis que d'autres auraient probablement bénéficié d'une approche différente.

Une contribution fondamentale de cette thèse est de proposer une approche formelle pour caractériser le phénomène de compression de la profondeur dans les écrans de réalité virtuelle, prenant en considération l'influence des facteurs d'affichage, tels que la distance focale, la condition de parallaxe stéréoscopique, les distorsions de la lentille et le champ de vision limité. Bien qu'il y ait plusieurs décennies d'études sur ses facteurs, aucune approche n'a tenté d'intégrer tous ces facteurs dans un cadre de référence unique. Une conclusion principale de cette thèse est que l'étude de la perception spatiale en réalité virtuelle, ainsi que toute tentative d'influencer cette perception, est tout aussi difficile que de l'étudier dans le monde physique. Les modèles proposés n'ont pas donné des résultats précis à 100%, puisque la perception humaine est incroyablement complexe. Par conséquent, d'autres études sont nécessaires sur l'interaction entre les signaux de profondeur, sur la façon dont nous percevons les distances et sur l'influence de l'action sur la perception.

Le prédicteur de performance linéaire est une bonne méthode pour caractériser le phénomène de compression de la profondeur dans l'écran, mais il n'est pas parfait, puisqu'il dépend de la confiance du sujet au moment d'estimer les distances. Plus il y a de variables dans le modèle, moins la prédiction de

7.7. CONCLUSION

la performance du sujet est précise. En outre, il n'est pas encore possible d'affirmer que les hypothèses utilisées pour le cadre de référence proposé sont vraies, puisqu'elles peuvent dépendre du type de la tâche et/ou de la méthode d'évaluation. Par conséquent, de nouvelles études et davantage de connaissances sur le phénomène sont nécessaires.

Enfin, s'agissant d'un cadre de référence, ce travail est continu. Nous devons réaliser d'autres études afin de valider les hypothèses ici posées, améliorer les techniques alternatives de projection, explorer différentes méthodes pour influencer la perception spatiale et considérer différents types de tâches spatiales. Nous considérons que cette thèse, ainsi que les progrès technologiques futurs, aideront à surmonter le problème de la sous-estimation des distances en réalité virtuelle.

7.7. CONCLUSION

Bibliography

- Andre, J. and Rogers, S. (2006). Using verbal and blind-walking distance estimates to investigate the two visual systems hypothesis. *Attention, Perception, & Psychophysics*, 68(3):353–361.
- Andrus, S. M., Gaylor, G., and Bodenheimer, B. (2014). Distance estimation in virtual environments using different hmds. In *Proceedings of the ACM Symposium on Applied Perception*, pages 130–130. ACM.
- Angelaki, D. E., Klier, E. M., and Snyder, L. H. (2009). A vestibular sensation: probabilistic approaches to spatial perception. *Neuron*, 64(4):448–461.
- Apress (2019). *Pimax brings VR 2.0 to Las Vegas at CES 2019* - <https://apnews.com/press-release/pr-businesswire/430ff90e3c0f4b2ca42e5c07d9325781>.
- Banakou, D., Groten, R., and Slater, M. (2013). Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes. *Proceedings of the National Academy of Sciences*, 110(31):12846–12851.
- Barrera Machuca, M. D. and Stuerzlinger, W. (2019). The effect of stereo display deficiencies on virtual hand pointing. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, page 207. ACM.
- Baumberger, B., FLÜCKIGER, M., Paquette, M., Bergeron, J., and Delorme, A. (2005). Perception of relative distance in a driving simulator 1, 2. *Japanese Psychological Research*, 47(3):230–237.
- Bolte, B., Bruder, G., Steinicke, F., Hinrichs, K., and Lappe, M. (2010). Augmentation techniques for efficient exploration in head-mounted display environments. In *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology*, pages 11–18. ACM.
- Bossard, M., Goulon, C., and Mestre, D. (2020). Viewpoint oscillation improves the perception of distance travelled in static observers but not during treadmill walking. *Experimental brain research*, pages 1–11.
- Bradshaw, M. F., Parton, A. D., and Glennerster, A. (2000). The task-dependent use of binocular disparity and motion parallax information. *Vision research*, 40(27):3725–3734.
- Bradshaw, M. F. and Rogers, B. J. (1996). The interaction of binocular disparity and motion parallax in the computation of depth. *Vision research*, 36(21):3457–3468.
- Bruder, G., Argelaguet, F., Olivier, A.-H., and Lécuyer, A. (2016). Cave size matters: Effects of screen distance and parallax on distance estimation in large immersive display setups. *Presence: Teleoperators and Virtual Environments*, 25(1):1–16.
- Buck, L. E., Young, M. K., and Bodenheimer, B. (2018). A comparison of distance estimation in hmd-based virtual environments with different hmd-based conditions. *ACM Transactions on Applied Perception (TAP)*, 15(3):21.
- Bülthoff, H. H. and van Veen, H. A. (2001). Vision and action in virtual environments: Modern psychophysics in spatial cognition research. In *Vision and attention*, pages 233–252. Springer.
- Bunch, D. A. (2014). The perception of distance on a real geographic slope.
- Campbell, F. W. (1957). The depth of field of the human eye. *Optica Acta: International Journal of Optics*, 4(4):157–164.

BIBLIOGRAPHY

- Channel, N. G. (2013). *National Geographic Channel, Brain games, Seeing is believing, Season 2 Episode 12*.
- Chessa, M., Maiello, G., Klein, L. K., Paulun, V. C., and Solari, F. (2019). Grasping objects in immersive virtual reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 1749–1754. IEEE.
- Creem-Regehr, S. H., Stefanucci, J. K., Thompson, W. B., Nash, N., and McCardell, M. (2015). Egocentric distance perception in the oculus rift (dk2). In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception*, pages 47–50. ACM.
- Creem-Regehr, S. H., Willemsen, P., Gooch, A. A., and Thompson, W. B. (2005). The influence of restricted viewing conditions on egocentric distance perception: Implications for real and virtual indoor environments. *Perception*, 34(2):191–204.
- Cruz-Neira, C., Sandin, D. J., and DeFanti, T. A. (1993). Surround-screen projection-based virtual reality: the design and implementation of the cave. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 135–142. Citeseer.
- Cutting, J. E. and Vishton, P. M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth. In *Perception of space and motion*, pages 69–117. Elsevier.
- Doi, M., Takai, T., and Chihara, K. (2000). Vr american football simulator with cylindrical screen. In *International Conference on Virtual Worlds*, pages 286–293. Springer.
- Dorward, F. M. and Day, R. H. (1997). Loss of 3-d shape constancy in interior spaces: The basis of the ames-room illusion. *Perception*, 26(6):707–718.
- El Jamiy, F. and Marsh, R. (2019). Survey on depth perception in head mounted displays: distance estimation in virtual reality, augmented reality, and mixed reality. *IET Image Processing*, 13(5):707–712.
- Febretti, A., Nishimoto, A., Thigpen, T., Talandis, J., Long, L., Pirtle, J., Peterka, T., Verlo, A., Brown, M., Plepys, D., et al. (2013). Cave2: a hybrid reality environment for immersive simulation and information analysis. In *The Engineering Reality of Virtual Reality 2013*, volume 8649, page 864903. International Society for Optics and Photonics.
- Foley, J. M. (1985). Binocular distance perception: Egocentric distance tasks. *Journal of Experimental Psychology: Human Perception and Performance*, 11(2):133.
- Fu, H., Wei, Y., Camastra, F., Arico, P., and Sheng, H. (2016). Advances in eye tracking technology: theory, algorithms, and applications. *Computational intelligence and neuroscience*, 2016.
- Fujimoto, K. and Ashida, H. (2020). Different head-sway responses to optic flow in sitting and standing with a head-mounted display. *Frontiers in psychology*, 11:2681.
- Fukushima, S. S., Loomis, J. M., and Da Silva, J. A. (1997). Visual perception of egocentric distance as assessed by triangulation. *Journal of Experimental Psychology: Human Perception and Performance*, 23(1):86.
- Gajewski, D. A., Philbeck, J. W., Wirtz, P. W., and Chichka, D. (2014). Angular declination and the dynamic perception of egocentric distance. *Journal of Experimental Psychology: Human Perception and Performance*, 40(1):361.
- Galati, G., Pelle, G., Berthoz, A., and Committeri, G. (2010). Multiple reference frames used by the human brain for spatial perception and memory. *Experimental brain research*, 206(2):109–120.
- Gärbling, T. and Golledge, R. G. (1989). Environmental perception and cognition. In *Advance in environment, behavior, and design*, pages 203–236. Springer.
- Gehring, W. L. and Engel, E. (1986). Effect of ecological viewing conditions on the ames’ distorted room illusion.
- Ghahramani, Z., Wolpirt, D. M., and Jordan, M. I. (1997). Computational models of sensorimotor integration. In *Advances in Psychology*, volume 119, pages 117–147. Elsevier.

BIBLIOGRAPHY

- Glennester, A., Tcheang, L., Gilson, S. J., Fitzgibbon, A. W., and Parker, A. J. (2006). Humans ignore motion and stereo cues in favor of a fictional stable world. *Current Biology*, 16(4):428–432.
- Gogel, W. C. and Da Silva, J. A. (1987). A two-process theory of the response to size and distance. *Perception & Psychophysics*, 41(3):220–238.
- Goodale, M. A. and Haffenden, A. (1998). Frames of reference for perception and action in the human visual system. *Neuroscience & Biobehavioral Reviews*, 22(2):161–172.
- Goodale, M. A. and Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in neurosciences*, 15(1):20–25.
- Grubert, J., Itoh, Y., Moser, K., and Swan, J. E. (2018). A survey of calibration methods for optical see-through head-mounted displays. *IEEE transactions on visualization and computer graphics*, 24(9):2649–2662.
- Haber, R. N. and Levin, C. A. (2001). The independence of size perception and distance perception. *Perception & psychophysics*, 63(7):1140–1152.
- Held, R. T., Cooper, E. A., and Banks, M. S. (2012). Blur and disparity are complementary cues to depth. *Current biology*, 22(5):426–431.
- Held, R. T., Cooper, E. A., O’Brien, J. F., and Banks, M. S. (2010). Using blur to affect perceived distance and size. *ACM transactions on graphics*, 29(2).
- Hershenson, M. (1999). *Visual space perception: A primer*. Mit Press.
- Hoffman, D. M., Girshick, A. R., Akeley, K., and Banks, M. S. (2008). Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of vision*, 8(3):33–33.
- Hollingsworth, K. P., Bowyer, K. W., and Flynn, P. J. (2008). The importance of small pupils: a study of how pupil dilation affects iris biometrics. In *2008 IEEE Second International Conference on Biometrics: Theory, Applications and Systems*, pages 1–6. IEEE.
- Hong, J., Kim, Y., Choi, H.-J., Hahn, J., Park, J.-H., Kim, H., Min, S.-W., Chen, N., and Lee, B. (2011). Three-dimensional display technologies of recent interest: principles, status, and issues [invited]. *Appl. Opt.*, 50(34):H87–H115.
- Huang, F.-C., Chen, K., and Wetzstein, G. (2015). The light field stereoscope: immersive computer graphics via factored near-eye light field displays with focus cues. *ACM Transactions on Graphics (TOG)*, 34(4):60.
- Hung, G. K., Ciuffreda, K. J., and Rosenfield, M. (1996). Proximal contribution to a linear static model of accommodation and vergence. *Ophthalmic and Physiological Optics*, 16(1):31–41.
- Ibbotson, M. and Krekelberg, B. (2011). Visual perception and saccadic eye movements. *Current opinion in neurobiology*, 21(4):553–558.
- IowaC6 (2014). *Iowa State University virtual reality system simulates the International Space Station* - <https://www.news.iastate.edu/news/2014/10/29/simiss>.
- Jacobs, R. A. (1999). Optimal integration of texture and motion cues to depth. *Vision research*, 39(21):3621–3629.
- Jo, D., Kang, H., Lee, G. A., and Son, W. (2006). Xsphere: A pc cluster based hemispherical display system. In *IEEE Virtual Reality Conference (VR 2006)*, pages 312–312. IEEE.
- Jones, G. R., Lee, D., Holliman, N. S., and Ezra, D. (2001). Controlling perceived depth in stereoscopic images. In *Stereoscopic Displays and Virtual Reality Systems VIII*, volume 4297, pages 42–53. International Society for Optics and Photonics.
- Jones, J. A., Swan, J. E., and Bolas, M. (2013). Peripheral stimulation and its effect on perceived spatial scale in virtual environments. *IEEE Transactions on Visualization & Computer Graphics*, (4):701–710.

BIBLIOGRAPHY

- Kehrer, B. (2016). *VR Distortion Correction using Vertex Displacement* - <https://www.gamasutra.com/blogs/BrianKehrer/20160125/264161>.
- Kellner, F., Bolte, B., Bruder, G., Rautenberg, U., Steinicke, F., Lappe, M., and Koch, R. (2012). Geometric calibration of head-mounted displays and its effects on distance estimation. *IEEE transactions on visualization and computer graphics*, 18(4):589–596.
- Kellnhofer, P., Didyk, P., Ritschel, T., Masia, B., Myszkowski, K., and Seidel, H.-P. (2016). Motion parallax in stereo 3d: Model and applications. *ACM Transactions on Graphics (TOG)*, 35(6):1–12.
- Kelly, J. W., Cherep, L. A., Klesel, B., Siegel, Z. D., and George, S. (2018). Comparison of two methods for improving distance perception in virtual reality. *ACM Transactions on Applied Perception (TAP)*, 15(2):11.
- Kelly, J. W., Cherep, L. A., and Siegel, Z. D. (2017). Perceived space in the htc vive. *ACM Transactions on Applied Perception (TAP)*, 15(1):2.
- Kelly, J. W., Donaldson, L. S., Sjolund, L. A., and Freiberg, J. B. (2013). More than just perception–action recalibration: Walking through a virtual environment causes rescaling of perceived space. *Attention, Perception, & Psychophysics*, 75(7):1473–1485.
- Kelly, J. W., Hammel, W. W., Siegel, Z. D., and Sjolund, L. A. (2014). Recalibration of perceived distance in virtual environments occurs rapidly and transfers asymmetrically across scale. *IEEE transactions on visualization and computer graphics*, 20(4):588–595.
- Kenyon, R. V., Phenany, M., Sandin, D., and Defanti, T. (2008). Accommodation and size-constancy of virtual objects. *Annals of biomedical engineering*, 36(2):342–348.
- Kenyon, R. V., Sandin, D., Smith, R. C., Pawlicki, R., and Defanti, T. (2007). Size-constancy in the cave. *Presence: Teleoperators and Virtual Environments*, 16(2):172–187.
- Kilpatrick, F. and Ittelson, W. (1953). The size-distance invariance hypothesis. *Psychological Review*, 60(4):223.
- King, M., Meyer, G. E., Tangney, J., and Biederman, I. (1976). Shape constancy and a perceptual bias towards symmetry. *Perception & Psychophysics*, 19(2):129–136.
- Klatzky, R. L. (1998). Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. In *Spatial cognition*, pages 1–17. Springer.
- Kline, P. B. and Witmer, B. G. (1996). 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*, volume 40, pages 1112–1116. SAGE Publications Sage CA: Los Angeles, CA.
- Knapp, J. M. and Loomis, J. M. (2004). Limited field of view of head-mounted displays is not the cause of distance underestimation in virtual environments. *Presence: Teleoperators & Virtual Environments*, 13(5):572–577.
- Knill, D. C. and Saunders, J. A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision research*, 43(24):2539–2558.
- Kolarik, A. J., Moore, B. C., Zahorik, P., Cirstea, S., and Pardhan, S. (2016). Auditory distance perception in humans: a review of cues, development, neuronal bases, and effects of sensory loss. *Attention, Perception, & Psychophysics*, 78(2):373–395.
- Kreylos, O. (2017). *Measuring the Effective Resolution of Head-mounted Displays* - <http://doc-ok.org/?p=1694>.
- Kuhl, S. A., Thompson, W. B., and Creem-Regehr, S. H. (2006). Minification influences spatial judgments in virtual environments. In *Proceedings of the 3rd symposium on Applied perception in graphics and visualization*, pages 15–19. ACM.
- Kuhl, S. A., Thompson, W. B., and Creem-Regehr, S. H. (2009). Hmd calibration and its effects on distance judgments. *ACM Transactions on Applied Perception (TAP)*, 6(3):19.

BIBLIOGRAPHY

- Kunz, B. R., Wouters, L., Smith, D., Thompson, W. B., and Creem-Regehr, S. H. (2009). Revisiting the effect of quality of graphics on distance judgments in virtual environments: A comparison of verbal reports and blind walking. *Attention, Perception, & Psychophysics*, 71(6):1284–1293.
- Lampton, D. R., McDonald, D. P., Singer, M., and Bliss, J. P. (1995). Distance estimation in virtual environments. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 39, pages 1268–1272. SAGE Publications Sage CA: Los Angeles, CA.
- Landy, M. S., Banks, M. S., and Knill, D. C. (2011). Ideal-observer models of cue integration. *Sensory cue integration*, pages 5–29.
- Landy, M. S., Maloney, L. T., Johnston, E. B., and Young, M. (1995). Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision research*, 35(3):389–412.
- LaValle, S. (2016). Virtual reality.
- Lederman, S. J. and Klatzky, R. L. (2009). Haptic perception: A tutorial. *Attention, Perception, & Psychophysics*, 71(7):1439–1459.
- Li, B., Nordman, A., Walker, J., and Kuhl, S. A. (2016). The effects of artificially reduced field of view and peripheral frame stimulation on distance judgments in hmds. In *Proceedings of the ACM Symposium on Applied Perception*, pages 53–56. ACM.
- Li, B., Walker, J., and Kuhl, S. A. (2018). The effects of peripheral vision and light stimulation on distance judgments through hmds. *ACM Transactions on Applied Perception (TAP)*, 15(2):12.
- Li, B., Zhang, R., and Kuhl, S. (2014). Minication affects action-based distance judgments in oculus rift hmds. In *Proceedings of the ACM Symposium on Applied Perception*, pages 91–94. ACM.
- Li, C., Lin, H., and Shi, J. (2004). A survey of multi-projector tiled display wall construction. In *Third International Conference on Image and Graphics (ICIG'04)*, pages 452–455. IEEE.
- Li, H., Mavros, P., Krukar, J., and Hölscher, C. (2021). The effect of navigation method and visual display on distance perception in a large-scale virtual building. *Cognitive Processing*, pages 1–21.
- Linton, P. (2020). Does vision extract absolute distance from vergence? *Attention, perception & psychophysics*, 82(6):3176.
- Lipton, L. (1982). *Foundations of the stereoscopic cinema: a study in depth*. Van Nostrand Reinhold.
- Lipton, L. (1997). Stereographics, developers handbook. *StereoGraphics Corporation*, 2(2.2):2–2.
- Livingston, M. A., Ai, Z., Swan, J. E., and Smallman, H. S. (2009). Indoor vs. outdoor depth perception for mobile augmented reality. In *2009 IEEE Virtual Reality Conference*, pages 55–62. IEEE.
- Loomis, J. M., Da Silva, J. A., Fujita, N., and Fukusima, S. S. (1992). Visual space perception and visually directed action. *Journal of Experimental Psychology: Human Perception and Performance*, 18(4):906.
- Loomis, J. M., Klatzky, R. L., and Giudice, N. A. (2013). Representing 3d space in working memory: Spatial images from vision, hearing, touch, and language. In *Multisensory imagery*, pages 131–155. Springer.
- Loomis, J. M., Knapp, J. M., et al. (2003). Visual perception of egocentric distance in real and virtual environments. *Virtual and adaptive environments*, 11:21–46.
- Luo, X., Kenyon, R., Kamper, D., Sandin, D., and DeFanti, T. (2007). The effects of scene complexity, stereovision, and motion parallax on size constancy in a virtual environment. In *Virtual Reality Conference, 2007. VR'07. IEEE*, pages 59–66. IEEE.
- Mansour, M., Davidson, P., Stepanov, O., and Piché, R. (2019). Relative importance of binocular disparity and motion parallax for depth estimation: A computer vision approach. *Remote Sensing*, 11(17):1990.

BIBLIOGRAPHY

- Marsh, W. E., Chardonnet, J.-R., and Merienne, F. (2014). Virtual distance estimation in a cave. In *International Conference on Spatial Cognition*, pages 354–369. Springer.
- Masters, R., Poolton, J., and Van Der Kamp, J. (2010). Regard and perceptions of size in soccer: Better is bigger. *Perception*, 39(9):1290–1295.
- Mather, G. (1997). The use of image blur as a depth cue. *Perception*, 26(9):1147–1158.
- Matsuda, N., Fix, A., and Lanman, D. (2017). Focal surface displays. *ACM Transactions on Graphics (TOG)*, 36(4):86.
- Memmert, D., Blanco, M., Merkle, V., et al. (2009). The effects of effort, performance, and expertise on apparent size perception in golf. *International Journal of Sport Psychology*, 40(2):270.
- Messing, R. and Durgin, F. H. (2005). Distance perception and the visual horizon in head-mounted displays. *ACM Transactions on Applied Perception (TAP)*, 2(3):234–250.
- Mittelstaedt, J. M., Wacker, J., and Stelling, D. (2019). Vr aftereffect and the relation of cybersickness and cognitive performance. *Virtual Reality*, 23(2):143–154.
- Mohler, B. J., Thompson, W. B., Creem-Regehr, S. H., Willemsen, P., Pick, Jr, H. L., and Rieser, J. J. (2007). Calibration of locomotion resulting from visual motion in a treadmill-based virtual environment. *ACM Transactions on Applied Perception (TAP)*, 4(1):4–es.
- Mon-Williams, M., Tresilian, J. R., and Roberts, A. (2000). Vergence provides veridical depth perception from horizontal retinal image disparities. *Experimental brain research*, 133(3):407–413.
- Montello, D. R. (1997). The perception and cognition of environmental distance: Direct sources of information. In *International Conference on Spatial Information Theory*, pages 297–311. Springer.
- Murgia, A., Sharkey, P. M., et al. (2009). Estimation of distances in virtual environments using size constancy. *The International Journal of Virtual Reality*, 8(1):67–74.
- Naceri, A., Chellali, R., Dionnet, F., and Toma, S. (2010). Depth perception within virtual environments: comparison between two display technologies. *International Journ. on Advances in Intelligent Systems*, 3.
- National, G. (2016). *National Gallery of Art London* - <https://www.youtube.com/watch?v=9KiVNIUMmCc>.
- National, G. (2018). *National Gallery of Art Italy* - <https://www.youtube.com/watch?v=FfLKWlak7zY>.
- Nawrot, M. and Stroyan, K. (2009). The motion/pursuit law for visual depth perception from motion parallax. *Vision research*, 49(15):1969–1978.
- Ogawa, N., Narumi, T., and Hirose, M. (2019). Virtual hand realism affects object size perception in body-based scaling. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 519–528. IEEE.
- Ooi, T. L., Wu, B., and He, Z. J. (2001). Distance determined by the angular declination below the horizon. *Nature*, 414(6860):197.
- Orlosky, J., Wu, Q., Kiyokawa, K., Takemura, H., and Nitschke, C. (2014). Fisheye vision: peripheral spatial compression for improved field of view in head mounted displays. In *Proceedings of the 2nd ACM symposium on Spatial user interaction*, pages 54–61.
- Patterson, R., Winterbottom, M. D., and Pierce, B. J. (2006). Perceptual issues in the use of head-mounted visual displays. *Human factors*, 48(3):555–573.
- Patterson, Z. (2019). Effects of avatar hand-size modifications on size judgments of familiar and abstract objects in virtual reality.
- Peer, A. and Ponto, K. (2017). Evaluating perceived distance measures in room-scale spaces using consumer-grade head mounted displays. In *3D User Interfaces (3DUI), 2017 IEEE Symposium on*, pages 83–86. IEEE.

BIBLIOGRAPHY

- Plumert, J. M., Kearney, J. K., Cremer, J. F., and Recker, K. (2005). Distance perception in real and virtual environments. *ACM Transactions on Applied Perception (TAP)*, 2(3):216–233.
- Pohl, D., Johnson, G. S., and Bolkart, T. (2013). Improved pre-warping for wide angle, head mounted displays. In *Proceedings of the 19th ACM symposium on Virtual Reality Software and Technology*, pages 259–262. ACM.
- Proffitt, D. R. (2006). Embodied perception and the economy of action. *Perspectives on psychological science*, 1(2):110–122.
- Proffitt, D. R., Stefanucci, J., Banton, T., and Epstein, W. (2003). The role of effort in perceiving distance. *Psychological Science*, 14(2):106–112.
- Reddit.com (2017). *Vive focus distance is 75 cm* - <https://www.reddit.com/r/Vive/comments/4gbhny>.
- Reichelt, S., Häussler, R., Fütterer, G., and Leister, N. (2010). 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*, volume 7690, page 76900B. International Society for Optics and Photonics.
- Renner, R. S., Velichkovsky, B. M., and Helmert, J. R. (2013). The perception of egocentric distances in virtual environments—a review. *ACM Computing Surveys (CSUR)*, 46(2):23.
- Richardson, A. R. and Waller, D. (2007). Interaction with an immersive virtual environment corrects users’ distance estimates. *Human Factors*, 49(3):507–517.
- Rieser, J. J., Pick, H. L., Ashmead, D. H., and Garing, A. E. (1995). Calibration of human locomotion and models of perceptual-motor organization. *Journal of Experimental Psychology: Human Perception and Performance*, 21(3):480.
- Sachse, P., Beermann, U., Martini, M., Maran, T., Domeier, M., and Furtner, M. R. (2017). “the world is upside down”—the innsbruck goggle experiments of theodor erismann (1883–1961) and ivo kohler (1915–1985). *Cortex*, 92:222–232.
- Sahm, C. S., Creem-Regehr, S. H., Thompson, W. B., and Willemsen, P. (2005). Throwing versus walking as indicators of distance perception in similar real and virtual environments. *ACM Transactions on Applied Perception (TAP)*, 2(1):35–45.
- Salomon, D. (2007). *Transformations and projections in computer graphics*. Springer Science & Business Media.
- Santillán, J. E. and Barraza, J. F. (2019). Distance perception during self-movement. *Human movement science*, 67:102496.
- Schreiber, K. M., Hillis, J. M., Filippini, H. R., Schor, C. M., and Banks, M. S. (2008). The surface of the empirical horopter. *Journal of Vision*, 8(3):7–7.
- Schuirman, D. J. (1987). A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *Journal of pharmacokinetics and biopharmaceutics*, 15(6):657–680.
- Seckel, A. (2006). *Optical illusions: The science of visual perception*. Firefly Books Ltd.
- Shibata, T., Kim, J., Hoffman, D. M., and Banks, M. S. (2011). The zone of comfort: Predicting visual discomfort with stereo displays. *Journal of vision*, 11(8):11–11.
- Siegel, Z. D. and Kelly, J. W. (2017). Walking through a virtual environment improves perceived size within and beyond the walked space. *Attention, Perception, & Psychophysics*, 79(1):39–44.
- Sinai, M. J., Krebs, W. K., Darken, R. P., Rowland, J., and McCarley, J. (1999). Egocentric distance perception in a virtual environment using a perceptual matching task. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 43, pages 1256–1260. SAGE Publications Sage CA: Los Angeles, CA.
- Stannard, I. (2010). *Ames room* - https://www.flickr.com/photos/silly_little_man/5132242358.

BIBLIOGRAPHY

- Steinicke, F., Bruder, G., Hinrichs, K., Kuhl, S., Lappe, M., and Willemsen, P. (2009a). Judgment of natural perspective projections in head-mounted display environments. In *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology*, pages 35–42.
- Steinicke, F., Bruder, G., Hinrichs, K., Lappe, M., Ries, B., and Interrante, V. (2009b). Transitional environments enhance distance perception in immersive virtual reality systems. In *Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization*, pages 19–26.
- Sun, Y. and Wang, H. (2010). Perception of space by multiple intrinsic frames of reference. *Plos One*, 5(5):e10442.
- Suryakumar, R., Meyers, J. P., Irving, E. L., and Bobier, W. R. (2007). Vergence accommodation and monocular closed loop blur accommodation have similar dynamic characteristics. *Vision research*, 47(3):327–337.
- Swan, J. E., Jones, A., Kolstad, E., Livingston, M. A., and Smallman, H. S. (2007). Egocentric depth judgments in optical, see-through augmented reality. *IEEE transactions on visualization and computer graphics*, 13(3):429–442.
- Szpak, A., Michalski, S. C., Saredakis, D., Chen, C. S., and Loetscher, T. (2019). Beyond feeling sick: the visual and cognitive aftereffects of virtual reality. *IEEE Access*, 7:130883–130892.
- Tam, W. J., Speranza, F., Yano, S., Shimono, K., and Ono, H. (2011). Stereoscopic 3d-tv: visual comfort. *IEEE Transactions on Broadcasting*, 57(2):335–346.
- Thompson, W. B., Willemsen, P., Gooch, A. A., Creem-Regehr, S. H., Loomis, J. M., and Beall, A. C. (2004). Does the quality of the computer graphics matter when judging distances in visually immersive environments? *Presence: Teleoperators & Virtual Environments*, 13(5):560–571.
- Todorovic, D. and Toskovic, O. (2012). The dependence of the perception of distance on the height of the observer’s vantage point. *Journal of Vision*, 12(9):900–900.
- Tresilian, J. R., Mon-Williams, M., and Kelly, B. M. (1999). Increasing confidence in vergence as a cue to distance. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 266(1414):39–44.
- Vienne, C., Masfrand, S., Bourdin, C., and Vercher, J.-L. (2020). Depth perception in virtual reality systems: Effect of screen distance, environment richness and display factors. *IEEE Access*, 8:29099–29110.
- Vienne, C., Plantier, J., Neveu, P., and Priot, A.-E. (2018). (disparity-driven) accommodation response contributes to perceived depth. *Frontiers in neuroscience*, 12:973.
- Viguiier, A., Clement, G., and Trotter, Y. (2001). Distance perception within near visual space. *Perception*, 30(1):115–124.
- Waller, D. and Richardson, A. R. (2008). Correcting distance estimates by interacting with immersive virtual environments: Effects of task and available sensory information. *Journal of Experimental Psychology: Applied*, 14(1):61.
- Wartell, Z. J. (2002). Stereoscopic head-tracked displays: Analysis and development of display algorithms. Technical report, Georgia Institute of Technology.
- Watt, S. J., Akeley, K., Ernst, M. O., and Banks, M. S. (2005). Focus cues affect perceived depth. *Journal of vision*, 5(10):7–7.
- Willemsen, P., Colton, M. B., Creem-Regehr, S. H., and Thompson, W. B. (2004). The effects of head-mounted display mechanics on distance judgments in virtual environments. In *Proceedings of the 1st Symposium on Applied perception in graphics and visualization*, pages 35–38. ACM.
- Willemsen, P., Colton, M. B., Creem-Regehr, S. H., and Thompson, W. B. (2009). The effects of head-mounted display mechanical properties and field of view on distance judgments in virtual environments. *ACM Transactions on Applied Perception (TAP)*, 6(2):8.
- Willemsen, P., Gooch, A. A., Thompson, W. B., and Creem-Regehr, S. H. (2008). Effects of stereo viewing conditions on distance perception in virtual environments. *Presence: Teleoperators and Virtual Environments*, 17(1):91–101.

- Williams, B., Rasor, T., and Narasimham, G. (2009). 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*, pages 7–10. ACM.
- Wired (2015). *A Pixar-Inspired Projector That Beams VR Into an Entire Room* - <https://www.wired.com/2015/02/pixar-inspired-projector-beams-vr-entire-room/>.
- Witmer, B. G. and Kline, P. B. (1998). Judging perceived and traversed distance in virtual environments. *Presence*, 7(2):144–167.
- Witmer, B. G. and Sadowski Jr, W. J. (1998). Nonvisually guided locomotion to a previously viewed target in real and virtual environments. *Human factors*, 40(3):478–488.
- Witt, J. K., Linkenauger, S. A., and Wickens, C. (2016). Action-specific effects in perception and their potential applications. *Journal of Applied Research in Memory and Cognition*, 5(1):69–76.
- Witt, J. K. and Proffitt, D. R. (2008). Action-specific influences on distance perception: a role for motor simulation. *Journal of experimental psychology: Human perception and performance*, 34(6):1479.
- Witt, J. K., Proffitt, D. R., and Epstein, W. (2004). Perceiving distance: A role of effort and intent. *Perception*, 33(5):577–590.
- Witt, J. K., Stefanucci, J. K., Riener, C. R., and Proffitt, D. R. (2007). Seeing beyond the target: Environmental context affects distance perception. *Perception*, 36(12):1752–1768.
- Wright, R. H. (1995). Virtual reality psychophysics: Forward and lateral distance, height, and speed perceptions with a wide-angle helmet display. Technical report, ARMY RESEARCH INST FOR THE BEHAVIORAL AND SOCIAL SCIENCES ALEXANDRIA VA.
- Wright, W. G. (2014). Using virtual reality to augment perception, enhance sensorimotor adaptation, and change our minds. *Frontiers in systems neuroscience*, 8:56.
- Zhang, R., Nordman, A., Walker, J., and Kuhl, S. A. (2012). Minification affects verbal-and action-based distance judgments differently in head-mounted displays. *ACM Transactions on Applied Perception (TAP)*, 9(3):14.

BIBLIOGRAPHY

Résumé : Nous proposons un cadre de référence pour influencer la perception spatiale dans des écrans de réalité virtuelle (RV). Il est difficile, en réalité virtuelle, d'obtenir une perception spatiale précise en raison du phénomène de compression de la profondeur, qui donne lieu à une sous-estimation systématique des distances. La solution de ce problème est complexe car la technologie actuelle ne permet pas de simuler, dans des conditions raisonnables de temps et de complexité, la perfection de l'œil humain. Le problème est d'autant plus compliqué qu'il existe une grande variété de systèmes d'affichage. Il est vrai que différentes techniques ont été proposées pour améliorer la perception à distance, mais une solution mise en place pour un écran peut ne pas fonctionner pour un autre. En ce sens, nous proposons un cadre de référence pour améliorer la perception spatiale en réalité virtuelle. Ce cadre de référence prend en considération la variété des facteurs d'affichage qui affectent les performances pour construire un "prédicteur de performances linéaire", une fonction de régression linéaire qui décrit le phénomène de compression de la profondeur à l'écran. Ce prédicteur peut ensuite être utilisé pour appliquer une technique de projection alternative qui exerce une influence positive sur la perception spatiale. Par conséquent, le cadre de référence peut être utilisé pour améliorer la perception des distances en réalité virtuelle.

Mots clés : perception spatiale,réalité virtuelle,perception de la distance,perspective

Abstract : We propose a framework to influence spatial perception in virtual reality displays. Accurate spatial perception in virtual reality (VR) is challenging due to the phenomenon of depth compression, where distances are underestimated. Solving this issue is complex due to the impossibility with the current technology of simulating, under a reasonable amount of time and complexity, the perfectness of the human eye. The issue becomes more challenging due to the explosion of system designs available in the market. Although some techniques have been proposed to improve the perception of distances, a solution implemented for a display is likely not suitable for others. In this sense, we propose a framework to characterizing and influencing spatial perception in VR. The framework takes into consideration the variety of factors that cause differences in performance to build a "linear performance predictor", a linear regression function that describes the depth compression phenomenon on the display. Then, this predictor can be used to apply an alternative projection technique that influences positively spatial perception by playing with the laws of perspective. Thus, the framework can be used to improve the perception of distances in VR.

Keywords : virtual reality, spatial perception, distance perception,perspective