Understanding and Improving the Quality of Experience in 3D media perception: Accommodation/Vergence conflict in Stereopsis

Cyril Vienne

Université Paris Descartes

Ecole doctorale 261 « Cognition, Comportements, Conduites Humaines »

Laboratoire de Psychologie de la Perception / Equipe Vision

Understanding and Improving the Quality of Experience in 3D media perception

Accommodation/Vergence conflict in Stereopsis

Par Cyril Vienne

Thèse de doctorat de Psychologie

Supervisée par Pascal Mamassian et Laurent Blondé

Présentée et soutenue publiquement le 27 Mars 2013

Devant un jury composé de :

Pr Paul Hibbard, Université de Saint-Andrews (UK) (Rapporteur)

Dr Paul Warren, Université de Manchester (UK) (Rapporteur)

Pr Jean-Louis de Bougrenet de la Tocnaye, Telecom Bretagne (Président)

Dr Pascal Mamassian, Université Paris Descartes (Directeur)

Dr Laurent Blondé, Technicolor (Co-Directeur)

Dr Mark Wexler, Université Paris Descartes

Résumé (français) :

Les technologies de relief tridimensionnel (3D) ont récemment reçu un nouvel attrait. Les raisons potentielles sont que ces technologies peuvent fournir une expérience visuelle plus riche et plus immersive. En effet, un observateur peut extraire les disparités binoculaires présentées entre les vues gauche et droite, pour retrouver plus efficacement la profondeur de la scène visuelle observée, et ainsi, trouver une nouvelle dimension dans le contenu. Cependant, tandis que la valeur de la profondeur est plutôt bien appréciée, un certain nombre de problèmes qui impactent la qualité de l'expérience dans les représentations 3D ont été identifiés. L'objective de cette thèse est d'étudier les principaux facteurs qui affectent la qualité de l'expérience en stéréoscopie dans le but de fournir des méthodes qui pourraient améliorer l'utilisation des systèmes stéréoscopiques. Trois aspects majeurs de la qualité de l'expérience sont adressés : (1) les sources et causes de la fatigue visuelle, (2) les distorsions perceptives et (3) l'amélioration de la qualité de l'expérience en 3D au travers de l'adaptation du contenu visuel. Pour étudier la fatigue visuelle, les mouvements de vergence étaient mesurés à la fois avec un écran 3D et avec un système à double écran qui permettaient la présentation de stimuli avec les informations de disparité et de flou présentés en congruence comme en incongruence. L'effet de la stéréoscopie sur les mouvements de vergence a été étudié dans le but de tester si la mesure oculaire pouvait être utilisée comme indicateur de fatigue visuelle. Le sujet suivant étudiait la consistance de la perception des formes 3D stéréoscopiques en fonction de distances virtuelles induites par la disparité et par le signal d'accommodation. Le rôle de la taille de la pupille et de la profondeur de champ en stéréoscopie étaient étudiés par la manipulation de la taille de la pupille avec deux conditions d'illumination contrôlée. Finalement, l'amélioration de la perception de la forme 3D est questionnée au travers de l'adaptation du contenu visuel en fonction de la mesure de seuils perceptifs individuels pour des stimuli se déplaçant en profondeur.

Mots clés (français) : perception, 3D, stereoscopy, qualité d'expérience, accommodation, convergence, mouvements oculaires, psychophysique

Title: Understanding and improving the Quality of Experience in 3D media perception

Abstract:

Stereoscopic 3-Dimensional (S3D) technology has recently received growing attraction, potentially because it provides a more informative and more immersive visual experience. Indeed, the viewer may extract the binocular disparities displayed between the left and the right views, more efficiently retrieve the depth of the observed visual scene, and thus, give visual content another dimension. However, while the additional value of depth is rather appreciated, a number of problems have been raised that impact the Quality of Experience (QoE) in S3D representations. The objective of this thesis is to investigate the main factors affecting QoE in stereopsis in order to provide guidelines towards the improvement and further use of stereoscopic systems. Three main aspects of QoE in S3D are addressed: (1) the sources and causes of visual fatigue, (2) the perceptual distortions arising in S3D and, (3) the improvement of S3D QoE through content adaptation. To study visual fatigue in S3D, vergence eye movements were measured both in S3D display and in dual-screen display that enables the presentation of matched disparity and defocus stimuli. The effect of stereopsis on vergence movements was studied so as to test whether vergence tracking can be used as indicator of visual fatigue. The next topic investigated the consistency in stereoscopic 3D shape perception as a function of vergence distance and accommodation distance. The role of the pupil size and the depth of focus in S3D were evaluated by manipulating the pupil aperture with two controlled lighting conditions. Finally, the improvement of 3D shape perception is addressed through content adaptation according to individual perception thresholds measurement for motion-in-depth stimuli.

Keywords: perception, 3D, stereopsis, quality of experience, accommodation, vergence, eye movements,psychophysics

To my son,

Remerciements

I am very indebted to my supervisor Dr. Laurent Blondé. First, because he offered me the opportunity to work with him and others engineers in the research lab in Technicolor. Second, he always trusted me in the advance of my research and provided me a great support and guidance that were very helpful all along the PhD. I think he is the person who taught me the most during these three last years. I know now that, more than a supervisor, he is a friend.

I would like to express my sincere gratitude to Dr. Pascal Mamassian. Your valuable guidance and scientific suggestions always helped me in my work. I would also like to thank you for your help in administrative steps and communication with the university staff.

These three years in Technicolor would have been absolutely different without all the great people I have met, colleagues and friends will recognize themselves in these words.

I would like to thank Mélanie Reversat for the support for helping me in the organization of the defense procedure.

I would like to thank Cécile and Quan for precious help in correcting some spelling details in the introduction.

Finally I would like to thank my family, my wife Alice and my son Stanislas.

Table des matières

REMERCIEMENTS 6

INTRODUCTION 10

CHAPTER 1: INTRODUCTION 11

1.1 VISUAL INFORMATION FOR DEPTH PERCEPTION IN HUMAN VISION 11
1.1.1 THE SOURCES OF INFORMATION 11
1.1.2 DEPTH FROM MONOCULAR AND BINOCULAR INFORMATION 13
1.1.3 BINOCULAR VISION & CORRESPONDENCE: GEOMETRY OF PERCEIVED DEPTH 16
1.1.4 AMBIGUITY OF DISPARITY INFORMATION & THE NEED FOR SCALING 20
1.1.5 CUES FOR SCALING HORIZONTAL BINOCULAR DISPARITY 21
1.1.6 INTERACTIONS BETWEEN DEPTH CUES: BRIEF OVERVIEW 23
1.1.7 SUMMARY: FROM MULTIPLE CUES TO ONE PERCEPT 25
1.2 PHYSIOLOGY OF DEPTH PERCEPTION 26
1.2.1 THE OCULOMOTOR SYSTEM: ELEMENTS AND RESPONSES 26
1.2.2 MODELING OF ACCOMMODATION AND VERGENCE ACTIVITY 34
1.2.3 THE OCULOMOTOR DYNAMICS 37
1.2.4 ADAPTIVE OCULOMOTRICITY AND DEPTH PERCEPTION 39
1.2.5 SUMMARY: ACCOMMODATION, VERGENCE AND PUPIL IN STEREOSCOPIC DISPLAYS 41
1.3 STEREOPSIS 41
1.3.1 BASIC PRINCIPLE 41
1.3.2 STEREOSCOPIC SYSTEMS 43
1.3.3 SOME EVIDENCES FOR NEURAL ORGANIZATION 45
1.3.4 LIMITATIONS: STEREOSCOPIC DISTORTIONS 46
1.3.5 THE STEREOSCOPIC DEMAND 48
1.3.6 LIMITATIONS: VISUAL FATIGUE & DISCOMFORT 50
1.3.7 SUMMARY: DISAPPOINTMENT OF STEREOPSIS, ANY CHANCE FOR IMPROVING? 54
1.4 THE QUALITY OF EXPERIENCE IN 3D MEDIA PERCEPTION 56
1.4.1 UNDERSTANDING VISUAL DISCOMFORT AND FATIGUE 58
1.4.2 EXPLAINING PERCEPTUAL DISTORTIONS OCCURRING IN STEREOPSIS 59
1.4.3 IMPROVING S3D QUALITY OF EXPERIENCE 61
1.5 OVERVIEW OF SUBSEQUENT CHAPTERS 62

CHAPTER 2: EXPLORATORY STUDY: VISUAL FATIGUE, VERGENCE EYE MOVEMENTS AND PERCEPTUAL BIAS OF STEREOSCOPIC 3D SHAPE 63

2.1 INTRODUCTION 63
2.2 MATERIALS AND METHODS 67
2.2.1 OBSERVERS 67
2.2.2 APPARATUS 67
2.2.3 STIMULI 68
2.2.4 PROCEDURE 70
2.3 RESULTS 71
2.3.1 SUBJECTIVE EXPERIENCE 71
2.3.2 PERCEPTUAL ESTIMATION 71

2.3.3 Heterophoria and Fixation Disparity 75
2.3.4 Vergence reaction to disparity-step 76
2.4 Discussion 80
2.5 Conclusion 86

Chapter 3: Ocular vergence in stereoscopic displays 89

3.1 Introduction 89
3.2 Experiment 1 92
3.2.1 Method 93
3.2.2 Results 96
3.2.3 Discussion 100
3.3 Experiment 2 101
3.3.1 Method 102
3.3.2 Results 104
3.3.3 Discussion 108
3.3.4 Analysis of vergence position difference 110
3.4 General Discussion 112

Chapter 4. An attempt to increase the depth-of-focus of the eyes for reducing discomfort in stereoscopic displays 115

4.1 Introduction 115
4.2 Method 118
4.2.1 Observers 118
4.2.2 Apparatus 118
4.2.3 Stimuli 119
4.2.4 Procedure 120
4.3 Results 120
4.4 Discussion 122

Chapter 5. The weight of accommodation in 3D shape perception with light-modulated depth-of-focus 126

5.1 Introduction 126
5.2. Experiment 1 132
5.2.1 Method 132
5.2.2 Results 136
5.2.3 Discussion 140
5.3 Experiment 2 142
5.3.1 Method 142
5.3.3 Discussion 147
5.4. General Discussion 150
5.5. Conclusion 152

CHAPTER 6. PERCEPTUAL BIAS IN STEREOPSIS & OPTICAL MODELS OF THE EYE 155

6.1 INTRODUCTION 155
6.2 ORIGIN OF THE SCALING FACTOR 156
6.2.1 OPTICAL EFFECT OF ACCOMMODATION AND THE "ZOOM LENS" HYPOTHESIS 156
6.2.2 ACCOMMODATION CONTRIBUTION FROM SENSORIMOTOR FEEDBACK 159
6.3 CONCLUSION 160

CHAPTER 7. DISTORTION FROM STEREOPSIS IN MOTION-IN-DEPTH WITH LINEAR PERSPECTIVE ..162

7.1 INTRODUCTION 162
7.2. EXPERIMENT 1 165
7.2.1 METHOD 165
7.2.2 RESULTS 169
7.2.3 DISCUSSION 171
7.3. EXPERIMENT 2 173
7.3.1. METHOD 175
7.3.2 RESULTS 177
7.3.3 DISCUSSION 180
7.4. GENERAL DISCUSSION 181

CHAPTER 8. CONCLUSION 184

8.1 VISUAL FATIGUE AND DISCOMFORT IN STEREOSCOPIC DISPLAYS 185
8.2 PERCEPTUAL DISTORTIONS IN STEREOPSIS WITH VIRTUAL VERGENCE DISTANCES 187
8.3 ON THE POSSIBILITY TO ADJUST 3D CONTENT ON THE BASIS OF PERCEPTUAL THRESHOLDS 188
8.4 UNDERSTANDING & IMPROVING QOE IN STEREOSCOPIC DISPLAYS 190

REFERENCES 192

APPENDICES 205

Introduction

The new wave of stereoscopic movies has stimulated a novel interest regarding the understanding of stereopsis. The most interesting contribution of stereoscopic movies is certainly the introduction and, somewhat trivialization of motion in depth in natural visual scene rendering (e.g. the famous commercial for fruit gums). In stereoscopic movies, motion-in-depth is reproduced on the basis of the combination of stereoscopic depth and simulated motion. These two techniques of simulation have been invented independently in the middle of the 19th century. This is only with the advance of motion pictures that the first stereoscopic movies are born in the beginning of the 20th century.

A second contribution of this new generation is the advance of technology. The new buzz of stereoscopic films in theater has strongly benefited from the disappearance of anaglyphs systems in spite of the persisting need to wear glasses. This enthusiasm had led the entertainment industry to transfer the stereoscopic cinema movies to formats adapted to 3DTV. Many modern TVs are now 3D compatible, even though the available number of broadcasting 3D media did not really extend.

At the beginning, this thesis work was thus part of a 3DTV project. This explained why we used 3DTV instead of stereoscope and why we placed observers at viewing distances larger than those of previous studies on stereoscopic depth perception. As the title suggests, the objective of this thesis was to improve the understanding of the specificities of stereoscopic perception and systems that may impact the visual experience. From this better understanding, it will thus possible to derive the guidelines that may improve the stereoscopic experience.

This thesis is a set of reports (sometimes of work submitted to scientific journals) preceded by an introduction which objective is to introduce the key concepts for the understanding of human factors involved in stereopsis and to explain what psychology as a science may do for improving the quality of experience in stereoscopic displays. A conclusion is proposed at the end of this thesis to summarize the findings and their possible use in stereoscopic imaging systems.

Chapter 1: Introduction

1.1 Visual information for depth perception in human vision

How humans see and understand their visual world and, especially, what information is needed to estimate the metrics used for visual perception is of major concern for who wants to address the quality of experience in 3D imaging displays. In order to understand perception, it is worth describing the visual environment and its constituents. This introduction part will describe the information available within the environment, what is monocular or binocular, the viewing geometry and, the special case of binocular disparity.

1.1.1 The sources of information

Information for perception is of pattern of stimulation from sensory input that is demonstrably used by an observer (Cutting, 1991). In this, a distinction can be made between the physics of natural situations and the pattern of stimulation *per se*. Koffka (1935) introduced the classical formal distinction between the energy emitted by or reflected from some object in the environment – the distal stimulus – and, the energy that reaches the senses – the proximal stimulus. The act of perceiving can thus be defined as the process of inferring the properties of the distal stimulus on the basis of the proximal stimulus (Koffka, 1935). Perception appears thus strongly based on the information from the structure of the environment. In nature, this flow of information may be considered as rich, organized and meaningful (Gibson, 1979).

The information related to the spatial extent of the surrounding world is certainly one of the most relevant for the human visual system. The potential information that is available for the visual system is often labeled a “cue”. This term is used in literature to make mention of the available information in the surrounding world, but implicitly assumes that the observer has prior knowledge about the features he/she encounters in his/her environment (Cutting and Vishton, 1995). According to Woodworth (1938) the cue “is a stimulus which serves as a sign or signal of something else, the connection having previously been learned”. For Brunswik (1956), a cue consists in a visual pattern or set of visual relations about which one have obtained experience. Each proximal depth cue is associated with a particular distal arrangement of objects whose association occurs with a given probability (Cutting, 1998). More specifically, Ernst and Bülthoff (2004) characterize the cue as “any sensory information that gives rise to a sensory estimate” which implies that a cue is a part of the proximal stimulus that has already been identified and pre-processed. It is important to note this term may be interchangeably used with others like signal, information, source of information and even *clues*.

Among the flow of information, the most important way by which humans obtain knowledge about the world is vision. As multiple cues can be projected onto the retina, the visual system can also make use of non-visual information, e.g., from oculomotor activity; this consists in the distinction between retinal and extra-retinal information. A number of visual information can be listed to account for visual perception of depth in three-dimensional space. For instance, Cutting and Vishton (1995) described several sources of information among others that are obviously useful for understanding layout and shape in the environment; their efficacy in mediating distance information are for many of them distance-dependent (Cutting, 1997). However, all of these cues do not provide the same sampling unit when sensed by the visual system and, therefore some provide ordinal information while other yields metric information about the world; these values are either relative or absolute. The redundancy of such sources of information provides the way by which individuals can efficiently interact within the environment.

It is broadly acknowledged that past experience can influence present perceptual experience; i.e., perceiving is also remembering (Gibson, 1979). As a matter of fact, prior knowledge or assumptions about the environment can influence the perceptual solution in various cases (Mamassian, 2006), i.e., when observers have to judge uncertain information or even when they have to combine multiple cues available in the rich environment (Hibbard et al., 2012). Hence, the perceptual systems enable individuals to see their surrounding world as stable, although the sensory information reaching the senses might be incomplete, uncertain and constantly changing. This is specifically the case when perceiving drawings and pictures, for example drawings on a two-dimensional flat surface that has to convey information about the three-dimensional structure (see Figure 1.1).

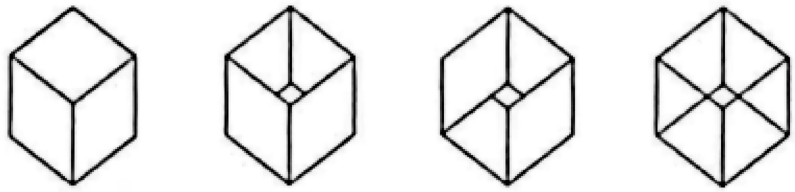

Figure 1.1: Depiction of four different cubes where occlusion is specified by different modes of intersection of the same line segments. From left to right: (1) Opaque surfaces cube, (2) Opaque sheets of a hollow body, (3) Anomalous occlusion leading to contradictory impression and, (4) Transparent cube leading to ambiguous impression due to reversibility. Figure from Gibson (1979, p290).

1.1.2 Depth from monocular and binocular information

Humans have frontally located eyes allowing the acute perception of shapes and object positions without requiring head or body movements – contrary to motion parallax. This frontal binocular information consists in two slightly different patterns of stimulation that lead to two overlapping points of view of the surrounding world. In that, information from the two eyes gives rise to both mismatched and matched optical information. Thus, binocularity provides both disparity and concordance of the visual information to the human visual system (Jones & Lee, 1981). In complement, even though it can be projected onto the retina of both eyes, there is monocular information available for depth perception. However, when speaking about monocular information, one used to think about the visual information from one eye. Indeed, there are numerous monocular sources of information that contribute to depth perception.

The most basic source of depth information is conveyed by the retinal image size that physically changes as the distance of the object changes from the observer. Using playing-cards of various sizes, Ittelson (1951) showed that relative size can operate as a cue to distance; any change in the assumed-size can be perceived as a change of apparent distance. As a matter of fact, the more the object becomes distant from the observer, the more the retinal size of the object of interest appears small. As a consequence, retinal image size becomes a relative cue to distance from the change in size of a same object or when two identical objects project different retinal image sizes. However, information from retinal size is strongly related to information of familiar size because objects of particular size often belong to a category of common objects. Also, relative density is a powerful cue to depth in the way that the projected number of similar textures or patches per solid visual angle varies progressively with distance (Cutting, 1997). This is a particular case of texture gradients that Gibson (1950) was the first to notice as providing visual information about surface orientation and depth. At close viewing distances, the blur present in the retinal image due to defocus can play a role in judging ordinal depth relation (e.g., Hoffman & Banks, 2010; Mather, 1996) but its efficient range of distances is very limited because the depth-of-field rapidly increases with distance. Blur can however be used in photographs or pictures so as to create potent illusion of depth as for example in the tilt-shift illusion, where pictures of natural scenes are perceived as miniature models. In addition, extra-retinal information from the activity of the ciliary muscles controlling the eye-lens accommodation could also yield ordinal information about distance for a range of distances between 10 cm (the punctum proximum) to 2 or 3 meters (Mon-Williams & Tresilian, 2000; Fisher & Ciuffreda, 1988) because only in this range the changes in accommodation lens are significant enough. Another monocular source of information that varies with distance is the height in the visual field (Cutting, 1997). Height in the visual field yields information about distance because objects further away are generally higher in the visual field. This refers specifically to the relation among the bases of objects in the environment as projected to the eye, moving from the bottom of the visual field to the top. The utility of such information is questionable as it depends on the presence of a ground plane. However it is apparently used in pictures perception (Dunn, Gray & Thompson, 1965). Generally at far distance, aerial perspective can also play a role because of the apparent reduction of contrast due to solar light scattering by atmospheric molecules or particles on the optical path to the object. Two final strong monocular sources of information about depth relation are occlusion and motion parallax. Occlusion (or interposition) appears when one object partially hides another from view and, therein only provides ordinal information but over a wide range of distances (Chapanis & McCleary, 1955). Motion parallax is the kind of observer movement that induces a corresponding movement of the image onto the retina. Generally in case of horizontal translation of the observer (leftward or rightward), the retinal motion for objects before the fixation point is opposed to that of objects behind the fixation point. In addition, there is a relationship between the retinal motion velocity and the distance of the object in the visual field (Gibson, 1950).

As it can be seen from this non comprehensive list, a lot of monocular cues for perceiving depth are available in the flow of information. To all these monocular cues must be added two types of potential binocular information, vergence and disparities. The former consists in the extra-retinal signal provided by the counter-rotational activity of the two eyes in bringing the gaze axes to the same fixation point. The measure of vergence is done through the angle between the visual axes of the two eyes; this angle increases for small viewing distance and decreases toward infinity. Contribution of vergence as extra-retinal cue was shown earlier (e.g., Tresilian, Mon-Williams and Kelly, 1999; Gogel, 1961) and its involvement could be rather precise for distances up to 2 meters (Gogel, 1961). The second class of binocular cues refers to binocular disparities. They are the slight horizontal and vertical differences that arise through the projection of the perspective of two different points of view. Figure 1.2 gives an example of two pictures taken from two slightly different points of view. When presented to the visual system, the two inputs provide both similar and dissimilar information. In addition, it can be seen that a lot of cues are present to render the natural scene depth. The use of binocular disparity in depth perception has been widely confirmed (e.g., Julesz, 1971) and the understanding of how the human visual system extracts the depth relations from two different monocular inputs has fascinated many researchers for a long time (Howard, 2002). The geometry of horizontal disparity is thus discussed in the following subsection.

Figure 1.2: How two pictures from slightly different points of view (upper left and right views) give binocular disparity. Upper left and right pictures were taken with a camera baseline of 64 mm, they are here arranged for cross fusion (i.e. left image for right eye and right image for left eye). The lower image is a representation of the two views by interlacing them line-by-line; this method can be used to display 3D content on a polarized display such

as the one used in this thesis. Acknowledgments to Jean-Jacques Sacré for the photographs of Châteaugiron (Brittany, France).

1.1.3 Binocular vision & correspondence: geometry of perceived depth

When we look at a particular place in the environment, both eyes receive optical information through the pupil reaching the retinas. This optical information is formed of photons, which before arriving on the retina will cross a thick optical system, namely a “thick dual lens system” formed by the cornea and the crystalline-lens. For that reason, the imaging lens system produces an anterior nodal point and a posterior nodal point on the optical axis. The nodal point is the location through which passes all the straight lines that join points in the object plane to their corresponding points in the image plane on the retina (Howard and Rogers, 1995). These lines are labeled visual lines when intersecting the center of the fovea (central region of the retina where visual acuity is the highest) or optical lines when meeting the optical centre of the eye. This implies that the visual axis in focus and the optical axis do not coincide (see Figure 1.2 below).

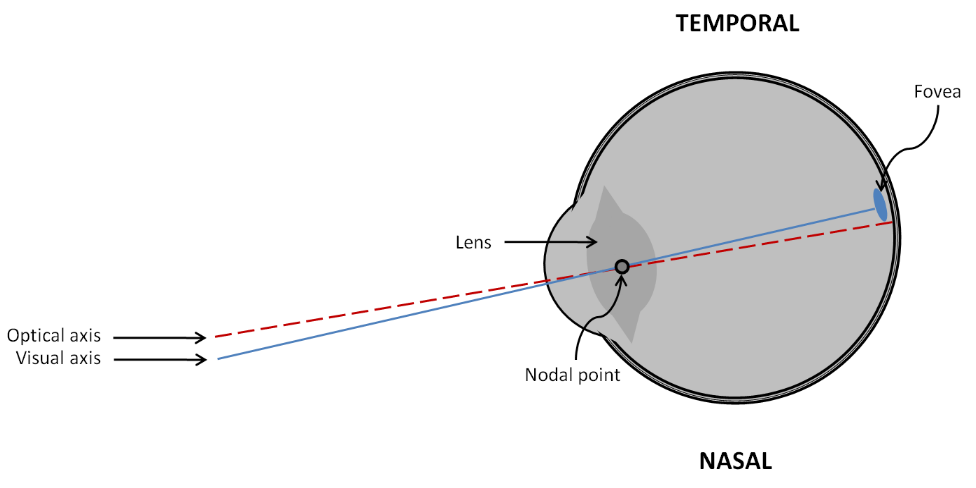

Figure 1.2: Schematic representation of the projection of the optical axis and the visual axis in the human eye. The angle between the visual axis and the optical axis is about 5° on average. The red dashed line represents the optical axis and the blue line represents the visual line projecting in the center of the fovea. The fovea is slightly shifted as compared to the projection of the optical center of the eye on the retina.

When an observer converges forward to a fixation point P, the vergence angle depends on both the fixation distance and the inter-ocular separation – on average across observers about 64 mm. When the convergence angle is equal to zero, meaning that the lines of sight are parallel, both eyes supposedly receive the same optical information; the retinal images are similar. Now, increasing the vergence angle by fixating an object located nearer than infinity will increase the differences between the two retinal images. Differences between the retinal images of the left and the right eyes are called binocular disparities. Binocular disparities are described in three types: horizontal disparity, vertical disparity and cyclo-disparity. For each point in the retina of one eye, there is a corresponding point in the retina of the other eye. Specifically, corresponding points in the two eyes are pairs of points located in the binocular visual field that are seen in the same direction (Hillis & Banks, 2001). Hence, images in both eyes falling on corresponding points have zero binocular disparity. There is a combination of points in the binocular visual field that give rise to corresponding points in both eyes and has no disparity. This zone is the Vieth-Müller circle – also called the theoretical horopter – passing through the fixation point and the nodal points of the eyes (see Figure 1.3 just below). Any points located out of this zone will produce disparity on the retinas and the object will be seen either at a different depth or seen double if located away from this zone – objects perceived as double refer to the term “diplopia”.

The Vieth-Müller circle symbolizes the theoretical horopter that can be computed as a function of the geometry of the eyes and a given viewing distance. The empirical horopter can be measured according to different methods in laboratory. For instance, Hering (1879) estimated empirical horopters using a technique called the apparent fronto-parallel plane horopter method. The procedure is as follows: the observer has to keep fixating on the same point and align a number of stimuli on either side of the fixation point so that they all are in a plane parallel to the subject’s face, i.e. a fronto-parallel plane. Other methods have been developed since then (e.g., the diplopia based-horopter, stereoacuity horopter and nonius horopter). The most accurate method is the nonius horopter that is based on the Vernier technique through dichoptic vision (e.g., using a stereoscope). Here the upper half part of the line is seen by one eye while the other sees the lower half part. While keeping the same fixation point, the task is to align the two lines without fusion (Ogle, 1964).

An object located nearer to an observer than the horopter (see example in Figure 1.3, point P4) creates crossed disparities in the two eyes because the visual lines intersect in front of the horopter. An object located farther than the horopter produces uncrossed disparities in the two eyes because the visual lines cross this time beyond the horopter. The angular disparity produced by a point P when converging at a fixation point F can be calculated by the difference between the convergent angle at F and the convergent angle at P where $\eta = \alpha_F - \alpha_P$

(see Figure 1.4). By convention, in this document, the angular binocular disparity is negative for crossed disparity and positive for uncrossed disparity.

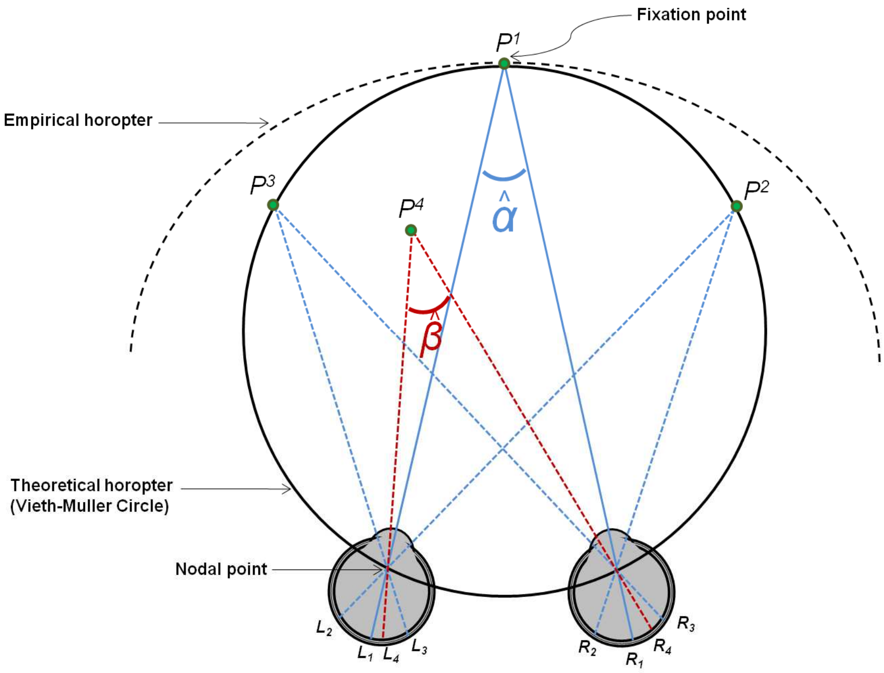

Figure 1.3: Schematic representation of the horizontal horopter. Here the two eyes converge on P^1 that is "fovealized". Fixation points P^1, P^2 and P^3 have zero retinal disparity as visual lines fall on corresponding areas in both retinas; these three points have identical convergence angles equal to α. P^4 being not on the horopter generates disparity in the two eyes; the angle of convergence is larger than the one of P^1, P^2 and P^3.

Retinal disparity can also be derived from the inter-ocular difference (*I*), the depth difference between point P and the fixation point F (Δ*d*) and the viewing distance between the observer and point P (*D*) such as:

$$\eta = \frac{I.\Delta d}{D^2 + D.\Delta d} \quad (1.1)$$

in radians (Howard & Rogers, 1995), assuming that P and F are on the sagittal plane and thus have no eccentricity.

One can also derive the following equations so as to reciprocally determine the theoretical depth distance (Adams, 1955):

$D_s = \frac{I.D}{I+s}$ for $D_s<D$ (crossed disparity) and, $D_s = \frac{I.D}{I-s}$ for $D_s>D$ (uncrossed disparity) (1.2);

s being the unsigned shift size (lateral displacement) on an imaged fronto-parallel plane situated at viewing distance *D* (see section about stereoscopy).

The slight difference produced in the retinal images of both eyes when converging can also be characterized in terms of disparity gradients. A disparity gradient G between two points in the binocular visual field is the difference in their angular disparities divided by their separation in visual angle unit. This concept of disparity gradients is particularly interesting because fusion can fail when this gradient exceeds a critical value of about 1 (Trivedi & Lloyd, 1985). Furthermore, the range of disparities that gives rise to object perception with singleness is called Panum's fusional area (Panum 1858). In other words, objects with small enough disparity that can be fused and seen as single fall in Panum's fusional area. Conversely, the points falling outside this area are seen as double, which is referred to as diplopia. The angular extent of Panum's fusional area depends on several factors, including retinal eccentricity, stimulus characteristics and surrounding stimuli (Howard & Rogers, 1995). For example, the limits of fusion have been shown to be influenced by the spatial frequencies of the stimulus; low spatial frequencies providing greater fusion capacities (Schor et al., 1984). Different images falling in Panum's zone fail to provide fusion but instead exhibit binocular rivalry, that is, the alternate perception of the two images.

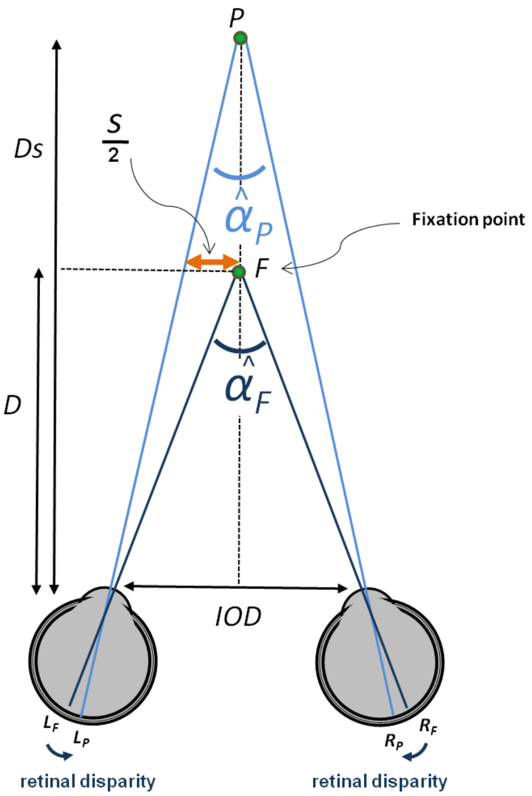

Figure 1.4: Schematic representation of convergence on fixation point F and the point P at disparity η (difference between convergence angle αF and angle αP). D represents the distance from the cornea of the eye to the fixation point F, DS represents the distance to the point P, IOD is the inter-ocular distance and, s is the lateral displacement on the fronto-parallel plane at distance D.

1.1.4 Ambiguity of disparity information & the need for scaling

Considering space perception as Euclidean and therefore that veridicality is possible, the perception of depth from stereopsis depends on viewing geometry and, for that reason is ambiguous when this information is considered alone. Indeed, the horizontal retinal disparity between two points depends on the distance between the observer and the points (Glennerster, Rogers & Bradshaw, 1996). Specifically, the magnitude of horizontal retinal disparity between any pair of points varies approximately inversely with the square of the viewing distance (Kaufman, 1974). Hence, knowing vergence angle and I, one can retrieve depth in 3D space only if the viewing distance that separates the observer from the pair of points is known. In order to see an object in stereopsis, the viewer thus has to determine an observation distance acting as a parameter in disparity disambiguation for depth perception.

According to the pinhole model – simple projection geometry in paraxial approximation –, the depth of a point relative to another one can be calculated using the following formula:

$$d = \frac{\eta . D^2}{I \pm \eta . D} \qquad (1.3)$$

where η is the relative disparity between the pair of points, D is the viewing distance (i.e., accommodation-vergence distance) and I is the inter-ocular distance. Note that we purposefully use the ± sign in the denominator to emphasize the dependence of the value of depth on the location of the simulated point relative to the fixation plane.

For a depth difference of two points located on the same z axis, disparity is given by the following formula:

$$\eta = \frac{I.d}{D(D \pm d)} \qquad (1.4)$$

assuming that depth difference d is the difference between two symmetric vergence angles.

Since d is usually small by comparison to D we get

$$\eta \approx \frac{I.d}{D^2} \qquad (1.5)$$

From equation 1.5, one can see the approximate inverse squared relationship between the disparity of an object and its distance from the observer. Rearranging the equation 1.5 gives:

$$d \approx \frac{\eta . D^2}{I} \qquad (1.6)$$

There it can be seen that in order to estimate the depth difference *d*, the visual system needs information about the viewing distance *D*. The information about the viewing distance can be perceived on the basis of visual cues under some conditions described in the following subsection.

1.1.5 Cues for scaling horizontal binocular disparity

It has been shown that observers can dramatically misperceive depth when judging 3D shape from binocular disparities (e.g., Johnston, 1991). According to the results of Johnston (1991), a disparity-defined hemi-cylinder is perceived circular at intermediate distance (107 cm) but its depth becomes overestimated (elongated) as the viewing distance decreases (53.5 cm) and underestimated (compressed) as the viewing distance increases (214 cm). This observation has been interpreted as the consequence of an incorrect estimate of scaling distance by the observer (Johnston, Cumming & Landy, 1994; Johnston, 1991). When asked to match the depth-to-height or depth-to-width ratio of a hemi-cylinder, participants had to compare the horizontal or vertical dimension (height or width) with the depth information, and both of these dimensions need estimate of the viewing distance to be properly used (Johnston, 1991). There is experimental evidence supporting the fact that retinal size and disparities can be scaled according to the same viewing distance (van Damme & Brenner, 1997).

The information from the state of extra-ocular muscles can theoretically provide an estimate of the vergence angle to the visual system so as to determine the distance separating the observer from the fixation point. There is some experimental evidence that supported this idea, for example when observers were required to judge distance in impoverished environment (Tresilian, Mon-Williams & Kelly, 1999; Rogers & Bradshaw, 1995; Foley, 1980) but, other authors found poor performances when vergence was the unique cue to distance (Brenner & van Damme, 1998; Gogel, 1972). Some authors thus concluded that vergence information, if involved in perceptual estimation, could only play a minor role (Brenner & van Damme, 1998). However, additional investigators have shown that vergence could contribute to distance perception in the absence of other cues for rather small fixation distances (Tresilian, Mon-Williams & Kelly, 1999; Mon-Williams & Tresilian, 1998). The study by Tresilian et al. (1999) also showed a strong contraction bias that had already been observed earlier (e.g., Gogel & Tietz, 1973). This specific distance tendency corresponded to distance overestimation for near target and distance underestimation for far target in reduced-cue environment (Foley, 1980; Gogel & Tietz, 1973), i.e., in the dark (Gogel, 1969). The specific distance tendency could be related to the resting tonic state of vergence (Owens & Leibowitz, 1976) and could have influenced distance estimates in the previous experiments.

The horizontal gradient of vertical disparity could also help observers to establish an estimate of viewing distance (Mayhew & Longuet-Higgins, 1982); this is the variation of vertical disparity with eccentricity. Vertical disparity exists because points that are not lying in the median plane (or sagittal plane) will be closer to one eye than the other, resulting in different vertical retinal projections in both eyes (Howard & Rogers, 1995). The horizontal size ratio (HSR) and the vertical size ratio (VSR) are represented for a surface patch in Figure 1.5. The HSR changes with the tilt of the surface but is ambiguous when considered alone because of the horizontal alignment of the eyes. Considering the VSR, from Figure 1.5 it can be seen that the vertical angular extent projected in the closer eye is larger than that projected in the farther eye. There is empirical evidence that showed that differential perspective (i.e. the pattern of vertical disparity is the consequence of perspective viewing) could be used to scale binocular disparities (e.g., Bradshaw, Glennerster & Rogers, 1996; Rogers & Bradshaw, 1995). Vertical disparity varies approximately linearly with eccentricity and its horizontal gradient decreases with observation distance, thereby explaining how such information may disambiguate the estimate of viewing distance (Gillam & Lawergren, 1983). However, this information becomes significant in situations using large field of view and, therefore requires sufficiently large display (Bradshaw, Glennerster & Rogers, 1996).

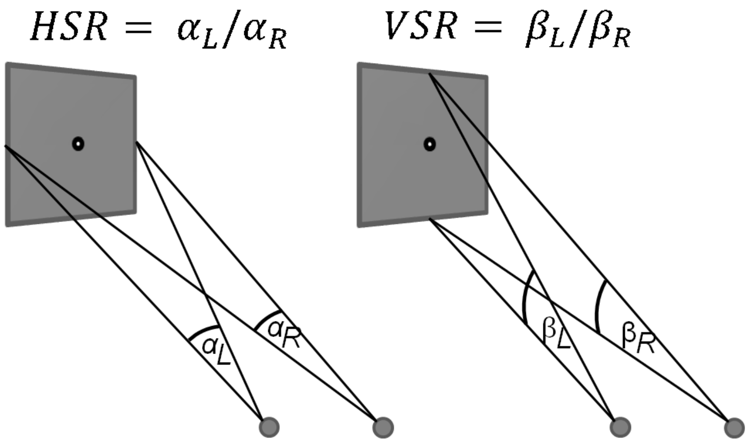

Figure 1.5: This figure represents the horizontal size ratio (HSR) and vertical size ratio (VSR). Angles α_L and α_R are the horizontal angles subtended by the patch for the left and right eyes. The horizontal size ratio (HSR) is defined as α_L/α_R. Angles β_L and β_R are the vertical angles subtended by the surface patch. The vertical size ratio (VSR) is defined as β_L/β_R. If the left eye is closer to the surface than the right eye, then $\beta_L>\beta_R$.

The ability to estimate the viewing distance would thus be dependent on the actual vergence and on the size ratios of horizontal and vertical disparities but, other factors such as some cognitive pictorial cues (familiar size or perspective, O'Leary & Wallach, 1980) and the actual state of accommodation (Watt et al., 2005) could play a role. It is also possible that horizontal disparities become disambiguated by the presence of other cues, i.e., by combining disparities with other depth cues such as texture or motion parallax that scale differently with the viewing distance (Frisby, Buckley, Wishart, Porrill, Garding, & Mayhew, 1995; Johnston, Cumming, & Landy, 1994; Johnston, Cumming, & Parker, 1993; Richards, 1985).

1.1.6 Interactions between depth cues: brief overview

As we have seen in the preceding paragraphs, the visual world includes multiple sources of information and, therefore one question that was raised was: how do perceptual systems deal with this multiplicity? Generally, studies investigating how the visual system combines two or more sources of information present the studied cues both in isolation and in combination. For example, Künnapas (1968) showed that distance judgments were more consistent and accurate when several sources were present as compared to reduced cues situations. There have been several proposed schemes of integration and selection to account for the use of multiple cues in perceiving depth (Howard & Rogers, 1995; Massaro & Friedman, 1990; Bülthoff and Mallot, 1988). For instance, in Howard & Rogers (1995), five types of interaction are described: cue averaging and summation, cue dominance, cue dissociation, cue reinterpretation and cue disambiguation. Cue summation refers to the combination that provides an improvement in depth sensitivity; the more cues are present the better observers are able to perceive depth. Cue averaging reflects the propensity to cancel one depth effect (e.g., binocular disparity) by another opposite depth effect (e.g., motion parallax). Cue dominance is present when two cues are opposed (in conflict) and when one effect of the two is suppressed. Cue dissociation is an interaction where two cues are interpreted as coming from different objects. Cue reinterpretation may arise to resolve contradictory cues by reinterpreting the stimulus context. Finally, cue disambiguation is the case of interaction where one cue comes to disambiguate another. It is worth noting that most empirical studies approached depth cue combination on the basis of cue averaging and summation.

The model described hereunder is a cue combination example, perhaps the most general model of weighted combination (Landy, Maloney, Johnston & Young, 1995). The combination process is also called fusion, and can be modeled as “weak” or “strong” fusion (Clark & Yuille, 1990). Weak fusion is defined as a modular combination based on the computation of separate estimates of depth for each depth cue in isolation, and then averaging the separate depth estimates so as to obtain an overall depth percept. The first major problem with weak fusion is that each depth cue is qualitatively distinct in nature and that all depth cues do not share the same unit. The second is that each depth cue does not confer the same reliability across the visual scene; the weak fusion should change the weights assigned to different cues to reflect their own reliability. Strong fusion refers to combination models in which depth cues estimates are not modular; there can be multiple non-linear interaction between depth cues and these interactions may not obey any rule. Strong fusion models become thus difficult to test because the essential distinction between cues necessary for experimental testing is rather tricky.

Considering the problems of both classes of fusion model, Landy et al. (1995) proposed the modified weak fusion model. In this model, the weak fusion is more flexible in the sense that additional interactions are now possible, such as the use of ancillary cues and the interaction type of “promotion” and “calibration”. Cue promotion is the way by which the conversion to common units is possible; this is the reaction to a second cue used to support the first cue. Cue calibration is described slightly differently in the way that the visual system would “learn” before providing the estimate. The linear combination is now performed on the basis of the weights of each involved cue, based on their reliability (Landy, Maloney, Johnston & Young 1995). This model of weighted average is based on the assumption that each single-cue estimate is unbiased but subject to independent Gaussian noise. The reliability of one cue is equal to the inverse of its estimated variance and the sum of the weights of each involved cue is equal to one (Landy, Maloney, Johnston & Young 1995). The reliability of a cue is thus lower with large variance and therefore such cue contributes less to the final estimate (Hillis, Watt, Landy & Banks, 2004, Landy, Maloney, Johnston & Young 1995). Cue weight depends on stimulus conditions and varies across observers. This model also integrates assumptions about prior knowledge to derive the rules applied to the weighted average (i.e., the depth maps) and cue promotion. Indeed, most sources of information cannot yield alone an accurate measure about depth information because they are either uncertain or ambiguous and, even though they may be disambiguated when considered together, experience that forms a priori knowledge can help explain how in some conditions (i.e., cue conflict) the percept can converge in a specific direction. In this model, Bayesian approaches thus permit the integration of prior knowledge into the estimation process (Mamassian, 2006). The Bayesian approach makes somewhat general predictions about the pattern of cue interaction encountered in case of novelty and, these predictions sometimes account for stimulus conditions (discrepant or contradictory sources of information) that may never occur in natural viewing conditions (Howard & Rogers, 1995). The Bayesian framework is however recognized by the scientific community as being an “elegant method” for explaining cue combination in intra- and extra-modality interaction.

1.1.7 Summary: from multiple cues to one percept

The visual perception of depth is the result of a complex process that can find its input among a vast flow of information called cues. Hence, numerous visual cues can contribute to depth perception. They can participate as independent sources of information or rather just to promote another cue. An example of promotion can be found in the combination of binocular disparity and motion information (Richards, 1985). Depth from disparity and depth from motion information scale differently with distance but can be combined together in order to recover correct 3D structure of the visual object. As a matter of fact, stereo cues are insufficient to provide an estimate of depth as the retinal projections can refer to various angular disparities that can be appreciated only once a fixation distance is estimated. In addition, structure-from-motion information cannot provide enough information for shape perception since each pattern of motion contains at least two possible interpretations due to stimulus reversibility. However, disparity can provide the sign to suppress reversibility and motion can yield the relation between the points to break the scaling distance dependency. Such theoretical considerations are supported by experimental evidence (i.e., Johnston, Cumming & Landy, 1994). Thus it seems that knowing what information is available for depth perception is not sufficient to explain how we can retrieve depth information but, knowing how these sources of information are combined to yield the final percept seems at least as important. It is possible that some of these important sources of information for depth perception can also be non-retinotopic, that is, they stem from the activity of the extra-ocular muscles. In line with this assertion, it is thus mandatory to understand how the physiological systems act in order to efficiently apprehend these significant sources of depth information. As a consequence, the next sub-section is devoted to the understanding of the oculomotor system and its implication in depth perception.

1.2 Physiology of depth perception

Depth perception is not just the outcome of information reaching retina's receptors; it requires the response of each part of the oculomotor system. For any given visual stimulus, there is a potential response of the near oculomotor triad (Schor, 1986; Semmlow & Hung, 1983; Burian & VonNoorden, 1974): accommodation, vergence and pupil responses work together in order to provide the best possible sampling of the information from the world around us. The following part of this chapter is devoted to the understanding of the oculomotor dynamics; components of the oculomotor system, their typical responses, models and limits of those systems will be presented. It is important to note that vergence and accommodation systems both present the same classes of response: reflex, cross-link, proximal and tonic. The reflex component is a way by which the oculomotor system can provide a quick and sustrained response, the cross-links ensure that both accommodation and vergence react in the same way, the proximal enables the observer to implement cognitive aspects in the response (i.e., proximity) and, the tonic component is the minimal muscle innervation ensuring the minimal system tonus under conditions without stimulation. It is Maddox (1893) who proposed this breakdown in order to describe how the oculo-motor responses attain unity in spite of the multiple influences it undergoes.

1.2.1 The oculomotor system: elements and responses

1.2.1.1 The vergence system

Vergence eye movements belong to a particular category of binocular eye movements. These disjunctive eye movements consist in simultaneous inward rotation of both eyes around the vertical axis; both eyes performing rotation in order to bring the visual axes on the object of interest. More specifically, this is foveation, the focusing of the optical information onto the retina. The term vergence can be used either for the angular state taken by the gaze of the eyes as well as for describing binocular eye movements. These opposed eye movements aim at providing a clear and single vision when fixating an object. Horizontal vergence is involved when the observer's eyes fixate a different distance – this is sometimes called convergent or divergent step as function of the sign of the change. Vergence is generally measured while visually stimulating with target vergence of disparity-steps, ramps or pursuit.

Vergence eye movements stem from the activity of six extraocular muscles receiving innervations from three cranial nerves (the abducens nerve, the trochlear nerve and the oculomotor nerve). These muscles are fastened to the eye's sclera on one side and are attached to the bony orbit of the eye on the other side. Through the control of the cranial lower motor neurons, contraction of the muscles produces movement of the eyes within the orbit. Horizontal vergence mainly commands the medial and lateral rectus that, when contracting either pulls the eye towards the nasal side or the temporal side as the result of the antagonistic muscular activity – the first muscle ought to relax whereas the antagonist contracts to perform horizontal eye movements.

Vergence response can be broken down into several components, such as the tonic component, the disparity component, the accommodative component and the proximal component. Different stimuli influence the overall response of the vergence system, namely: retinal disparity, accommodation change and target proximity.

The disparity component of vergence (reflex component) occurs when an observer changes fixation from object A to object B (i.e., fixation points P1 and P4 on Figure 1.3, respectively), i.e., two objects located on two different horopters that do not intersect in Panum's fusional zone. For a long time, disparity vergence was called "fusional component" because of Maddox' theory stating that diplopia was the vergence stimulus and that this mechanism aimed at restoring fusion (Maddox, 1893). Experimental evidences now strongly suggest that this component is driven by retinal disparity and, this is the reason why the name changed to "disparity vergence". This disparity vergence component is also considered to be the outcome of two neural components: the fast preprogrammed transient component and the slow feedback-control sustained component (Semmlow, Yuan & Alvarez, 2002). This dual-mode behavior is supported by an experiment showing that ramp velocities less than 2 $deg.s^{-1}$ provided smooth vergence response, ramp velocities greater than 9 $deg.s^{-1}$ entailed step responses and, intermediate ramp velocities yielded step-ramp and multiple-step responses (Semmlow, Hung & Ciuffreda, 1986). These results suggest that the transient component aims at initiating vergence response and seems to be preprogrammed (inaccurate) and is a response of an open-loop system. This initial response is rather stable, certainly because no feedback intervenes. On the contrary, the sustained component is involved at the end of the movement and is responsible for maintaining fusion within Panum's zone. These two components contribute to performing vergence response precisely whereas it remains a small constant fixation error called fixation disparity (Ogle, 1954).

Disparity vergence usually provides latency of about 100 to 200 ms. Several factors however exist that are involved and can change the temporal dynamics of this component, i.e., direction, initial starting position, age and predictability (Alvarez, Semmlow & Pedronoo, 2005; Heron, Charman, & Schor, 2001; Hung, Zhu, & Ciuffreda, 1997; Semmlow & Wetzel, 1979). For instance, convergence and divergence exhibit different temporal dynamics presumably because they use two different neural pathways (Hung, Zhu, & Ciuffreda, 1997).

In laboratory, disparity vergence is measured through the use of pinholes or with stimuli that do not contain large spatial frequencies (e.g., difference of Gaussian), (Kotulak & Schor, 1987; Tsuetaki & Schor, 1987). Indeed, even though the disparity component of vergence seems to dominate the vergence response, the accommodative component (i.e. the crosslink component) is the potent secondary information for the vergence system (Semmlow & Wetzel, 1979). Accommodative vergence is specifically driven by the blur in the stimulus producing also accommodative responses. As a matter of fact, vergence movements can be observed consecutively to a change in blur stimulation whereas only one eye is stimulated (Hung, Semmlow & Ciuffreda, 1983). Accommodative vergence was however shown to be relatively slower than disparity vergence (Heron et al., 2001; Semmlow & Wetzel, 1979).

Contrary to the reflex component and the crosslink component, proximal vergence is rather “cognitive”. This component is driven by stimuli that can be perceived as nearer or farther in absence of disparity or blur cues (Howard & Rogers, 1995). More specifically, proximal vergence reflects the influence of the perceived distance of the object of interest that is derived from various stimuli related to depth such as, relative size, linear perspective, occlusion, shading and texture gradient, as well as motion parallax, looming and motion in depth (Schor et al., 1992). However, it is important to note that the sensitivity of vergence to these categories of stimuli (also called “spatiotopic”) is generally lower compared to retinotopic stimuli (disparity and blur).

Tonic vergence reflects the minimal innervation of the extra-ocular muscles; this is the state of vergence in the dark or its propensity to return back to this state – also called dark vergence (Howard & Rogers, 1995). This is roughly the mechanism by which the transition from anatomical resting (relaxed muscle state) to physiological resting states (innervated muscle state) is possible. The distance for dark vergence is between 0.62 and 5 m with a mean value of about 1.2 m (Owens & Leibowitz, 1980).

Other authors dissociated voluntary vergence from proximal vergence and, also added the component of “vergence adaptation” to the list of vergence components (i.e., Daum &

McCormack, 2006). According to Patel & Firth (2003), vergence adaptation is an actual part of the vergence system and serves the system to work in a comfortable way. Vergence adaptation is a compensation mechanism for ocular misalignment arising in different gaze positions. This compensation can remain active several minutes after the eyes moved to a new fixation point. The retinotopic stimulus is only indirect. Voluntary vergence is the vergence that one can perform consciously. For example, when trying to fuse a stereogram, one has to adjust vergence in and out to complete fusion. Another example is the act to intentionally cross or uncross the eyes and, temporarily reach diplopia.

1.2.1.2 The accommodative system

Accommodation impacts the way light casts onto the retina. This is the change in optical power of the crystalline lens that is responsible for focusing the light in order to keep a sharp retinal image as fixation distance varies. The function of accommodation is to minimize blur in retinal image for a certain range of distances. The depth of focus is the perceptual tolerance to retinal defocus (Wang & Ciuffreda, 2006). Change in the curvature of the crystalline is achieved through the activity of the intraocular muscles called the ciliary muscles. Under the control of the parasympathetic system, when the ciliary muscles contract, this causes the decrease in zonular tension and, as a result, the lens becomes more convex in order to focus light issued from a closer distance to the observer. Conversely, the sympathetic innervations lead to the flattening of the lens for far vision. As for many thick optical systems, when the crystalline lens shape flattens, the focal length increases and therefore, the optical power of the system decreases. Conversely, accommodation to a near object will decrease the focal length and the optical power of the thick system increases. During steady fixation, the accommodative state of the eye exhibits fluctuation of about 0.3 diopters at a frequency between 1 and 2 Hz (Campbell et al. 1959). The magnitude of these fluctuations increases with target proximity and could be involved in the control of the accommodation response (i.e., Yao et al., 2010).

The primary stimulus for accommodation is out-of-focus blur, even though it is also acknowledged that the accommodation system reacts to chromatic aberration and size change (Kruger & Pola, 1986). While the eye optical system forms an image of the observed scene, blur information is the resultant of optical imaging in front of or behind the retina. Blur depends on the distance of observed points in the scene relative to the point in focus, and on the light beam cross-section as defined by the eye pupil. Blur is perceived when the cross section of a light beam in the retina plane (its point spread function) is larger than the circle of confusion. The circle of confusion is an area of a light sensor (e.g. a small group of cones on the retina) where different incident light rays cannot be separated. Knowing the circle of confusion in the retinal plane and the size and position of the eye pupil, one can define two beams representing the limits between sharpness and blur. Each of these two beams has a distance of smallest cross-section in the object space defining the closest and the farthest sharp plane, respectively in front and behind the point in focus. The distance positions of these two planes define the depth of field, a region where all objects points are seen sharp as their image on the retina is smaller than the circle of confusion. Related to the depth-of-field is the depth-of-focus; the depth-of-field is the dioptrical projection of the depth-of-focus in image space. The depth-of-focus confers the advantage that observers do not have to focus precisely to perceive a sharp image. In other words, the depth-of-focus provides the neural and perceptual tolerance for relatively small focusing error (Wang & Ciuffreda, 2006). Hence, as long as the object of interest remains in the depth-of-field, its retinal image will similarly remain within the depth-of-focus (Wang & Ciuffreda, 2006). The depth-of-field is about 0.30D with a pupil size of 3mm and in optimal viewing conditions (Ciuffreda, 2006). According to Wang & Ciuffreda (2006), several factors may influence the depth-of-focus of the eye. They can be categorized as external – properties of the visual target or test environment (e.g., luminance, contrast, spatial frequency) – and, internal factors – optical and neurological attributes of the observer (e.g., visual acuity, pupil size, age). The depth-of-focus would be at the basis of accommodative lead or lag. These are possible errors when focusing on an object; when instead of focusing right on the plane of the object, the eye is focusing on a point behind it or in front of it, there are respectively the lag and the lead of accommodation (Wang & Ciuffreda, 2006).

The subjective depth-of-field of the human eye can be estimated using a Badal optical system (i.e., Atchison, Charman & Woods, 1997). The Badal principle is as follows, an eye placed at the back focal point of a lens will perceive the angular size of the target independently of its position – and constant – while the target vergence perceived by the eye is proportional to the distance of the target from the front focal plane (Atchison, Bradley & Smith, 1995). Hence, thanks to this specific optical arrangement, it is possible to vary the position of the target and, therefore the defocus blur on the retinal image, without perceiving any change in the relative size of the target. While varying the target distance, the observer may compare the just noticeable difference in defocus of the target as compared to a similar standard (Legge et al., 1987). It is worth noting that the estimate of the subjective depth-of-field is presumably realized for one eye only. The objective measure of the depth-of-field is also possible by using an autorefractor and the prescription of some amount of cyclopegic agent with the Badal system: while changing the target distance and the observer is instructed to keep clear his retinal image, the refractive error of the eye is measured. Any change in the refractive error indicates the limits of the depth-of-field. The comparison between the subjective and objective depth-of-field has shown that the objective depth-of-field is significantly smaller than the subjective one (Yao et al., 2010).

Similar to the vergence system, accommodation can be broken down into four components, such as, phasic accommodation, convergence accommodation, proximal accommodation and tonic accommodation. Different stimuli influence the final response of the accommodation system, namely: retinal blur defocus, vergence change and target proximity.

Similarly to disparity vergence, phasic accommodation – or blur-driven accommodation – is stimulated by a retinotopic signal, namely the blur. The human eye focuses as a result of the control process of retinal image defocus serving as stimulus for the ciliary body (Khosroyani & Hung, 2002). In this sense, accommodation can be seen as a negative feedback control system in which the error is the retinal image defocus, the brain is the controller, and the ciliary body is the system plant (Hung, 1998). As for vergence, accommodation would be directed by two components, the fast preprogrammed transient component and the slow feedback control sustained component (Hung & Ciuffreda, 1988). This dual-mode behavior is supported by an experiment showing that ramp stimuli below 2.5 $D.s^{-1}$ provided a smooth response, whereas beyond this level, step-ramp and staircase-like multi-step responses were observed (Hung and Ciuffreda, 1988). The authors concluded that the dual-mode accommodation could be represented by fast and slow components. The fast component consists in the initial open-loop preprogrammed movement that is responsible for the rapid increase in the accommodative response. The slow component provides the closed-loop movement in the accommodative step response that accounts for the stable and accurate focusing of the eyes due to the feedback system. Blur-driven accommodation exhibits a mean latency of about 300 ms (Campbell & Westheimer, 1960) but can be relatively smaller with target predictability (Ciuffreda, 2006). There are other factors that can change the temporal dynamics of this component, in particular age (Anderson et al., 2010). Disparity-driven accommodation usually provides smaller latency than the one driven by blur (Suryakumar et al., 2007) but remains longer than the one for disparity vergence.

Accommodation also responds to vergence change; this is called vergence accommodation. Hence, the change in horizontal vergence is followed by an appropriate change in accommodation. It is often measured as the change in convergence accommodation per unit change in convergence or CA/C ratio. Vergence accommodation can be measured with a change in vergence in absence of change in blur-driven accommodation. The observer has thus to watch through pinholes during vergence stimulation in order to maximize the depth-of-field so that there is no perceived blur whatever the fixation distance or the accommodative state of the lens. Then the vergence demand can be varied using prisms.

Proximal accommodation is somewhat similar to proximal vergence and, therefore will not be further described here. Proximal accommodation also concerns spatiotopic stimuli and, as a consequence, links the perception of space in a cognitive way with the accommodative response. For instance, accommodative responses have been shown with stimuli like relative size (Alpern, 1958) and linear perspective (Ittleson & Ames, 1950).

Tonic accommodation reflects the minimal innervation of the ciliary muscles; this is the state of accommodation in the dark or its propensity to return back to this state – also called dark focus (Howard & Rogers, 1995). This is the mechanism by which the transition from anatomical resting (relaxed muscle state) to physiological resting states (innervated muscle state) is possible. It is interesting to note that the distance for dark focus is not correlated with dark vergence and is about 0.76m (Owens & Leibowitz, 1980).

1.2.1.3 The pupillary system

The pupil is the circular opening in the iris of the eye and it plays the important role of aperture stop for the optical system of the eye (Charman, 1995). The pupil system controls the amount of light entering in the ocular globe but also the quality of the projected retinal image through its influence on diffraction, aberration and depth-of-focus. Two iris muscular components control the pupil aperture: the circular muscles called the sphincter pupillae and the radial muscles called the dilator pupillae. The first group is responsible for pupil constriction (miosis) while the later is involved in dilation (mydriasis). The iris musculature receives its motor innervation from the autonomic system.

There are three possible pupillary responses defined according to the stimulus: the light reflex, the near response and the emotional response (Andreassi, 2000). The light reflex depends on lighting conditions and produces constriction due to intense light or dilation with dim light. The near reflex results in responses due to change in fixation distance and, specifically the activity of the accommodation-vergence system. A few changes in pupil diameter (often less than 0.5 mm) can also be observed during cognitive processing or in response to emotional/excitation states. Normal pupil diameters range from 1.5 to 9 mm and as said, are strongly dependent on environmental lighting conditions. The latency of pupil response is beyond 200 ms (Andreassi, 2000) with an average value of about 260ms (Kawahata, 1954) whereas longer reaction time can be observed when focus cue is the only available stimulus (Campbell & Westheimer, 1960). Also, pupil response is under a consensual reflex: stimulation of one eye affects both eyes equally.

Hence, the pupil response to light is the predominant and specifically due to the stimulation of retinal receptors (rod, cone and melanopsin ganglion cells) when light is cast onto the retina (Viénot, Bailacq & Rohellec, 2010). Even though different pathways for color and luminance reaction could not be put forward, there is some evidence that the pupillary reflex could be driven by both a tonic and a phasic component (i.e., Young, Han & Wu, 1993).

The near pupil response shows constriction with the subject convergence of the eyes on a near object and dilation with divergence on a far object (Kasthurirangan & Glasser, 2005). The pupillary light reflex follows an “uncomplicated pathway” involving the pretectal area whereas the near response follows a central pathway common to that of the accommodation and convergence reaction (Campbell & Westheimer, 1960). Pupil responses have been shown to be associated with accommodation and dependent on initial aperture (i.e., Semmlow, Hansmann & Stark, 1975), but some studies suggested that in absence of cues other than blur, pupil near responses can be reduced or even absent (Phillips, Winn & Gilmartin, 1992; Stakenburg, 1991) suggesting that pupil near response predominantly reacts to proximity cues. However, the magnitude of pupillary near response is dependent on subject age and apparently absent in young subjects (Schaeffel, Wilhelm & Zrenner, 1993). The absence of response with only accommodation cue in the two previous studies could be explained by the relatively young pool of participants (Mays & Gamlin, 1995).

Pupil size changes, in response to a change in fixation distance, affect the size of the point spread function and change the retinal image quality (Bharadwaj, Wang & Candy, 2011). The depth-of-focus increases with smaller pupil diameter and decreases with larger pupil diameter. It is worth noting that increasing the depth-of-field will reduce the demand on accommodation in order to sustain clear vision (Ward & Charman, 1987). The change in the depth-of-focus (measured objectively with a laser speckle) with pupil size is nonlinear but has been shown to increase significantly for pupil sizes smaller than 3 mm in diameter (Charman

& Whitefoot, 1977). Also, the human eye does not behave as predicted by geometrical optics and, therefore cannot be compared to optical systems such as cameras. On the basis of the work by Stiles & Crawford (1933), Campbell (1957) proposed an empirically determined formula that showed the depth-of-focus of the human eye not decreasing proportionally with increasing pupil aperture. Rather, these observations proved that large pupils behave like smaller ones. The radius of the pupil can be corrected to account for the depth-of-focus of the eye by: $r_c = \sqrt{(1 - \exp(-0.105 r^2)) / 0.105}$, r_c being the corrected radius and r being the apparent pupil radius. However, this effect is minimal for pupil diameter under 4mm.

1.2.1.4 Specificity of accommodation-vergence relationship

Ocular-motor systems have been categorized in two different subgroups, but as it was discussed previously, the cross-coupling between these two groups presents several ways by which accommodation and vergence react. Other examples of cross-linkages embrace the cross-talk between saccade-vergence (Zee, Fitzgibbon, & Optican, 1992) and between horizontal-vertical vergence (Schor, Maxwell, McCandless, & Graf, 2002). The most detailed cross-linkage is certainly the coupling between accommodation and vergence (i.e., Schor, 1992; Hung, Semmlow & Ciuffreda, 1983; Fincham & Walton, 1957). Hence, accommodation and vergence interact together through reciprocal neural cross-link interactions: vergence accommodation and accommodative vergence. However, it is interesting to note that the eyes always start to converge before they start to accommodate; this phenomenon would be related to the difference between the neural circuit latency of the ciliary body and the one of the extra-ocular muscles (Howard & Rogers, 1995). Moreover, they also provide different and independent resting states in the dark (Owens & Leibowitz, 1980).

1.2.2 Modeling of accommodation and vergence activity

Accommodation and vergence have been described by several models explaining the functions of both systems within a feedback driven loop system (i.e., Hung, Ciuffreda & Rosenfield, 1996; Hung & Semmlow, 1980). Figure1.2.1 presents an example of this kind of static model (redrawn from Hung et al., 1996). The model describes the two parallel systems of accommodation and vergence that interact with each other through the cross-links accommodative vergence (AC) and vergence accommodation (CA). A third system is added to account for the proximal system effect on accommodation and vergence systems. Each system is composed of four different components: the controller gain, the cross-links, the tonic input and the proximal input.

The modeling of biological systems was broadly influenced by Control theory (Toates, 1975) with the strong interest that it is possible to refine any model on the basis of experimental data. These models are generally conceived as feedback systems that can be described in engineering control system terminology in terms of a controller, a plant and a feedback loop (Eadie & Carlin, 1995). In modeling the oculomotor system, the controller is the central nervous system, the plant is the muscles organization that changes lens shape or eye direction. The controller responds (with a certain delay) following the error signal provided by the feedback to change the plant in a direction in order to minimize the upcoming error between the desired and actual outputs (for details see Eadie & Carlin, 1995).

Now considering the accommodation system, the first step is to know whether the switch of the feedback loop system (see Figure 1.2.1) is opened or closed given that it controls the negative feedback to accommodation. When the switch is opened, the input to the accommodative dead-space operator (ideal mathematical function that approximates the detection threshold of a variable in a physiological system, see Hung, 1998) that represents a depth-of-focus equal to zero. When the switch is closed, the defocus blur or the accommodative error becomes input to the dead-space operator. The output of dead-space operator is multiplied with the accommodative controller gain so as to provide the accommodative controller output. The controller then responds as a reflex to the stimulus presented through the loop. The controller output becomes also input to an adaptive element, which in turn controls the time constant of the accommodative controller. The basis of proximal accommodation is first represented by its stimulus (the distance of the object from the observer) and is input to the perceived distance gain element that represents the subjective apparent distance estimate. The subsequent output then reaches the accommodative proximal gain element representing the contribution from target proximity. The outputs from the accommodation controller and the proximal controller are summed at the summing junction and are also cross-linked to the vergence system via the accommodation vergence gain. Tonic accommodation (accommodation bias) is also summed at the summing junction together with the cross-link signal from the vergence controller output via the vergence accommodation. Finally, these four signals will give the accommodative response by appropriate stimulation of the ciliary body. See Figure 1.2.1 for details.

Considering the vergence system, the switch state reflects whether binocular or monocular vision is ongoing. When the switch is closed, fixation disparity or the difference between the target vergence and the vergence response is the input to the dead-space operator. The output is multiplied with the vergence controller gain in order to yield the vergence controller output. The controller output is then input to an adaptive element which in turn controls the time constant of the vergence controller (this could reflect adaptation i.e., to prisms). The distance stimulus reaches the perceived distance gain element, which represents the subjective apparent distance estimate. It then goes through the vergence proximal gain element, which represents the contribution from target proximity. The outputs of the vergence controller and proximal controller are summed at the summing junction and are also cross-linked to the accommodation system via the vergence accommodation gain. Tonic vergence is also summed at the summing junction together with the cross-link signal from the accommodative controller output via the accommodation vergence. Finally, these four signals will generate the vergence response by appropriate stimulation of the extra-ocular muscles.

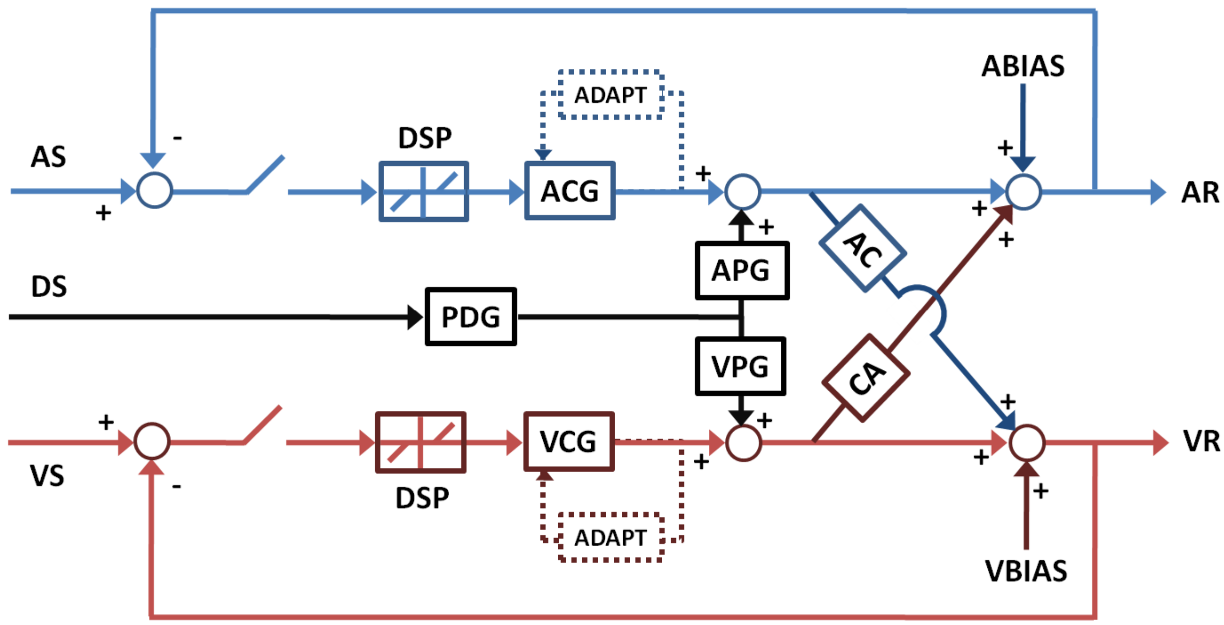

Figure 1.2.1: Static interactive dual-feedback model of the accommodative and vergence systems. AS, DS and VS represent the accommodative, distance and vergence stimulus respectively. DSP are the dead-space elements: the depth-of-focus for accommodation and Panum's fusional area for vergence. With the switch opened, the input to the dead-space operator becomes zero. Conversely, with the switch closed, the difference between the stimulus and the response (the error) is now input to DSP. The controller gain of each system is denoted by ACG, PDG (then APG and VPG) and VCG for the accommodation, perceived

distance and vergence stimulus respectively. The controller output is input to an adaptive element (ADAPT), which in turn controls the time constant of either the accommodative controller or the vergence controller. The cross-links between accommodation and vergence are represented by AC for accommodative (con)vergence and by CA for (con)vergence accommodation. The tonic components are depicted by the arrows with ABIAS and VBIAS for the accommodation bias and the vergence bias respectively. The four signals of each oculomotor system are added together to provide the overall responses: AR for the accommodation and VR for the vergence system.

1.2.3 The oculomotor dynamics

It was previously discussed that blur-driven accommodation and disparity-driven vergence are composed of two components. The static models of accommodation and vergence are useful to understand the overall response of each oculomotor system but, in order to encompass the dynamics of each system, dynamical models must specify the components of each reflex element. The study of the time varying properties (or dynamics) of the oculomotor responses allows us to identify and differentiate the components of each system. Figure 1.2.2 represents an ocular trace with its target stimulus and an example of the dynamical parameters that can be detected in the ocular response. Several oculomotor parameters can be defined from this figure; some are based on the response position and others based on the velocity profiles. In the case of vergence, the response time series are obtained by subtracting the left and the right gaze position. The velocity profile can be extracted by computing the constant velocity over the time series or by computing a weighted moving average to suppress noise (Engbert & Kliegl, 2003). The measure of time delay separating the onset of the stimulus and the onset of the response is called latency (L) and this is roughly the reaction time of the system. The time delay between the onset and the end of the response is referred to the movement time (MT), also called response duration. The amplitude (A) of the initial phase is the absolute difference in the position of the response between the event onset and its end. From the velocity trace, the maximum value among the velocity profile is defined as the peak velocity (PV). The response time constant (TC) is the time value from the onset of the transient response required to reach 63% of the maximal amplitude. The time to peak velocity (TPV) is defined as the time from the onset of the movement to the time position of the peak velocity. These measures in oculomotor responses can unravel the properties in oculomotor dynamics and in the neural basis of the response. In addition, the knowledge of the normal response dynamics could help the categorization of anomalous responses that may occur due to pathology, specific viewing conditions (i.e., stereoscopic) or other reasons.

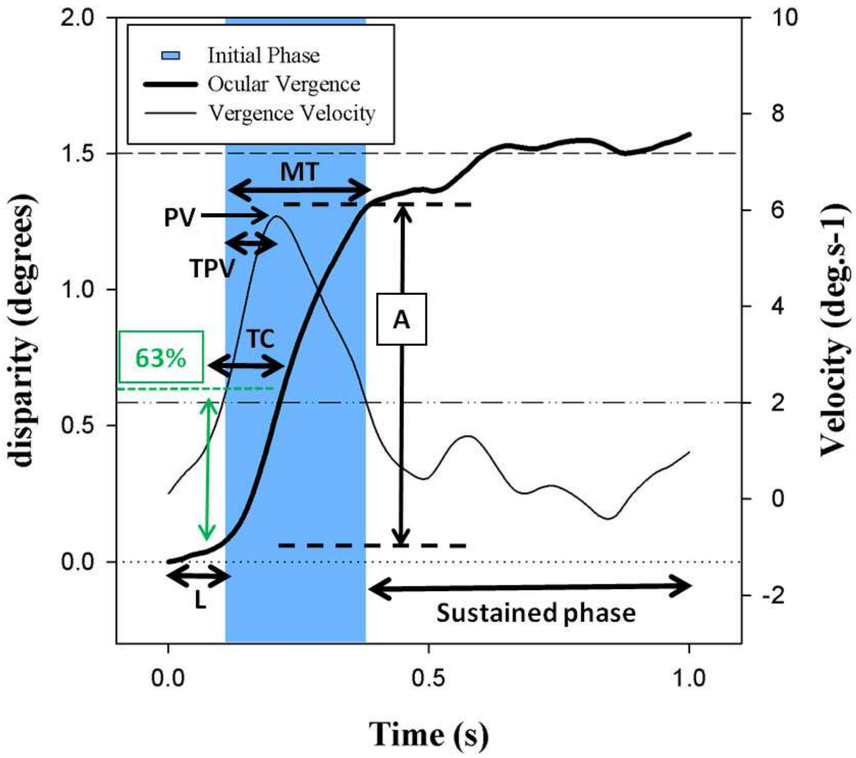

Figure 1.2.2: Example of ocular trace and velocity for stimulus-step of 1.5 degree. Ocular trace is represented by the thick solid line and velocity by the thin solid line. The blue region is the part of the time series where velocity exceeds the velocity threshold (the dashed line on the second vertical axis). L is the latency, A is the amplitude, PV the peak velocity, MT the movement time, TC the time constant and TPV the time to peak velocity (see descriptions in the text). The initial phase is represented in the blue zone and followed by the sustained phase, therein emphasizing the dynamical component.

A potent dynamical indicator in oculomotor system can be found in main sequence analysis. The main sequence was first used in astronomy so as to study the relationship between luminosity and mass of stars (Bahill, Clark & Stark, 1975). It began to become a tool for studying eye movements after Bahill et al. (1975) presented analysis of the plot of the peak velocity vs. response amplitude. Hence, it was described for the saccadic ocular motor system (Bahill, Clark, & Stark, 1975) and then used for vergence eye movements (i.e., Semmlow, Hung & Ciuffreda, 1986) and accommodation (i.e., Ciuffreda & Kruger, 1988). The main sequence plot for eye movements shows strong relationship between peak velocity and amplitude and, therein provides an idea of how the dynamic responses of a system can change with increasing amplitude (Kasthurirangan, Vilupuru & Glasser, 2003).

Main sequences of vergence dynamics are useful to check and validate that the detection criterion is satisfied and to guarantee that eye movements correspond to reproducible patterns of dynamics with acceptable variance, therein showing normal behavior of the visual system. The relationship between the velocity peak and the initial amplitude of each vergence response thereby can show that vergence dynamics is comparable in different conditions (i.e. if they exhibit similar distributions). For example, for a given stimulus amplitude, it is possible to compare the velocity distribution of different vergence responses. That way it was possible to see that divergence and convergence responses exhibit different timing properties (e.g. latency, velocity peak and time constant), suggesting that both mechanisms do not share the same neural pathways (see Hung et al., 1997). Another potential use of the main sequence could be to draw individual pattern of the performances of the visual system and, thereby characterize individual from a performance perspective and specifically facing stereoscopic stimulus target. This approach is supported by the work presented in this thesis (see chapter 3 in particular) and could offer new perspective to understand the origin of visual fatigue and to adjust visual content as a function of individual characteristics.

1.2.4 Adaptive oculomotricity and depth perception

When base-out prisms are placed in front of the eyes, the amount of convergence required when fixating a given object is increased; with base-in prisms, the amount of convergence required to fixate the object is decreased. This amount of convergence or divergence required to fixate the object through prisms is the vergence demand. In these prismatic viewing conditions, the state of tonic vergence must adapt to the vergence demand so as to maintain fusion. When these viewing conditions are maintained over time, even a few minutes, a shift in tonic vergence is observed that can last some minutes after the prisms are removed (Alpern, 1946). This phenomenon, observed by measuring the position of dark vergence, phoria or fixation disparity, is one type of vergence adaptation that can be realized using a limited amount of prisms (Henson & North, 1980) but only up to a certain deviation value, beyond which diplopia arises. Another way to observe adaptation in vergence is through wearing lenses because they change the demand for accommodation that is accompanied by a change in the resting state of vergence (Schor, 1983). It has also been shown that viewing the world through a telestereoscope (Wallach, Moore, & Davidson, 1963) can lead to change in the vergence resting state since such a device produces an increase in the inter-ocular distance. However, the resulting effects are different because prisms change all the visual field and in this way change constantly the vergence demand, whereas a telestereoscope affects the vergence demand proportionally to the viewing distance.

Hence, the perceived depth of a rotating wire forms was significantly reduced after adapting with a telestereoscope system that increased the inter-ocular distance of about 117 percent (Wallach, Moore, & Davidson, 1963). This observation leads to the question of knowing whether it occurs from changes in the vergence tonic state, cue conflict between vergence and other sources of information or, a recalibration of the apparent viewing distance. Wallach and Halperin (1977) found that the hypothesis of tonic vergence changes did not suffice to explain the effect. However, Fisher and Ebenholtz (1986) showed similar effects in the absence of conflict between disparity and monocular cues. In addition, Fisher & Ciuffreda (1990) used mirror spectacles that doubled the inter-ocular distance and found consistent changes in perceived depth and distance and tonic vergence in a way that was opposed to that predicted by the cue conflict hypothesis. They concluded that vergence adaptation changed perceived depth and distance due to changes in tonic vergence state.

Another indirect effect of the extra-ocular muscles state on depth perception is atropine micropsia (Howard & Rogers, 1995). This can be observed by prescribing cyclopegic agent to the ciliary body of the crystalline lens and asking observer to judge the size of an object. Objects appear smaller and nearer, even though they are physically unchanged. This observation was explained as being due to accommodative vergence in attempting to accommodate in spite of the lens muscles paralysis. This crosslink vergence response could signal that the object is nearer than it actually is and, therefore its size is interpreted with a biased scaling distance that leads to overestimate the object size.

While the involvement of vergence as a cue to distance is well accepted, some authors disagree as accommodation state could have an implication in distance perception. Indeed, the problem is difficult to test experimentally as vergence and accommodation show strong interactions in every-day life and, therefore it is difficult to vary one while keeping constant the other. Some experimental evidences however reported an indirect effect of accommodation on depth perception (Watt et al., 2005) and an adaptive change in accommodative response after exposition to stereoscopic stimulus (e.g., Eadie, Gray, Carlin & Mon-Williams, 2000). Accommodation may thus affect the vergence response through the accommodative vergence component and, therefore impacts depth and distance perception (Mon-Williams & Tresilian, 2000).

1.2.5 Summary: accommodation, vergence and pupil in stereoscopic displays

The oculomotor system plays an essential role in the sampling of the visual information and in keeping a stable retinal image a critical first step for the use of the multiple sources of information about depth. Moreover, the oculomotor system is well organized as it responds to various stimulus components; each sub-system of the near triad reacts to specific dynamical aspects of the stimulation. In addition to the responses that optimize the sampling of information, the oculomotor system may integrate further information coming from cognitive or emotional aspects of the viewing situation. In this sense, it can be forcefully efficient since it can be driven by conscious intentions. In order to optimize its performance, the oculomotor system rapidly performs early neural crosslinks between its sub-systems, thereby ensuring that both accommodation and vergence systems react to the same stimulation. However, even though the visual system appears to cover adaptive mechanisms that allow us to deal with unexpected events, it appears very plausible that the strength of such linkage is at the basis of major problems in stereoscopic systems.

1.3 Stereopsis

1.3.1 Basic principle

Wheatstone was the first to introduce the stereoscope in 1838. Its basic principle was presented in "Contributions to the Physiology of Vision.—Part the First. On some remarkable, and hitherto unobserved, Phenomena of Binocular Vision". It was composed of an arrangement of mirrors and two pictures of the same object taken from different points of view so as to make the object stand out in the third dimension. Wheatstone (1838) made some – mainly geometric – stereoscopic drawings that when observed from the stereoscope views were fused and gave rise to a 3D solid object.

Research on binocular disparity rapidly gathered new attention with the introduction of a special kind of stereogram, the random-dot stereogram (RDS) by Bela Julesz in the years 1960 (see for example Julesz, 1971, and Figure 1.3.1). An RDS contains two dense arrays of random dots, one for each eye. When the two views are identical, the viewer sees a unique fused planar surface composed of random-dots, while by shifting dots in one or the other view, horizontal disparity can be created so as to yield dots with a depth relative to the planar surface. The direction of this shift creates crossed horizontal disparity (negative disparity) or uncrossed horizontal disparity (positive disparity) and the depicted depth is perceived either in front or behind the reference surface. The advantage of RDS compared to other stereograms is that they only displayed binocular disparity information to mediate a sense of depth. These stimuli show that depth perception is possible on the basis of disparity information alone and, therefore can be used in experimental psychology together with other sources of information or in isolation in order to study the relative contribution of such information for judging object layout.

Stereopsis is function of visual acuity and, depending on individuals, can be precise up to a few seconds of arc – with a mean value of about 25 seconds of arc. As a consequence, the maximum distance at which stereopsis can be used to discriminate layout is limited by stereo-acuity. This maximal distance (D) can be computed, assuming an inter-ocular distance of IOD=64 mm and a stereoacuity threshold α of 25 arc seconds, with $D = IOD / rad(\alpha)$, D ≈ 530 m. It would be thus impossible to judge the relative depth of an object relative to another when they are located beyond distance D from the observer. However, this mean value of stereo-acuity concerns stable laboratory viewing conditions; generalization of this assertion should carefully take into account the contextual viewing conditions.

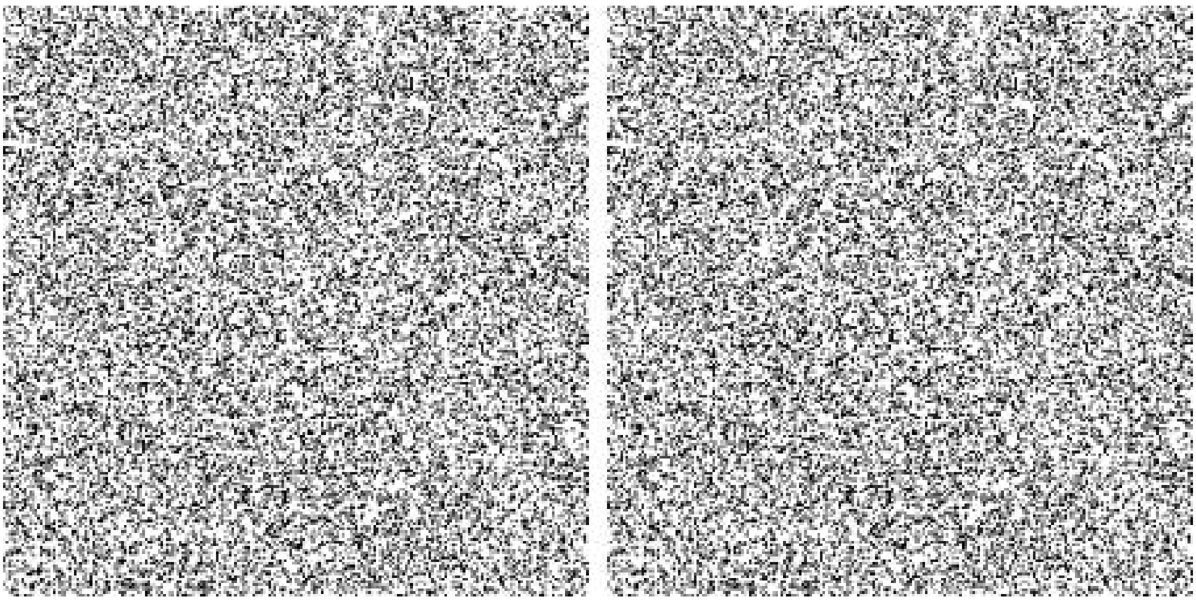

Figure 1.3.1: Figure representing a random-dot stereogram (RDS) as can be found in Julesz’s famous book “Foundations of cyclopean perception”. When fused, these two dense arrays of

random dots provide the sensation of a solid square appearing in front of the dotted background for cross-fusion or, behind the background with parallel fusion.

The disparity generated by a point out of the horopter can be expressed as (horizontal) angular disparity; this is the angle formed between the cyclopean eye – the virtual eye placed midway between the left and the right eyes – and the two points formed by the intersection of each visual line and any fronto-parallel plane (see Figure 1.3.2, angular disparity is denoted by angle η). This has the advantage of defining binocular disparity as a function of viewing distance and regardless of the inter-ocular distance. This is the main nomenclature that will be adopted in the next part.

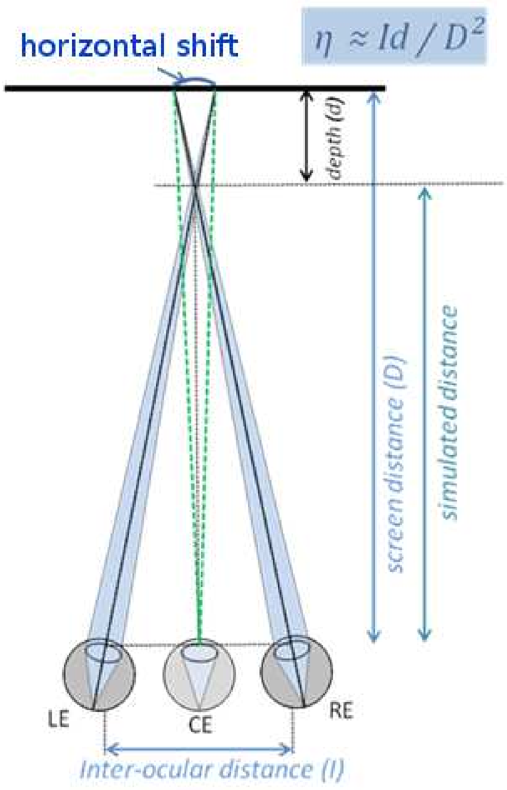

Figure 1.3.2: Schematic representation of a stereoscopic configuration. Both left and right eyes (LE and RE) here converge to a point in front of the screen plane (crossed disparities). The cyclopean eye (CE) is represented as being in the middle between the left and right eyes. The disparity η is the difference between the angles of convergence for the two points of fixation being considered. The angular extent from the cyclopean eye is a possible rough approximation of the measure of disparity.

1.3.2 Stereoscopic systems

Nowadays, there are mainly three types of 3D displays: stereoscopic, holographic and volumetric displays. As it was discussed above, stereoscopic systems principle is to render depth on the basis of a couple of 2D plane images. Rather, volumetric and holographic display systems attempt to display real 3D models for which the user could move around without altering its 3D structure. It is obvious that these three technologies do not present the same level of maturity. Given the greater complexity of holographic and volumetric displays, technical improvements are necessary, i.e., for rendering colors as well as for data generation, when compared to stereoscopic systems. Only the case of stereoscopic systems is addressed here.

In the case of stereoscopic displays, only two views are then displayed. Image parallax (defined by the disparity between the left and the right views) can be achieved with or without glasses. There are three categories of stereoscopic displays: stereoscopic, auto-stereoscopic and auto-multiscopic. The first category is the most popular because of its simplicity; it only requires the displaying of two images, one for each eye. These stereoscopic systems mainly differ by the type of glasses used to create binocular parallax. In anaglyph system, the left and right views are separated by the adjunction of two different color filters for each view – typically red and cyan. Even though this provides a relatively low cost solution for 3D rendering, this has dramatic outcomes because of the asymmetry of color between the two eyes entailing significant loss of color in the perceived scene. On the contrary, active shutter glasses provide a solution for a full colored definition. These glasses are used to present alternatively the left image to the left eye and the right image to the right eye. This alternate displaying solution offers a full spatial definition but suffers from potential loss of temporal frequency, loss of luminance, presence of image crosstalk and possible Pulrich induced effect due to the temporal separation between each view. Another displaying method (the method used in this thesis) consists in using linear or circular polarized filters in front of the screen that splits the two views that will be projected in the intended eye thanks to the corresponding filter apposed on the glasses that viewer wears while viewing the stereoscopic content. Circular filters gather the advantage that the viewer can within some tolerance lean his/her head to the right or left sides. This method also provides the advantage that polarized glasses are low-cost and lighter than active shutter goggles. However, it is worth noting that this method also suffers from loss of spatial definition (the screen resolution is divided by two) and that the quality of the polarized filter strongly affects the appearance of image crosstalk between the two views; spatial crosstalk can however be reduced when viewers are aligned with the perpendicular to the middle of the screen plane and, in laboratory conditions, when the content displayed grayscale image with poor color contrast. Finally, a special case of stereoscopic system is the head-mounted display (HMD). HMD are generally composed of two miniaturized displays and optic elements. They are highly immersive but require that the viewer maintain near vision and, this potentially quickly tires the stereoscopic viewer.

In the case of auto-stereoscopic systems, each image is spatially oriented to the right or the left eye using appropriate filters associated to the display set-up; there is thus no need to wear any glasses. The most common filter is called parallax barrier: a filter composed of a layer of material with a series of precision slots, allowing each eye to see a different set of pixels. Other displays use optical filter composed either of micro cylindrical lenses (lenticular sheet) or of micro spherical lenses (integral imaging). The main disadvantage of these kinds of display is the loss of spatial definition and the presence of image flipping when, while moving the head from side to side, viewers look at the transition between the stereo-pairs. Auto-multiscopic systems refer to way by which auto-stereoscopic displays are used to provide more than one stereoscopic view of the depicted 3D scene. In order to increase the number of views of the visual scene, mainly two solutions can be used. The former consists in tracking the observer head position and thus re-computes the left and right images corresponding to the current viewing position. The second concerns prior generation of multiple views with regard of a number of viewing angles that includes multiple couples of stereoscopic views.

1.3.3 Some evidences for neural organization

A large number of inter-individual differences can be identified when studying stereopsis. Individuals must exploit correctly their two eyes for extracting binocular disparities but also, need to be efficient in performing vergence eye movements and, must have disparity detectors. It is worth noting that 5 to 10 percent of the population do not meet these requirements (for example, in squint) and, as a consequence, these individuals cannot perceive depth from disparity alone. Richards (1971, 1970) found that people could lack the three possible classes of disparity detectors: crossed, uncrossed and zero disparity. These results suggest that some people can perceive disparity-defined depth for crossed disparity (or in front of a 3D display) while they are unable to do that for uncrossed disparity (or behind the 3D display) or *vice versa*. In addition, there is also one category of persons experiencing normal stereo acuity, however unable to see stereoscopic motion-in-depth, although they can efficiently observe static stereoscopic depth (Regan, Giaschi, Sharpe & Hong, 1992); this condition is called as stereo-motion blindness and is accompanied by normal binocular vision.

Two physiological components are said to be involved in stereoscopic perception; they are labeled "coarse stereopsis" and "fine stereopsis" (Tyler, 1990). The physiological bases for coarse and fine stereopsis are probably the magnocellular and parvocellular subsystems of the visual pathways (Tyler, 1990). The fine stereopsis mechanism responds to higher spatial-frequency patterns, smaller retinal disparities (30 minarc), and to stationary or slowly moving targets. Fine stereopsis supports high stereoscopic acuity and could also be involved in the control of fine disparity vergence. The coarse stereopsis mechanism responds instead to lower spatial-frequency patterns, large retinal disparities (30' to 10 degrees) and moving or flashed targets. It dominates peripheral vision and could also be involved in the vergence response to large retinal disparities.

1.3.4 Limitations: stereoscopic distortions

Several studies suggested that distortions in the stereoscopic image could be the prominent cause of fatigue once the visual system was adapted to the uncommon mismatching between accommodation and vergence (Kooi & Toet 2004; Meesters et al. 2004; Woods et al. 1993). Stereoscopic distortions refer to “ways in which a stereoscopic image of a scene differs from actually viewing the scene directly” (Woods et al. 1993) and, therefore constantly impact 3D realism. Lambooij et al. (2009) distinguished two categories of distortions: the generation-related and display-related distortions. The former category results from inadequacy in the process of creating stereo-pairs. For instance, in order to create left and right views, the camera configuration must be determined. As humans perform rotational eye movements when changing of fixation distance, the most natural option considered is to converge the optical axis of the left and right cameras toward the object of interest. However, this solution results in the major problem in stereoscopic viewing of introducing vertical disparities, and thus the so-called keystone distortion (e.g., Allison, 2004). Convergent camera configuration of real or virtual stereoscopic views generates inadequate vertical disparities potentially giving rise to difficulty in fusion and strongly affecting depth perception. As a matter of fact, vertical disparities are known to be used in the estimate of the egocentric distance being then used to scale binocular disparities (Rogers & Bradshaw, 1995). However, the sampling of vertical disparities by the visual system is a process specifically occurring at the retina level of the left and right eyes and, as a consequence, stereo-pairs from convergent camera will present vertical disparities from camera acquisition and result in perceptual depth distortion. Moreover, these distorted stereo-pairs generally display vertical disparities that have a range beyond the vertical fusion capacities and thus can generate visual fatigue. As a consequence, when using convergent camera configurations to generate stereoscopic images, a rectification technique should be used so as to weaken the influence of incorrect vertical disparity and render stereoscopic depth. However, as re-sampling of the images is not always practical, one can prefer the use of parallel camera configuration that does not entail incorrect vertical or horizontal disparities and preserve relative disparity information. Content generation related distortions also include the depth-plane curvature, puppet theater effect, cardboard effect, shear distortion and perspective distortion (see Meesters et al., 2004; Woods et al., 1993 for details). Conventional 3D displays – except multi-view – provide only one view of the 3D scene and, therefore do not compensate for head-movements. As a result, when observers move on the side, objects perceived farther away (with uncrossed disparity) appear moving in the same direction whereas objects perceived nearer (with crossed disparity) appear moving in the opposite direction. This shear distortion produces a perceptual effect that runs counter to the principle of motion parallax information naturally providing potent depth information. In natural viewing, near objects appear to move in a direction opposite to the observer movement and far objects in the same direction as the observer movement (Gibson, 1979). As such, the mismatching between motion parallax and observer head movement has been shown to be related to the occurrence of visual symptoms (Nojiri et al. 2006). Another interesting perceptual distortion is that stereoscopic viewers often report that the scene observed is similar to an animated puppet theater where characters seem to be miniaturized (Meesters et al., 2004). This phenomenon is potentially due to the fact that the specified object distance in the stereoscopic content sometimes mismatches with the intended object size on which observer holds prior knowledge and expectation when viewed in real world conditions – the familiar size information (e.g., Hochberg & Hochberg, 1952).

The second category of stereoscopic distortions identified by Lambooij et al. (2009) is display-related distortions, including for example, the picket fence effect, image flipping and crosstalk. To various degrees, all 3DTV displays suffer from crosstalk, consisting in a stereoscopic artifact due to the leakage of one eye view into the other view (Khaustova, Blondé, Huynh-Thu, Vienne & Doyen, 2012). The presence of such distortion can lead to a decrease of both visual quality and visual comfort, and also affects perception of depth. There are numerous reasons for crosstalk in 3D stereoscopic displays, depending on the technology. Among them, one can cite the temporal leakage that is related to the use of shutter eye glasses, and the spatial crosstalk that is dependent on viewer position relative to the screen plane. These two types of crosstalk are inherent to physical limitations and it is well acknowledged that they appear to strongly affect the user quality of experience (Woods, 2010). Picket-fence effect and image flipping uniquely concern multi-view autostereoscopic display and, as the shear distortion, these artifacts are dependent on the lateral movement of the viewer's head. According to Meesters et al. (2004), the picket-fence effect refers to "the appearance of vertical banding in an image due to the black mask between columns of pixels in the liquid crystal display" (LCD) and, image flipping indicates "the noticeable transition between neighboring viewing zones which leads to discrete views and is experienced as unnatural compared to the continuous parallax experienced in the real world". These two distortions could be minimized by increasing the number of sub-images with a small angle between them, but the number of images is constrained by the possible image resolution and barrier screen.

1.3.5 The stereoscopic demand

This section is devoted to what is referred to as the stereoscopic demand and its relation to the study of visual fatigue and discomfort. Assuming an observer with an inter-ocular difference (IOD) of 65 mm, an observer-screen distance (D) of 3m and a positive (divergent) disparity limit expressed as an image shift of about 30 mm, the far limit can be expressed in meters using the following formula: $D_S = IOD.D/(IOD - s)$ for uncrossed disparity. Here DS ≈ 5.57m. This limit can thus be expressed in terms of dioptric demand, where:

$1/D - 1/D_S = Limit$ (in diopters (D)).

In this case, the approximate limit in diopters is: $1/3 - 1/5.57 = 0.15$ D. Note that this limit is below the limit of the depth of field (for a pupil aperture of 3 mm, ±0.2-0.25 diopters). This way of characterizing the stereoscopic demand (in diopters) allows us to compare the comfort limits across a wide range of distance (see Figure 1.3.2 below).

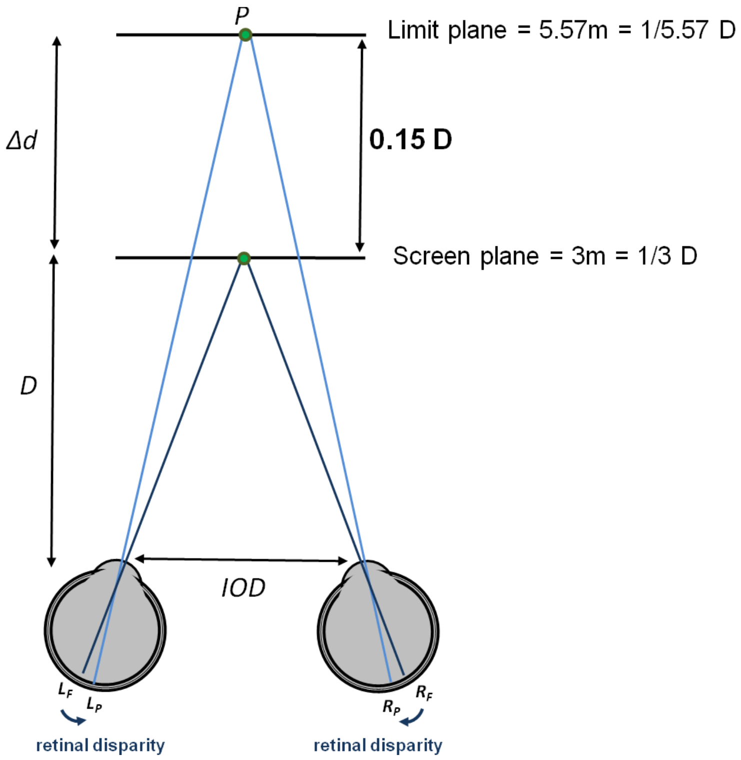

Figure 1.3.2: Schematic representation of two convergent configurations depicting the stereoscopic demand. Both left and right eyes (LE and RE) here converge to a point in front of the observer on the subject's midline at distance *D* where the screen plane is placed. Another point (*P*) behind the screen plane provides a stereoscopic demand that can be computed as the difference of the inverse distances in meters (the unit is thus in diopter but is equivalent in meter angle).

Determining when a stereoscopic viewer feels discomfort or fatigue is thus particularly relevant for an optimal use of stereoscopic viewing systems. The most common statement for eye strain occurrence is that stereoscopic displays yield incongruent stimulation forcing viewers to decouple accommodation and vergence and create strong stereoscopic demand. Previous studies generally used very large conflicts between accommodation and vergence so as to demonstrate the adverse effect. The rationale was that conflicts outside the zone of clear and single binocular vision (ZCSBV) would entail much more discomfort and fatigue than conflicts in this area. As a consequence, the study of visual fatigue in stereopsis using large demand showed significant results for many subjects while a consistent part of viewers cannot fuse the stimuli. It is worth noting that, even though disparity remains in the comfort zone (e.g. that is defined as the middle third of the ZCSBV), viewers may experience discomfort or fatigue especially in presence of motion-in-depth (i.e. Speranza et al., 2006). Indeed, rapid and repeated changes in vergence from crossed to uncrossed disparity or the inverse seems to strongly alter the visual system. As a consequence, the previous studies' need of large conflicts between accommodation and vergence could be related to the absence of motion-in-depth. In the present study, we manipulated stimuli depicting rather small conflicts between accommodation and vergence – sometimes slightly exceeding the theoretical zone of comfort (see Figure 1.3.3) – when compared to other studies but the viewing conditions of the study in this thesis appear for many participants sufficiently disturbing.

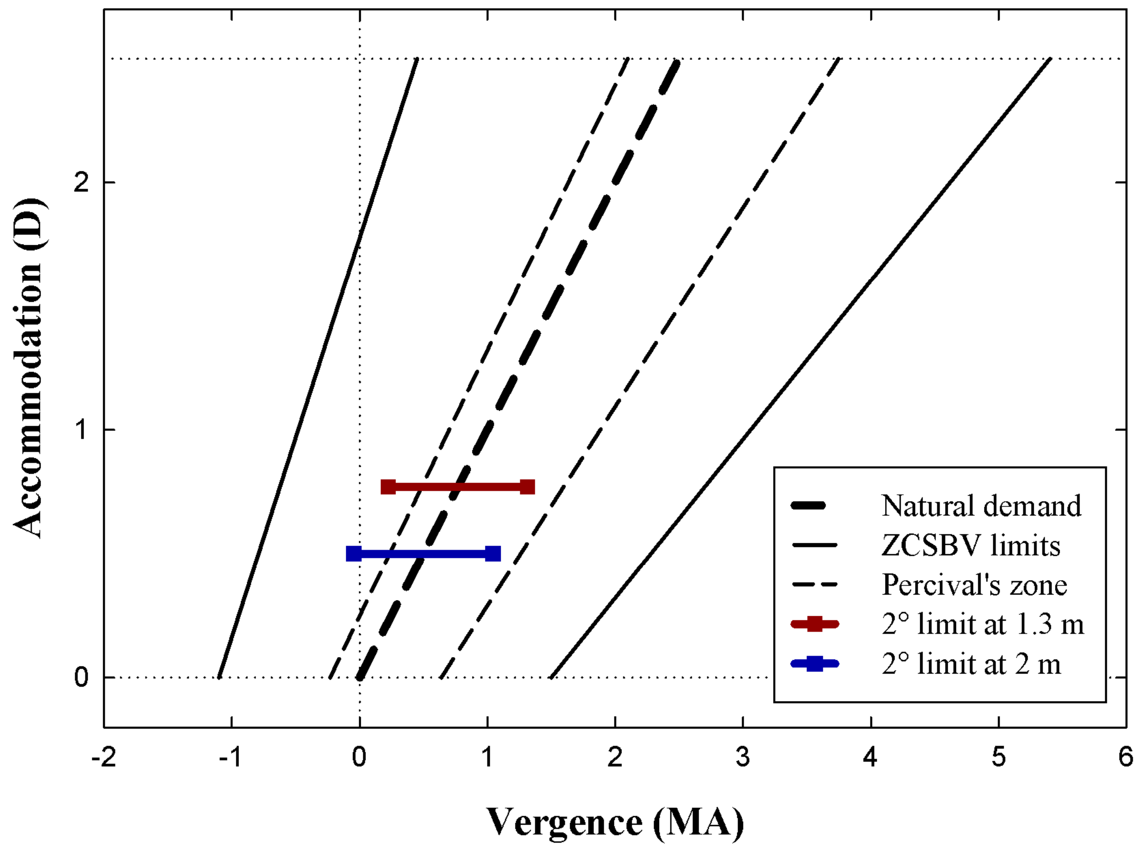

Figure 1.3.3: Schematic representation of the zone of clear and single binocular vision (ZCSBV), the Percival's zone of comfort and some example of stereoscopic demand used in this thesis work. The width of the ZCSBV whose limits are the solid back lines represents the range of vergence within which accommodation is acceptable. The Percival's zone of comfort is defined as the middle third of the ZCSBV and is bounded by the dashed thin black lines. The thick dashed black line represents the accommodation-vergence demand in natural viewing conditions. The blue and red horizontal lines represent the stereoscopic demand at viewing distance 2 and 1.3 meters for ± 2 degrees of angular disparity. The unit of accommodation is in Diopter (D, inverse of distance between the observer and the accommodation locus in space) and the unit of vergence is in Meter Angle (MA, unit of prismatic deviation, inverse of distance between the observer and the intersection of the visual lines).

1.3.6 Limitations: visual fatigue & discomfort

Visual fatigue may emerge from the prolonged observation of stereoscopic contents that include too large binocular parallax (Wöpking, 1995). The experience of visual symptoms in these conditions has been attributed to the outcomes for observers of decoupling accommodation and vergence (i.e. Hoffman et al. 2008; Okada et al. 2006; Emoto, Niida & Okano, 2005; Yano, Emoto & Mitsuhashi, 2004), both systems being naturally coupled. This explanation supports the sensory conflict hypothesis in comparing real-world and stereoscopic viewing conditions. As noted by Hoffman et al. (2008), accommodation state and blur specify the position of the display in layout rather than the depicted scene. This discrepancy between accommodation and vergence is considered to trigger a variety of visual symptoms (Shibata et al. 2011). As a whole, these visual symptoms are usually grouped under the term "asthenopia" so as to refer to both its objective (visual fatigue) and subjective aspects (Lambooij et al. 2009). Asthenopia thus characterizes a decreased efficiency of the visual system and is commonly accompanied by a number of signs such as, for example, Eyestrain, Eyes' soreness, Eyes' dryness, Headache, Blurry vision (Blehm et al. 2005; Sheedy et al. 2003). The mismatch between the sensory inputs could bring potential difficulties in (1) fusing the two monocular inputs, (2) maintaining a sharp image projection onto both retinas and (3) leading to pain and fatigue of the visual system.

Several studies investigated the stereoscopic factors that could potentially be involved in generating visual fatigue. For instance, Emoto et al. (2005) manipulated vergence demands using base-in or base-out prisms in order to strongly stimulate the vergence system on the basis of the vergence signal. These changes in vergence loads were accompanied by a theoretical fixed (screen-constrained) accommodation demand. Observers' changes in visual functions (relative vergence, accommodation response, delay in the P100 latency) increased when the vergence load was heavy or changing over time, suggesting that when the vergence demand exceeded the accommodation demand, repeated vergence adaptation plays a potent role in fatiguing the visual systems. Another investigation driven by Yano et al. (2004) has emphasized the consequence of displaying images disparity out of the area of the depth-of focus, thereby creating situations of conflict between accommodation and vergence. They observed that the degree of experienced symptoms increased with the amount of accommodation vergence mismatching. Hoffman et al. (2008) used a multiple planes display that enables variation of the amount of conflict between accommodation and convergence for stereoscopic pictures viewing. They showed with appropriate control conditions, that stereoscopic fusion was more quickly achieved, observers had greater stereo acuity, distortions effects were reduced, and viewer fatigue and discomfort were reduced when the conflict between accommodation and convergence was reduced. From these results, it appears that accommodation-vergence mismatch can strongly affect the visual system activity and therefore produces visual fatigue and symptoms following prolonged stereoscopic viewing.

Limits of sensory fusion can be particularly small; extending to 27 arc min for crossed disparity and 24 arc min for uncrossed disparity when the head is immobilized and vergence movements are absent (Yeh & Silverstein 1990). Depending on individuals, disparity magnitude beyond these limits can entail relative vergence demand, incapacity to fuse or diplopia. These fusion limits have been shown to be considerably affected by several factors such as eye movements, blur, presentation time, amount of illumination, objects size and motion (i.e. Patterson & Martin 1992). Wöpking (1995) noted that observers judged their preferred disparity magnitude in terms of viewing comfort in the range of 60-70 arc minutes between the left and right images, while this limit increased as critical parts of the depicted scene was blurred. However, even though the range of fusion can be enlarged by blurring the parts in the stereoscopic scene, it is worth noting that viewer can pay attention on these regions and, therefore experience discomfort due to the inability to explore these parts. As a consequence, disparity magnitude should be considered as a potent parameter that must satisfy both requirements of the 3D scene realism and observer visual comfort.

Stereoscopic viewing conditions are recognized as conditions presenting conflicting information for the senses. As other conflicting cues condition can affect perception (e.g. conflicting perspective and disparity cues, Gillam 1968), the mismatch between accommodation and vergence cues is also considered to affect depth perception. However, some authors argued that sensory conflict cannot lead to pain or fatigue (Riccio & Stoffregen, 1991). On the contrary, sensory conflict could lead to impairment in adjusting visual mechanisms to its stimulation. The distinction between motor and sensory aspects of vergence and accommodation in relation to visual fatigue is very important to explain the sources of adverse effects. Motor aspects concern the responses of vergence, eye rotations in opposite directions and of accommodation, adjustments of the power of the eye-lens. Sensory aspects concern the stimuli that drive these responses: for vergence, the stimulus is binocular disparity and for accommodation, the stimulus is blur. The assumption of this thesis work is that instead of sensory conflicts, incongruence in motor responses could entail visual discomfort and fatigue. If the visual system adapts to stereoscopic viewing conditions and, therefore does not perform unnecessary movements (i.e. accommodative responses) will result in no discomfort. In addition, causal factors must generally share similar characteristics with the effects they entail so as to make meaningful the causal relation. Sensory conflicts cannot be considered as a causal factor of asthenopia, otherwise all sensory conflict should lead to visual pain and fatigue. A possible candidate causing asthenopia in viewing stereoscopic content would thus be found out in oculomotor responses.

Interestingly, in clinical optometry, visual symptoms are commonly acknowledged as resulting from the occurrence of vergence errors in the gaze control. Oculomotor errors (phoria, vergence and fixation disparity) have been shown to be correlated with symptoms (Sheedy & Saladin, 1978). Two kinds of vergence errors can be described: a constant error in fixation between the intersection of the visual axes and the fixation target, and the variability of vergence eye positions over a specified time interval (Ukwade, Bedell & Harwerth, 2003). Constant and variable horizontal vergence errors obtained in normal observers under optimal conditions are typically less than 10’, and can be as low as 2’; but, according to Dowley (1989), this fixation disparity misalignment lies within Panum’s areas and is in the order of 0 to 20 min of arc. Hence, it is generally acknowledged that fixation disparity (the vergence error during fixation) is a typical candidate for causing asthenopia (Kommerell et al., 2000) or, at least, that it appears under conditions of visual stress (Dowley, 1989). Fixation disparity (FD), characterized by fusion without perfect alignment of both eyes toward the same fixation point, can be corrected through the use of prismatic lenses. This treatment has proved to be effective in treating asthenopic symptoms in patients with errors of vergence in fixation (London & Crelier, 2006). The same efficacy has been put forth in treating fusional vergence insufficiency in patients complaining of asthenopia and, forced-vergence fixation disparity curves have proved to be in close correspondence with visual symptoms before and after fusional training (i.e. Cooper et al., 1983). Fixation disparity has been shown to be closely related to fine stereopsis (Saladin, 2005) and is still considered as a way to predict stereoacuity (Zaroff, Knutelska & Frumkes, 2003). Hence, it is well acknowledged that (1) FD can determine performance in stereoacuity and, (2) FD is a factor associated with fatigue of the visual system. Given that depth perception is strongly dependent on vergence response, it is interesting to assess how the vergence system reacts in stereoscopic viewing conditions.

The measure of the zone of clear and single binocular vision (ZCSBV) is also another clinical measurement that can serve as an indicator of visual discomfort. This zone can be individually estimated either with blur points (first noticeable change in sharpness) or with break points (first disparity value at which diplopia occurs). The first measurements of the ZCSBV were published by Donders (1864; see Hofstetter, 1945). A clinical application of these experimental results was performed by Percival (1892); his objective when investigating the ZCSBV was made for understanding the relationship between optical correction and visual discomfort. He proposed the zone of comfort as being the middle third of the ZCSBV. Previous research in stereoscopic viewing conditions already stated Percival’s area of comfort as a guideline for creating comfortable content (Shibata et al., 2011; Hoffman et al., 2008;

Emoto et al., 2004). For example, Shibata et al (2011) aimed at identifying subjective zones of comfort on the basis of questionnaires, but found more reliable indicators through the assessment of the ZCSBV. Their objective was to predict visual discomfort along a range of viewing distances and found correlations between visual discomfort and clinical estimates of the ZCSBV. However, they observed that subjective reports of symptom severity alone were not sufficient to categorize susceptibility to visual adversity. It is also important to note that a few aspects of these studies somewhat revealed that Percival's area of comfort cannot so easily be applied to stereoscopic viewing conditions. More specifically, Percival's zone of comfort is measured through prisms whereas stereoscopic depth perception using this method strongly differs from the one in stereoscopic displays (Lambooij et al., 2009; Yano et al 2004). Prism deviations require viewers to perform vergence eye movements because the demand is always beyond the sensory fusion threshold and, more importantly, change the whole visual field whereas stereopsis in 3D displays requires fusion demand for only some part of the visual scene. Moreover, this zone of comfort does not take into account the presence of the screen frame that should impact the viewer fusion capacity at least for 3D displays. Two other points are highlighted by Lambooij et al. (2009). For example, while the comfort zone limits depend on viewing distance, current estimations of the ZCSBV do not represent the comfort range for long and for short viewing distances. However, this drawback can be solved by experiments where new individual data are now gathered (e.g. in Shibita et al., 2012). Another aspect is that it is difficult to compare how people might adapt to changes in prism load and how adaptation to stereoscopic viewing conditions can take place. Indeed, as prisms change the whole visual field, the oculomotor components that can adapt and the effect size when adapting cannot be compared to stereoscopic conditions.

1.3.7 Summary: disappointment of stereopsis, any chance for improving?

Stereoscopic imaging systems give the possibility to introduce a new dimension in common flat two-dimensional video content. Stereoscopic viewers may extract binocular disparities displayed between the left and the right views, retrieve depth of the observed visual scene, and thus, give visual scene content another dimension in the visual scene content. This additional value of depth is rather appreciated among the stereoscopic user. However, a number of problems have been raised that impact the Quality of Experience (QoE) in S3D representations. As we have seen in this sub-section, the list of problems can be summarized with two major aspects. The first concerns visual discomfort or fatigue that stereoscopic viewer can experience even with minimal exposure duration to 3D content. The second aspect deals with the perceptual dimension because stereoscopic systems may create strong geometrical distortions impacting the realism in the 3D visual scene. These two categories of problem resulted in a consistent piece of studies that aimed at solving or overcoming the difficulties arising with stereoscopic systems. This thesis is about understanding and improving the quality of experience in 3D media perception and, therefore will also address how to better render the perceptual experience in stereopsis. The following part will describe more aspects about quality of experience in 3D media perception and the topics addressed in this thesis.

1.4 The Quality of Experience in 3D media perception

Quality of Experience (QoE) is often used in information technology and consumer electronics domain to indicate the overall satisfaction with the service users receive. Its main objective is to model the impact of the factors involved in quality feeling. Three-dimensional video sequences provide additional experience of depth as compared to two-dimensional video sequences. 3D display systems give one particular additional cue to their viewer enhancing the viewers' perception of depth in a video scene. This particular cue is binocular disparity and is obtained by providing two views of the same scene, from slightly different perspectives, to each eye of a viewer. In order to understand the sensitivity of the human visual system to stereopsis, it is worth addressing the specific aspects that viewer will directly perceive as a quality parameter.

Stereoscopic Quality of Experience results from the combination of different influencing factors (Häkkinen, 2008) such as, the contents characteristics, camera baseline, display systems, perceptual and physiological limitations, cognitive and emotional state of the viewer. A major difference has been introduced recently between the methods to approach the QoE in stereoscopic viewing; they are labeled bottom-up and top-down because the subjective parameters are studied for the former on the basis of content characteristics and for the later on the basis of perceptual experience (e.g. Xing, 2011). For instance, the bottom-up approach concerns the study of camera baseline configurations for content generation whereas the top-down approach explores topics such as crosstalk visibility or 3D strength effect.

The principal method used in investigating the QoE is to vary the parameters expected to be involved while undertaking subjective tests. From these subjective tests, it is then possible to understand how these factors can impact the QoE and how to adapt the technology for providing an improved satisfying experience. Previous studies have easily identified the factors impacting QoE in stereoscopic viewing. The positive points are the possibility to infer binocular depth and layout of objects in the visual scene and the sensation to be part of the visual experience (immersivity). The negative points relate to visual discomfort (e.g. loss of fusion) and eye fatigue, the presence of artifacts and the statement that viewers can perceive geometry distortions degrading QoE.

The QoE parameters for stereoscopic systems are well understood; customers want impressive, immersive and eye safety experiences of stereoscopic viewing. Figure 1.4.1

summarizes the major QoE dimensions that are at stake for improving quality of experience in stereoscopic displays. In this thesis, the subjectivity of stereoscopic viewers will be approached through the objective methods of psychophysics and eye-tracking. Indeed, rather than directly question the subjective experience possibly biased in questionnaires, the quality of experience is studied indirectly. For instance, the ability to perceive depth and the resulting 3D strength effect will be determined through psychometric procedures allowing us to compute perceptual thresholds. However, the study of visual discomfort and fatigue will always be supported by subjective questionnaire. Figure 1.4.1, again, summarizes the proposed solutions and their relationship to the different QoE dimensions. Indeed, the major QoE dimensions for stereoscopic systems can be broken down into three categories: the viewer discomfort and fatigue, the 3D strength effect and the immersivity. The combination of these dimensions is supposed to strongly impact the satisfaction of the stereoscopic viewer. The proposed solutions stem from several statements presented in the following subparts.

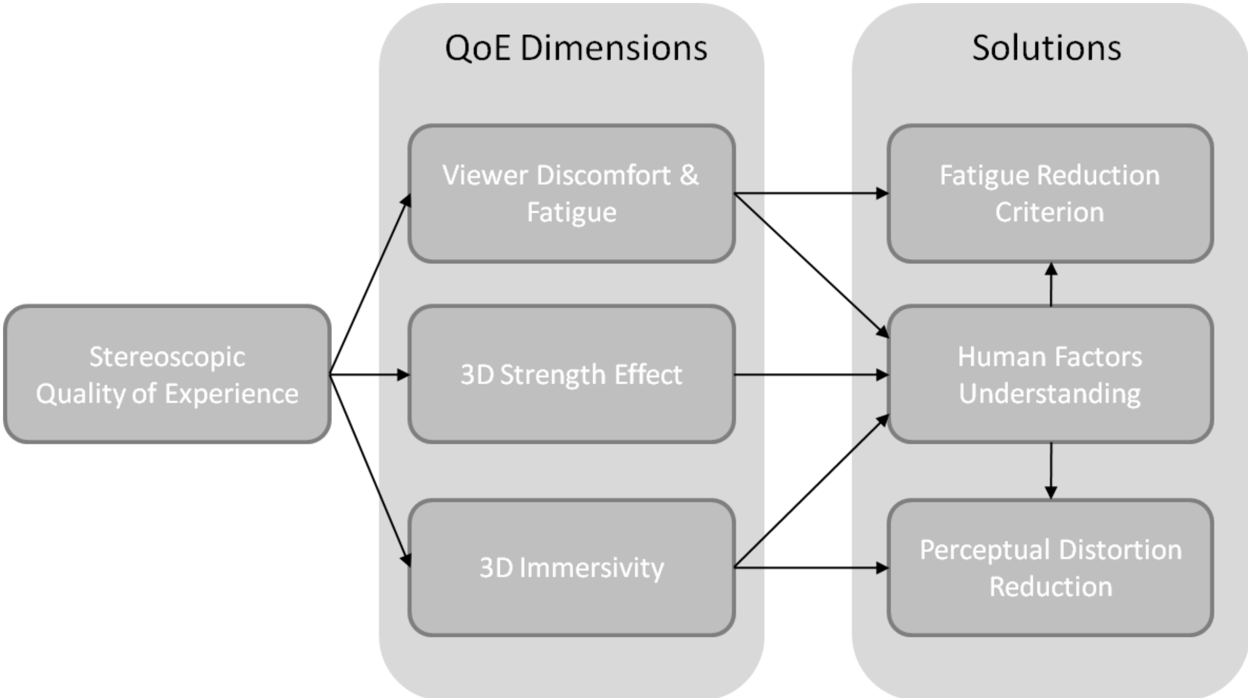

Figure 1.4.1: Quality of Experience Dimensions for stereoscopic viewing and proposed solutions.

1.4.1 Understanding visual discomfort and fatigue

One fundamental issue of quality of experience in stereoscopic displays is the understanding of the sources and causes of visual fatigue. It has been proposed that the major problem of stereoscopic systems is the propensity to place the viewer with mismatching accommodation and vergence distances (Howarth, 2011; Kim et al., 2011; Hoffman et al., 2008; Ukai & Howarth, 2008; Akeley et al., 2004; Rushton & Riddell, 1999; Wann, Rushton, & Mon-Williams, 1995). This proposal has been supported by a large set of studies (Shibata et al., 2011; Hoffman et al., 2008; Emoto, Nojiri, & Okano, 2004). However, these studies emphasized more the fact that accommodation-vergence conflict can lead to viewer discomfort and fatigue rather than the relationship between stereoscopic systems and discomfort. Indeed, in order to observe strong adverse effect on viewer comfort, the stereoscopic conditions of observation were generally composed of large conflict between accommodation demand and vergence demand (more than 1 diopter).

The first solution proposed to reduce the viewer discomfort resulting from prolonged exposition to stereoscopic content was a significant decrease of the range of disparities; the limits of comfort were evaluated around 1 degree of angular disparity (e.g. Wöpking, 1995). Consequently, several studies attempted to find out the limits of comfort (e.g. Shibata et al., 2011) and nowadays, content creators pay attention to these human factors limitation and try to respect these advised limits. It is worth noting, however, that this first solution can strongly impact depth perception in stereoscopic systems simply because any reduction in the disparity range will also reduce the 3D strength effect.

A second proposed solution was to process the displayed content parts of the visual scene which intended depth location are induced by disparities beyond the recommended limits, e.g. by blurring these annoying stereoscopic parts (e.g. Leroy, Fuchs & Moreau, 2010). This proposition makes sense because, as discussed previously, the fusion limits can be extended using stimuli with low spatial frequencies (Schor et al., 1984). However, even though this solution seemed satisfying, two major problems could have weakened the use of such a strategy. The first concerns the artistic choice of the content creator. Indeed, seeing objects pop-out from the screen is certainly one of the most interesting visual effects provided by stereoscopic systems and, as such, stereoscopic viewers could prefer see the object rendered sharp. The second is related to the fact the blurring process concerns a priori discomfortable zones in the content and that stereoscopic viewers could want to explore the scene freely.

However, a processing procedure could easily provide an improving 3D experience for single user system with a gaze contingent blurring of the stereoscopic content.

Predicting visual discomfort in stereoscopy is thus of major concern for an optimal use of stereoscopic viewing systems. The proposed solution in this thesis is to establish one or more objective criterion that could serve to detect fatigue in the viewer visual system. To this end, the responses of the vergence system can be objectively measured using eyetracking equipment and, therefore changes in the response to disparity step could be used as indicator of visual fatigue consecutive to prolonged observation of stereoscopic content. The advantages are that vergence step responses can be easily implemented using the same imaging method as the stereoscopic content. Two studies will be dedicated to unravel whether objective indicators of visual fatigue induced by stereoscopic viewing could be observed in the responses of the vergence system.

1.4.2 Explaining perceptual distortions occurring in stereopsis

As seen in the previous sections, stereoscopic systems suffer from numerous imperfections that can impact the stereoscopic QoE in creating strong geometrical distortions that can perceive the viewers (Woods et al. 1993). These stereoscopic distortions have been categorized in two specific classes: the generation-related distortions and the display-related distortions (Lambooij et al., 2009). Another class, often omitted, is the perception-related distortions (or human-factor-related). This category of distortions is particularly salient in stereo estimation of 3D shape, specifically when disparity cues dominate over perspective cues (Buckley & Frisby, 1993). For instance, perceptual distortions may appear from incorrect distance of observation or too large viewer eccentricity from the screen median plane. In the case of distortion from wrong viewing distance, when observers are asked to match the depth of an object according to its height, they show systematic biases that are related to the location of the plane of reference used to estimate depth (Johnston, 1991). More specifically, S3D viewers tend to underestimate the depth of stereoscopic objects when the viewing distance increases while they overestimate depth when viewing distance is reduced. This phenomenon has been related to the viewers' propensity to scale binocular disparity as a function of an established egocentric distance.

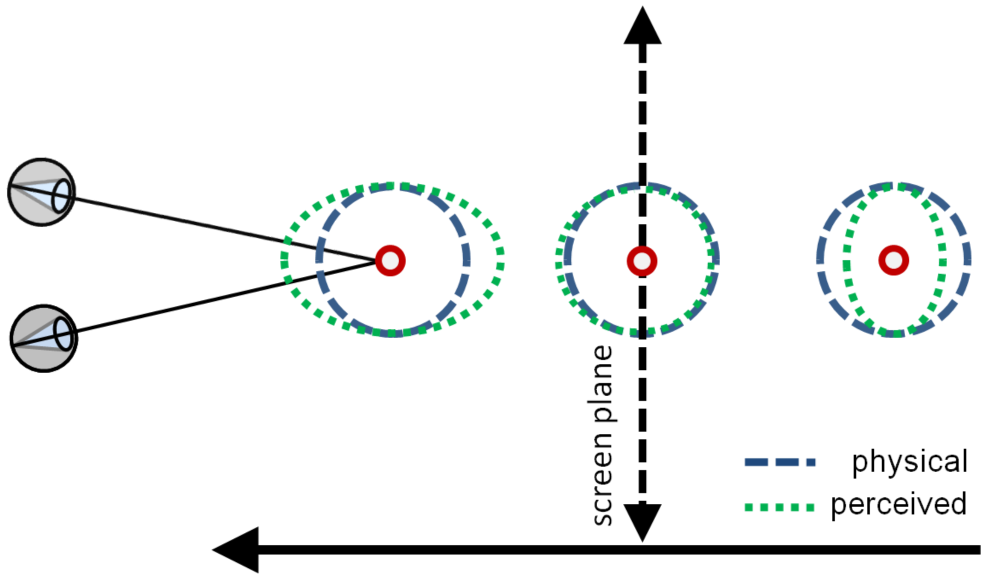

Figure 1.4.2: Schematic representation of the vergence distance effect on the estimation of a physically circular cylinder. The observed gaze points are represented by the three red circles. The green circles represent the cylinders to-be-displayed for the viewers to perceive them as circular (represented with dark blue dashed circles).

In the present thesis work, we focus on another particular perception-related distortion that has been put forward at least in Watt et al. (2005). In this study, the authors varied the accommodation distance and the vergence distance independently or in congruence, and explored the resulting effects on 3D perception. Viewers had to judge whether a vertical dihedral angle – displayed in an open-book configuration – was smaller or larger than 90 degrees. In this classical experimental design, the stimulus is composed of random-dots and only contains a weak but congruent perspective cue. The main result linked to the present thesis work was that perceived depth was strongly affected by changing the vergence distance that was induced by binocular disparity. Specifically, objects displayed behind the screen (with uncrossed disparities) appeared flattened while objects displayed in front of the screen plane (with crossed disparities) were perceived elongated in depth (see Figure 1.4.2). Moreover, the effect of disparity-induced distances was shown to be weakened by the indirect effect of the accommodation distance, as shown by the tendency for observers to base their 3D shape estimations on a scaling distance close to a compromise between the vergence and accommodation distances (Hoffman et al., 2008; Watt et al., 2005). This perceptual phenomenon will be approached in two chapters of the present study with the aim to unravel the origin of such a bias in depth perception in stereoscopic displays. The possibility to adjust perceived depth from changes in content disparity will be investigated (considering the specific case of motion-in-depth).

1.4.3 Improving S3D Quality of Experience

Improving quality of experience in stereoscopic systems will naturally result from the success in solving discomfort and fatigue related issues and from the ability to better render a strong 3D effect. Here the question addressed deals with the possibility to improve the quality of experience in stereoscopic systems by adjusting the content disparities on the basis of perceptual estimation done with psychophysics procedures. One of the causes accounted for to lessen the viewer's stereoscopic QoE is the perception of unnatural scene content. As we have seen just above, the way by which depth is rendered in stereoscopic displays can lead to significant perceptual biases in recovering the 3D metrics. However, using psychophysical tools, one can precisely estimate the geometrical distortions a viewer experiences and, therefore use these psychometric data to change content disparities. This idea seems considerably feasible for static stereoscopic pictures because the distortions are stable. However, one problem can arise if the content to-be-adjusted consists in moving pictures. This is the reason why the last study of the present thesis will take up this topic.

1.5 Overview of subsequent chapters

Following this introductory chapter are six experimental chapters (except chapter 6) followed by a concluding chapter in which the experimental research constituting the thesis is discussed as a whole. In chapter two the effect of vergence demand on 3D shape perception and vergence pursuit is investigated. Specifically, we analyzed the results on initial vergence eye movements using an analysis over time blocks. Chapter three investigates vergence eye movements to disparity-step from combined disparity and defocus information in stereoscopic and real-world viewing conditions and the consequences of stereoscopy on such disparity-step vergence movements. Specifically, the emphasis is put on whether the oculomotor system is able to face up the stereoscopic demand after watching moving in depth disparity target, for example, by adapting the components of the oculomotor system. The fourth chapter details an exploratory experiment which objective was to modulate the pupil size aperture of the eyes and thus the depth-of-focus with lighting conditions in order to observe the resulting change on visual discomfort and in performances in a stereoscopic visual search paradigm. Chapter five presents an experiment using a similar method to increase the depth-of-focus through lighting conditions and, therefore the decrease of the pupil diameter, for the estimation of 3D shape and the role of accommodation in stereopsis. In chapter six, is discussed the optical characteristics of the human eye in stereoscopic viewing condition in order to analyze the link between perceptual bias in stereopsis and eye optical effect. Chapter seven is about how to improve 3D shape perception in stereoscopy and, especially when content displays motion-in-depth. The concluding chapter, chapter eight, provides discussion about the research presented in this thesis as a whole.

Chapter 2: Exploratory study: visual fatigue, vergence eye movements and perceptual bias of stereoscopic 3D shape

2.1 Introduction

Stereoscopic viewing of simulated scenes for long durations has been shown to be a potential source of visual fatigue and discomfort. Complaints of a variety of visual symptoms as well as the decreasing efficiency of the oculomotor system have been reported by several studies (Shibata et al. 2011; Hoffman et al., 2008, Emoto et al. 2005; Yano et al., 2004). The first studies have shown that stereoscopic visualizations generate better quality of experience and comfort when the three-dimensional (3D) contents display relatively small binocular disparities (Wöpking, 1995) and, investigations of critical limits have been conducted to find out the zones of comfort. Hence, the issues are to unravel what could be the sources of visual fatigue and to identify what human factors could lead to discomfort.

The origin of visual symptoms following prolonged exposure to stereoscopic content was discussed since the introduction of the stereoscope (Wheastone, 1838) or, at least, since its use in theaters for entertainment. The most common form of rendered three-dimensional layout is to simulate two points of view for both eyes, using either a toed-in field of view configuration or a parallel eyes-axes configuration (Woods et al., 1993). In these conditions, the binocular disparities between the two views are detected by the observers and can specify a realistic 3D structure, but only for specific viewing distances and lateral displacement relative to the screen plane (Vishwanath et al. 2005; Woods et al., 1993). Therefore, incorrect viewing position can produce perceptual distortions that will alter the realism of the 3D shape. Together with these possible geometrical distortions of the Stereoscopic 3-Dimensional (S3D) layout, it is generally acknowledged that a major problem is the mismatch between vergence and accommodation cues (Hoffman et al., 2008; Wann et al., 1995). This discrepancy between accommodation and vergence is considered to trigger a variety of visual symptoms (Shibata et al. 2011). Globally, these visual symptoms are generally grouped under the term “asthenopia” so as to refer to both its objective (visual fatigue) and subjective aspects (Lambooij et al. 2009). Asthenopia thus characterizes a decreased efficiency of the visual system and is commonly accompanied by a number of signs such as, for example, Eyestrain, Eyes' soreness, Eyes' dryness, Headache, Blurry vision (Blehm et al. 2005; Sheedy et al. 2003).

In natural viewing conditions, accommodation and vergence responses have been shown to be reciprocally coupled (i.e. Schor, 1992). The function of accommodation is to minimize the blur in the retinal image for a certain range of distance, the depth of focus, which is the perceptual tolerance to retinal defocus (Wang & Ciuffreda, 2006), whereas the goal of vergence is to direct the fovea of both eyes towards the object of interest. While blur is the signal for eye-accommodation and disparity is the signal for vergence, some interactive neural processes ensure that both eye-movements react to either signal (Semmlow et al. 1981). These interactions are commonly named accommodative vergence (the vergence response to blur) and convergent accommodation (the accommodation response to disparity). For example, monocular blur stimulation can elicit vergence response when only one eye is actually stimulated. Interestingly, this link between accommodation and vergence can already be observed at an early age of human development (Turner et al., 2002; Aslin & Jackson, 1979), thus attesting that this interaction is strongly sturdy. Even though accommodation and vergence do not share the same unit, both systems depend on object distance. In classical stereoscopic displays, different vergence angles are induced by the binocular disparities displayed and give rise to the perception of objects layout while the stimulus-to-accommodation specifies a unique range of distances around the screen plane position (the depth of field). In varying both focus and vergence cues, several studies have shown that observers can experience visual fatigue (Hoffman et al., 2008; Emoto et al., 2005; Yano et al., 2004). For instance, Emoto et al. (2005) manipulated vergence demands using base-in or base-out prisms in order to create two conditions: one in which the vergence signal was varied and then fixed for a given time, and the other in which the vergence signal changed periodically. These changes in vergence loads were accompanied by a theoretical fixed accommodation demand. Observers' changes in visual functions (relative vergence, accommodation response, delay in the P100 latency) increased when the vergence load was heavy or changing over time, suggesting that when focus-vergence cues were inconsistent, repeated vergence adaptation play a potent role in fatiguing the visual systems. Hoffman et al. (2008) used a novel 3D display that allowed them to manipulate the difference magnitude between accommodation and vergence in binocularly-viewed pictures. They showed that binocular fusion was faster, stereoacuity thresholds were reduced, and scores on a visual comfort scale were increased when accommodation and vergence cues were consistent.

The reason why visual fatigue occurs in S3D may be that observers fail to decouple accommodation and vergence (Hoffman et al. 2009; Okada et al. 2006; Inoue et al. 1997; Wann et al. 1995). Indeed, the divergence between these two systems takes place when the simulated depth of a displayed point exceeds the projection of the dioptric interval of the depth of focus in space, because the accommodation response is then saturated (Hiruma, 1990, cited by Pastoor, 1991). According to several authors (Hoffman et al. 2008; Torii et al. 2008; Okada et al. 2006), the discrepancy is characterized by the opposing direction taken by the blur-driven accommodation and by the vergence-driven accommodation. There are actually two temporal components in the accommodative response: the primary accommodation response reacts quickly to disparity, and then accommodation is regulated by the amount of blur present in the stimulus. Thus, in the case of binocular vision, the first accommodative response is elicited by vergence: the lens deformation follows the direction in agreement with the vergence distance. The second component of the accommodative response is driven by the spatial frequency content of the stimulus and then, in the case of binocular vision, the lens deformation evolves to an adequate state of refraction. In other words, there is an accommodative response following the vergence response that is then reduced by spatial frequency information. Hence, according to Okada et al. (2006), asthenopia follows from an oculomotor imbalance that would essentially involve accommodation. When the simulated stereoscopic depth remains within the focus range, accommodative responses are possible beyond the screen plane, and, in that case, fluctuations of the lens are too small to generate a blurred vision (Koh & Charman, 1998). This is the reason why it is generally advised to keep the displayed vergence distance in the zone of comfort which is approximately the middle third of the zone of clear single binocular vision (Emoto et al., 2005). Otherwise, the required effort of the oculomotor system related to accommodation error correction will produce visual fatigue (Lambooij et al. 2010). This anomalous oculomotor response that provokes asthenopia would be an accommodative overshoot induced by the vergence demand that requires an accommodative correction (Fukushima et al. 2009; Torii et al. 2008; Okada et al. 2006).

Accommodation-vergence mismatching appears to affect strongly the visual system, but the resulting effects on depth perception are still relatively unexplored. As such, it is worth remembering that the inference of depth from stereopsis suffers from misestimation of the binocular disparity information (e.g. Adams & Mamassian, 2004; Johnston, 1991). Indeed, binocular disparity is supposed to provide an unambiguous cue to depth if it is correctly scaled by a proper estimate of the egocentric viewing distance (Garding et al. 1995). Interestingly, this estimate of viewing distance is presumably relying on the actual vergence and accommodative states (Watt et al. 2005; MonWilliams et al. 2000), and on the size ratios of horizontal and vertical disparities (Rogers & Bradshaw, 1995), but the accuracy of this estimate and the information used for scaling are distance-dependent (Domini et al., 2006; Cutting & Vishton, 2000). These considerations are very important when creating stereoscopic content and displays because the estimate of viewing distance could be biased in systems where accommodation and vergence specify discrepant distances. Hence, given that binocular-induced visual fatigue modifies the relationship between accommodation and vergence (Emoto et al., 2005; MonWilliams et al., 1993), we expect a decrease in the ability to properly infer the correct viewing distance. Nonetheless, effects of visual fatigue are usually considered to affect stereopsis rather than distance perception (i.e. a decrease of stereo-acuity). Some evidence has been provided by Hoffman et al. (2008) who showed that the time required to fuse random-dot stereograms increases with increasing discrepancy between accommodation and vergence distances. Watt et al. (2005) have also shown that stereoscopic depth constancy is affected by the accommodation signal. They asked participants to judge whether a dihedron angle was smaller or larger than a right angle for three possible viewing distances and found that accommodation could cause a flattening or an increase of the perceived hinge angle depending on the vergence distance relative to the screen.

The literature reviewed so far suggests that misestimations of the information provided by binocular disparities in stereoscopic contents could be the result of an accommodation-vergence conflict. If we consider visual fatigue to be a state resulting of a prolonged immersion in this cue conflict situation, the perception of 3D shape from disparities should also be affected. However, as people may subjectively experience visual discomfort, they do not necessarily exhibit any decrease in the performance of the binocular system due to visual fatigue. The reason is twofold: subjective and objective measures of visual strain do not covary (Lambooij et al. 2009), and visual fatigue is limited by potential adaptation processes of the binocular system (Schor, 2009). As such, it was suggested that this possible adaptation of the binocular system could be achieved through the neural crosslinks between accommodation and vergence (Wann et al., 1995), changes in the gain of the phasic accommodation and the phasic vergence, and changes in the tonic components (Rushton & Riddell, 1999). People may thus experience subjective visual symptoms either with or without changes in the efficacy of the visual system because of inter-individual variation in the susceptibility to visual fatigue (Lambooij et al., 2009); this possibly reflects a failure for visual mechanisms to adapt to effortful stimuli. Prolonged exposure to stereoscopic content may thus engender the sensation of fatigue and discomfort in conditions where visuo-motor processes are ineffective as well as when these processes suitably adapt to viewing conditions. The investigation of the bias direction in the 3D perception after strong stereoscopic demand should then enlighten the effect of visual fatigue on stereopsis. The objective of this study is to explore the potential effect of visual fatigue on depth perception. In addition, the relationship between the perceptual bias due to fatigue and the vergence system performance is explored during sessions of binocular fixation with different amounts of binocular disparity. We used a procedure with two interleaved tasks: the first was a binocular tracking task with changing disparities for tiring the visual system, and the second was stereo-estimation of a dihedron angle displayed as random-dot stereograms. We also made pre- and post-clinical tests to unravel potential links between the adaptation of the vergence system and the reported bias direction in stereo-estimation.

2.2 Materials and Methods

2.2.1 Observers

A total of 12 observers took part in the entire study. Four other participants were discarded because one was stereo-anomalous, one was diagnosed as suffering from presbyopia and the two others reported large difficulties to fuse the stimuli during the experimentation. Our selection provided a full counterbalanced group of participants (on average 34.5 years old, 6.6 SD). All except one author were unaware of the experimental hypotheses. All had normal stereoacuity as assessed by the stereo-test of the vision tester (Visiolite, FIM-Medical). All had normal vision possibly thanks to their usual optical corrections. They gave their informed consent before beginning the experiment.

2.2.2 Apparatus

Participants sat comfortably in a chair within a chinrest facing a table on which the binocular eye-tracker was placed (Eyelink 1000 with 2000 Hz upgrade, SR-Research). The viewing distance was 2 meters from the display (Hyundai S465D 46’’ HDTV LCD Polarized monitor). To minimize crosstalk, we placed participant eyes on the axis perpendicular to the center of the screen. We used two computers in this experiment, the first displayed 3D content using Matlab and the Psychtoolbox extensions (http://psychtoolbox.org) and the second monitored the gaze tracking. Participants wore polarized glasses to fuse left / right views and see the content in depth.

2.2.3 Stimuli

Vergence effort

Observers saw an apparent fixation cross whose disparities changed over time. Its theoretical displacements, as predicted by disparity information, were either back to front or front to back relative to the screen plane. This fixation cross was surrounded by a frame composed of small black and white squares to help maintain stereoscopic fusion. Observers were requested to fixate the cross in its center and were asked to refrain from moving eyes toward the surrounding frame. The size of the fixation cross was 40 arc minutes and the size of the surrounding frame was 2 degrees. We created two conditions of cross displacement: the high vergence demand and the medium vergence demand, and one condition where the cross remained static. The moving-in-depth cross velocity was always sinusoidal, that is, changes in binocular disparity corresponded to a sinusoidal waveform over time. In order to obtain equivalent velocities in the two conditions where the cross moved, we started and finished all trials by a period of fixation, matching the intended disparity. Final disparities were opposite in sign to starting disparities (± 90' for the high vergence demand and ± 45' for the medium vergence demand). We selected these disparity values because beyond 60 arc minutes, disparities lead to severe discomfort (Wöpking, 1995) and because of the commonly reported limit of binocular fusion in stereopsis (Lambooij et al., 2009). Therefore, the conditions of vergence effort were equivalent in terms of velocity demand for tracking ability and of presentation duration. That is why we refer to these two conditions of vergence effort by the maximal disparities at the onset and the end of the trial. A 0' condition was introduced as a control condition. In this condition the cross and its frame were static on the screen during the entire test. We will label these three conditions as the static vergence condition, the medium vergence effort condition and the high vergence effort condition (respectively, zero disparity, +/-45' disparity and +/-90' disparity).

Stereo-judgment task

After each run of vergence effort, performances in judging the inside angle of a vertical dihedron (as in an open-book configuration) were assessed using random-dot stereograms.

This stimulus has already been used elsewhere and showed high reliability in accounting for the capacity to perceive 3D shape (Hoffman et al. 2008, Watt et al. 2005). The task was to focus in the depth plane of the monitor on a small fixation cross (25 arc min) which was placed in the center of the dihedron and then to judge whether the displayed dihedron angle was smaller or larger than 90° (see figure 2.1). In order to ensure that the vergence effort did not influence directly the 3D judgment by the way of gaze position, we asked participant to accurately fixate a target in the middle of the screen and then to press a button when ready. The dot density for the dihedron was 0.87 dots/deg, with a width of 14 degrees of visual angle and a height of 7 degrees; the dot density of the background was 5.6 dots/deg. The fixation cross was always in the monitor plane while the parts of the hinge appeared in front of the screen. Before each experimental session, we carried out a learning task (with audio feedback) during which observers had to obtain a score of 90% in judging the dihedron angle. For the learning task, we used the method of constant stimuli and presented dihedron angles in the range of 81 to 99° (10 stimuli intensities separated by 2 degrees and the whole repeated 4 times until participants reached the attended score). Each dot was placed relative to the cyclopean eye by projective geometry. For the experimental session, angles of each dihedron were displayed according to an adaptive staircase procedure (Accelerated Stochastic Approximation, Kesten 1958). This procedure allowed us to estimate the point of subjective equality (PSE) and the slope at this point for each block and session. The maximal number of repetitions was set to 15 for each staircase. During one block, observers participated to 3 threshold estimations (50%, 25% and 75%) using 2 staircases (one ascending & one descending) per threshold estimation for a total of 90 trials. This procedure was repeated 3 times so as to obtain 3 estimations of the psychometric function in each session.

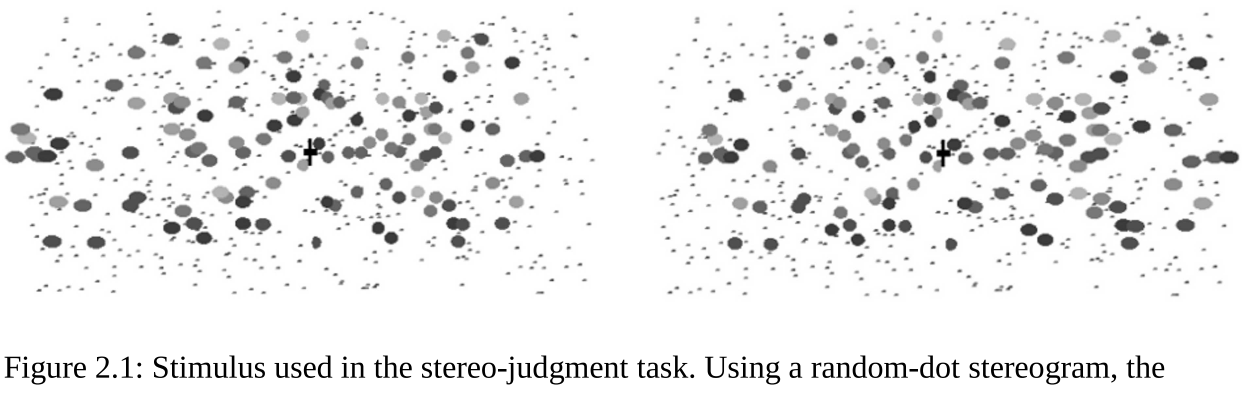

Figure 2.1: Stimulus used in the stereo-judgment task. Using a random-dot stereogram, the stimulus depicted a dihedron (large dots) in front of a random-dot background (small dots). The shape of the dihedron was similar to that in Hoffman & al. 2008, except that we added (1)

a random-dot background and (2) a gray level for each slanted dot (i.e., (2) was supposed to improve fusion speed and was appreciated by the participants in pre-tests).

2.2.4 Procedure

Two sessions of vergence pursuit and their control condition were designed according to different vergence demands (± 90, 45 and 0 arc minutes of initial/final disparity). Each session was run independently, on three different days, so as to avoid any contamination effect of effort & fatigue across trials. During each session, the vergence task and the stereo task were interleaved. We split each session into three consecutive time blocks of 90 trials each (block 1 from the first to the 90th trial, block 2 from the 91th to the 180th trial and block 3 from the 181th to the 270th trial) in order to test for potential effects over time. The sessions of vergence demands were counterbalanced across participants.

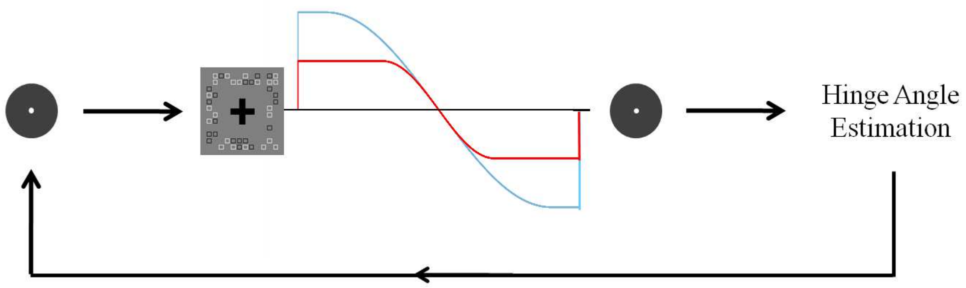

Figure 2.2: Experimental Schedule of the two interlaced tasks (vergence effort & stereo-estimation)

Pre- and Post-tests

Before and after each individual session, we asked participants to fill in the Simulator Sickness Questionnaire (SSQ, Kennedy & al. 1993) and we assessed their vertical and horizontal near heterophoria (with the method described in Corbett & Maples, 2004) and fixation disparity using the Saladin Near-point Balance Card (fixation disparity was measured without prism).

The SSQ is useful to establish a profile of symptoms following prolonged viewing of stereoscopic content. It can be used to evaluate scores for three possible components: nausea, oculomotor deficits and disorientation. These scores can be combined to produce a global SSQ score. Participants have to rate among 16 symptoms the intensity they experienced on a 4 points scale (None, Slight, Moderate & Severe). Symptoms are the following: General discomfort, Fatigue, Headache, Eyestrain, Difficulty focusing, Increased salivation, Sweating, Nausea, Difficulty concentrating, Fullness of head, Blurred vision, Dizzy (eyes open or closed), Vertigo, Stomach awareness & Burping. Scores are then computed using a weighting table (see Kennedy & al. 1993).

2.3 Results

2.3.1 Subjective Experience

From the responses obtained for each of the 16 items of the SSQ, we computed the scaled scores for the three factors (nausea, oculomotor and disorientation) and the total scores of visual sickness (see Kennedy & al. 1993). The analyses carried out on the weighted scaled scores of the SSQ revealed an effect of the vergence effort condition (0, 45', 90') on the component of the SSQ ($p<0.0001$). First, the oculomotor factor was predominantly more involved in the experience of symptoms in the condition of high vergence demand as compared to the other two ($F(2, 33) = 3.87$, $p= 0.03$). Second, we observed that the global scores (sum of the scores for the 3 components) were also significantly more important in the high vergence demand condition than in the other two ($F(2, 66) = 15.845$; $p < 0.0001$). However, the interaction effect between the condition of vergence demand and the components of the SSQ on the scores was not significant ($F(4, 66) =1.61$, $p >0.05$). In comparing the responses for each question of the SSQ between the conditions of vergence demand, we observed that the SSQ scores were mainly influenced by the following items: “General Discomfort”, “Fatigue”, “Eyestrain”, “Difficulty focusing” and “Blurred vision”. They were significantly more rated in the high vergence demand condition as compared to the other two ($p < 0.05$).

2.3.2 Perceptual Estimation

We extracted points of subjective equality (PSE, the 50% point on the psychometric function) in the dihedron angle estimation. The present analyses did not separate the condition in which trials could either start with a crossed disparity step or an uncrossed disparity step. We estimated the PSE for each participant and each session and block using the maximum likelihood procedure applied on the raw responses (coded as 0 and 1 for “smaller” and “larger” than 90 degrees) and displayed as a cumulative distribution. The analysis performed on the PSEs for the 12 participants between each session of vergence effort did not show any significant effect ($F(2,20)=0.716$, $p=0.339$); but a large inter-individual variability was observed in the zero disparity and the 90' disparity session (the standard deviation was 4.1 degrees in the zero disparity group, 3.12 degrees in the 45' disparity group and 5.9 degrees in the 90' disparity group). Likewise, we did not observe a significant effect of the variable time blocks ($F(2,20)=0.259$, $p=0.775$). We also extracted the slope of the psychometric function at the PSEs but the analysis did not reveal any significant effect of vergence effort ($F(2,20)=0.167$, $p=0.847$); in other words, the variability of the responses at the PSE could not be differentiated between the three conditions of vergence demand.

The slope at the PSE of a psychometric function is usually regarded as an indicator of uncertainty in the response because it reveals the variability around this point. We expected that visual fatigue could affect 3D shape perception in two different ways: an increase of the uncertainty that would reflect a state of fatigue without adaptation, and a decrease or unchanging uncertainty that could reflect fatigue with adaptation. By comparing the slope at the PSEs between the session that was not supposed to entail visual fatigue (zero disparity condition) and the one that was supposed to introduce a tiring of the visual system (90' disparity condition), we observed two different tendencies across participants: a decrease of the slope or an increase of the slope across the vergence effort conditions (group1 and group2 respectively). We assigned participants to two groups according to the variation of the psychometric slope (sensitivity) over the three conditions of vergence effort (6 participants in each group). Now, the evolution of the slope between the static vergence effort condition and the high vergence effort condition was significantly different ($F(1,10)=24.88$, $p<0.0001$). We found two tendencies in the bias direction (interaction effect: $F(2,20)=6.034$, $p=0.009$). Post-hoc analysis revealed that Group1 showed a decrease of the perceived angle with the vergence effort (the angle is judged smaller in the high vergence effort condition than in the static vergence effort condition, $p<0.05$) and Group2 exhibited instead an increase of the perceived hinge angle with the vergence effort (the angle is judged larger in the high vergence effort condition than in the static vergence effort condition, $p<0.01$; see figure 2.3). In other words, for the first group, the higher the level of vergence demand, the less is the over-estimation of the dihedron angle while for the second group, the higher the level of vergence demand, the more is the over-estimation of the dihedron angle. The variability for the high vergence effort condition in the group1 is due to one participant who provided marginal estimation in spite of the match with the grouping criterion (i.e. based on the sensitivity of the psychometric function). Unfortunately, the analyses realized on the PSEs between the time blocks did not show any significant effect (F(2,20)=0.259,p>0.05). The estimation of depth from stereopsis is likely dependent on the scaling distance observer must consider for perceiving depth and, as a consequence, authors converted perception thresholds in equivalent/scaling distance (e.g. Hoffman et al. 2008, Watt et al. 2005, Johnston, 1991). The scaling distance enlightens what depth extent a given pattern of disparity is supposed to mediate at a particular distance. Therefore, we computed these distances which matched with the pattern of disparities that observer considered as corresponding to the intended 3D structure (see formula 2.1 in appendices). The averaged results can be found in figure 2.4 for the two groups. The analysis performed on the scaling distances did not reveal any effect when the scaling distances of both groups where taken together (F(2,22)=0.47, p>0.05), but again the vergence effort condition gave rise to two tendencies considering the grouping criterion (interaction effect: F(2,20)=6.51, p=0.007). Group1 tended to underestimate the scaling distance with the increase in vergence effort (p<0.05) while group2 increasingly overestimated the scaling distance with the increase in vergence effort (p<0.01).

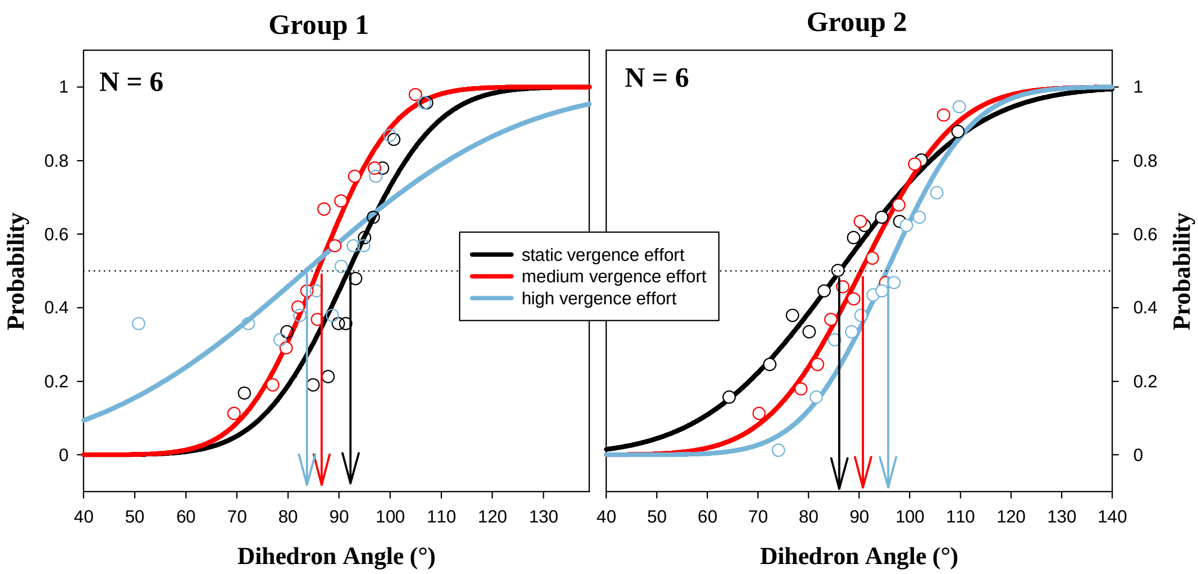

Figure 2.3: Plot of the cumulative Gaussian fitting for the two sets of participants in the 3 conditions of vergence effort (static, medium, high). Figure 2.3 shows the results from the stereo-task (for all observers) as well as fits by a cumulative Gaussian function. This figure shows both trends in the progression of the PSE between the 3 possible conditions of vergence effort (zero disparity, 45’ disparity condition & 90’ disparity condition).

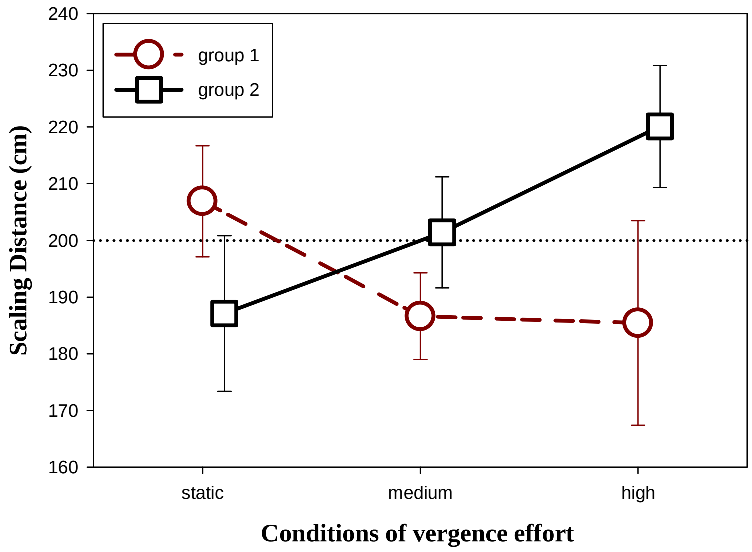

Figure 2.4: Effect of the vergence effort sessions on the scaling distance considered to be used in perceptual estimation. Scaling distances for group1 and group2 are represented in dark red and black respectively. The error bars denote standard errors of mean.

The effect of the vergence effort condition on the reaction time (for all trials) was significant (F(2, 2134)=5.569, p=0.004), the time required to judge the angle size was smaller in the static vergence effort condition than in the medium and the high vergence effort conditions (p<0.002 and p<0.01, respectively). The effect of the block session on the reaction time was also significant (F(2, 2134)=33.78, p<0.0001), the time required to judge the angle was larger in block 1 than in blocks 2 and 3 (p<0.001 and p<0.001, respectively). However, the interaction effect between the condition of vergence effort and the condition of time block was not significant (p>0.05). Figure 2.5 sums up the results for the reaction time.

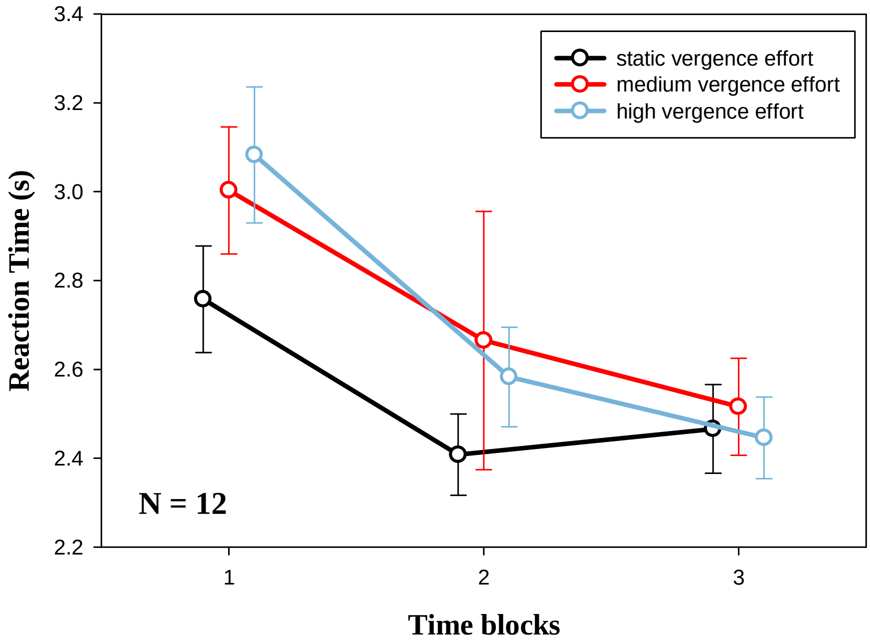

Figure 2.5: Effect of the vergence effort sessions and of the time blocks on the Reaction Time in the perceptual estimation. The first time block concerns the first to the 90th trial, the second block the 91th to the 180th trial and the third the 181th to the 270th trial. The reaction times for the static vergence effort (zero disparity), the medium vergence effort (45') and the high vergence effort are represented in black, blue and red respectively. The error bars denote confidence intervals of mean.

2.3.3 Heterophoria and Fixation Disparity

We computed the correlations for heterophoria in pre-tests for each condition of vergence effort. For horizontal phoria, we found correlations being of .93 between static and high vergence effort, .63 between static and medium vergence effort and .64 between medium and high vergence effort (all correlations significant, $p<0.05$) and, for vertical phoria, correlations were .79 between static and high vergence effort ($p<0.05$), .29 between static and medium vergence effort ($p>0.05$) and .67 between medium and high vergence effort ($p<0.05$). The difference between heterophoria in pre-tests for each condition of vergence effort did not show to be different both for vertical and horizontal phoria (Friedman ANOVA: Chi-square=0.7, $p>0.05$ and Chi-square=5.6, $p>0.05$). The repeatability of this clinical test sounded thus quite good. We computed the difference between the pre- and post-test for each vergence effort condition but no significant effect of horizontal and vertical phoria was observed ($F(2,22)=0.05$, $p>0.05$ and $F(2,22)=2.12$, $p>0.05$, respectively). Due to a lack of results, we analyzed the effect on absolute values of horizontal and vertical heterophoria, but still there was no effect ($F(2,22)=1.84$, $p>0.05$ and $F(2,22)=1.44$, $p>0.05$, respectively). The variability among the participants about the direction of the lateral phoria shift was important: 3 showed no effect, 5 were exophoric and 4 were esophoric.

The difference between fixation disparity in pre-tests for each condition of vergence effort did not show to be different both for vertical and horizontal phoria (Friedman ANOVA: Chi-square=0.4, $p>0.05$ and Chi-square=2, $p>0.05$). We computed the difference between the pre- and post-test for each vergence effort condition but no significant effect of horizontal and vertical fixation disparity was observed ($F(2,22)=0.05$, $p>0.05$ and $F(2,22)=2.12$, $p>0.05$, respectively). However, horizontal fixation disparity showed an increasing esophoric shift with vergence effort when considering only the post-tests (Friedman ANOVA: Chi-square=6.75, $p<0.05$).We converted the scaling distance (see above) and horizontal near heterophoria and fixation disparity in meter angles so as to compute the correlations between estimated fixation distance and clinical results for each condition of vergence effort but we did not found any consistency ($p>0.05$).

2.3.4 Vergence reaction to disparity-step

In order to extend our analysis of vergence eye movements during the fixation task, we investigated how the vergence system responded to transient changes of binocular disparity. This kind of analysis has already been performed in pre- and post-tests after a stereoscopic visually guided task using head-mounted displays (Hasebe et al., 1996). We first extracted the peak velocity and amplitude of the ocular vergence at the onset of each fixation trial in the high vergence effort condition (90’ of disparity) and in the medium vergence effort condition (45’ of disparity), and checked for correlations between initial phase and velocities (Hung et al., 1997, 1994). Correlations between velocity peak and amplitude of the initial phase are useful to validate that the detection criterion is satisfied and to ensure that eye movements represent reproducible patterns of dynamics with acceptable variance, therein showing normal behavior of the visual system. We dissociated initial phase and amplitude of ocular vergence because amplitude of vergence may contain more than one phase characterized by their pattern of velocity (see figure 2.6). The peak velocity was measured as the slope of the first degree polynomial curve fitted to each initial phase. It is worth noting that the detection algorithm discarded a substantial part of the trials because of the presence of blinks and saccades during the fixation task (almost 60% of the trials were removed from the analysis).

The correlation coefficients between the velocity peak and the amplitude of the intial phase ranged from 0.70 to 0.76. It should be noted that these correlations coefficient appear to differ from those reported in previous studies (e.g., Hung et al., 1997, 1994) for two possible reasons. First, the range of the tested disparities was limited to 0.75 and 1.5 degrees, whereas that of Hung et al (1997, 1994) extended from 2 to 16 degrees. Second, the measured vergence was solicited using two different methodologies (real world vs. stereoscopy), with matched accommodation and vergence distances in Hung's studies and decoupled distances in our study. Nevertheless, the correlations we found between both variables were strong enough to analyze the overall effects. Indeed, correlations between velocity peak and initial phase of the vergence responses are represented in Figure 2.7. As it was observed in real world (Hung et al., 1997), the slope of the amplitude-velocity regression line is in general larger for the convergent movement (crossed disparity-step) than for the divergent movement (uncrossed disparity-step). But these considerations are subject to clarification considering that participants may differ in their fusion abilities either for crossed or uncrossed disparities.

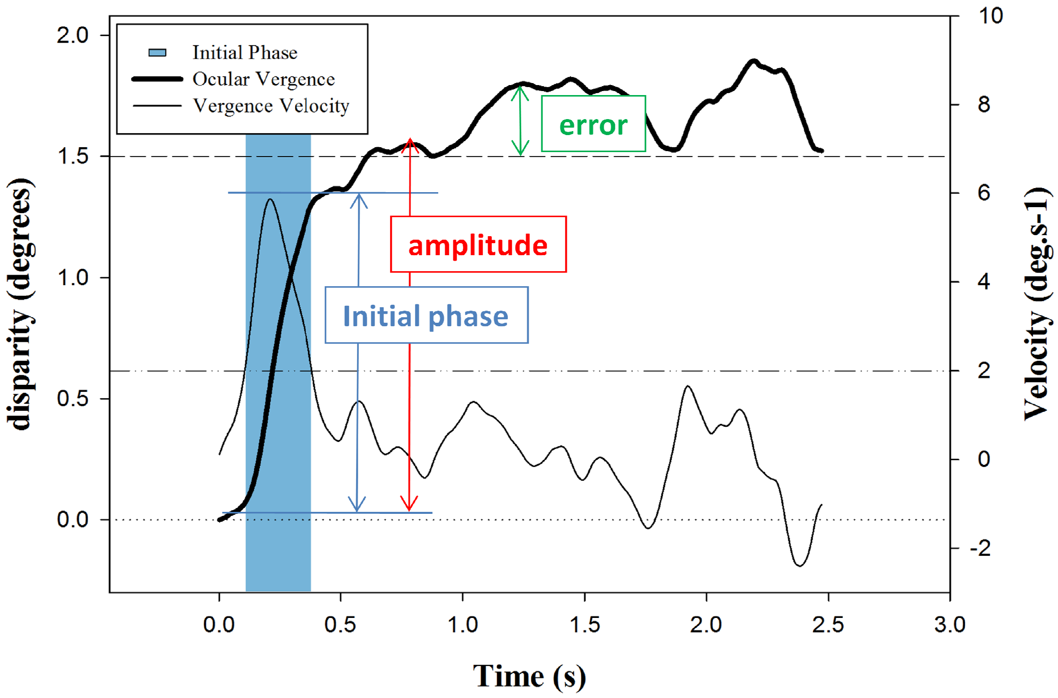

Figure 2.6: Ocular vergence trace and vergence velocity of a binocular reaction to convergent disparity-step of 90 arcmin. The ocular vergence is represented by the thick black line; this is a typical response to a disparity-step of 1.5 degree. The vergence velocity is represented by the thin black line. The initial phase of the vergence movement is indicated with the blue arrow and is defined by the blue region where the velocity goes beyond the fixed threshold.

The mean amplitude of the vergence movement is signaled with the red arrow. The error corresponds to the mean error during the maintenance of the vergence effort and is showed in green. The dashed line shows the intended amount of disparity. The dotted-dashed line shows the velocity threshold used for the detection of the initial phase of the vergence movement.

Hung et al (1997) reported velocity-amplitude relation slopes as being contained between 3.3 to 4 s-1 for convergence and between 1.3 to 2.1s-1 for divergence movements. The values we observed in the present study agree with the measured slope for divergence (2.3 s-1) but not quite for convergence (2.7 s-1) – the values of the slopes are respectively for the block1, block2 and block3: s1=2.7, s2=2.6 and s3=2.7 for convergence, and s1=1.9, s2=2.3 and s3=2.6 for divergence.

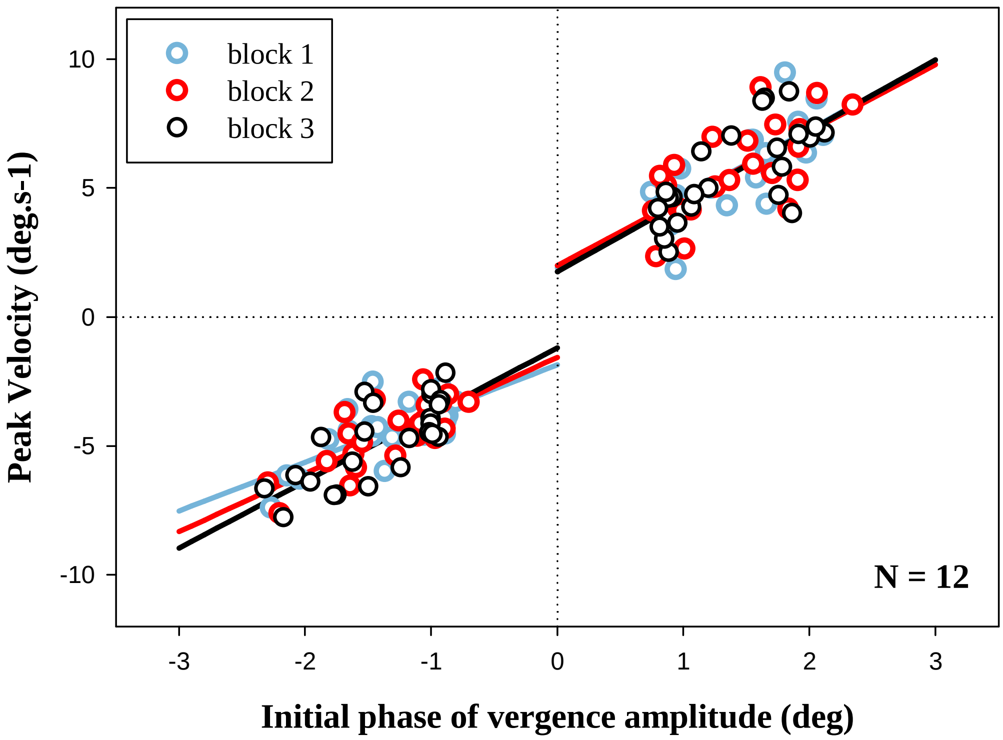

Figure 2.7: Correlations in vergence dynamics for convergent and divergent disparity-step (peak velocity versus amplitude) over the 3 blocks of time session. Coefficients of correlation range from 0.70 to 0.76 across vergence signs and blocks (all correlations are significant, $p<0.05$). The fitted lines are linear regressions (slopes ranged from 1.9 to 2.7 s^{-1}).

Overall effect (N=12)

For the disparity step of 45 arcmin, we did not find any effect of the block sessions on the components of the vergence dynamics for the uncrossed and crossed disparity step (for the velocity peak, the amplitude and the latency, $p>0.05$). Hence, the following analyses will discard the data for the condition with disparity-step of 45 arcmin.

For the disparity step of 90 arcmin, in order to investigate the relationship between convergent and divergent movements, we computed the correlations of the velocity peak between the crossed disparity-step and the uncrossed disparity-step conditions and the correlations of the amplitude between the crossed disparity-step and the uncrossed disparity-step conditions. We found significant correlations with a coefficient correlation of -0.55 for the amplitude and of -0.40 for the velocity peak ($p<0.001$ and $p<0.02$, respectively). As a consequence, we analyzed the absolute values of the velocity peak and the amplitude with crossed and uncrossed conditions taken together. We observed an effect of the time blocks on the velocity peak (Kruskal-Wallis test: $H=12.81$, $p<0.01$, ANOVA: $F(2,1531)=5.24$, $p<0.01$) and on the amplitude (Kruskal-Wallis test: $H=6.51$, $p<0.04$, ANOVA: $F(2,1531)=2.61$, $p<0.05$). The peak velocity was larger in block3 than in the other two ($p<0.05$) and the amplitude was larger in block3 than in group1 ($p<0.05$).

When considering convergent and divergent movements separately, an effect of the block sessions puts forth a time effect only for uncrossed disparity-step (leading to divergence responses) and that, for the velocity peak and the amplitude of the initial phase ($F(2,664)=10.52$, $p<0.001$ and $F(2,664)=4.38$, $p=0.013$, respectively) but not for the latency ($F(2,664)=1.31$, $p>0.05$). More precisely, the pairwise comparison analyses revealed that the velocity peak (absolute value) was higher in block 3 than in blocks 1 and 2 ($p<0.05$) and the amplitude (absolute value) was also greater in block 3 than in blocks 1 and 2 ($p<0.05$). We did not find any effect of the block sessions on the vergence dynamics components for the crossed disparity step (leading to convergence responses) (for the velocity peak, the amplitude and the latency, $p>0.05$).

The velocity peak was greater for the crossed disparity-step than for the uncrossed disparity-step (Rank Sum test: $U=205923$, $p<0.001$; t-test: $t=10.849$, $p<0.001$). However, the amplitude of the initial phase could not be statistically contrasted between the crossed and uncrossed disparity-step conditions ($p>0.05$).

Grouping analyses

Given the heterogeneously shared effect of the blocks on the observers, we tried to put forth a few differences thanks to the grouping of participants. We used the previous criterion based on the confidence in the response of the stereo-task that was used for the analysis of the psychophysical data. For convergent movements, we observed that the amplitude could not be differentiated between groups 1 and 2 ($F(1,765)=1.96, p=0.16$) but that the velocity peak was larger in group 1 than in group 2 ($F(1,765)=13.7, p<0.001$). For divergent movements, we found that the amplitude was larger in group 1 than in group 2 ($F(1,588)=101.32, p<0.0001$), and that the velocity peak was more important in group 1 than in group 2 ($F(1,588)=24.47, p<0.0001$). We then tested the effect of the time blocks. For the crossed disparity-step condition, both groups did not show any difference in the velocity peak and in the amplitude over the three blocks (respectively for group1 and group2: $F(2,428)=1.246$, $p>0.05$, $F(2,337)=1.273$, $p>0.05$ for the velocity peak, $F(2,428)=0.031$, $p>0.05$, $F(2,337)=0.768$, $p>0.05$ for the amplitude). For the uncrossed disparity-step condition, group1 showed only a significant effect for the velocity peak ($F(2,319)=8.209$; $p<0.001$), the absolute values of the velocity peak were smaller in block 1 than in blocks 2 and 3 ($p<0.05$ and $p<0.001$, respectively) but the absolute values of the amplitude could not be differentiated between the blocks ($F(2,319)=0.872$; $p>0.05$). For group2, the velocity peak could not be differentiated between the blocks ($F(2,269)=2.747$, $p>0.05$)) but an effect of the block on the amplitude was revealed ($F(2,269)=6.232$, $p<0.01$), the absolute values of the amplitude was significantly larger in block 3 than in block 1 ($p<0.001$). In summary, the first group showed an effect on the velocity peak while the second group showed an effect on the maximum amplitude.

2.4 Discussion

In the present study, participants tracked a fixation target whose binocular disparities changed over time. Interestingly, they were unable to report any displacement in depth – either the large initial step or the motion-in-depth – but a few observers sometimes mentioned a kind of shaking of the cross. It is worth noting that the size of the cross remained unchanged while disparity actually changed. In so doing, we generated a cue-conflict between the vergence angle specifying a particular depth-position and the relative change in its size it had to undergo. Moreover, it is also possible that the depicted disparity conveyed different perceived sizes of the fixation cross (for example, between crossed and uncrossed disparities) since it is well known that perceived size of objects cast on the retina increases with decreasing vergence angle and decreases with increasing vergence angle (Adams, 1955; Hermans, 1937). This phenomenon of "zooming" is described by the Emmert's law. The change in perceived size observed with change in disparity has been considered to result from a binocular size constancy mechanism (Julesz, 1971) so as to provide size constancy in an attempt to stabilize the size of the percept. Heinemann et al. (1959) have investigated the role of oculomotor adjustments on apparent perceived size and showed that in absence of any change in eye accommodation or pupil size, the variations exerted on the vergence angle lead to modification in the apparent object's size. On the basis of the position of the gaze alone, it seems that the apparent size may vary independently of the retinal size. Given that the potential change in size was perceptual, this is actually the vergence mechanism that causes the size scaling (Julesz, 1971) and, as a consequence, no motion in depth was perceived during the vergence effort part (see for example, Wismeijer & Erkelens, 2009). The perception of motion-in-depth can be achieved in the presence of congruent and sufficient cues and is not effective in presence of binocular disparity alone (Brenner et al., 1996; Erkelens & Collewijn, 1985). The apparent task of the participants in the vergence effort part of the experiment was thus akin to a fixation task of a cross whose displacements in depth were only "unconscious".

The aim of the present study was to test the effect of visual fatigue generated through the presentation of changing horizontal binocular disparities on the perception of 3D shape. We looked for possible signs of visual fatigue or discomfort as a subjective measure (the Simulator Sickness Questionnaire, Kennedy et al. 1993) so as to check that participants really felt fatigue after watching our stimuli. The originality of this study was the use of a vergence effort task (i.e. in order to fatigue the binocular system) that was interleaved with a stereo-judgment task (i.e. to observe potential effects on depth perception). In this way, we found that the main symptoms were largely dominated by oculomotor factors rather than the components of nausea and disorientation. We observed that global scores for the SSQ were indeed larger in the condition of highest vergence effort as compared to the others. Unfortunately we failed at observing an interaction effect between the scores of each symptom category of the SSQ and the conditions of vergence effort. From these subjective results, one can conclude that possible symptoms related to stereoscopic visualization focus around the oculomotor sphere. This is not really surprising because the head of subjects was restrained in a chin rest to optimize the eye-tracking precision and thus, stereoscopic induction of motion sickness was minimized. However, this confirms that a wide variety of symptoms may be experienced following long and demanding observation of 3D stereoscopic content on classical displays (i.e. eyestrain, difficulty of focusing, blurred vision), and that our experimental procedure was strong enough in terms of propensity to produce visual fatigue.

3D shapes were depicted using random-dot stereograms so as to isolate the effect on the disparity processing mechanism only. Using a stereo-categorical-judgment paradigm, the findings reported in the present study revealed that participants, on average, did not provide any overall tendency in the bias direction following the vergence effort task – but showed an effect on the reaction time. However, by grouping participants according to the variation of their uncertainty in the response (as a function of the slope at the point of subjective equality), we discover two tendencies. The rationale was that visual fatigue could affect the perception of a 3D dihedron in two different ways: an increase of the uncertainty that would reflect a state of fatigue without adaptation, and a decrease or unchanging uncertainty that could reflect fatigue with adaptation. Hence, two trends could be observed after grouping participants in these two groups. Before examining the results on the bias direction in the estimation of the 3D shape, it is worth defining what corresponds to over- and under-estimation of the dihedral angle. As a matter of fact, when an observer estimates the right angle as being smaller than 90 degrees (PSE less than 90°), s/he under-estimates the angle. In this case, the amount of disparities mediating the 3D shape required to perceive a right angle is increased as compared to a correct estimation, and so, the observer under-estimates the binocular disparities and s/he is less sensitive to horizontal binocular disparities. In contrast, when the participant judged the dihedral angle as larger than 90 degrees, s/he over-estimates the angle and the binocular disparities. The group analysis revealed that the first group (increase of uncertainty with vergence effort) showed a gradual underestimation of the dihedron angle with the increase of the vergence effort. This group showed the misestimation of the dihedral angle with a bias to underestimate depth from binocular disparities. As a matter of fact, observers judging the angle smaller than 90° did require more disparities to perceive a right angle when the vergence system was more solicited. In this case, the dihedral angle was judged as if observers were less sensitive to binocular disparities. Participants were placed in group2 because of an increase or no change in the slope of their psychometric function as the demand on the vergence effort increased. This group showed a tendency which was the opposite of the first group. The dihedron angle was overestimated (judged larger) since its shape was perceived more expanded in depth.

One kind of perceptual changes reported in 3D estimation has already been shown in previous studies (e.g. in Blakemore & Julesz, 1971). Hence, adaptation to binocular disparities after fixation could lead to stereoscopic after-effects where random-dot stereograms used as adaptation stimuli influence the perceived depth of next similar stimuli. These after-effects have been explained as a consequence of fatigue or decrease in responsiveness of disparity detectors (Domini et al. 2001). A broad collection of studies suggests the existence of disparity detectors in the human visual cortical system, specifically selective for a category of binocular disparities (i.e. Ohzawa et al. 1990; Regan & Beverley, 1973; Fiorentini & Maffei, 1970). For example, adaptation to crossed disparities with a stimulus in front of the monitor plane will lead to an after-effect to perceive an object at zero disparity behind the fixation plane, presumably because of a two-channel organization of the cells sensitive to binocular disparities (Long & Over, 1973). According to Long & Over (1973), this situation could create a brief imbalance in the two-channel mechanism because of the tiring of the group of disparity detectors that are tuned to this category of disparities. According to Howard and Rogers (1995), even a multiple-channels organization for disparity detection would exist in the visual system. Misinterpretations of binocular disparities might be the consequence of the fatigue of the disparity detectors. Therefore, in case of strong demand for the binocular system, it could be possible that the decreasing efficiency in discriminating stereoscopic patterns will be due to the fatigue of these disparity detectors.

Central to the subject of the present paper, the fact that these results on the two possible shifts in the mean PSEs over the vergence effort conditions are really triggered by visual fatigue is corroborated by the results on the reaction time. As it was previously shown (Hoffman et al. 2008), observation conditions with decoupled accommodation and vergence may affect the time required to fuse a random-dot stereogram. Our study shows that a strong stimulation of the vergence system under stereoscopic conditions affects the time to respond when watching a random-dot stereogram as a function of the intensity of the vergence effort and of the time elapsed.

In order to enlighten the effect of visual fatigue on the perceptual estimation, we also converted the estimated thresholds in scaling distances; the distances at which the pattern of disparity matching with the perceived depth should specify a right vertical angle. This kind of analysis is often realized to account for the estimated fixation distance that observer must consider for disambiguating the disparity signal (e.g. Watt et al., 2005; Johnston, 1991). The rational was that the oculomotor adjustments could be reflected in the misetimation of the egocentric fixation distance, as for example a change in the accommodation state. For instance, it was already suggested that accommodation could adopt a state approaching its resting position or a "specific distance tendency" in open-loop conditions and affect perceptual estimation (Frisby et al., 1995). One possible explanation of the results of the present study is that the two tendencies found in perceptive estimations are related to biases in estimating the fixation distance used to scale disparity (overestimation or underestimation). These misestimations of the fixation distance could be related to the effect on the oculomotor system. As a matter of fact, any adaptation of the oculomotor system could at least indirectly influence either accommodation or vergence or both together and lead to fluctuations in the estimation of the fixation distance used to scale disparity. As such, it was already shown that distance perception can be influenced by adaptation of the oculomotor system (Wallach et al., 1972). Therein, the absence of consistent effect on lateral heterophoria disagrees with that as it is related to the tonic component (e.g. Owens & Tyrell, 1992). However, an effect of the vergence effort on fixation disparity was observed, which however could not explain the bias directions in stereo-estimation. Specifically, even though tonic vergence is set to the average disparity of the stimulus (Rushton & Riddell, 1999), this does not explain why a part of the participants estimated their fixation distance beyond the screen plane. In any case, changes in the accommodation state could also be responsible for the bias direction found in the stereo-estimation. As a matter of fact, it is well known that the thickness of accommodative lens varies as a function of the fixation distance. Specifically, for far viewing, the lens is relaxed and less rounded, the focal length tend to be maximal whereas in near viewing, the crystalline lens becomes more rounded because more refraction is needed for focusing. It is worth noting that the accuracy of the accommodation response also varies among observers and that a certain amount of accommodative error can be tolerated as long as the focusing distance remains in the depth of field. As a consequence, the bias direction in 3D estimation reported in this study could be related to the changes in the oculomotor system in such a way that vergence or accommodation errors affected the estimation of the fixation distance.

One of our main objectives was to identify aspects in eye movements characterizing stereoscopic activity and its effect on the visual system. The most common method used to evaluate visual fatigue in stereoscopic visualization involves subjective questionnaires, while a lot of objective methods are known to measure or to infer visual fatigue (e.g. error in accommodative refraction, pupil size variation, and blink rate). Here we investigated the components of vergence dynamics in the vergence reaction to disparity at the onset of each trial as they may relate to vergence error (Hasebe et al. 1996). As such, we dynamically measured vergence movements during the vergence effort parts of our experiment. We also examined phoria and fixation disparity in pre- and post-test so as to find out any possible relationship between the bias in the stereo-estimation and the adaptation of the vergence system. We failed at observing any consistent effect on near heterophoria whereas the literature previously provided evidence for an esophoric shift (e.g. MonWilliams et al, 1993). The lack of result on near heterophoria in a stereoscopic configuration could be explained by the fact that the visual system quickly recovers its normal state when it is immersed in a natural world. However, we cannot exclude that the present clinical test was not biased by accommodative vergence and, some previous studies showed such effect on phoria measurement (e.g. Owens & Tyrell, 1992). If it is the case, the absence of results on phoria could have been due to accommodative convergence which suggests no crosslink adaptation. However, we found an effect on fixation disparity when considering only the post-tests – that may be due to the lack of repeatability of the clinical set-up. Moreover, our results demonstrated that vergence demand in general modified the components of vergence movements; the velocity peak and the amplitude of the initial phase of the vergence reaction for crossed and uncrossed disparity-step were changed over time. It is worth noting that the study of these two components across the vergence stimulation offered two important indicators of the vergence system efficiency. Indeed, the velocity peak of the vergence reaction is better related to the participants' ability to quickly fuse the stereoscopic stimulus; in contrast, the amplitude of the initial phase is more connected to the error of the vergence system. Hence, the increase in the velocity peak of the vergence reaction will be informative of the increasing ability in performing the task while the increase in the amplitude will mediate an increase in the error of the vergence system. It is important to note that the correlations obtained between the velocity peak and the amplitude of the initial phase were slightly different than those reported for real world conditions (see Hung et al. 1997). Specifically, we observed that the slopes of the regression lines of the velocity-amplitude relationship were consistent with those obtained in real world for uncrossed disparity-step but not for crossed disparity-step, the slopes were diminished for the latter. This result should be added to the fact that the effects on the velocity peak and on the amplitude were broadly due to the variation in the divergent movements as opposed to that of convergent movements. Although the slopes of the regression line were reduced in the stereoscopic viewing conditions as compared to those of real world, they were also more stable over time, and this calls for a more sturdy mechanism. The components of vergence dynamics were only affected in the block sessions for uncrossed disparity-step, the vergence reaction tended to be faster and of greater amplitude over time. It is possible that the convergent system showed more stability than the divergent system as a consequence of its adaptive value in detecting objects that quickly approach and threaten the peri-personal space (Hung et al. 1997).

The individual analysis of the vergence dynamics demonstrated clear individual differences in the fluctuations in the components of vergence. As a consequence, we were interested in comparing the results of the vergence dynamics for the grouping we previously made in the analysis of the perceptual estimation. We observed that the first group showed an effect on the velocity peak while the second group showed an effect on the maximum amplitude. Clearly, group1 (classified according to the tendency of a decrease in stereo-estimation sensitivity, i.e. decrease in the slope at the PSEs) showed an increase in the velocity peak over time blocks while the group2 (with absence of change or increase in precision) showed an increase in the amplitude of the vergence response over the blocks. It should be noted that although group2 exhibited an increase in the maximum amplitude of the vergence, the amplitude was on average smaller to that of group1. The group1 was overall less sensitive to binocular disparity after stimulation of the vergence system and, showed consistently more important amplitude and greater velocity peak in the vergence reaction than group2.

In summary, vergence dynamics evaluated after vergence effort in the present study appeared interestingly faster, but the amplitude of the vergence reaction appeared increase over time. The first group presented an increase of the reaction time to disparity but with spatial overshoot, that is, amplitude response that exceeded the stimulus amplitude. This overall error could be related to the decreasing ability to interpret binocular disparities for the judgment of the 3D dihedron. In general, the findings of the present study support the conclusion that vergence disturbance may, at least indirectly, alter higher level stereo-processing leading observers to misestimate the 3D shape of stereoscopic objects.

2.5 Conclusion

In the present study, we showed that both the perceptual estimation in judging a 3D shape and specific components of eye movements were influenced by the extent to which the vergence system was solicited. Depending on the participants, two distinct trends in the bias direction (over-estimation or under-estimation of binocular disparities) were observed when participants underwent large demands and changes of disparity, and the reported symptoms were largely directed toward oculomotor factors. These two possible trends could be connected to the inter-individual difference in the capacity to respond to stereoscopic stimuli. Additionally, we found that the components of the vergence reaction were changed over time when the vergence effort was sufficiently large for the visual system. An analysis of subjects grouping helped us argue that two trends in the bias direction may be connected to fluctuations in the components of vergence. Hence, when participants underwent large demands and changes of disparity, the reported visual symptoms were largely increased and, perception of the 3D shape changed according to individual capacities. Overall, these findings suggest that visual fatigue may affect stereopsis as function of individual adaptive capacities.

♦

The present study was an exploratory one. The first objective of this study was to investigate the potential effect of prolonged stereoscopic demand on the ability to perceive 3D shape from stereopsis. The second was, by tracking eye movements during vergence effort, to find out possible changes in the oculomotor system that could reflect the decrease of its efficacy that is accompanied by visual symptoms. To address these questions, we used a stereoscopic ocular fixation task to strongly stimulate the visuomotor system and an interlaced 3D shape estimation task to observe subsequent effect on 3D perception. Results did not reveal an overall tendency when the whole pool of participants was considered. However, after ranking participants in two groups according to their uncertainty in doing the task, we found two main tendencies. The perception of 3D shape was progressively affected by the strain of the visuomotor system. One group exhibited increasing perceived dihedral angle whereas the other showed decreasing perceived angle. The analyses of the components of the vergence reaction for each trial also revealed several changes in vergence skills. Hence, when participants underwent large demands and changes of disparity, the reported visual symptoms were largely increased and perception of the 3D shape changed according to individual capacities. As a consequence, these results suggest that visual fatigue may affect stereopsis depending on individual adaptive capacities.

This first study showed some changes on vergence eye movements after the prolonged exposition to stereoscopic vergence effort. In that, this first experiment motivated the followings presented in the next chapter and where, this time, vergence dynamics were compared both in stereoscopic and real-world conditions and, in pre- and post-tests to evaluate the effect of a maintained stereoscopic activity.

Chapter 3: Ocular vergence in stereoscopic displays

3.1 Introduction

The binocular visual system holds a potent role in bringing the gaze axes toward the object of interest. Using vergence, the disjunctive rotational eye movements, one can detect relative depth and distance information referring to the surrounding world. In natural vision, vergence eye movements are rather slow (up to 40° per second) as compared to saccadic movements (up to 700° per second). Even though very short latency have been observed in a somewhat machinelike way (Busettini, Fitzgibbon & Miles, 2001), vergence eye movements usually show latency of about 100 to 200 ms depending on the direction, the initial starting position and the predictability of the target vergence (Alvarez, Semmlow & Pedronoo, 2005; Heron, Charman, & Schor, 2001; Hung, Zhu, & Ciuffreda, 1997; Semmlow & Wetzel, 1979). In binocular vision, vergence eye movements are dominated by the retinal disparity information, the slight differences projected onto the retinas of the left and right eyes. However, vergence eye movements are composed of several components: the disparity, blur-driven, proximal and tonic vergences (Howard & Rogers, 1995). The accommodative vergence is the potent secondary information for the vergence system. It can be easily demonstrated that monocular blur stimulation can elicit vergence response when actually only one eye is stimulated (Hung, Semmlow & Ciuffreda, 1983). This component together with the disparity-driven accommodation is acknowledged to be the outcome of interactive neural processes ensuring that both eye-movements respond to either signal (Semmlow & Hung, 1981). Semmlow and Wetzel (1979) investigated the dynamics of vergence eye movements in response to step stimulation defined both by disparity and defocus information presented together or separately. The target vergence steps were between 0 and 3D (Diopters) for accommodative vergence and between 0 and 3MA (Meter angles) for disparity vergence. They showed that both components of vergence responses were additive even if the disparity vergence dominated. As a matter of fact, by comparing vergence responses with or without the accommodative stimulus, accommodative vergence showed to be involved in the sustained vergence response but much more in the dynamics response.

Here the focus is on the dynamic activity of the two major vergence components produced either in real world or stereoscopic viewing conditions. Translational motion of a target in depth normally provides a stimulus for both the disparity-driven and the blur-driven vergence. In natural vision, these two components are combined to provide an accurate and fast response of the oculomotor system. The nature of this combination in stereoscopic displays is rather unknown, but is very important for content design, considering that such displays create ambivalence between the two stimulations (Howarth, 2011; Kim et al., 2011; Hoffman et al., 2008; Ukai & Howarth, 2008; Akeley et al., 2004; Rushton & Riddell, 1999; Wann, Rushton, & Mon-Williams, 1995). In conventional stereoscopic displays, different vergence angles are induced by the binocular disparities displayed and give rise to the perception of objects layout while the stimulus-to-accommodation specifies a unique range of distances around the screen plane position (linked to the depth of field). These stereoscopic viewing conditions thus may alter the normal execution of the visual system. For instance, Hoffman et al. (2008) used a multiple focal planes 3D display so as to manipulate the difference magnitude between accommodation and vergence in binocularly-viewed pictures. They showed that binocular fusion was slower, stereoacuity thresholds were worse, and scores on a visual comfort scale were decreased when accommodation and vergence cues were increasingly unbalanced.

Predicting visual discomfort in stereoscopy is thus of major concern for an optimal use of stereoscopic viewing systems. Several studies have put forth significant changes in oculomotor functions after prolonged exposition to stereoscopic contents (i.e. Neveu et al., 2012; Shibata et al., 2011; Ukai & Howarth, 2008; Emoto, Nojiri, & Okano, 2004; Eadie et al., 2000). Hence, the most usual explanation is that stereoscopic displays yield incongruent stimulation that forces viewers to decouple accommodation and vergence. As stereoscopic content engenders numerous repeated vergence responses, disparity-driven accommodation also reacts to the displayed content (Fukushima et al., 2009). The inability for viewers to weaken the frequency of these superfluous accommodation responses and thus to adapt their visual system could be the predominant cause of visual fatigue in stereoscopic 3-dimensional conditions. Previous studies generally used very large conflicts between accommodation and vergence so as to show the adverse effect. The rationale was that conflicts outside the zone of clear and single binocular vision (ZCSBV) would entail much more discomfort and fatigue than conflicts in this area. As a consequence, studying the effect of stereoscopy on the visual system showed changes in oculomotor responses with potent conflicts between accommodation and vergence for many subjects while a consistent part of viewers cannot fuse the stimuli. It is worth noting that, even though disparity remains in the comfort zone (e.g. that is defined as the middle third of the ZCSBV), viewers may experience discomfort or fatigue especially in presence of motion-in-depth (i.e. Speranza et al., 2006). Hence, the previous studies' need of large conflicts between accommodation and vergence was possibly related to the absence of motion-in-depth. In the present study, we manipulated stimuli depicting rather small conflicts between accommodation and vergence – sometimes slightly exceeding the zone of comfort – but, requiring potent vergence changes for fixating the disparity-defined target oscillating in depth.

Several methodologies have been employed so as to objectively estimate the subjective visual annoyance that observers may experience after watching stereoscopic content. These methods range from questionnaires to clinical measurements, and binocular eye-tracking (i.e. Neveu et al., 2012; Vienne, Blondé & Doyen, 2012; Lambooij et al., 2009; Shibata et al., 2011; Hoffman et al., 2008; Mon-Williams & Wann, 1998, Peli, 1998). For example, Shibata et al (2012) aspired to identify subjective zones of comfort on the basis of questionnaires but they found more reliable indicators in the individual estimate of the ZCSBV. Their objective was to predict visual discomfort along a range of viewing distances and they found correlations between discomfort and clinical estimates of the ZCSBV, but they observed that subjective reports of symptom severity alone did not suffice to characterize susceptibility to visual adversity. Recently, Neveu et al. (2012) measured oculomotor performance during stereoscopic viewing of in-depth oscillating target and found an overall decline in the amplitude of the vergence response over time. The eleven participants however reported very few complaints of visual annoyance. Emoto et al. (2004) found that one hour of stereoscopic viewing of content with various disparities produced significant changes in fusional limits (a decrease, measured through break points) but not in AC/A ratio (accommodation-induced change in vergence per change in accommodation in the absence of disparity stimulus). In Mon-Williams et al (1993), signs of binocular stress – changes in distance heterophoria and associated phoria – were also observed thanks to clinical measurements following ten minutes of stereoscopic presentation on a HMD (head mounted display). Almost all of these methods however did not directly relate to the activity of the vergence system which is at stake when viewers have to fuse stereoscopic content. The study of the vergence response should allow establishing a relationship between the outcomes of stereoscopic-induced visual fatigue and depth perception. To this end, we investigated whether vergence response to disparity step could be used as indicator of visual fatigue consecutive to prolonged observation of stereoscopic content. The advantages are that vergence step responses can be easily implemented using the same imaging method as the stereoscopic content.

Previous studies on vergence eye movements to disparity-step are generally dealing with large vergence disparity-steps (i.e. Hung, Zhu & Ciuffreda, 1997; Hung et al., 1994), therefore providing strong significant differences and correlations when analyzing the main sequence. In the present study, the range of target vergence step was much smaller, because the objective was to investigate the possible vergence step in stereoscopic 3D, that is, target vergence that does not give rise to diplopia. For most viewers, motoric fusion becomes very limited beyond 1.5 degree of disparity. Our objective was that the tested vergence demands fall within the range of the zone of clear single binocular vision for any observer. The first experiment was conducted in order to investigate the potential difference between vergence responses in both stereoscopic and real-world conditions of observation. The second experiment was designed to evaluate the stability of vergence dynamics following the observation of a specific stereoscopic stimulation.

3.2 Experiment 1

This first experiment was conducted in order to investigate the potential difference between vergence disparity-step in stereoscopic and real-world conditions of observation. Participants performed disparity-step vergence movements of various amplitude both for stereoscopic stimuli and for real-world stimuli. On one hand, stereoscopic stimuli provided incongruent disparity and blur information for the second fixation position (in front or behind the screen plane). On the other hand, real-world stimuli provided matched disparity and blur information of the to-be-fixated target. Hence, two components were supposedly involved in vergence reactions: the disparity-vergence and the accommodative vergence. In order to compare stereoscopic and real-world conditions of observation, we analyzed the main sequences and correlations between amplitude and peak velocity of vergence responses. The two conditions were expected to provide significant differences that could be the results of the accommodative vergence component that was shown to occur later than the disparity-driven component (Semmlow & Wetzel, 1979).

3.2.1 Method

3.2.1.1 Participants

A total of 14 observers took part in the study. Two participants were discarded, both because they revealed very poor performances in judging relative disparities (under the chance level) and because of their difficulty in fusing the content, presenting very long and somewhat marginal reaction times. Two more were discarded because they had potent difficulties in undertaking the task both in S3D and R3D (Realistic 3D) conditions (less than 50% of trials were valid). The ten remaining participants were tested according to a full counterbalanced order; they were on average 29.3 years old. All had normal or corrected vision and presented stereoacuity threshold at least inferior to 30 arc minutes as assessed by the Randot Stereo Test. They gave their informed consent before beginning the experiment.

3.2.1.2 Apparatus

For the purpose of the experiment, we designed a specific apparatus depicted in Figure 3.1. The system was composed of a beam splitter (360x255 mm, Edmund Optics) tilted 45 degrees to the sagittal plane in front of the eyes of the participants and, perpendicular to the sagittal plane, a long optical bench (2.8 meters) used to move a 2D-screen (Dell 1908FP 19'' LCD monitor) at the desired vergence distances. Participants' head was placed in a chinrest that was placed 1.3 meters away from the 3D display (Hyundai S465D 46" HDTV LCD Polarized monitor) on an optical table (Newport, 120x90 cm) which served as firm mechanical connection for all elements of the system. The center of both displays was carefully aligned along the subject's midline using visible light. For each vergence distance, we visually checked that alignment was correct by displaying a set of vertical and horizontal lines crossing at the center of each screen. In so doing, we verified that the two displays were aligned in the horizontal and vertical dimensions. Binocular eye-movements were recorded using a binocular eye-tracker (Eyelink 1000 with 2000 Hz upgrade, SR-Research). In order to minimize crosstalk, we placed participant eyes on the axis perpendicular to the center of the screen. We used two computers in this experiment, the first displayed 3D content using Matlab and the Psychtoolbox extensions (Brainard, 1997), the second monitored the gaze tracking. Participants wore polarized glasses to fuse left / right views according to each corresponding eye and could perceive the content in depth; the stereoscopic displaying method was to present left/right views interleaved line-by-line. Participants were tested in total darkness, such that only our display-generated stimuli were available for vergence responses.

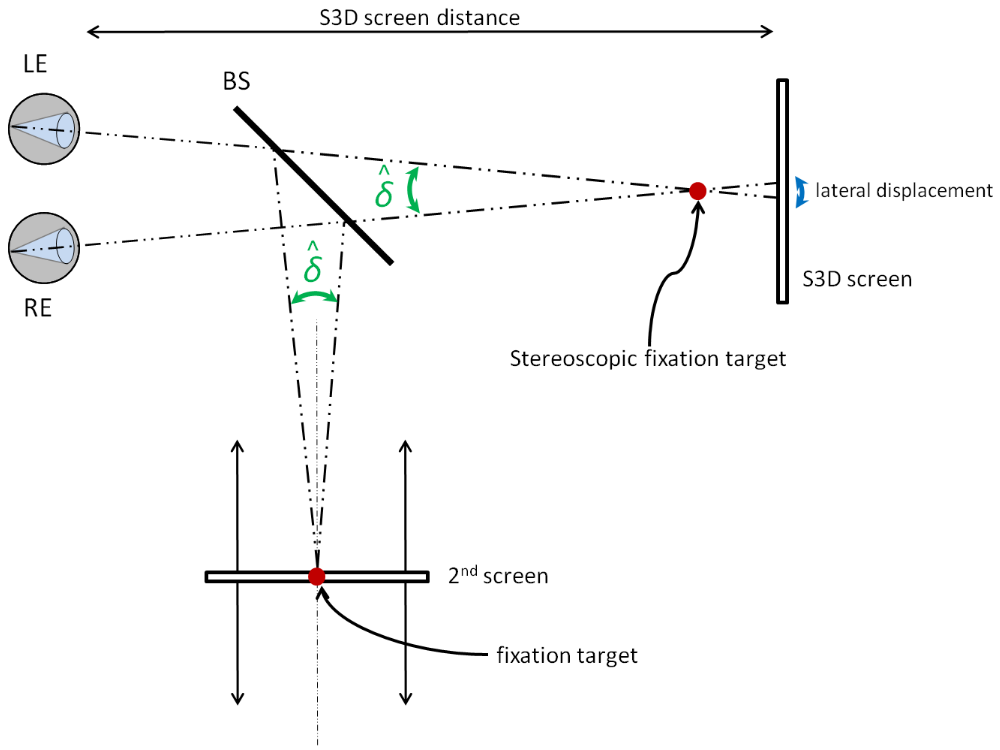

Figure 3.1: Schematic representation of the experimental apparatus used in the study. S3D refers to stereoscopic 3D display, BS to Beam splitter. The second display presents the fixation target for the R3D. The second display distance was adjusted on the bench to match the 3D stereoscopic distance used in the S3D condition. Drawing is not to scale.

3.2.1.3 Procedure & Stimuli

Vergence responses were measured under two viewing conditions of combined disparity-blur stimulation: (1) screen-constrained blur stimulus (S3D) and, (2) congruent disparity-blur stimulus (R3D; the "R" stands for Realistic because this condition is closer to reality than the purely stereoscopic condition). For each participant vergence response was measured for convergent and divergent disparity-step stimuli of 0.75, 1.0, 1.25 and 1.5 degrees. Each level of this condition was repeated ten times in blocks presented in counterbalanced order. Hence, 80 trials were performed in each S3D and R3D conditions (a total of 160 trials). We used a fixation target (35 arc minutes radius) formed of a white fixation cross (18 arc minutes) surrounded by a frame composed of small squares of various shades of grey, approximately 5 by 5 degrees to help and maintain stereoscopic fusion (Vienne et al., 2012). The starting vergence position was located in the middle of the S3D screen plane 1.3 meter away from the observer (convergence angle of 2.8 deg). Hence, each trial began with a zero-screen-disparity fixation target whose presentation duration was randomly either 0.5, 1 or 1.5 seconds so as to limit any anticipation behavior. This pre-vergence step period was also used to weaken post-saccadic enhancement effect on vergence (Busettini et al., 1994). For the S3D condition, the left and right views of the fixation target were then shifted according to the eight possible disparity values. For the R3D condition, the central portion of the fixation target was moved on the second 2D display so as to simulate the disparity step with appropriate defocus signal. The vergence step duration was 2.5 seconds. The fixation target was then again presented with zero-disparity so as to control the vergence has returned back to its normal position (see Figure 3.2).

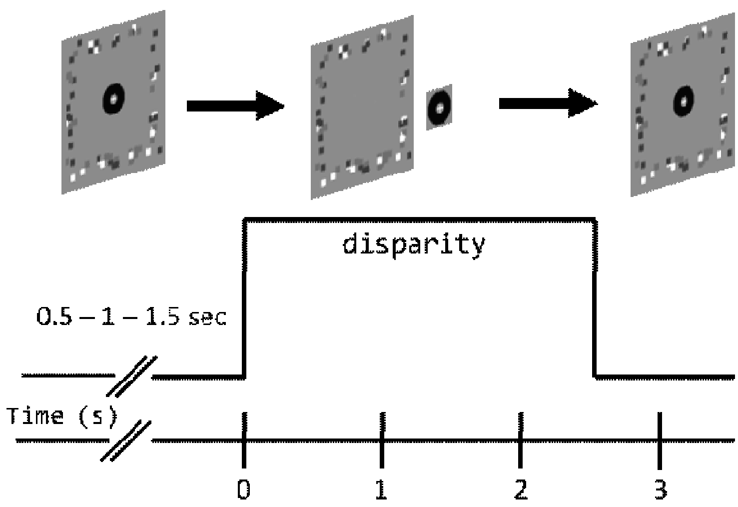

Figure 3.2: Experimental schedule of the disparity-step phase used in measuring vergence eye movements.

3.2.1.4 Data Analysis

Before each analysis, we discarded all data containing blinks or missing samples. Our analysis focused on the amplitude of the initial phase, the velocity peak, the latency of the vergence eye movement, the position of the maximum velocity peak and the time constant (see Figure 3.3). The detection of the dynamics components was carried out on the basis of the velocity time series, as usually done in the eye movement literature (i.e. Engbert, 2005). We developed a semi-automatic algorithm for the detection of each movement in the vergence data (Left-eye gaze – right-eye gaze) which requires user interaction only when it did not converge according to critical requirements. The vergence movement must be detected between the stimulus onset time t0 and t0 + 500ms, with the same direction as the target vergence and, having amplitude within ±20 arc minutes relative to the target amplitude. The time series of eye positions was transformed to velocities and smoothed using a weighted moving average over ten data samples adapted to our high sampling rate. We computed the velocity threshold as being twice the standard error based on the median of the velocity time series to both protect the analysis from possible noise and to take into account the nature of the signal in the events detection. A valid vergence eye movement was defined as having a minimum magnitude of 30 arc minutes and a minimum duration of 100 ms. The visual checking of the algorithm detection ensured binocular vergence was efficiently detected among the time series. The total number of trials used in the analysis was 1454 (about 91% of the entire dataset). It is important to note that, in this study, convergence is presented as negative and divergence as positive because, by convention, target vergence is defined by left eye position – right eye position.

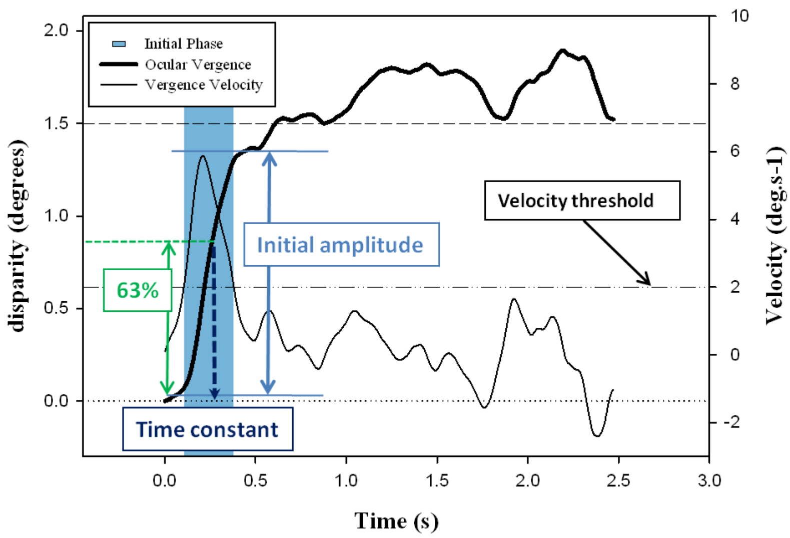

Figure 3.3: Ocular Vergence trace and vergence velocity of a binocular reaction to divergent disparity-step of 90 arc minutes. Ocular vergence is represented by the thick solid line and vergence velocity by the thin solid line. The blue region is the part of the time series where velocity exceeds the velocity threshold (the dashed line on the second vertical axis). The initial phase of the amplitude and the response time constant are represented with the blue and dark blue arrows.

3.2.2 Results

The relationship between the velocity peak and the initial amplitude of each vergence response (also called main sequence) is represented in Figure 3.4 and, therein it can be seen that vergence dynamics is comparable in the two conditions as they exhibit similar distributions. For the R3D condition, coefficients of correlations are 0.82 for divergence and 0.79 for convergence and the slopes are 3.03 s-1 and 2.27 s-1 respectively (all correlations are significant $p<0.001$). For the S3D condition, coefficients of correlations are 0.78 for divergence and 0.69 for convergence and the slopes are 2.71 s-1 and 2.05 s-1 respectively (all correlations are significant $p<0.001$). It should be noted that these correlation coefficients appear slightly smaller compared to those reported in previous studies. The reason could be that the range of the tested disparities was limited from 0.45 to 1.5 degrees, whereas the ones of Hung et al (1997, 1994) extended from 2 to 16 degrees.

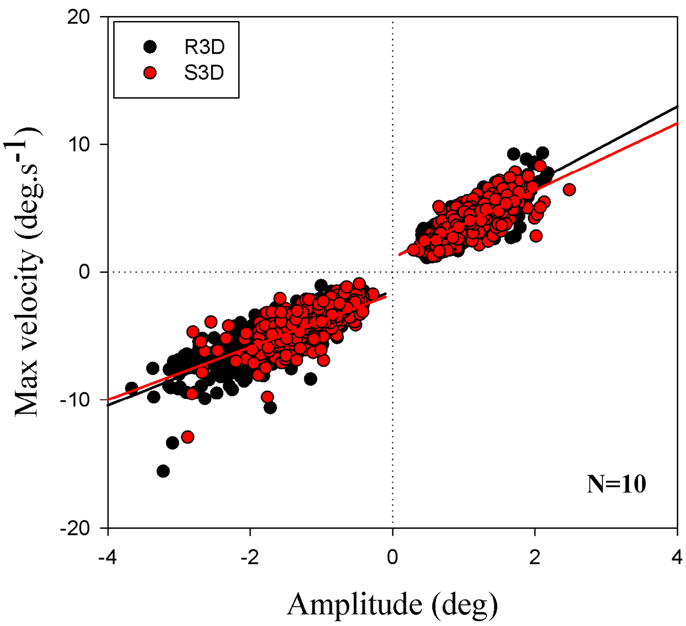

Figure 3.4: Plot of main sequence of peak velocity vs. amplitude data points, for symmetric vergence step responses for all participants combined. The figure also represents linear regression line (solid red and black lines) for convergence (negative amplitudes and peak velocities) and divergence (positive amplitudes and peak velocities) in both the R3D (real world) and S3D (stereoscopic) conditions. The plot contains 1454 observations for ten participants.

Independent ANOVA performed on latency for the ten participants revealed a significant effect of the viewing conditions ($F(1,1438)=29.05$, $p<0.0001$) and of the range of disparities ($F(7,1438)=3.65$, $p<0.001$). The latency was 106 ms (±54.3 SD) in the R3D condition and 121 ms (±51.8 SD) in the S3D condition. Additionally, an interaction effect was present between the viewing conditions (R3D vs. S3D) and the range of disparities ($F(7,1438)=4.76$,

$p<0.0001$) and between the viewing conditions and the vergence direction (convergence/divergence) ($F(1,1438)=25.09$, $p<0.0001$). ANOVA realized on the (unsigned) velocity peak for the ten participants showed a significant effect of the viewing conditions ($F(1,1438)=47.36$, $p<0.0001$) and of the range of disparities ($F(7,1438)=61.54$, $p<0.0001$). Moreover, an interaction effect was observed between the viewing conditions (R3D vs. S3D) and the range of disparities ($F(7,1438)=4.71$, $p<0.0001$). ANOVA carried out on the (unsigned) amplitude provided a significant effect of the viewing conditions ($F(1,1438)=61.67$, $p<0.0001$) and of the range of disparities ($F(7,1438)=85.19$, $p<0.0001$). Furthermore, an interaction effect was put forward between the viewing conditions (R3D vs. S3D) and the range of disparities ($F(7,1438)=17.98$, $p<0.0001$). The analysis performed on the time constant for the ten participants did not reveal any significant effect of the viewing conditions ($F(1,1438)=0.35$, $p>0.05$) but an effect of the range of disparities was observed ($F(7,1438)=11.25$, $p<0.0001$). The time constant was 252 ms (± 100 SD) in the R3D condition and 247 ms (± 104 SD) in the S3D condition. ANOVA performed on the velocity peak position did not reveal any significant effect of the viewing conditions ($F(1,1438)=1.27$, $p>0.05$) and of the range of disparities ($F(7,1438)=1.473$, $p>0.05$). The velocity peak position was around 295 ms (±120 SD) for both viewing conditions.

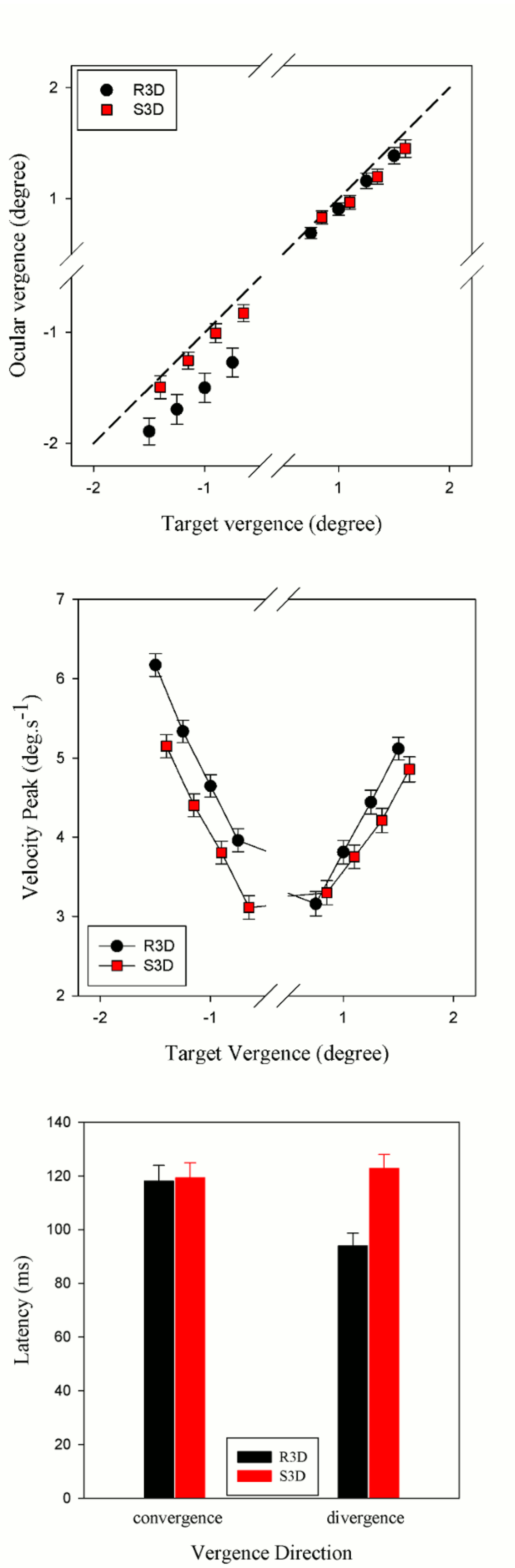

Figure 3.5: Main components of vergence dynamics in R3D (real world) and S3D (stereoscopic) viewing conditions for each possible value of disparity (± 1.5, 1.25, 1 and 0.75 degrees). Upper panel represents ocular vergence (measured amplitude relative to the screen position) as a function of target vergence for both R3D and S3D conditions. The dashed line represents matched ocular and target vergences. Vertical error bars denote confidence intervals of mean. The middle panel represents the vergence step velocity peak as a function of the target vergence in both R3D and S3D conditions. The figure represents unsigned velocity peak so as to easily compare convergence (negative step) and divergence (positive step) responses. Vertical error bars denote standard error of mean. The lower panel represents the vergence latency as a function of convergence and divergence movements and in both viewing conditions. Vertical error bars denote standard error of mean.

3.2.3 Discussion

The objective of this experiment was to find out potential differences in vergence step response in two experimental viewing conditions. The first condition was called real-world (R3D) because it provided congruent disparity and defocus stimuli while the second was a stereoscopic condition (S3D) where a conventional stereoscopic display was used to provide the disparity-step target. This study was motivated by the fact that stereoscopic displays are considered presenting conflicting stimulus information between accommodation and vergence (e.g. Hoffman et al., 2008). Depth perception being directly dependent on vergence, we therefore wanted to explore potential effects of such a presentation method.

The main sequences of velocity peak and amplitude of the vergence responses were found to be meaningfully comparable. They indeed provided comparable coefficients of correlation and slopes of linear regression (in the expected range). However, as mentioned above, the correlation coefficients reported in the present study differ slightly from those reported in previous studies (i.e. Hung, Zhu, & Ciuffreda, 1997; Hung et al., 1994). As a matter of fact, the range of the tested disparities was smaller than in Hung et al (1997, 1994). Moreover, the decrease of correlation coefficients in the S3D condition as compared to the R3D condition could be related to an increase of response variability which naturally goes along with increased task difficulty. The slope of the linear regression on the amplitude-vergence relationship seemed unchanged for convergence but appeared reduced in the S3D condition as compared to the R3D condition for divergence (see Figure 3.4). While two elements of the vergence dynamics (the time constant and the time to velocity peak) were not significantly changed, we found a few significant differences as shown in Figure 3.5. More specifically, vergence velocity peak and amplitude were larger in the R3D condition as compared to the S3D condition for convergence movements (i.e., they presented overshoots); divergence movements were, on average, rather precise in both conditions. However, divergence movements showed longer latency in the S3D condition as compared to the R3D condition. Latency for convergent movements was similar in the two conditions but, the mean latency for divergence movements was indeed much shorter in R3D as compared to S3D conditions. Such distinctions between divergent and convergent eye movements are not new (e.g. Alvarez et al., 2005). However, it is interesting to note that in our study, the latency for convergence in R3D was in the common range while the latency for divergence was very small and, for the largest disparities, approached the ultra-short latency observed elsewhere (Busettini et al, 2001). This observation may be related to the fact that disparity-step trials were presented in blocks and, therefore possibly provided larger predictability. Vergence latency has been shown to be smaller in convergence than in divergence in some studies (Zee, Fitzgibbon, & Optican, 1992; Hung et al., 1994; Hung, Zhu, & Ciuffreda, 1997) whereas other studies found the inverse (Alvarez et al., 2002; Krishnan, Farazia, & Stark, 1973). Alvarez et al (2005) found that the latency of vergence eye movements – and especially for divergence – is dependent on the initial starting position. In this study, we found that latency was roughly equal in convergence and divergence responses to disparity-step in stereoscopy. Hence, given that the fusion range is smaller for divergence than for convergence in stereoscopic displays (e.g. Yeh & Silverstein, 1990), this absence of difference between convergence and divergence could be related to a directional misestimation due to incorrect accommodative stimulus. The short latency observed in the R3D condition (shorter by about 20 ms as compared to S3D) is possibly the result of a sufficiently large disparity-step that can significantly provide good vergence responses. The velocity peak and the amplitude of vergence were also larger in the R3D than in the S3D conditions but this effect was really marked for convergence. This observation replicated for the small disparity-step used in our study for the R3D condition, the fact that convergence exhibits faster velocity than divergence (Hung et al., 1997). However, this benefit seems lost in the stereoscopic condition.

The main dynamics components of vergence could thus be affected in stereoscopic viewing conditions as compared to real-world viewing conditions. These changes depended on the direction and the angle of the target vergence; vergence was overall slower in stereoscopy than in real-world condition. However, it could be noted that vergence response to stereoscopic target could also benefit from the block presentation with R3D vergence targets. Greater differences in vergence dynamics are thus expected between stereoscopic and real-world viewing conditions when targets are displayed without blocking.

3.3 Experiment 2

Considering that vergence response timing components may be affected in stereoscopic display (i.e. different from those in real-world), we were interested in studying the outcomes of such a presentation method in a vergence-step pre- and post-tests experimental paradigm. As the visual system may adapt to this particular viewing situation (Neveu et al., 2012, Rushton & Riddell, 1999), we expected to find significant changes in vergence dynamics following the use of a visual fatigue simulator (VFS) which will be described just below.

3.3.1 Method

The same ten observers participated in this second experiment. They gave their informed explicit consent before beginning the experiment. We also used the same experimental apparatus except that we removed the beam splitter and the second screen.

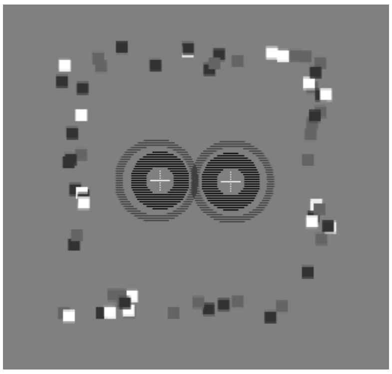

Figure 3.6: The fixation target of the Visual Fatigue Simulator (VFS). The Fixation target is surrounded (1) by a frame composed of small white-to-black squares to help maintain fusion and (2), by an annulus oscillating in depth.

Procedure & Stimuli

Participants were invited for two experimental sessions during which they were presented with a set of pre- and post-tests in order to estimate the outcomes of the Visual Fatigue Simulator (VFS). The two sessions were run on two different days. The Visual Fatigue Simulator was designed in order to mimic the possible ocular displacements experienced by a spectator of a stereoscopic video including several sequences. This experimental phase was composed of a fixation target identical to the one of experiment 1. A typical trial began with the appearance of the fixation target at a random planar location in a virtual frame of a 7.5 by 11 degrees centered in the middle of the screen and with a random angular disparity in a given range, defining a “weak” or “strong” session. The VFS differed between the two sessions according to the possible values of the presented fixation target angular disparities. We defined the “weak session” the one where the fixation target could be displayed with crossed or uncrossed disparities of 0, 10, 20, 30 or 40 arc minutes and the “strong session” with crossed or uncrossed disparities of 80, 90, 100, 110 or 120 arc minutes. Only the fixation target and oscillating annulus were displayed with disparity; the surrounding frame composed of small white-to-black squares remained in the screen plane. For each session, ten possible values of disparity were used and repeated until the duration reached 30 minutes. Trial duration was between 4 and 15 seconds. During the fixation target presentation, a gray annulus (47 arc minutes radius) was presented around the fixation target and oscillating back and forth from +6 arc min to -6 arc min relative to the fixation target position with a mean velocity of 18 arc minutes per second (oscillation of 1.5Hz). During the VFS, we did not track vergence.

In order to ascertain that participants really fused and tracked the fixation target, we asked them to report as quickly as possible whether the annulus was seen in front of or behind the fixation target. The task was as following: during one second, the annulus stopped, the fixation target became yellow and the participant used the response box to provide the response. A visual feedback was provided (green for a correct response or red for a wrong response).

In order to investigate the possible changes due to the Visual Fatigue Simulator, several pre- and post-tests were performed. Before and after each session of VFS, participants were asked (1) to fill in a symptoms questionnaire, (2) to report phoria and fixation disparity and, (3) to perform disparity-step vergence eye movements.

The Symptoms Questionnaire was inspired by the one used by Hoffman et al. (2008) where participants indicated the severity of their symptoms on a 5 points-scale with intensities ranging from Very Tired to Very Fresh. There were 5 questions: (1) How tired are your eyes? (2) How clear is your vision? (3) How tired and sore are your neck and back? (4) How do your eyes feel? (5) How does your head feel?

Phoria and fixation disparity measurements were carried out using the Saladin Near Point Balance Card (Bernell Corporation, see i.e. Corbett & Maples, 2004). Horizontal phoria was measured using the Maddox rod and the Saladin Card held at 40 cm. Horizontal phoria was measured three times for each VFS session: a first pre-test at the onset of the experiment, a second pre-test after the vergence step phase and, a last time after the VFS session. The second measure of horizontal phoria was performed so as to assess whether the vergence step phase could modify even slightly the vergence resting state. Fixation disparities were measured with a set of base-in and base-out prisms (0, 4, 8 and 12 diopters) following the procedure explained in the manual and taking care to avoid vergence adaptation using maximal exposure of 15 seconds.

In order to measure vergence eye movements, we presented crossed/uncrossed (simulated) disparity-step stimuli of 0.75, 1.0, 1.25 or 1.5 degrees. Each level of this condition was repeated nine times in a counterbalanced order. Hence, a total of 72 trials were performed in each pre- and post-test for one session. We used the same fixation target as in experiment 1 and in the VFS except that the oscillating annulus was absent. The same procedure as in experiment 1 was used for the S3D condition. The left and right views of the fixation target were shifted according to the eight possible disparity values during 2.5 seconds. The fixation target was then again presented with zero-disparity so as to control that vergence had returned back to its normal position.

3.3.2 Results

3.3.2.1 Questionnaire

All participants reported their degree of fatigue/discomfort at the onset and the end of both sessions thanks to the symptom questionnaire (previously used in Hoffman et al., 2008). Figure 3.7 illustrates the average reported symptoms intensity. The higher the plotted bar, the smaller the symptoms intensity. Repeated measures nonparametric analysis revealed that only questions 1 and 4 were significantly different across the pre- and post-test (Wilcoxon matched pairs test, $p<0.05$). Supplementary one-tailed analysis revealed that symptom questions 1 and 4 were more severely judged by participants in the post-test as compared to the pre-test in the strong session ($p<0.025$).

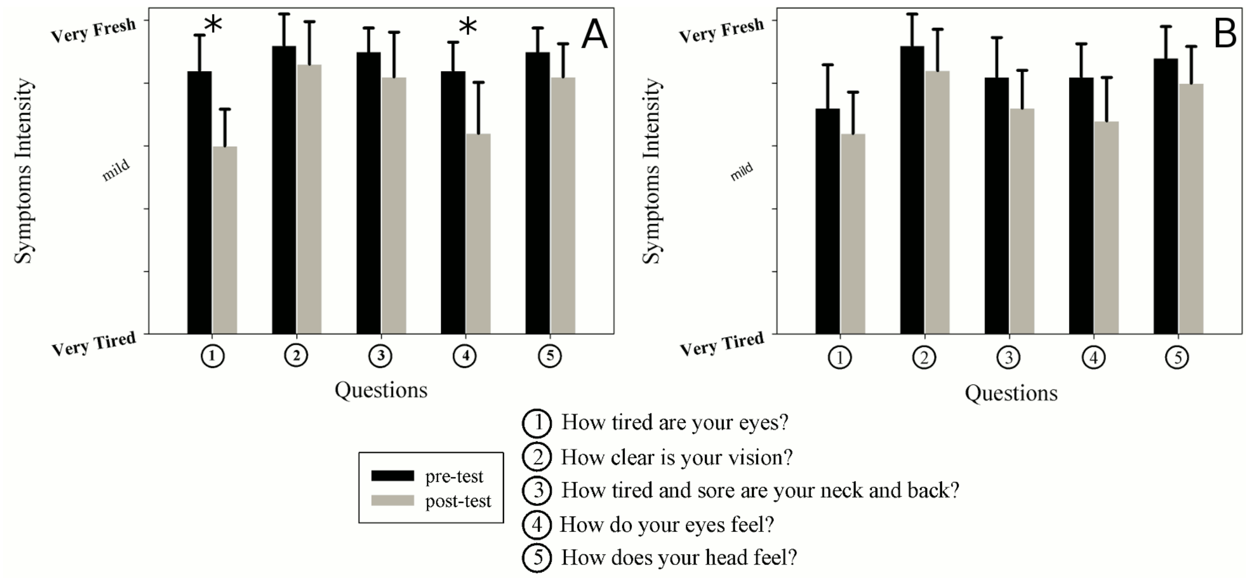

Figure 3.7: Results from the symptom questionnaire. This figure represents the average severity (from very tired to very fresh) reported for the five questions of the questionnaire for the strong and the weak sessions of the visual fatigue simulator. Left panel (A) represents the strong session and right panel (B) corresponds to the weak session. Black bars represent pre-

test and gray bars represent post-test. Vertical error bars denote 95% confidence intervals. Asterisks represent significant differences between pre- and post-test ($p<0.05$, Wilcoxon matched pairs test).

3.3.2.2 Performances in the VFS sessions

In order to ensure that participants really fused the stimuli in the VFS session, they had to perform a relative depth judgment task. Using an ANOVA, we compared the results (success/failure) obtained in the two VFS sessions. The analysis revealed an effect of the session, the success rate was larger in the weak condition (80% ±40 SD) than in the strong condition (64% ±48 SD) ($F(1,5006)=105.6$, $p<0.0001$). Additionally, we found an effect of the range of disparities ($F(16,5006)=21.4$, $p<0.0001$); the larger the disparity, the greater the error ($p<0.0001$). Also, target displayed with uncrossed disparity provided overall less success than target displayed with crossed disparity ($p<0.0001$).

3.3.2.3 Phoria and fixation disparity

Our analyses carried out on horizontal phoria and fixation disparity did not reveal any significant effect of the treatment or of the condition of the VFS. Except the correlations obtained between measured phoria before and after the vergence step phase, we do not report them there. Thus, significant between-subjects correlations were obtained between the measured phoria before and after the vergence step pre-test in the weak condition ($r=0.84$ $p<0.01$) and in the strong condition ($r=0.96$, $p<0.001$). The correlation between measured horizontal phorias in both VFS session was also significant ($r=0.61$, $p<0.01$). Although pre- and post-vergence step phase measured phoria were significantly correlated, we did not observe any significant difference ($p>0.05$).

3.3.2.4 Vergence step

Plots of the main sequence for each vergence response for the pre- and post-tests of the two VFS sessions (strong and weak) are represented in Figure 3.8. In the strong session (left panel), correlations coefficients are 0.67 for divergence and 0.62 for convergence and the slopes are 2.38 s-1 and 1.73 s-1 for the pre-test and, are 0.68 for divergence and 0.67 for convergence and the slopes are 2.27 s-1 and 1.97 s-1 for the post-test (all correlations are significant $p<0.001$). In the weak session (right panel), correlations coefficients are 0.68 for divergence and 0.66 for convergence and the slopes are 2.3 s-1 and 1.82 s-1 for the pre-test and, are 0.72 for divergence and 0.63 for convergence and the slopes are 1.99 s-1 and 1.71 s-1 for the post-test (all correlations are significant p<0.001).

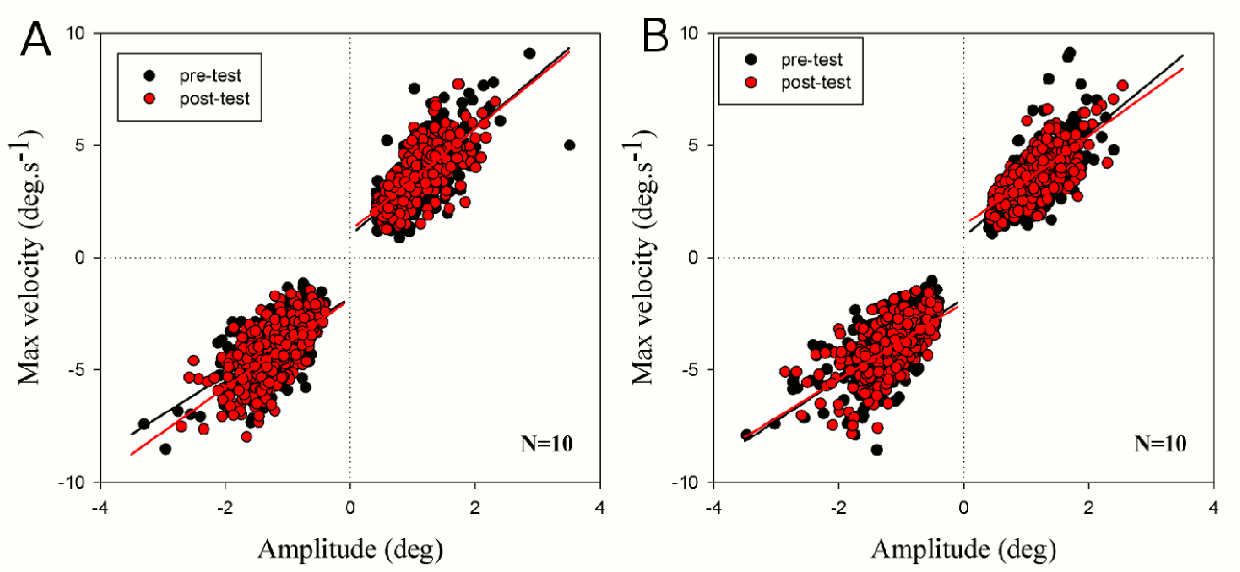

Figure 3.8: Main sequences for both conditions of the visual fatigue simulator. Left panel (A) represents the strong session and right panel (B) corresponds to the weak session. The figures also represent linear regression lines (solid red and black lines) for convergence (negative amplitudes and peak velocities) and divergence (positive amplitudes and peak velocities) in both the pre-test and the post-test (black and red color coding). The plot contains 1216 observations for the strong session and 1271 observations for the weak session.

Repeated measures ANOVA performed on latency for the ten participants did not reveal any significant effect of the VFS session (i.e. “strong” vs. “weak”) ($F(1,1141)=1.31$, $p>0.05$) or for the range of disparities ($F(7,1141)=1.13$, $p>0.05$). The latency was 153 ms (±101 SD) in the strong condition and 159 ms (±111 SD) in the weak condition. The effect of the treatment (pre- vs. post-test) was also not significant ($F(1,1141)=0.69$, $p>0.05$). The analysis did not show any significant interaction effect between the main variables ($p>0.05$). Supplementary analysis realized on the (unsigned) velocity peak for the ten participants did not show any significant effect of the VFS session ($F(1,1141)=0.5$, $p>0.05$) but revealed an effect of the range of disparities ($F(7,1141)=48.73$, $p<0.0001$). More importantly, the analysis yielded a main effect of the experimental treatment (pre- vs. post-test) on the unsigned velocity peak; the velocity peaks were increased following the experimental treatment ($F(1,1141)=23.11$, $p<0.0001$). A second order interaction effect was obtained between the treatment, the VFS session and the range of disparities ($F(7,1141)=2.41$, $p<0.02$); the analysis revealed that the treatment affected the velocity peak in the strong session as compared to the weak session for medium to large crossed disparities (see Figure 3.9, $p<0.05$). Post-hoc analysis revealed that the velocity peak were increased in the strong session as compared to the weak session for medium to large crossed disparities (-1, -1.25 and -1.5 degrees) following the VFS treatment ($p<0.05$). Repeated measures ANOVA performed on the (unsigned) amplitude for the ten participants did not provide a significant effect of the VFS session ($F(1,1441)=1.01$, $p>0.05$) but an effect of the range of disparities ($F(7,1141)=133.41$, $p<0.0001$). Furthermore, no effect of treatment could be put forth between the pre- and post-test ($F(1,1141)=1.01$, $p>0.05$). ANOVA performed on the time constant for the ten participants did not reveal any significant effect of the VFS session ($F(1,1141)=0.01$, $p>0.05$) but an effect of the range of disparities was observed ($F(7,1141)=19.13$, $p<0.0001$). Moreover, the analysis showed a main effect of the experimental treatment (pre- vs. post-test) on the time constant ($F(1,1141)=27.09$, $p<0.0001$); the response time constant was smaller after the treatment ($p<0.0001$). The analysis performed on the velocity peak position did not reveal any significant effect of the VFS session ($F(1,1141)=0.002$, $p>0.05$) but showed an effect of the disparity range ($F(7,1141)=2.11$, $p<0.05$). A statistical tendency was found for the treatment effect between the pre- and the post-test ($F(1,1141)=0.002$, $p<0.06$).

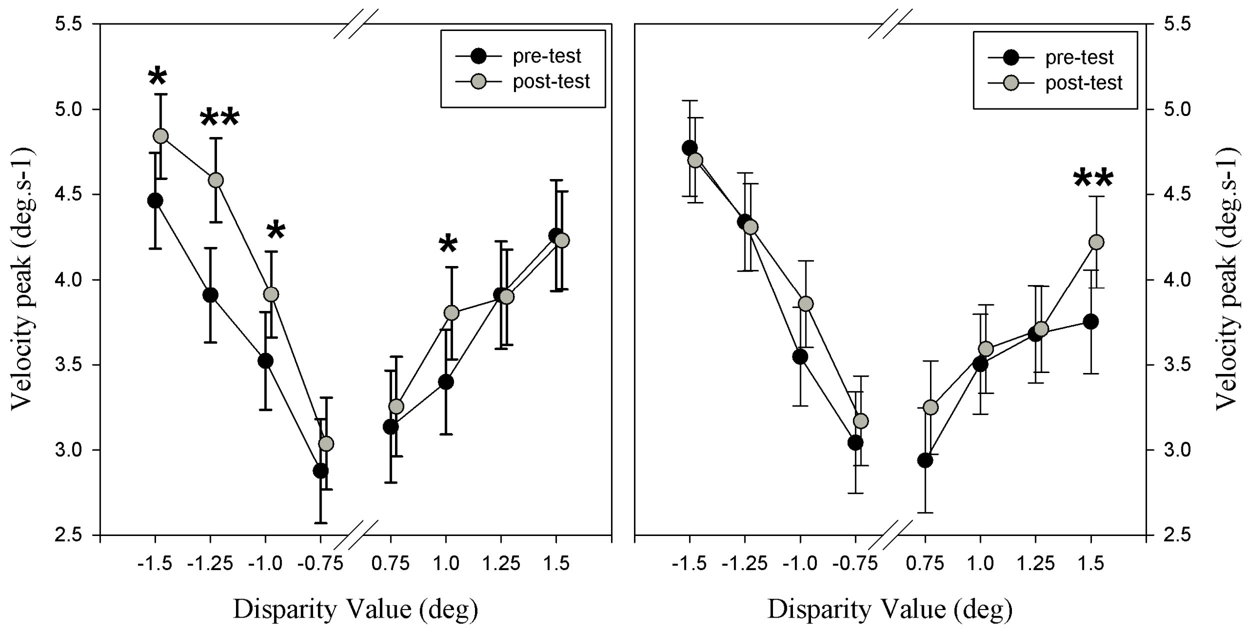

Figure 3.9: This figure represents the vergence step velocity peak as a function of the target vergence in both strong (left panel) and weak (right panel) conditions of the visual fatigue simulator. The figure represents the unsigned velocity peak so as to easily compare convergence (negative step) and divergence (positive step) responses. Vertical error bars denote 95% confidence intervals of mean. Asterisks represent significant differences between

pre- and post-test ($p<0.05$ for single asterisk and $p<0.01$ for double asterisks, post-hoc analysis based on Fisher LSD test). The second order interaction between the variable treatment, VFS session and range of disparity is significant ($p<0.02$).

3.3.3 Discussion

The main objective of this experiment was to study potential changes in vergence dynamics for small disparity-step after a prolonged stereoscopic activity. We designed an experimental session called the visual fatigue simulator that was expected to strongly stimulate the vergence system. Participants were tested for two sessions of the visual fatigue simulator for which two independent ranges of disparities were used to display the stimulus. The “strong” session contained stereoscopic target with disparity up to ±2 degrees while the “weak” session was limited to disparity of ±40 arc minutes. These ranges of disparity were chosen with respect to the revisited Panum’s fusional area and the comfort limit of about 1° (Lambooij et al., 2009). We thus expected that the strong session would be much more tiring than the weak session. This first statement seemed to be supported by the results on the subjective questionnaires (see Figure 3.7). Questions related to eye strain were significantly more severely judged by participants after the strong session than after the weak session. Moreover, there were no significant changes in the symptom severity after the weak session. During each experimental session (“weak” vs. “strong”), we also asked participants to perform a relative depth judgment task about a stereoscopic annulus that oscillated in depth relative to the fixation target. The success rate was significantly higher in the weak condition as compared to the strong condition, indicating greater difficulties in doing the task in the strong session as compared to the weak session. These results together with the subjective reports of visual symptom indicated that on the basis of the range of disparities used during the visual fatigue simulator, stereoscopic activity could entail visual fatigue and/or discomfort in our experimental conditions, though disparities remained in the zone of clear and single binocular vision.

Clinical pre- and post-tests were also performed in order to measure any change in phoria or fixation disparity. We did not report the results here because we did not find significant changes. However, we measured phoria before and after the pre-test of the vergence step so as to evaluate whether such exposure to small vergence steps (of about 5 minutes duration) could affect significantly the resting vergence position. We did not observe any significant difference in phoria after the vergence step phase whereas good inter-subject phoria correlations were found. Though an absence of effect should be interpreted cautiously, these observations support the conclusion that the vergence step exposure had on average, little or no effect on the global vergence state.

Central to the topic of the present paper, the dynamics of vergence eye movements in response to disparity step were in general significantly affected between the pre- and post-test according to the range of disparities used in the visual fatigue simulator. The main sequences (relationship between velocity and amplitude of the vergence movements) were comparable across the treatment and the VFS sessions (see Figure 3.8). The coefficients of correlation were significant and were from 0.62 to 0.72, therefore relatively smaller than those estimated in experiment 1. However, in comparing the pre- and post-test between the two VFS session, the slope of the post-test in the strong session seemed higher than the slope of the pre-test and, this observation was not true for the weak session. This graphical result was supported by the results obtained on the peak of velocity. We observed significant increases of peak velocity in the strong session for crossed disparity targets (-1, -1.25 and -1.5 degrees) in the post-test as compared to the pre-tests (see Figure 3.9); this effect was not present in the weak session (except for the largest uncrossed disparity). The fact that vergence velocity could increase after stereoscopic activity with a range of large disparities supports the occurrence of an adaptation process that responded to the vergence demand. In addition, we found a main effect of the VFS treatment on the time constant; the response time constant significantly decreased between the pre-test and the post-test and, a tendency was found for the time to peak velocity. The effects on these two components are however not statistically distinguishable between the two VFS sessions.

The major dynamics elements of vergence could thus be affected after stereoscopic activity. The response time constant was significantly decreased after the VFS sessions and the time to velocity peak showed a tendency, indicating that these temporal components could be reduced after the treatment. Moreover, the vergence velocity peak increased much more when the range of disparities of the stereoscopic activity was “strong” as compared to “weak”. Overall, these results argue for the emergence of an adaptation process that copes with the stereoscopic vergence demand. This adaptation of the visual system occurs here for a rather small range of disparities – sometimes slightly exceeding the zone of comfort – and this effect could stem from the decreasing efficacy of the accommodative stimulus for the vergence system.

3.3.4 Analysis of vergence position difference

In this section, vergence eye movements in response to disparity step are averaged and plotted for one observer and a specific value of disparity for the conditions of experiment 1 (R3D vs. S3D) and of experiment 2 (pre- and post-test in strong and weak conditions of the visual fatigue simulator). The objective here is to illustrate the effect found in the previous section. The contribution of the accommodative vergence can be observed using a subtraction analysis (Semmlow & Wetzel, 1979) where the position difference between a set of two averages of ocular vergence are subtracted. The results of this subtraction analysis can be found in Figure 3.10: The first panel (A) represents an example of time series averaged over ten trials; the standard error is also plotted. The second panel (B) shows the averaged vergence time series extracted from experiment 1 and, for one participant, the target vergence was -1.5° in both the R3D (real-world) and S3D (stereoscopic) conditions. The curve for the R3D condition (in black) appeared to have a greater velocity peak and amplitude than the curve for the S3D condition (in dark red). The dark-blue curve just below represents the subtraction of previous R3D and S3D observations (vergence position difference). It can be seen that this subtraction curve contains a first peak that concurs with the vergence peak in both vergence time series and, then, this position difference remains roughly constant below zero during the maintained vergence phase (larger vergence for R3D). The third panel (C) represents the averaged vergence time series extracted from experiment 2 and, for the same participant, for both the pre- and post-test of the “weak” session. In this figure, the curves of the pre-test (in black) and post-test (in dark gray) seem to have similar velocity peak and the vergence difference (in dark blue) is very small. The last panel (D) shows exactly the same features as in panel C, except that the curves correspond to the “strong” session of the visual fatigue simulator. This time the pattern of results is closer to that of the panel B: the curve of the post-test has a greater velocity peak and amplitude, the vergence difference contains a slightly more marked peak coinciding with the vergence peak, and also the vergence difference is maintained during the sustained vergence phase. These graphical results are consistent with the fact that the potential vergence adaptation in stereoscopy resulted in responsivity close to that observed in real-world conditions.

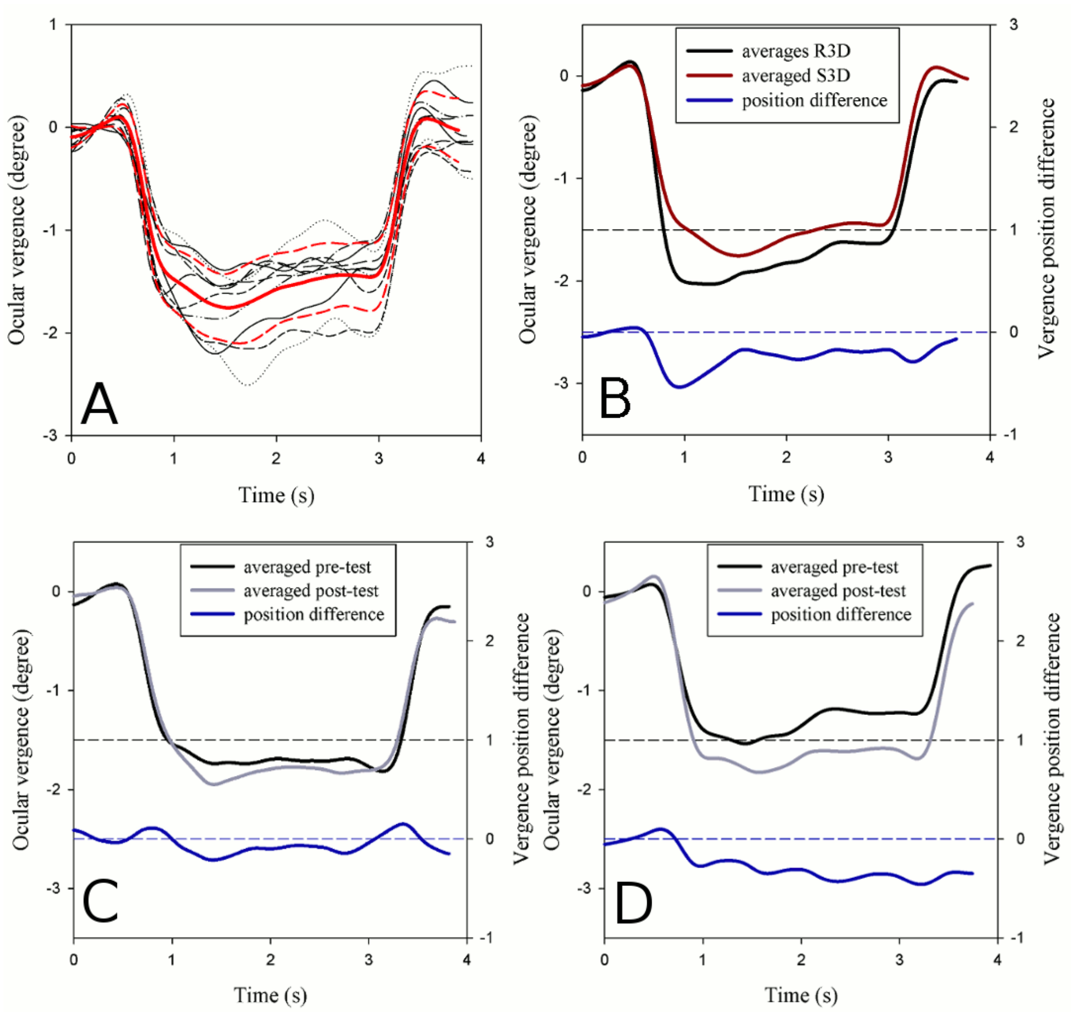

Figure 3.10: This figure represents the time series of ocular vergence in response to a disparity step of 1.5° of convergence for one observer. Panel A represents an example of time series average for ocular vergence response to a disparity-step of -1.5°. Black lines correspond to each vergence responses (n=10) that are averaged to provide only one curve (the red line). The red dashed lines represent ±1 SD. Panel B represents the averaged time series of one participant in experiment 1 for a disparity step of -1.5° both for the R3D condition (black curve) and for the S3D condition (dark red curve). On the second vertical axis is represented the vergence position difference which is the subtraction of both curves (dark blue curve). Panel C represents the averaged time series of one participant in experiment 2 for a disparity step of -1.5° both for the pre-test (black curve) and for the post-test (gray curve) of the "weak" session of the VFS. On the second vertical axis is represented the vergence position difference which is the subtraction of both curves (dark blue curve). Panel D follows the same representation as in C but for the "strong" session of the VFS. Note that the beginning of the time series on the figures does not coincide with the stimulus onset.

3.4 General Discussion

Previous studies on vergence eye movements provided evidence that the vergence positional control of the eyes is under the influence of several components (e.g. Hung et al., 1997; Hung et al., 1994; Hung, Semmlow & Ciuffreda, 1983; Semmlow and Wetzel, 1979). The vergence components can be dissociated according to their driving stimuli: accommodative vergence, driven by blur on the retinal image, and disparity-vergence, driven by the spatial differences between the two retinal images (Semmlow & Venkiteswaran, 1976). Hence, the vergence position in natural vision is mainly the result of a combination of accommodative and disparity components, even though it is acknowledged that the proximal component and the tonic vergence can participate to the vergence response (see Howard & Rogers, 1995).

The two experiments in the present report were designed in order to investigate the potential differences that may arise in vergence response to disparity step when disparity and blur are combined, providing either congruent or incongruent stimulation. In experiment 1, we designed a specific experimental apparatus that was used to compare vergence responses in real-world (R3D) and stereoscopic (S3D) conditions. The main finding of this report is that vergence response provided overall slower temporal components in S3D as compared to R3D. In experiment 2, we tested the outcomes of a visual fatigue simulator on several pre- and post-tests. The most relevant indicator was certainly the vergence step period that affected vergence dynamics after strong stimulation through stereoscopic activity with large target disparities. Stereoscopic activity induced a change in vergence dynamics that is presumably related to the adaptation of the vergence system. The idea that the visual system may adapt to stereoscopic conditions is indeed not new (Eadie et al., 2000; Rushton & Riddell, 1999; Wann, Rushton & Mon-Williams, 1995). As an example, the direction of this adaptation is suggested in Figure 3.10 from the analysis of the vergence time series of one observer. This analysis suggests that when participants had to fuse relatively large disparities in stereoscopic activity, the vergence system adapts in order to solve the ambiguity between the disparity and defocus stimuli.

This adaptation mechanism following the use of stereoscopic display could be possible if crosslinks between accommodation and vergence become ignored by the visual system. Some studies have shown that changes in the gain of the cross-links are possible, for example, in viewing through periscopic spectacles (Judge & Miles, 1985), and haploscopes (Eadie et al., 2000). As stereoscopic displays induced many variations on vergence but constant demand on accommodation, a crosslink adaptation is necessary because subsequent variations of vergence accommodation produce fluctuations of accommodation. It was suggested that the gain of vergence-accommodation crosslinks would need to be decreased to avoid changes in accommodation driven by changes in vergence (Wann, Rushton & Mon-Williams, 1995). Eadie et al. (2000) observed the decrease in both crosslinks between accommodation and vergence after exposure to a virtual reality system in two participants. Rushton & Riddell (1999) also argued that this decrease in the gain of the crosslinks would be accompanied by an increase in the gain of the disparity vergence. The results presented here on vergence eye movements to disparity step could thus be the outcomes of an increase in the disparity vergence gain in disfavor of the accommodative vergence gain.

In the second experiment, the visual fatigue simulator was expected to produce asthenopic signs in participants. Reports on the subjective questionnaire were significantly more severely rated for questions related to eyestrain and eye fatigue in the strong condition as compared to the weak condition. However, the severity of signs was not as important as expected; participants did not complain of severe visual trouble or headache perhaps because they could still fuse all stimuli. Moreover, it is worth noting that changes in vergence response such as an increase in latency or decrease in velocity are rather expected to concur with significant asthenopic signs. In the present report, the vergence velocity peak was increased after stereoscopic activity. This observation is presumably related to the occurrence of adaptation that allows the oculomotor system to counteract visual fatigue (e.g. Schor, 2009). However, the relationship between oculomotor changes and the subjective experience of visual fatigue does not appear evident (Lambooij et al., 2009) and, viewer’s sensitivity could presumably limit the study of such linkage.

Overall, our results highlight the importance of appropriate accommodation information in combination to binocular disparities for repetitive vergence eye movements. Given that accommodation cues and disparities are often processed separately in three-dimensional media, our results may be related to the source of discomfort sometimes reported while watching three-dimensional displays for long periods of time.

♦

The present study had two main objectives: the first was to compare vergence eye movements in stereoscopic and real-world viewing conditions and, the second was to generate fatigue in the observers' visual system so as to measure the resulting changes in subjective discomfort and in temporal vergence dynamics. The first class of results concerns the fact that vergence dynamics is overall slower in stereoscopic viewing conditions as compared to real-world viewing conditions. In fact, accommodation and vergence signals do not exactly specify the same distance and, therefore the visual system has to deal with a sensory ambiguity (i.e. due to accommodative vergence) that has to be solved, to the cost of a greater reaction time. The second class of results shows that prolonged stereoscopic stimulation can lead to significant changes in vergence eye movements even with relatively small range of disparities. In this situation, the resulting changes on vergence response were more related to a beneficial effect on the visual system, specifically because of the overall increase in vergence velocity. However, negative changes are expected when the stereoscopic visual scene engenders larger disparity range, causing strong eyestrain.

Beyond the theoretical considerations, this study also brings a possible new indicator of visual fatigue in stereoscopic displays. Vergence dynamics could be used as an indicator of fatigue as it informs about the performance of the visual system in stereoscopic conditions. Moreover, it is worth noting that, compared to other signs of visual fatigue (e.g. eye-blink rate), vergence dynamics is directly related to stereoscopic depth perception. However, supplementary work is needed to investigate the intra- and inter-variance of vergence dynamics in predicting adaptation or fatigue of the visual system in stereoscopic displays.

Hence, stereoscopic displays showed to be potentially sources of discomfort for prolonged use and are depending on inter-individual differences. The sources of visual discomfort and fatigue seem to occur when the disparity conveyed by the displayed content induced a stereoscopic demand that exceeds the limits of the depth-of-focus. For that reason, the following part is devoted to an experimental study of the role of the depth-of-focus in stereopsis and its relation to visual discomfort.

Chapter 4. An attempt to increase the depth-of-focus of the eyes for reducing discomfort in stereoscopic displays

4.1 Introduction

As stereopsis is getting more and more popular following its use in entertainment media, the question of visual discomfort and fatigue remain at stake. As a matter of fact, visual fatigue may emerge from the prolonged observation of stereoscopic contents containing too large binocular parallax, entailing strong and repetitive stereoscopic demand. A significant proportion of spectators complain about a variety of visual symptoms such as, e.g. eyestrain, headache and blurred vision after prolonged exposition to stereoscopic media. The experience of visual symptoms in these conditions has been attributed to difficulties for observers to decouple accommodation and vergence (i.e. Hoffman et al. 2008), both systems being naturally coupled (Hung and Semmlow, 1980). Hence, the vergence system can respond both to binocular disparity (disparity vergence) and to retinal defocus (accommodative vergence); the latter results from the activity of a neural cross-link. Indeed, a vergence response can be observed by varying the stimulus spatial frequency while only one eye is stimulated (Semmlow et al. 1981). Conversely, the accommodation system reacts both to defocus stimulation (blur-driven accommodation) and to disparity information (vergence accommodation or disparity-driven accommodation). As a consequence, in configurations with strong stereoscopic demand, the visual system may deal with a viewing situation where accommodation and vergence specify incongruent signals. In this situation, the activity of both neural crosslinks (accommodative vergence and vergence accommodation) should be decreased in order to provide efficient oculomotor responses. The possible adaptation of the crosslinks after stereoscopic activity has thus been discussed (Rushton & Riddell, 1999). Stereoscopic viewing conditions therefore induce an oculomotor demand related to accommodation error correction that is supposed to generate visual fatigue consecutively (Lambooij et al. 2010). This anomalous oculomotor response entailing visual symptoms could be an accommodative overshoot induced by the vergence demand that requires accommodative correction (Okada et al. 2006; Torii et al. 2008; Fukushima et al. 2009).

Visual fatigue would however take place only when the visual system cannot adapt to this viewing condition (Schor, 2009).

Stereoscopic depth perception is actually based on binocular disparity that induces significant vergence responses when the vergence demand exceeds the individual limit of Panum's fusional area, the zone around the fixation point where single vision is possible. As for vergence, there is a range of distances for which the accommodation system does not perform re-adjustment because no retinal defocus is detected; this is the depth-of-field (Campbell & Westheimer, 1959). The depth-of-field is the projection of the depth-of-focus of the eye in object space. In mathematical models of dual interaction between vergence and accommodation, the depth-of-field is the dead space of the accommodation system feedback loop (Hung, 1998). In other words, this component acts as a threshold function and allows the accommodative system to send motor commands to adjust the shape of the eye lens for minimizing retinal defocus on the region of interest. As a consequence, one way to minimize the activity of both the accommodative vergence and the blur-driven components is to increase the depth-of-focus of the eyes or, alternatively, to provide stimulation without fine spatial frequency information. It is thus possible to render the accommodation system open-loop by presenting stimuli with image blurring for which observers cannot perform precise accommodative responses (Okada et al., 2006). Stereoscopic viewing conditions presenting low spatial frequencies have furthermore shown to provide greater fusion capacities (Schor et al., 1984). Another way is to place observers in pinhole viewing conditions because wearing pinhole removes defocus information contained in the projection onto the retina (Rosenfield, Ciuffreda & Hung, 1991). It is however difficult to use such pinhole system in binocular vision where vergence and version movements are required because the holes must remain aligned with the visual axes. As an attempt to circumvent this, some researchers used "aerobic glasses" to allow the eyes to perform eye movements (Frisby, Buckley & Horsman, 1995). The uses of pinholes or image blurring do not however directly manipulate the depth-of-focus of the eyes (i.e. it depends on pupil aperture) and, therefore strongly impact the perceived visual information. For practical use, image blurring highly suffers from the dramatic change of sharpness in the visual scene and may be annoying to see when performed on the fovealized parts. Pinhole viewing endures both the imprecision of the holes alignment with the visual axes and the dramatic reduction of light entering through the pupil aperture. In addition, as the available light decreases on the retina, the pupil size aperture may also increase in reaction to the loss of luminance. Pinholes could also place the accommodation system in resting state that does not necessarily reflect the normal behavior of the oculomotor system. As a consequence, the objective of the present study is to manipulate the depth-of-focus of the eyes so as to increase the blur detection threshold leading the accommodation system to tolerate a larger range of disparity in stereoscopic viewing conditions.

Pupillary responses may arise following changes in the lighting conditions, the proximity of the target and the emotional state of observers (Andreassi, 2000). The light reflex depends on lighting conditions and produces constriction due to intense light or dilation with dim light. The pupil response to light is predominant and is specifically due to the stimulation of retinal receptors (rod, cone and melanopsin pigments) when light cast onto the retina (Viénot et al., 2010). Pupil aperture size is the predominant factor over the depth-of-focus of the eye. Geometrical optics predicts that the depth-of-focus varies proportionally with the pupil aperture size but, however, empirical works have shown that large pupil sizes in fact behave as smaller pupil sizes in terms of depth-of-focus (Campbell, 1957). In general, as the circle of confusion grows linearly with the pupil size, it is expected that the larger the pupil size, the smaller is the depth-of-focus.

Using a more ecological method, the present study aims at varying the depth-of-focus of the eye so as to observe the resulting changes on visual discomfort in stereoscopic displays. Variations in lighting conditions therefore affect the focus range of accommodation and, therein, may influence viewing comfort of stereoscopic stimuli. Here is investigated the general benefits of increasing the depth-of-focus of the eyes for improving visual comfort. As such, we used a stereoscopic visual search paradigm to put forth the effect of changing the depth-of-focus limits and evaluated those limits with different lighting conditions. In order to vary the pupil aperture, we used a peripheral white frame located in front of the screen which was illuminated by a video-projector. The visual field containing the stimuli was unchanged so as to avoid any attention or dazzling effect. As within the depth-of-field, accommodation response is unnecessary and that the less the pupil size, the more is the depth-of-field, a well lit environment was supposed to entail less visual discomfort as less occurrence of conflict between accommodation and vergence should occur.

4.2 Method

4.2.1 Observers

A total of seven observers took part in the entire study. One participant was discarded because of large difficulties in fusing the stimuli during the experimentation. Our selection provided a full counterbalanced group of participants (on average 32.8 years old, SD = 10.1). All except one author were unaware of the experimental hypotheses. All had normal stereoacuity as assessed by the stereo-test of the vision tester (Visiolite, FIM-Medical). All had normal vision possibly thanks to their usual optical corrections. They gave their informed consent before beginning the experiment.

4.2.2 Apparatus

Participants sat comfortably in a chair within a chinrest which was placed at two meters from the display (Hyundai S465D 46’’ HDTV LCD Polarized monitor). In order to vary the pupil size, a video-projector (Optoma EP747) was set elevated just behind the observer (see Figure 4.1). The light stemming from the projector illuminated a large white panel (90 x 100 cm) placed in front of the observer. This white panel included an elliptical aperture (about 15 cm, cut in the paper sheet) in order to leave the center of the visual field unaffected by the luminance.

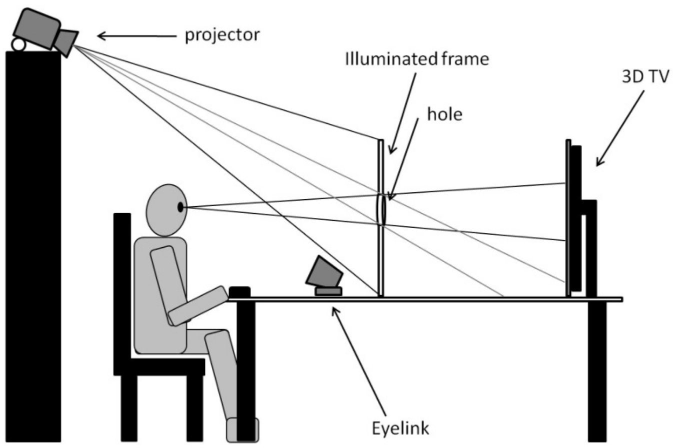

Figure 4.1: Schematic representation of the experimental apparatus used in the study. Observers seated facing the 3D display. The illuminated frame was placed in between. Drawing is not to scale.

4.2.3 Stimuli

Stimuli were vertical grating patterns (three vertical white bars) that were rotated in opposite vertical directions in each eye and, thereby when fused providing the sensation of an apparent rotation around the horizontal axis (e.g. positive disparity on top, negative at the bottom). The grating patterns were rotated on the vertical axis by amount of 1, 2, 3, 4 or 5 degrees. For each trial, five stimuli were simultaneously presented according to a pentagonal spatial distribution (see Figure 4.2). Each grating pattern was displayed at one of the five possible random depth locations. The visual scene was thus composed of five stimuli randomly assigned to five simulated distances which were selected from two possible ranges of disparities: the large range of disparities {-80 ; -40 ; 0 ; +40 ; +80 } arcmin and the small range of disparities {-60 ; -30 ; 0 ; +30 ; +60 } arcmin. These two ranges of disparity were selected because the lower remained in the advised “one degree of comfort limit” (Lambooij et al., 2009) and the other was expected to be more annoying for the visual system.

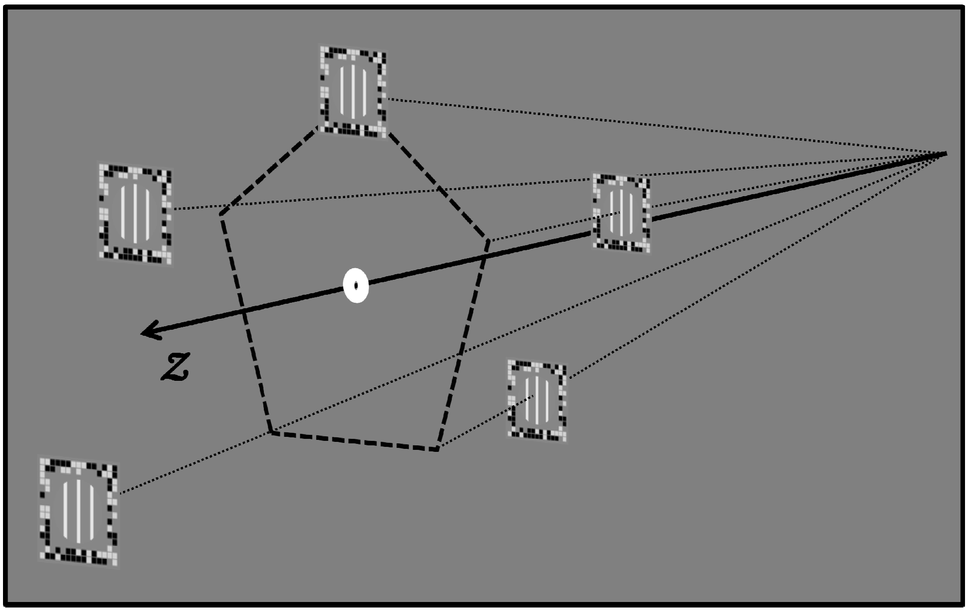

Figure 4.2: Schematic oblique view of the stereoscopic visual scene. The scene contains five grating patterns located according to a pentagonal layout. They are assigned to five possible depth locations relative to the screen plane where the central fixation target (the white disc) lies.

The participants were instructed to “spot the odd one out”, in other words, they had to find as quickly as possible which of the five stimuli was slanted in the opposite depth direction.

Stimuli were viewed under tow different lighting conditions: the high luminance condition where the projected white light on the frame had a luminance of 400 $cd.m^{-2}$ and, the low luminance condition where the frame was not illuminated. The luminance of the display was also decreased so as to limit the lighting effect from the display and to reach large value of pupil diameter in the low luminance condition. Each trial was repeated ten times leading to a total of 100 trials (2 blurring x 2 ranges of disparity x 5 angles x 2 luminance levels).

4.2.4 Procedure

Increasing luminance and decreasing pupil size was expected to increase the depth-of-focus of the eye. As with a larger depth-of-focus the conflict between accommodation and vergence is minimized, we expected find significant decrease of viewing discomfort and increase in the performances in the visual task. The two sessions of luminance were run independently, on two different days, so as to keep the pupil size approximately similar during the same session of illumination. The two levels of illumination were counterbalanced across participants. At the onset of each experimental session, the variation of the pupil size aperture in response to illumination was measured. Observers looked at a fixation cross that was displayed in the middle of the screen plane while the pupil size of the right eye was measured through a camera capture. The measurement was made after a fixation duration of two minutes so that the pupil size was stabilized and adapted to the projected diffused light. After some training trials, the participants began the experimental session. Each trial was followed by the rating of visual discomfort on a five points Likert scale (very comfortable – comfortable – moderate – uncomfortable – very uncomfortable).

4.3 Results

The analysis performed on the pupil aperture size between the low luminance condition and the high luminance condition for the seven participants revealed a significant effect of lighting condition ($t=10.78$, $p<0.001$); the pupil size was decreased in the high luminance condition (on average 3.6 mm, SD=0.46) as compared to the low luminance condition (on average 6.2 mm, SD=0.97). This result is presented on Figure 4.3.

The analyses performed on the discomfort rating showed a slight effect of the lighting condition ($F(1,345)=20.2$, $p<0.001$); the discomfort decreased of about 4.3 % in the high luminance condition as compared to the low luminance condition. The disparity range did not provide significant differences except a tendency: the discomfort appeared to increase with large disparities as compared to small disparities (F(1,345)=3.1, p=0.08). The analyses did not provide any effect of interaction between the variables and no effect of the rotation angle of the gratings.

The analyses performed on the stereo performances revealed a strong effect of the rotation angle of the gratings (F(4,345)=19.6, p<0.0001) showing that the percentage of correct responses decreased when the rotation angle (and, therefore the depth difference) was decreased. No effect of lighting condition was observed (F(1, 345)=1.2, p>0.05) but an interaction effect between the luminance condition and the rotation angle showed a decrease in the percentage of correct response for the rotation angle of one degree (F(4, 345)=2.7, p<0.05). No effect of the disparity range was obtained (F(1, 345)=0.3, p>0.05) and no effect of interaction could also be put forth with the rotation angle (F(4, 345)=2.1, p>0.05).

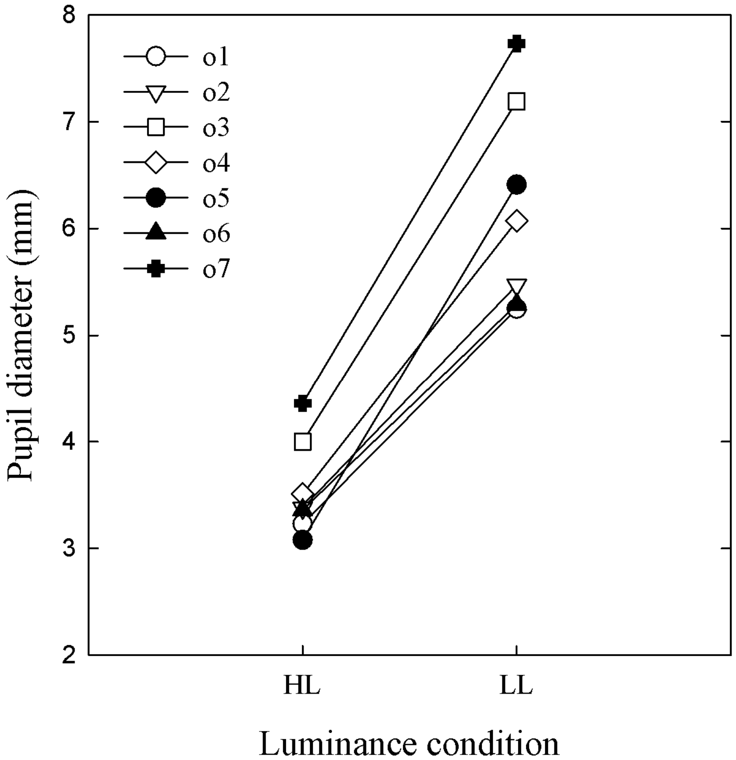

Figure 4.3: This figure represents the averaged pupil diameters of both eyes measured at the onset of each luminance condition for the seven participants (represented by different symbols (o)). HL is the results for the high luminance condition and LL for the low luminance condition.

Supplementary analyses performed on the reaction time provided a significant difference between the possible values of the rotation angle of the gratings (F(4,345)=49.1, p<0.0001); the less the angle, the longer was the reaction time. A significant difference was also observed between the lighting condition ($F(1,345)=7.3$, $p<0.01$); the reaction time was decreased in the high luminance condition as compared to the low luminance condition. However, no significant effect was observed for the disparity range ($F(1,345)=0.7$, $p>0.05$).

4.4 Discussion

The main objective of this exploratory study was to investigate the general benefits of optically increasing the depth-of-focus of the eyes through the use of peripheral illumination and its potential relation to the displayed depth locations and the subjective discomfort. Hence, we used a stereoscopic visual search paradigm to put forth the effect of changing the limits of the depth-of-focus and assessed those limits with two different lighting conditions. In order to vary the pupil aperture size, we designed a specific experimental apparatus that contained a peripheral white frame located in front of the screen that was illuminated by a video-projector. By changing the luminance of the reflected light, the pupil size could be varied and because a circular part was not illuminated in the center of the visual field, observers could perform the task precisely, avoiding dazzling effect. Participants thus had to find one stimulus among five, the four others being slanted in the opposite direction, and then to report the subjective discomfort felt when exploring the visual scene.

The results obtained on the discomfort rating, even if they provided some significant effects, did not show strong significant differences. Indeed, such types of subjective reports usually show large intra- and inter-subject variances and therefore require a consistent number of participants so as to provide stronger and more reliable statistical effects. However, we surprisingly found that the effect of the disparity range was not much more important. The reason may be that the range of disparity used in this study was not sufficiently large to provide strong effect on discomfort. Moreover, inter-individual variance could have weakened the effect. This inter-individual variability may arise from inter-individual differences in the ability to fuse both views during the experiment. Some small pools of participants have already shown that the use of small range of disparities did not powerfully affect visual comfort in stereoscopic displays (e.g. Peli, 1998). The difficulty of choosing the disparity range resides in the need for a compromise between a set of disparity fusable for all participants and a sufficiently large set of disparities so as to produce discomfort or fatigue of the visual system. But, as it was mentioned earlier, the definition of “large” and “small” stereoscopic demand is idiosyncratic to the individual (Howarth, 2011) and this is what makes difficult the study of subjective discomfort in stereopsis.

A second category of results was the effects obtained on the stereoscopic visual search task. The difficulty of this task was shown to increase with finer depth difference of the slanted gratings as it could be observed from the decrease in the percentage of correct responses and the increase of the reaction times. When observers were facing fine depth rotations, the luminance condition appeared to have a positive effect on the percentage of correct responses with high peripheral luminance. However, this benefit could be due to an attentional effect; high luminance could have forced observers to maintain intensely their attention on the task as compared to low luminance. This possibility is also supported by the benefits found on the reaction time.

Finally, the lack of effects we obtained by manipulating the disparity range could also be explained by the presence in each luminance condition of the frame that was used to reflect the projected light intended to vary the pupil aperture. This rectangular white frame included a circular aperture that allowed the participants to see the display in order to perform the task. However, due to the limited size of this circular aperture, it is possible that participants experienced a pinhole effect on accommodation as that can be observed when they perform a spatial task by looking through distant holes (e.g. Henessy & Leibowitz, 1971). As a consequence, if accommodation was affected both in low luminance condition and in high luminance condition by the peripheral frame used to reflect the light that had to change the pupil size; this may be the reason why we failed at observing a strong effect of the range of disparities. To conclude, even though the experimental set-up really affected the pupil size aperture, this study does not provide strong enough empirical elements to identify an effect due to the increase of the depth-of-focus.

♦

The study of visual discomfort and fatigue in stereoscopic displays seems to require a consistent number of participants if the major experimental variable measured is highly subjective. Even though this experiment did not provide strong significant results and suitable control conditions, this was a first step to manipulate the pupil size aperture in order to investigate the role of the depth-of-focus in depth perception. Indeed, in the following chapter, the presented studies used a similar approach to vary the pupil size but the experimental set-up was optimized. For example, the loss of contrast due to miosis with high luminance was compensated, the non-illuminated aperture in projected light was increased to avoid some pinhole effects and, the light was projected onto a diffuser that was located on the side of the participants, reflected by a beam splitter. The following chapter reports the study of 3D shape perception based on binocular disparity in relation to the depth-of-focus of the eyes.

Chapter 5. The weight of accommodation in 3D shape perception with light-modulated depth-of-focus

5.1 Introduction

A well known fact about stereoscopic vision is that perceived depth from stereopsis suffers from misestimation of binocular disparity information (e.g. Adams & Mamassian, 2004; Johnston, 1991). Binocular disparity can provide an unambiguous cue to depth if it is correctly scaled by a proper estimate of the egocentric viewing distance (Garding, Porrill, Mayhew & Frisby, 1995). This estimate of viewing distance is presumably relying on the actual vergence and accommodative cues (Watt, Akeleyn Ernst & Banks, 2005; MonWilliams & Tresilian, 2000), and on the size ratios of horizontal and vertical disparities (Rogers & Bradshaw, 1995), but the accuracy of this estimate and the information used for scaling are distance-dependent (Domini, Caudek & Tassinari, 2006; Cutting & Vishton, 1995). The main distinction between stereoscopic and real-world conditions of observation is considered to be the mismatch between the stimulus-to-accommodation and the stimulus-to-vergence. This gives rise to at least two consequences: one directly affects the retinal signal whereas the other indirectly influences the extra-retinal signal and thus the oculomotor activity. By presenting disparate information for the left and right eyes on a flat display, a 3D structure may be specified but, the corresponding retinal blur gradient is discrepant with the stimulation. This fact has been considered as a potential origin of the misestimation of the 3D structure simulated in stereoscopic depth (Hoffman, Girshick, Akeley & Banks, 2008; Watt et al., 2005). A few studies revealed that focus cues could bring more than range information about distance and even provide quantitative information about the egocentric distance (e.g. Vishwanath & Blaser, 2010). The second consequence comes from the indirect influence of the incorrect blur information in stereoscopic systems on the synkinesis (associated movements) between accommodation and vergence systems (Frisby, Buckley & Horsman, 1995). In natural viewing conditions, accommodation and vergence responses have been shown to be reciprocally coupled (Schor, 1992); some interactive neural processes ensure that both eye-movements react to either signal (Semmlow & Hung, 1981). As a matter of fact, changes in the activity of the blur-driven and the disparity-driven components mechanisms could alter the performance of vergence and lead either to perceived distortions and/or to visual fatigue and discomfort.

Perceiving precise 3D structure from stereopsis has been questioned in a wide literature investigating stereoscopic depth constancy (e.g. Johnston, 1991; Ritter, 1977; Wallach & Zuckerman, 1963), exploring to which extent the perceived depth from binocular disparities remains the same when the viewing distance varies. Poor depth constancy was observed when the construction of a 3D perceptual structure was involved as opposed to when observers performed a disparity matching task (Glennerster, Rogers & Bradshaw, 1996; Johnston, 1991). A concrete example was provided in a study using stereoscopic elliptical half-cylinder (Johnston, 1991), in which was varied the distance of the screen plane depicting the stereograms. It was found that percepts were elongated at close viewing distance, accurate at intermediate distance and flattened at far distance. These perceived distortions were interpreted as the consequence of using an erroneous measure of egocentric distance for scaling horizontal disparities because observers misperceived the 3D shape specified only by disparity information in a way that was related to the viewing distance. Watt et al. (2005) reported other effects of perceptual distortions linked to the variations of the vergence distance (the distance induced by horizontal disparities) with accommodation distance fixated. When the simulated (vergence) distance was less than that of the accommodation distance (crossed disparities), observers over-estimated the depth of the 3D shape, whereas they under-estimated depth when the simulated distance was larger than the accommodation distance (uncrossed disparities). More specifically, a right angle displayed in an open-book configuration with an intended distance nearer than the projection plane appeared smaller than 90 degrees (acute) because the percept was elongated in depth. When the right angle is now displayed for an intended distance located farther than the projection plane, the angle is then seen flatter (obtuse). At this stage, it is worth noting that 3D estimations for congruent accommodation and vergence distances were drastically more precise than those with disparate accommodation and vergence distances. This phenomenon has been attributed to the indirect effect of focus cues throughout the coupling of accommodation and vergence; focus cues could influence vergence through the accommodative vergence mechanism. These results have also been reproduced in an experiment using a multiple-focal-plane display enabling appropriate control conditions (Hoffman et al., 2008). The authors also observed longer reaction times with incongruent accommodation and vergence distances, suggesting that the decoupling between accommodation and vergence systems lead to greater difficulties to fuse. By converting the results of 3D estimation in equivalent distances, the authors showed that the distance of objects displayed stereoscopically for an intended location nearer than the accommodation distance is overestimated and underestimated for locations beyond the accommodation distance plane.

The literature reviewed so far suggests that recovering depth from stereopsis requires the establishment of an egocentric distance mainly thanks to oculomotor factors. In the case of stereoscopy, people would misestimate depth because one of the sources of information conveys flatness and, because of the existence of crosslink between vergence and accommodation. The first assertion was labeled the blur-cue hypothesis and supports the idea that the lack of correct blur gradient information is critical for perceiving depth from stereopsis with veridicality. The second statement stems from the modeling of the synkinesis between accommodation and vergence and was labeled the indirect effect of focus cues as it may impact the vergence system (Watt et al., 2005). The idea suggested in order to test these hypotheses was to constrain accommodation by asking people wearing pinhole during a stereoscopic task (Hogervorst & Eagle, 2000; Frisby, Buckley & Horsman, 1995). By looking through a pair of pinholes, the retinal projection of the visual object becomes sharp, and the blur circles information is suppressed or diminished (Ward & Charman, 1987). If accommodation thus would participate in the systematic distortions, the pinhole condition of observation should cancel the effect. Hogervorst & Eagle (2000) tested this condition of observation on one observer and did not find any effect; they concluded accommodation has no effect on 3D estimation of dihedral angle. Frisby et al. (1995) dedicated a set of three experiments so as to investigate the effect of pinhole viewing of real 3D and stereograms shapes. In a previous study, Buckley and Frisby (1993) had showed a surface-orientation anisotropy in the perception of 3D ridges and, more important in the scope of our study, that disparity cues dominated texture cues in all real stimuli whereas they did not do so in all stereograms. From these results, the motivation of researchers was then to investigate whether this anisotropy and the weakness of stereo cues in stereoscopic conditions as compared to real world conditions of observation were related to focus/blur cues (Frisby et al., 1995). In the first experiment, in order to render accommodation open-loop, they used “aerobic glasses” as pinhole viewing apparatus that consisted of aligned binocular arrays of pinholes of 0.75 mm diameters set approximately 3 mm apart. The authors found some evidence that the surface-orientation anisotropy (change in cue weighting between disparity and texture as a function of stimulus orientation) in the perception of real ridges could be related to the inappropriate blur cues. Such incorrect cues would weaken the efficacy of the disparity cue to convey depth and would leave a greater role to texture cues. In another experiment, they tested the pinhole arrays for the perception of stereoscopic simulated ridges. The prediction was that disparity signal should be stronger with pinhole viewing of stereograms, assuming that the accommodation open-loop conditions are thereby attained, and the blur cues removed. It was shown that pinhole viewing neither affected the anisotropy nor the perceived depth of simulated ridges and, as a consequence, in this case the results were interpreted against the blur-cue hypothesis. According to the authors, considering the possible role of resting accommodation however still supports the latter hypothesis because the reference egocentric distance taken into account for estimating stimuli could approximate the fixated accommodation distance related to dark focus. Later, Watt et al. (2005) found a significant contribution of focus cues (accommodation and retinal blur) in the estimation of simulated right angle. As a consequence, they re-interpreted the results by Frisby et al. (1995). They concluded that the method of pinhole is not suitable for studying blur cues effect. Indeed, the decreased perceived depth of real ridges with pinhole viewing could have been due to an increase of the depth of focus – the focus range for which vision remains sharp – causing a greater signal to flatness.

As it was indicated in the beginning of this introduction, some neural crosslinks exist that allow accommodation and vergence to react both to retinal disparity and to defocus (Hung, Ciuffreda & Rosenfield, 1996; Schor, 1986; Semmlow & Hung, 1981). In this way, in the case of stereopsis, when the depth-of-focus is increased, the conflict between accommodation and vergence should be reduced; accommodation response should be unnecessary. In Frisby et al. (1995), the lack of results in the stereograms condition is surprising because an indirect effect could thus have resulted from the state of the accommodative lens that could be used as a signal for estimating the scaling distance. However, the absence of effect of pinhole viewing and the slight decrease in perceived depth on the simulated stimuli could have also been due to the pinhole apparatus. As a matter of fact, several points should be put forth concerning the use of such pinhole arrays. The misalignment of the holes with the visual axis could have had some consequences on eye movements and could either make invisible some parts of the visual stimulus or make the light coming from several holes contributing to the retinal projection. The misalignment of the holes as well creates a chief ray not aligned with the eye optical axis, displacing the image on the retina due to e.g. spherical aberration. Another point to note is the fact that if the lighting level is not compensated, the eye pupil will increase behind the partially occluding surface. The pupil aperture of the eye may be larger because placed behind the array of pinholes which diminished the light entering in the eye, and so the amount of light casting the melanopsin receptors responsible for the pupil light reflex (Viénot, Bailacq & Rohellec, 2010). As a consequence, the fact that observers were rather placed in quasi-open-loop conditions must be taken into account in explaining the absence of results for stereoscopic simulated ridges.

The use of pinholes shows that accommodation is rendered open-loop and that the accommodation system may consequently adopt its resting state (Ward and Charman, 1987). This lessens the interest in using pinholes because, rather than simply removing the blur cue, pinhole viewing could affect the oculomotor contribution to distance estimate (Frisby et al., 1995). As a consequence, a more relevant and ecological way to affect the accommodation response is to directly manipulate the depth-of-focus of the eyes. One way to change the depth-of-focus of the eyes without affecting normal binocular eye movements is to vary the pupil size of the eyes. The pupil aperture is directly connected to the depth-of-focus of the eyes in such a way that decreasing the pupil size will optically increase the depth-of-focus (see figure 1 for example). By extending the depth-of-focus through the pupil size manipulation, the tolerance for the accommodation system is increased and the blur cue is rendered less discrepant with the viewing situation. In addition, as the accommodation system does not take its resting position, the vergence system is unaffected.

Even though the depth-of-focus of the human eye depends on many parameters, it is possible to consider two simple parameters for approximation, namely the pupil diameter and the blur circle. Considering the eye as a simple thin optical system, the depth of focus can be defined as the difference between the plane of perfect optical focus on the retina and the plane that produces the just-detectable blur circle on the retina (Green, Powers & Banks, 1980).

Hence, the depth-of-focus:

$$\Delta D = \pm(n / f - n/(f + \delta)),$$

where n is the refractive index of the eye, f is the focal length of the eye and δ is the amount required to move the focal plane on the out-of-focus image point. In addition, from Figure 5.1, on can derive the following relation:

$$\Delta D = \pm bc.n / P.f,$$

where bc is the extent of the blur circle and P is the pupil diameter. It can be seen that the depth-of-focus is inversely proportional to the pupil aperture size. Decreasing twice the pupil aperture will theoretically increase twice the depth-of-focus. However, according to Campbell (1957), the eyes' pupil gives larger depth-of-focus as if the pupil diameter is smaller and, therefore larger depth-of-focus can be expected as compared with the one that can be estimated using this optical model.

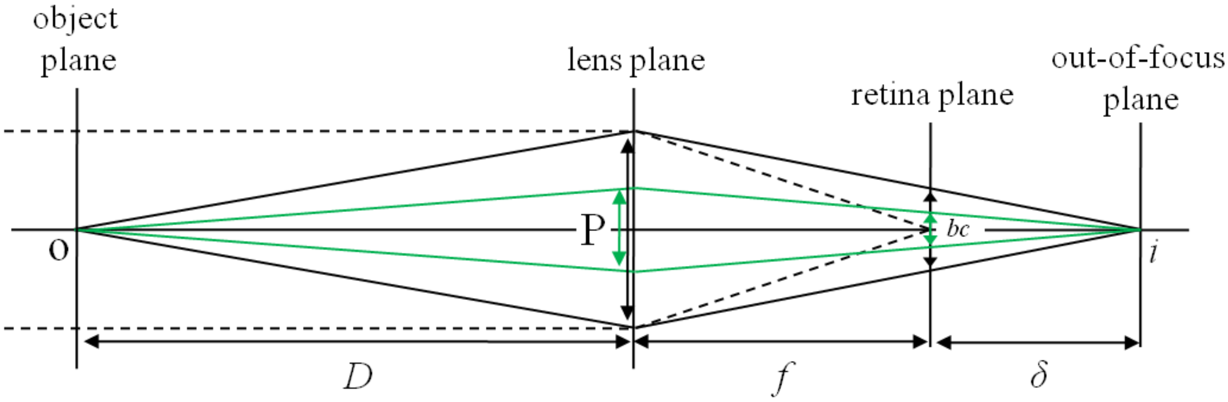

Figure 5.1: Optical representation of the effect of changing the pupil aperture (P) on the retinal projection, the blur circle (bc). In this configuration, an object at theoretical infinity is perfectly focused on the retina (retina plane). Object O is at an out-of-focus distance (D) and creates a blur circle (bc) on the retina plane because its conjugate projection (i) lies on another plane than the one of the retina. Comparing the out-of-focus blur circle size with the size of the blur spot on the retina defines the depth-of-focus. The green lines configuration represents the effect for a pupil aperture that is twice as small as the black lines configuration: the blur circle is half.

Contrary to the aim of Frisby et al. (1995), the objective of the present study does not consist in inferring cue-integration model of real world vision conditions but to account for the misestimations that are usually reported in perceiving shape from stereopsis. Here the focus is put on the contribution of the accommodation signal in relation to the depth-of-focus of the eye in the estimation of stereoscopic 3D shape. Watt et al. (2005) found that the bias direction in judging 3D shape with congruent but weak texture signal is dependent on the vergence distance with an indirect effect of the accommodation distance. Thus, stereo-object depicted at a vergence distance induced by crossed disparities is perceived elongated in depth while object presented at a vergence distance induced by uncrossed disparities is seen flatter. The results of Frisby et al. (1995) did not show this tendency as they did not vary the vergence distance. Here, the contribution of accommodation in 3D shape perception is investigated by observing whether people can correctly estimate the 3D shape of a displayed stereogram as a function of the vergence distance and the possibly extended depth-of-focus. In a first experiment, we investigated the effect of decreasing the pupil size with lighting conditions on 3D shape estimation using stereograms of dihedral vertical angle. In a second experiment, we used the same type of apparatus to study the effect of increasing the depth-of-focus through lighting conditions when the accommodation and vergence distances were varied. The modulation of the depth-of-focus with different lighting conditions has the advantage of removing a consistent part of blur cue in the visual signal and, contrary to the use of pinhole, will not place the accommodation system in its resting state (or dark focus), thereby leaving the vergence system unaffected. Considering the double contribution of focus cues to depth perception (direct vs. indirect), changes in accommodation would produce two potential effects on the viewing distance estimate: (1) depth-of-focus could in itself act as a distance cue and, (2) increasing depth-of-focus could diminish accommodation-vergence conflict. The effect of focus cues should be larger with stronger stereoscopic demand (i.e. distances induced by disparity should impact judgment of 3D shape) and, decreasing the pupil size should decrease this vergence distance effect because of the increase of the depth-of-focus (i.e. due to decreasing pupil size) and the decrease of the mismatch between focus and vergence cues. The overall objective of this experiment is to study whether an increase of the depth-of-focus of the eyes could reduce the weight of the accommodation signal and thus affect the perception of 3D shape from binocular disparity. Conflict between focus cues and vergence cues could be directly dependent on the depth-of-focus which manipulation could have the potential to overcome the difficulties of 3D displays.

5.2. Experiment 1

5.2.1 Method

5.2.1.1 Observers

A total of 20 observers took part in the entire study. All had normal stereoacuity as attested using the Randot StereoTest (Stereo Optical Co) and normal vision possibly thanks to their usual optical corrections. Three other participants were discarded because of either large difficulties in fusing the stimuli during the experimentation or inability to perceive the 3D structure. Our selection provided a full counterbalanced group of participants (on average 32.8 years old, SD = 10.1 years old). They gave their informed consent before beginning the experiment.

5.2.1.2 Apparatus

Participants sat comfortably in a chair within a chinrest which was placed at 1.3 meters from the display (Hyundai S465D 46'' HDTV LCD Polarized monitor). To minimize crosstalk, we placed participant eyes on the axis perpendicular to the center of the screen. We used two computers in this experiment, the first displayed 3D content using Matlab and the Psychtoolbox extensions (Brainard, 1997), the second allowed us to control the progress of the experiment. Participants wore polarized glasses to fuse left / right views and see the content in depth. In order to vary the pupil size, a video-projector (Optoma EP747) was set just beside on the right of the observer (see Figure 2). The light stemming from the projector went across a diffuser (36.4 x 31.5 cm) in front of it. Then, the diffused light was reflected by a beam splitter whose center was placed at 50 cm from the eyes of the observer and the orientation was 45 degrees so that the diffuser reflected light appeared to come from the front. In the condition of high luminance, the projected light consisted of a white rectangle (35 x 30 cm) in which a black elliptical aperture (24.8 x 14.8 degrees) was superimposed to leave the center of the visual field unaffected by luminance. The size of this aperture was chosen large enough to avoid the pinhole effect on accommodation (e.g. Henessy & Leibowitz, 1971).

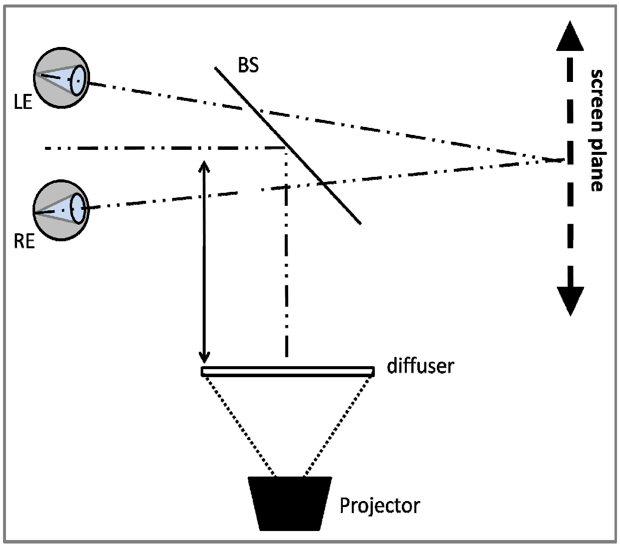

Figure 5.2: Schematic representation of the experimental apparatus used in the study. In this set-up, white light is projected onto a diffuser; the subsequent diffused light is reflected on the beam splitter to illuminate the entrance pupil of the two eyes. Note that the projected light contained a non-illuminated central elliptical part for observers to be able to see with adequate contrast the stereoscopic display placed in front of them. LE and RE refer to Left Eye and Right Eye, BS to Beam Splitter. Drawing is not to scale.

5.2.1.3 Stimuli

Participants had to judge the angle of a vertical dihedron (as in an open-book configuration) which was depicted using random-dot stereograms. This stimulus has already been used elsewhere and showed high reliability in accounting for the capacity to perceive 3D shape (Hoffman et al. 2008, Watt et al. 2005). The task was to fixate a small fixation cross (53 arc min) which was placed in the center of the dihedron and then to judge whether the displayed dihedron angle was smaller or larger than 90°. Participants were tested during different days for two conditions of luminance and had to give judgment both for a sharp stimulus and for a blurred stimulus. The level of blur of the dihedron angle was randomly displayed and generated by applying an uniform averaging filter (8-by-8 filter containing equal weights); we could therefore analyze and compare the effect of presenting stimuli without high frequencies to stimulus conditions with large depth-of-focus (i.e. with a decreased blur gradient). In the blur condition, defocus could not be used as distance cue because it was uniform. The dot density for the dihedron was 0.7 dots/deg2, with a width of 9.2 degrees of visual angle and a height of 4.2 degrees. Dihedron dots angular size was 0.26 degrees; the dot density of the background was 6.7 dots/deg2 background dots angular size was 0.1 degrees. In the experimental part, the fixation cross could be displayed at three possible intended depth locations from the screen plane (0, 200 and 340 mm) while the parts of the hinge always appeared in front of it. As a consequence, the position in depth of the fixation cross created stereoscopic demand of 0, 0.14 and 0.27 Diopters (D) and, because the averaged peak-to-through of the angle was 0.07 D, the maximum stereoscopic demand could attain 0.34 D. Hence, considering a 0.22 D depth-of-focus estimated for a 3 mm pupil aperture (Campbell, 1957), the two first depth locations of this experiment remained in the depth-of-field instead of the third one. Before each experimental session, we carried out a familiarization task (with physical model and audio feedback and then only with the audio feedback) during which observers had to obtain a score of 90% in judging the dihedron angle. For the familiarization task, we used the method of constant stimuli and presented dihedron angles in the range of 81 to 99° (10 stimuli intensities separated by 2 degrees and the whole repeated four times until participants reached an expected score). Each dot was placed relative to the cyclopean eye by projective geometry. For the experimental session, the angles of each dihedron were displayed according to an adaptive staircase procedure (Accelerated Stochastic Approximation, Kesten 1958). This procedure allowed us to estimate the point of subjective equality (PSE) and the slope at this point for each block and session. The maximal number of repetitions was set to

15 for each staircase. During one illumination level session, observers participated to 2 threshold estimations (25% and 75%) using 2 staircases (one ascending & one descending) per threshold estimation for a total of 60 trials either for the sharp and the blurred stimuli (a total of 120 trials). This procedure was repeated 3 times so as to obtain 3 estimations of the psychometric function for each possible depth location.

5.2.1.4 Procedure

Two conditions of illumination were set up in order to vary the pupil size; we labeled them the high luminance and the low luminance condition (2000 cd/m² for the white light reflected on the diffuser periphery and 0 cd/m² in absence of illumination). It is worth noting that changing the participant's pupil aperture affected the perceived contrast of the stimuli simply because less light enters in the eye with a decreased pupil size. This is the reason why we carefully adjusted the luminance of the stimuli according to a standard RGB model of the display EOTF response. The objective of this transform was to keep the same perceived visual signal between the high and low luminance conditions for the central visual field. The maximum luminance on the display of black, gray and white points was measured using the Minolta Colorimeter (CS-200) and, considering that in this setup, the mean pupil size is half in the high luminance condition as compared to the low luminance condition, the luminance needed for the low luminance condition were computed with a gamma curve transfer function adapted to our display. We subjectively checked the outcomes of these transforms by asking several participants to compare the perceived contrast in the two conditions.

Before each condition of illumination, observer carried out the training regime as described above. Each session was run independently, on two different days, so as to keep the pupil size approximately similar during the same session of illumination. The two levels of illumination were counterbalanced across participants. At the onset of each experimental session, the variation of the pupil size aperture in response to illumination was measured. Observers looked at a fixation cross that was displayed in the middle of the screen plane while the pupil size of the right eye was measured through a camera capture. The measurement was made after a fixation duration of two minutes in such a way the pupil size was stabilized and adapted to the projected diffused light.

5.2.2 Results

The condition of illumination had a significant effect on the pupil size aperture ($t=17.9$, $df=19$, $p<0.001$). On average, the pupil size was halved in the high luminance condition (1.4 mm, SD=0.17) as compared to the low luminance condition (3 mm, SD=0.44). Psychometric functions were fitted with cumulative Gaussian curves on the probability distribution and the points of subjective equality of the dihedron angle estimation were derived using probit analysis (PSE, the 50% point on the psychometric function). An example of one participant can be found in Figure 3. The analysis performed on the PSEs for the twenty participants revealed a significant effect of vergence distance ($F(2,38)=47.62$, $p<0.0001$); as the displayed hinge angle gets closer, it appeared more stretched in depth. A significant effect of the lighting condition was also observed ($F(1,19)=8.65$, $p<0.01$) leading to a decrease in the displayed angle required to perceive a right angle in the high luminance condition as compared to the low luminance condition. The interaction effect between the vergence distance condition and the lighting condition was not significant ($F(2,38)=0.67$, $p>0.05$). A slight blur effect was also obtained ($F(1,19)=8.02$, $p<0.05$); the perceived hinge angle was perceived flatter when it was blurred as compared with the sharp stimulus. The interaction between the blur condition and the vergence distances was not significant ($F(2,38)=0.23$, $p>0.05$). However, we found an interaction effect between the blur condition and lighting condition ($F(1,19)=24.76$, $p<0.001$); the lighting condition effect was greater (the dihedral angle was perceived flatter for high luminance relative to low luminance condition) when the stimulus was blurred. No second order interaction effect was observed between the blur condition, the lighting condition and the vergence distances ($F(2,38)=0.54$, $p>0.05$). These results are summarized in Figure 4.

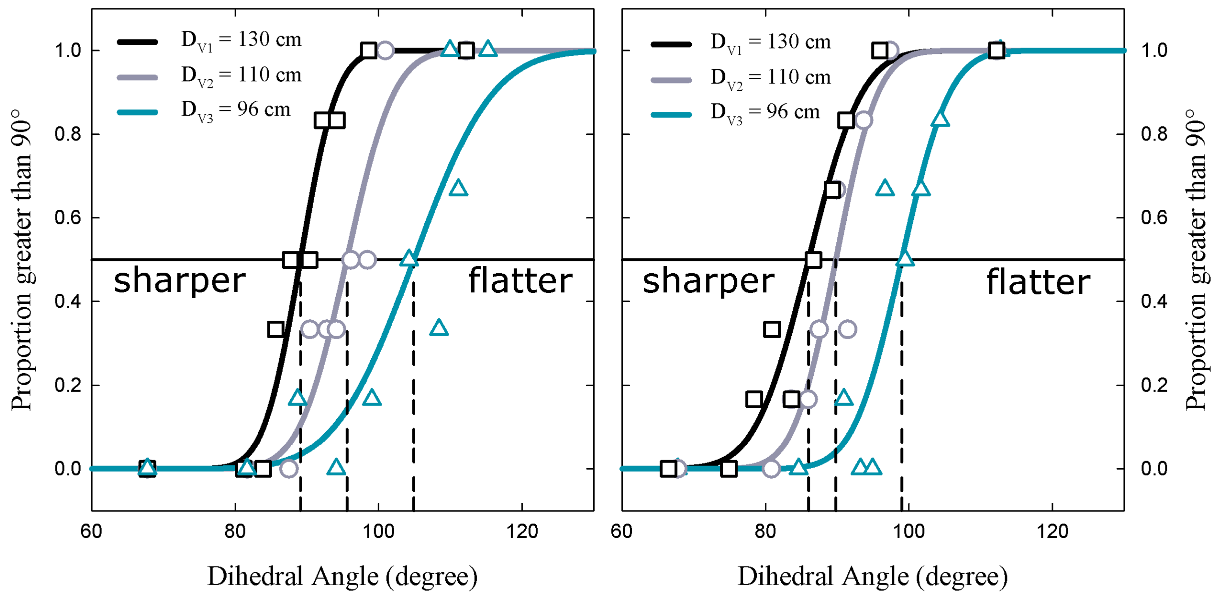

Figure 5.3: Example psychometric functions for one observer at an accommodation distance of 130 cm and for three vergence distances (D_V = 130, 110 or 96 cm). The left figure corresponds to the results for the low luminance condition and the right figure concerns the results for the high luminance condition. The proportion of trials in which the observer responded "greater than 90°" is plotted as a function of the dihedral angle. Vertical dashed lines correspond to the estimated PSEs.

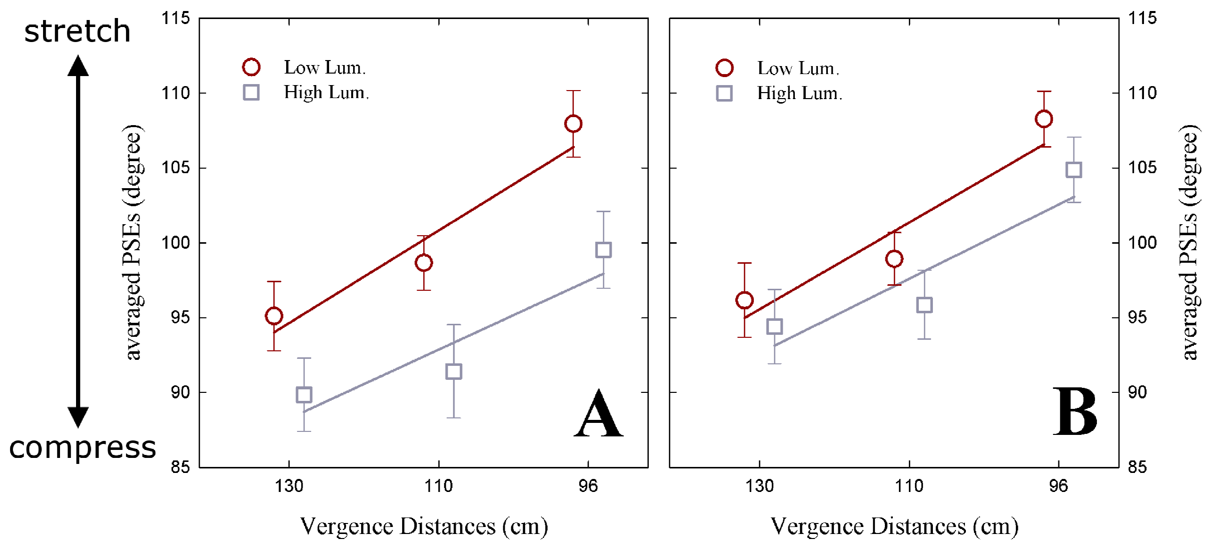

Figure 5.4: This figure represents the averaged PSEs of the twenty participants as a function of the vergence distances induced by disparity (130, 110 or 96 cm) both for the low luminance condition (corresponding to large pupil, in dark red) and the high luminance (corresponding to small pupil, in dark gray). The left panel (A) represents the results for the blurred stimulus and the right panel (B) represents the results for the sharp dihedral angle. The

left figure corresponds to the results for the low luminance condition and the right figure concerns the results for the high luminance condition. PSEs going down means perceiving fixed dihedral angle stimuli as flatter as distance increases. Solid diagonal lines are linear regression lines performed on the averaged PSEs over the three possible vergence distances. Error bars denote the standard error of means.

The slope of psychometric function describes the sensitivity of the participant to the effect of changing stimulus intensity; the steeper the slope, the better the discrimination between the possible values of the stimulus. Conversely, a shallow slope will indicate coarse discrimination of stimulus intensity changes. Hence, supplementary analyses performed on the slope of the PSEs (defined as the slope at the tangent of the PSE) for the twenty participants revealed no effect of vergence distances on the slope but a tendency to decrease ($F(2,38)=2,68$, $p=0.08$). Conversely, an effect of the blur condition showed an increase in slope value when the stimulus was blurred as compared to the non blurred stimulus ($F(1,19)=7.50$, $p<0.05$) and an effect of the lighting conditions showing an increase in slope value in high luminance as compared to low luminance condition ($F(1,19)=10.38$, $p<0.005$). As for the analysis on the PSEs, the interaction effect between the blur condition and the lighting condition was significant ($F(1,19)=5.19$, $p<0.05$). Other analyses performed between the position of the maximum reaction time and the PSE showed interesting results presented in Figure 5.5. For the two conditions (Luminance vs. Blur), the position of the maximum reaction time (relative to the stimulus intensity in degree) is plotted as a function of the estimated PSEs for each participant and the three vergence distances. In normal viewing conditions, the position of the maximum reaction time and the PSE are correlated because the difficulty to make a decision is normally maximal for the stimulus intensity for which the participant cannot say if the angle is smaller or larger than 90 degrees. The coefficients of correlation are reported and show to be decreased for the third vergence distance (i.e. the maximum stereoscopic demand) specifically when the depth-of-focus is small (low luminance condition: panels A and B) as compared to when the depth-of-focus is large (high luminance condition: panels C and D).

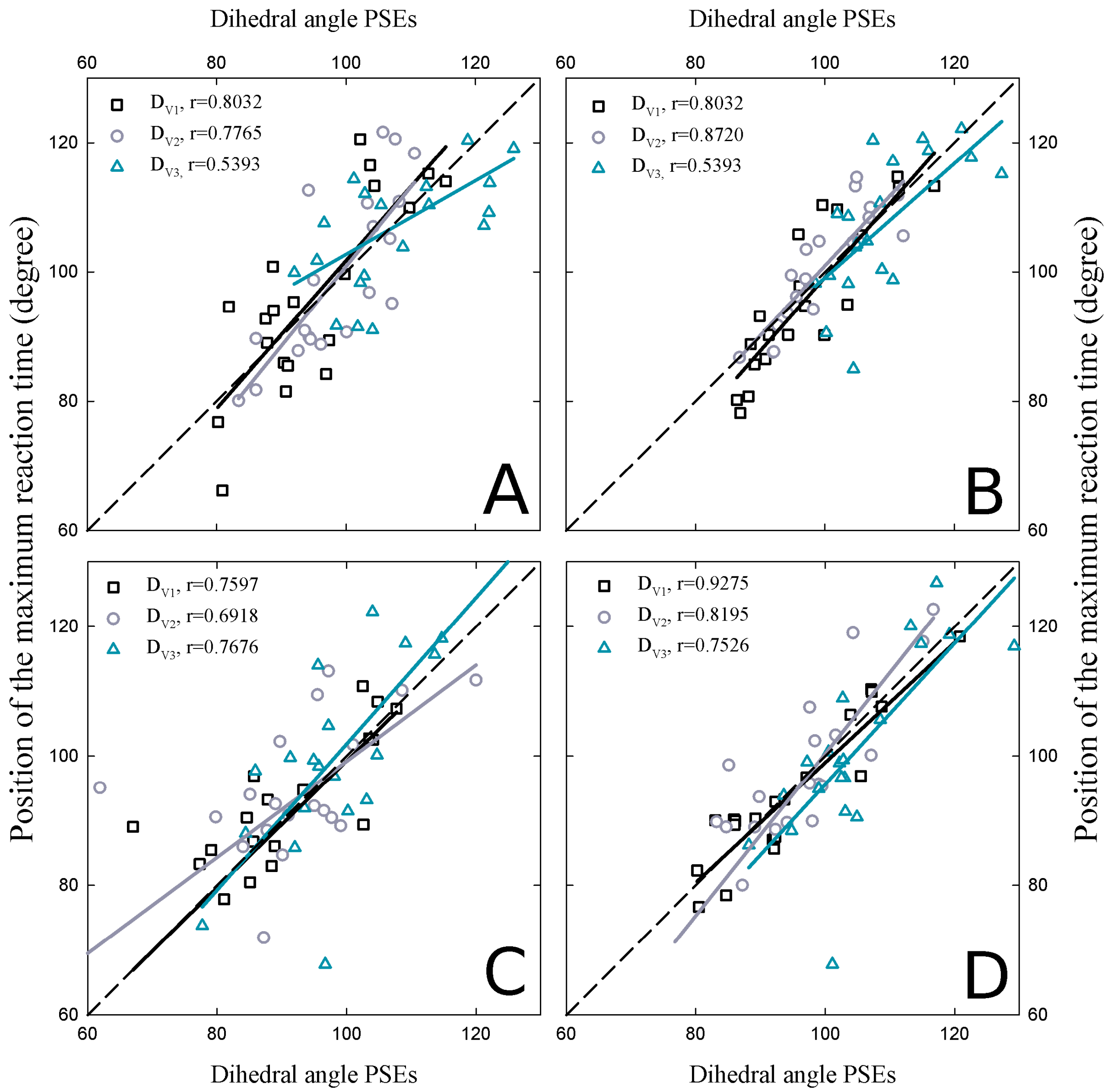

Figure 5.5: This figure represents the position of the maximum reaction time (RT) as a function of the estimated PSE of the twenty participants for each vergence distance induced by disparity: D_{V1} = 130 (in black), D_{V2} = 110 (in gray) or D_{V3} = 96 cm (in dark cyan). Panel A represents the results for blurred stimulus in the low luminance condition, panel B represents the results for the sharp stimulus in the low luminance condition, panel C represents the results for the blurred stimulus in high luminance condition and panel D represents the results for the sharp stimulus in the high luminance condition. Dashed diagonal lines are perfects matches between maximum reaction time position and PSE. The figure also provides for each distance the Spearman coefficient of correlation (all correlations significant, $p<0.05$).

5.2.3 Discussion

The objective of this first experiment was to investigate whether a change in pupil size could affect the perception of 3D shape in stereoscopic viewing conditions where the vergence distances induced with crossed disparities were varied. Indeed, several pieces of experimental evidence suggest that the source of information from the extra-ocular vergence muscles and accommodation muscles can be used in the estimate of viewing distance (i.e., Watt et al. 2005; MonWilliams & Tresilian, 2000). In addition, the accommodation system is known as being under the control of feedback loop and thus reacts to defocus blur when its amount becomes detectable in spite of the visual system tolerance (i.e. the depth-of-focus). Because the use of pinhole can affect both the vergence state of the eyes due to misalignment of the visual axes with the pinhole eyelet and the tonic state of accommodation, the present study proposed to change the pupil aperture size using luminance changes. In order to weaken the potential dazzling effect, only the peripheral part of the visual field was enlightened. Hence, two lighting conditions were used to provide either a large pupil aperture (an average of 3 mm) or a small pupil aperture (an average of 1.4 mm), therefore providing a change in pupil size with a factor close to two. The loss of contrast of the observed stimulus was consequently compensated by increasing the luminance level of the displayed content after measuring the normal luminance levels of the display so as to present closely similar perceived contrasts. In addition, we also manipulated the amount of image blur in the stimulus to investigate the potential changes given its propensity in rendering the accommodation system open-loop (Schor et al., 1984).

In the present experiment, participants were asked to judge whether a dihedral angle was acute or obtuse for three possible virtual distances in front of the screen plane. It was expected an effect of the distances induced by disparity on the ability to judge 3D shape as it was previously reported (e.g. Hoffman et al., 2008). A second expectation was that decreasing the pupil size by illuminating the peripheral field should affect 3D shape perception because of the role of accommodation in interpreting 3D structure (Watt et al., 2005). The vergence distance had a significant effect on observers' ability to judge the dihedral angle as a right angle, and also showed a tendency for decreasing the sensitivity in judgment with the increase in virtual distance from the screen plane. Specifically, the dihedral angle was perceived more and more compressed (i.e. the PSE is stretched) when the observer-stimulus distance induced by disparity (vergence distance) was reduced (i.e., the stereoscopic demand is increased). The depth was overestimated because, as the angle was judged flatter to mediate a right angle, the amount of disparity was smaller and sufficient to provide the intended perception. Hence, the results obtained on the PSEs indicated a general decrease in the underestimation of the dihedral angle with the increase in luminance of the peripheral visual field. This effect was accompanied by an increase in sensitivity in the high luminance condition as compared to the low luminance condition, attesting that the high luminance condition effect could not be merely reduced to a contrast effect on stereopsis. Interestingly, these results indicated that increasing the depth-of-focus of the eyes by illuminating the peripheral field of view potentially reduces the overestimation of depth that is observed for crossed disparity-induced objects and is moreover accompanied by an increase in sensitivity. Even though an interaction effect between the vergence distance condition and the lighting condition could have reinforced this interpretation, we did not observe such an effect.

In addition, image blurring showed an effect on the perceived hinge angle in such a way that observers perceived the angle flatter (less underestimated). This effect was accompanied by an increase in sensitivity (slope at PSE) for blurred stimuli as compared to sharp ones, supporting that blurring could not be reduced to a simple issue of contrast effect on stereopsis. Interestingly, an interaction effect was reported between the image blurring condition and the luminance condition both for the PSE and the sensitivity. In the high luminance condition as compared to the low luminance condition, dihedral angles were perceived flatter, observers were more likely to detect a fine difference between the stimulus variation and these effects were larger when the stimulus was blurred. This therefore suggests that image blurring and high peripheral luminance were complementary for reducing the overestimation of depth from binocular disparity with an increased certainty in the response due to the increased sensitivity to the stimulus variation.

A graphical analysis performed between the position of the maximum reaction time and the PSEs provided results supporting the fact that the luminance condition really affected the extent of the depth-of-focus. In two forced-choice tasks where observers are asked to respond as quickly as possible, when the reaction time distribution and the probability of response are plotted against the stimulus intensity, the position of the maximum reaction time and the PSE coincided and are normally correlated. From Figure 5, it can be seen that overall coefficients of correlation between the position of the maximum reaction time and the PSE are high at least for the first and the second vergence distances (respectively, 130 and 110 cm). A striking difference can be noted for the third vergence distance (the strongest stereoscopic demand).

Indeed, the coefficients of correlation between the position of the maximum reaction time and the PSE decreased in the two low luminance conditions whereas they were more stable in the two high luminance conditions. This graphical analysis supports the idea that the increased luminance expected to increase the depth-of-focus of the eyes seemed to have an effect on the way observers responded to the task. This result thus suggests that increasing the depth-of-focus probably improves the way observers estimate 3D shape.

In summary, decreasing the pupil size by increasing peripheral luminance affects the estimation of 3D shape such that overestimated angles in front of the screen tend to be less overestimated and observers become more sensitive to stimulus variation. This general result could be related either to depth-of-focus or contrast issues, but the results obtained on observers' sensitivity (on the slopes at PSE) and reaction times suggest that the increase of the DOF could better be the reason for this change. Image blurring on one side and light-modulated depth-of-focus on the other side can thus lead to the same effect but, however, the present study does not provide any result that would be consistent with the fact that both effects share the same cause. The results of this first experiment need to be attested by further experimental evidences because increasing the depth-of-field through lighting conditions may have also engendered a dazzling effect, thereby providing this bias direction. As a consequence, we conducted a second experiment that examines the effect of manipulating vergence and accommodation distances and varied light-modulated depth-of-focus.

5.3 Experiment 2

5.3.1 Method

5.3.1.1 Observers

A total of five participants: three out of the twenty previous observers and two new took part in the entire study. They were on average 36 years old (SD = 11 years old). All had normal stereoacuity as attested using the Randot StereoTest (Stereo Optical Co) and normal vision possibly thanks to their usual optical corrections. They gave their informed consent before beginning the experiment.

5.3.1.2 Stimuli

Participants had to judge the same stimuli as in Experiment 1. Participants were identically tested on different days for two conditions of luminance and provided judgment only for sharp stimuli. In this experiment, we varied both the accommodation distance (screen-constrained distance) and the vergence distance (distance induced by binocular disparity) and tested all possible combinations. The accommodation distance was manipulated by displacing the screen display distance and the corresponding vergence distance was then computed for this distance. The three possible viewing distances (relative to the fixation cross) were 96, 130 and 201 cm and, there were therefore nine possible stimulus conditions. Angles of each dihedron were displayed according to an adaptive staircase procedure (Accelerated Stochastic Approximation, Kesten 1958). This procedure allowed us to estimate the point of subjective equality (PSE) and the slope at this point for each block and session. The maximal number of repetitions was set to fifteen for each staircase. During one illumination level session, observers participated to two threshold estimations (25% and 75%) using two staircases (one ascending & one descending) per threshold estimation for a total of sixty trials. This procedure was repeated nine times so as to obtain three estimations of the psychometric function for each possible viewing distance. The cyclopean angular size of the dihedron was kept constant across distances.

5.3.1.3 Procedure

Each lighting condition was run independently, on two different days, so as to keep the pupil size approximately similar during the same session of illumination. The experiment was run by block of accommodation distance, but the order of presentation of the vergence distance and the accommodation distance was randomly selected. 5.3.2 Results

5.3.2.1 Effect of lighting conditions over different viewing conditions

From the PSE estimates, we calculated the equivalent/scaling distance for each participant and each vergence distance (see, i.e., Glennerster, Rogers & Bradshaw, 1996). The results can be observed in Figure 6. The analysis performed on these equivalent distances for the five participants were made independently for the two lighting conditions. In the low luminance condition (Figure 6 – panel A), the analysis revealed a significant effect of vergence distance ($F(2,8)=74.26$, $p<0.0001$); the estimated scaling distances increased significantly with the vergence distance. The analysis also provided an effect of the accommodation distance on the scaling distance ($F(2,8)=10.85$, $p<0.01$); the scaling distance increased with the screen-constrained accommodation distance. More interestingly, we obtained an interaction effect between the accommodation distance and vergence distance conditions ($F(4,16)=3.24$, $p<0.05$). Multiple pairwise comparisons were performed and provided results that can be summarized as follows: when considering a given accommodation distance, dihedral angle displayed in front of the screen plane is perceived more acute (i.e. the scaling distance is shorter than the accommodation distance) and angle displayed behind the screen plane is perceived more obtuse (i.e. the scaling distance is larger than the accommodation distance) as compared to the one displayed close to the screen plane. From Figure 6 – panel A, it can also be seen that the slope of the linear regression lines for the accommodation distances (D_A = 96, 130 or 201 cm have slope values of: 0.79, 0.47 and 0.81 cm^{-1}, respectively) are more shallow than for matched accommodation and vergence distances ($D_A = D_V$, slope value: 1.08 cm^{-1}), therein providing the accommodation distance effect in addition to the vergence distance effect. In the high luminance condition (Figure 6 – panel B), the analysis revealed a significant effect of vergence distance ($F(2,8)=33.73$, $p<0.0002$); the estimated scaling distances increased significantly with the vergence distance. However, the analysis did not provide an effect of the accommodation distance on the scaling distance ($F(2,8)=0.12$, $p>0.05$). An interaction effect was observed between the accommodation distance and the vergence distance conditions ($F(4,16)=6.37$, $p<0.01$). Pairwise comparisons did not provide this time any effect between the accommodation distances for the three possible vergence distances as for the low luminance condition but the vergence distance effect affected unequally the scaling distance in the three accommodation distance levels. This result can also be observed from the slope of the linear regression lines for the accommodation distances (D_A = 96, 130 or 201 cm have slope values of: 0.87, 0.38 and 0.51 cm^{-1}, respectively) that are closer to the one of the matched accommodation and vergence distance ($D_A = D_V$, slope value: 0.68 cm^{-1}).

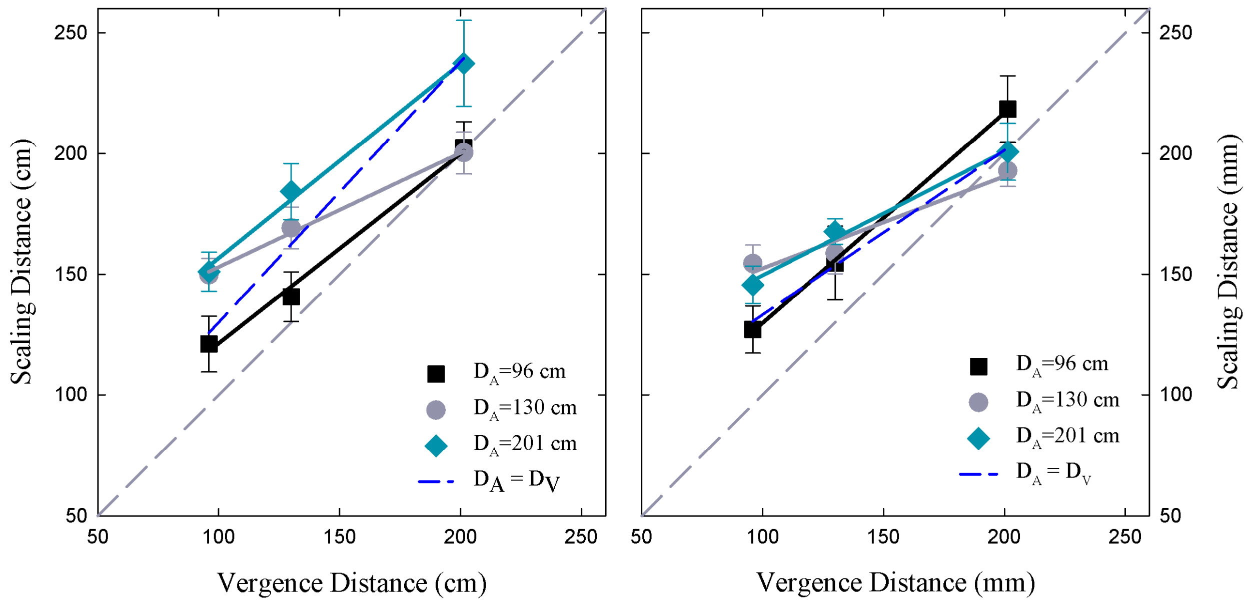

Figure 5.6: Scaling Distance as a function of vergence distance induced by disparity for the averaged results of the five observers who participated in the experiment. The left panel (A) represents the results for the low luminance condition (large pupil) and the right panel (B) represents the results for the high luminance condition (small pupil). The data for the closest accommodation distance (96 cm) are represented in black, the data for the intermediate accommodation distance (130 cm) are represented in grey and the data for the largest accommodation distance (201 cm) are represented in dark cyan. The dashed gray line represents perfect depth constancy. The diagonal dashed blue line ($D_A = D_V$) represents the averaged estimation for pseudo-matched accommodation and vergence distances.

Supplementary analyses performed on the slope of the PSEs (defined as the slope at the tangent of the PSE) for the five participants revealed no effect of vergence distances on the slope in the low luminance condition and in the high luminance condition but, a trend for both conditions ($F(2,8)=3{,}58$, $p=0.077$ and $F(2,8)=3{,}95$, $p=0.064$). No effect of the accommodation distance was observed in the low luminance condition ($F(2,8)=0.10$, $p>0.05$) but again, a trend was found in the high luminance condition ($F(2,8)=4.21$, $p=0.056$). No interaction effect was observed between the accommodation distance condition and the vergence distance condition neither in the low luminance condition nor in the high luminance condition ($F(4,16)=0.39$, $p>0.05$ and $F(4,16)=1.21$, $p>0.05$ respectively).

It is possible to describe the extent to which the visual system used the distance signal (i.e. accommodation and vergence) from the slopes of the fitted lines of each observer as for example the regression lines that can be found for the mean in Figure 6. A value of one refers to perfect depth constancy and value above one corresponds to over-constancy while value below one indicates under-constancy. The results on depth constancy for each observer in each condition are shown in Figure 7. The analyses realized on the slopes of the fitted lines for each participants did not reveal an effect of viewing condition ($F(3, 16)=0.25$, $p>0.05$) but showed an effect of lighting condition ($F(1, 16)=6.76$, $p<0.02$); depth constancy decreased in high luminance condition as compared to low luminance condition. The interaction effect between the viewing conditions and the lighting conditions was not significant ($F(3, 16)=2.20$, $p>0.05$).

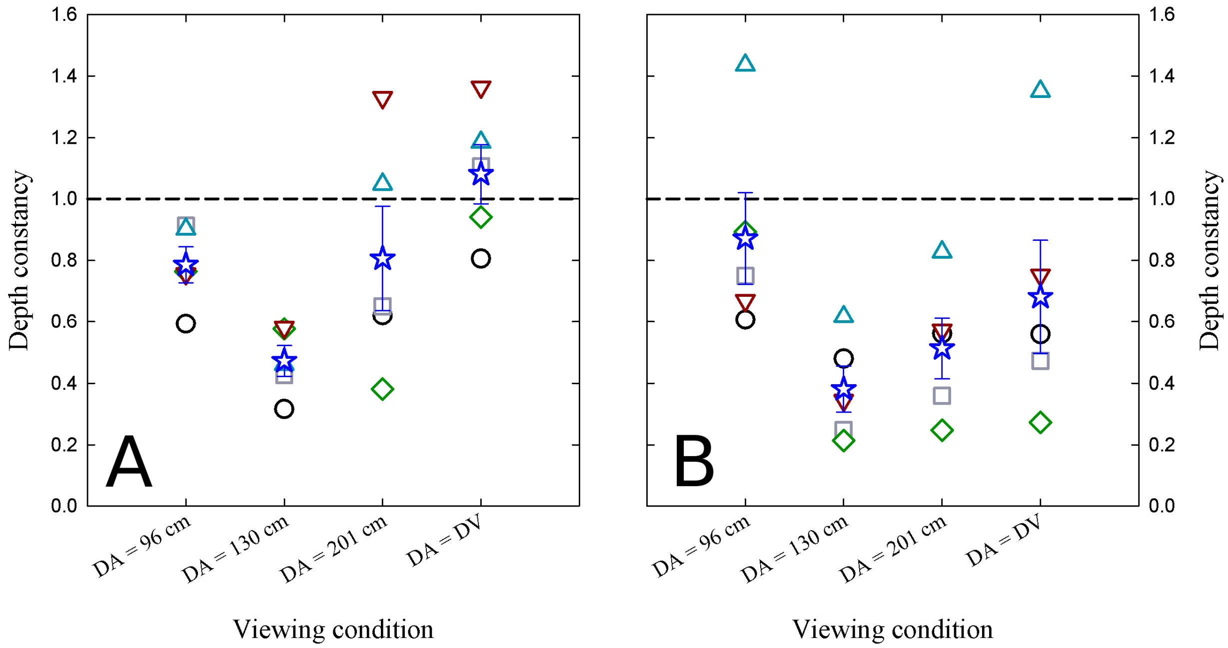

Figure 5.7: Stereoscopic depth constancy as a function of the possible viewing conditions (D_A = 96, 130 or 201 cm and $D_A = D_V$). The left panel (A) represents the results for the low luminance condition (large pupil) and the right panel (B) represents the results for the high luminance condition (small pupil). On abscissa are represented the values for the different viewing conditions. On ordinate are represented the values for the slopes of the lines fitting the relationship between the scaling distances to the vergence distances (induced by disparity) for each participant. The different colored symbols indicate different observers and the stars correspond to the averaged values. Vertical error bars denote the standard errors of the means.

5.3.2.1 Effect of pinhole viewing

A qualitative analysis is here proposed on the basis of the graphical analysis obtained for two out of the five participants. Hence, two participants participated a new time in the experiment but wore the aerobic glasses (binocular arrays of pinholes of 0.75 mm diameters set approximately 3 mm apart) during a 3D judgment task that was used in a past study (see Frisby, Buckley & Horsman, 1995).

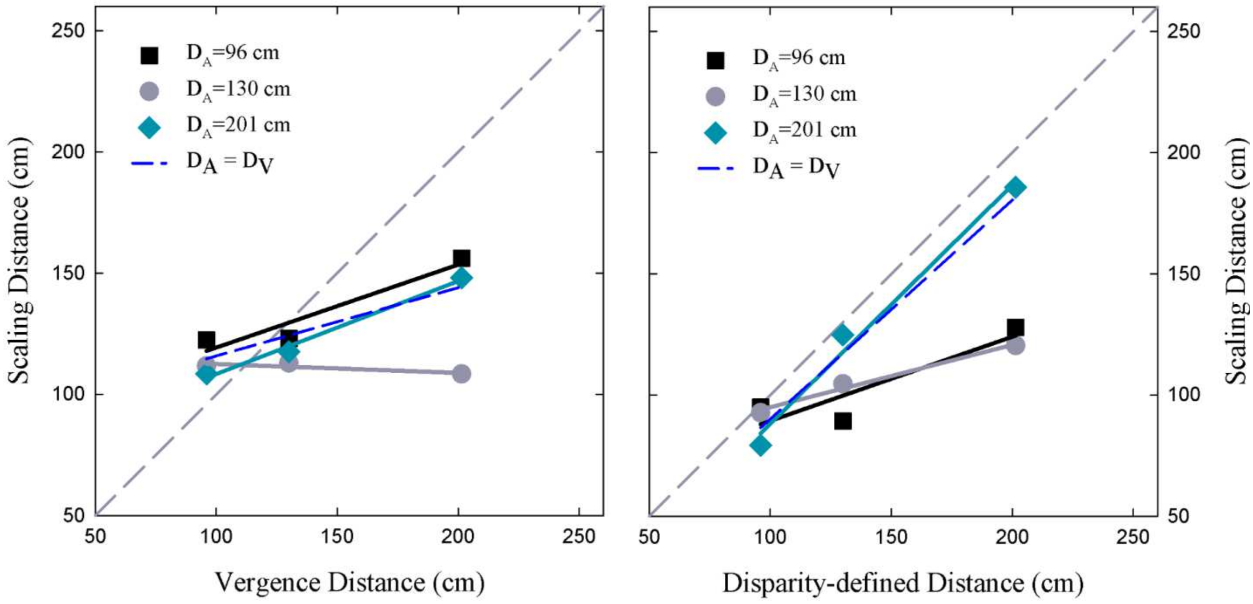

Figure 5.7: Scaling Distance as a function of vergence distance induced by disparity for two observers who participated in the experiment. The data for the closest accommodation distance (96 cm) are represented in black, the data for the intermediate accommodation distance (130 cm) are represented in grey and the data for the largest accommodation distance (201 cm) are represented in dark cyan. The dashed gray line represents perfect depth constancy. The diagonal dashed blue line ($D_A = D_V$) represents the averaged estimation for pseudo-matched accommodation and vergence distances.

What can be seen from this graphical analysis, in comparison to the averaged results presented in Figure 5.6 that the regression lines represented here show slopes that seem much smaller than that in the low luminance condition (except for the second participant for the estimations with the third accommodation distance, D_A = 201 cm). In addition, the results show that these two observers, by wearing aerobic glasses, estimated the dihedral angle with a scaling distance that was on average around 1.1 meters.

5.3.3 Discussion

The objective of this second experiment was to investigate whether a change in pupil size aperture could affect the perception of 3D shape in stereoscopic viewing conditions that varied both the accommodation distance and the vergence distance with crossed and uncrossed disparities. The major hypothesis of this experiment was that an increase of the depth-of-focus of the eyes could reduce the weight of the accommodation signal and thus affect the perception of 3D shape from binocular disparity. It was also expected that estimations would not be affected by the effect of the resting position of accommodation on vergence. The PSEs estimated from this experiment were converted in scaling distance (the distance at which the disparity pattern specify a right angle) both to study what is the possible distance used to scale disparity and how accommodation and vergence signals contribute to its estimate. Indeed, the linear regression lines obtained by fitting the PSEs for each accommodation distance can be used to interpret the contribution of vergence and accommodation in the final estimate of scaling distance. For example, on Figure 5.6, the closer the slope of the regression line is to one, the more the visual system relied on the vergence information. Conversely, the closer the slope of line was to zero, the more the accommodation signal was used to interpret the scaling distance.

When considering the low luminance condition (panel A of the Figure 5.6), the results showed that the scaling distance used for the estimation of the pseudo-matched accommodation and vergence distances ($D_A = D_V$) was on average overestimated but the slope for this condition was close to one, meaning that observers reached depth constancy in this condition. However, the results for the three other conditions (accommodation distance at 130, 110 and 96 cm) showed slopes being smaller than one, indicating the effect of accommodation that counteracted the vergence distances effect. Indeed, the effect of vergence distances was observed in such a way that the larger the vergence distance, the larger was the scaling distances. In addition, accommodation and vergence interacted such that the stimuli displayed with crossed disparity tended to be overestimated in depth (the hinge angle is perceived more acute because the scaling distance is underestimated) and those displayed with uncrossed disparity appeared underestimated (the hinge angle is perceived more obtuse because the scaling distance is overestimated) as compared with the matched accommodation-vergence signals. More specifically, the estimations made when accommodation and vergence distances are incongruent resembled a compromise between the two sources of information. Even though the effect is rather slight in this experiment due to the use of small stereoscopic demands, the present results replicated for larger viewing distances those obtained by Hoffman et al. (2008) and Watt et al. (2005) for short viewing distances.

Now, considering the high luminance condition (panel B of the Figure 6), the results showed that the slope of the linear regression line performed on the scaling distances used for the estimation of the matched accommodation and vergence distances ($D_A = D_V$) was more shallow than in the low luminance condition, indicating for observers a kind of loss in depth constancy. However, the results for the three other conditions (accommodation distance at 130, 110 and 96 cm) showed slopes being this time closer to the condition with matched accommodation and vergence distances (da = dv), indicating that the effect of accommodation was minimized. There was moreover no more significant accommodation effect whereas the effect of vergence distances remained. In this condition, the estimations made when accommodation and vergence distances were incongruent did not strongly differ from that when both signals were paired. The increase of the depth-of-focus seemed to reduce the weight of the accommodation signal in the estimate of the distance used to scale binocular disparity because the linear regression lines for each set of accommodation distances nearly lie on the same diagonal line (i.e. the pairwise comparisions revealed no accommodation effect for each vergence distance).

The results in Figure 5.7 and the accompanying statistical data indicated that, on average, depth constancy was more easily reached with a small depth-of-focus (low peripheral luminance) than with a larger one (high peripheral luminance). More specifically, these results showed that performances were poorer with large depth-of-focus, suggesting that accommodation does not contribute anymore to the estimate of scaling distance or at least that the contribution of one signal is weakened. The decrease in these performances can be explained by the fact that depth constancy is difficult to obtain on the basis of vergence information alone (Tresilian, Mon-Williams and Kelly, 1999) – reducing the available cues have often shown to decrease performances related to distance perception in a number of task, specifically in depth perception (Künnapas, 1968). These results overall support the conclusion that accommodation and focus cues play a key role in perceiving 3D shape with precision even for viewing distances up to two meters.

A quick comparison was also made between the results in the low luminance condition and the estimations made during the wearing of aerobic glasses. We observed that the slopes of the regression lines performed on the scaling distances for each accommodation distance were smaller with the aerobic glasses and for some accommodative distances were very close to zero. This observation highly supports the fact that these two observers, by wearing aerobic glasses, estimated the dihedral angle with a scaling distance that was on average around 1.1 meters. However, as it was suggested in the introduction of this report, the loss of luminance due to these glasses may have reduced the perceived contrast in such a way, the pattern of depth that was required to perceive a right angle had to be increased, even in the condition of matched accommodation and vergence distances. However, as these glasses did not appear to us a suitable way to test our hypothesis, it is worth noting that they indeed appeared to remove the blur information in the perceived visual scene in spite of other inconvenient (i.e. the adopted resting state of accommodation).

5.4. General Discussion

The two experiments in this report were designed in order to investigate the perception of 3D shape of random-dot stereograms for several accommodation distances and vergence distances displayed with crossed or uncrossed disparities. The objective was twofold: to replicate previous results on the impact of interacting vergence and accommodation signals on 3D shape perception for larger viewing distances and, to study the effect of increasing the depth-of-focus of the eyes on the ability to infer a scaling distance for binocular disparity. For this purpose, we presented stimuli whose perceived distances were varying by changing the vergence distance and/or the accommodation distance. The first experiment only varied the vergence distances but obtained interesting results by studying a large pool of participants. The second experiment provided a complete set of viewing conditions for three possible vergence/accommodation distances. In these two experiments, the peripheral luminance projected in the visual field was varied so as to change the pupil aperture size and, therefore the depth-of-focus of the eyes.

As accommodation and vergence are involved in estimating viewing distance, the motivation of experiment 1 for detailing the effect on 3D estimation for a set of vergence distances displayed within or beyond the depth-of-field comes from the question of whether the assigned weights to accommodation may change with vergence distance. Models of sensory integration of the sources of information for depth perception describe the cue-combination as a weighted linear model (Landy, Maloney, Johnston & Young, 1995). Perceived depth from stereopsis depends on the combination of at least two types of cue, namely the disparity and the accommodation sources of information. When incongruent with the vergence distance, accommodation leads to misestimation in the final 3D shape percept because spatial frequency specifies the display region. Perceived 3D shape from binocular disparity was thus consistently affected by the variations of the vergence distance induced by disparity in both of our experiments. Specifically, as the vergence distance became smaller the estimations of depth were overestimated as if the scaling distance was overestimated. In the first experiment, we observed particular effect on the perceived 3D shape, the observer sensitivity and the reaction times distribution as a function of image blurring and peripheral luminance conditions, thereby suggesting that the distance used to scale binocular disparities was strongly influenced by the accommodation system. In the second experiment, we found in the low peripheral luminance condition (large pupil and small depth-of-focus) that when vergence and accommodation did not specify the same distance, they led to estimations reflecting a possible compromise between the two signals. In the high peripheral luminance condition (small pupil and large depth-of-focus), this “compromise effect” was minimized and consistent with the vergence signal and, depth constancy was reduced. This latter observation suggests that, in this case, the weight attributed to the accommodation signal was weakened as compared to the first condition.

It has been supposed that when the accommodation system is under open-loop viewing condition, observers may use the accommodative resting state (e.g. the resting position in dark) in conjonction to other signal (e.g. the vergence signal and focus cues) to depth perception (Frisby, Buckley & Horsman, 1993; Owens & Leibowitz, 1976) and, somehow in the estimation of the egocentric viewing distance. Indeed, when no defocus information is available for the accommodation system, it may derive to an intermediate resting distance that is on average around 67 cm (Leibowitz & Owens, 1978), even though large inter-individual differences exist. As compared to studies that used pinhole viewing in 3D shape estimation task (Hogervorst & Eagle, 1998; Frisby, Buckley & Horsman, 1993), our study aimed at increasing the depth-of-focus of the eyes in order to observe the potential consequences on depth perception. Even though the use of pinholes and the use of peripheral luminance lead to a decrease in blur circle, two main differences can be found out in relation to the visual system.

The first is that the use of pinhole may increase the pupil size of the eyes due to a loss of perceived luminance. The second is that pinhole viewing suppresses all defocus information and, therefore may lead to render the accommodative system open-loop, affecting the neural-crosslinks between accommodation and vergence. On the contrary, the use of peripheral luminance seems more ecological in the sense that the interactions between accommodation and vergence are preserved (i.e. there is no effect of the resting position of accommodation on vergence); however, this could be the reason why the accommodation effects presented here are rather small.

5.5. Conclusion

The effect of peripheral luminance on pupil size and, thereby on the depth-of-focus, suggests an important role played in stereoscopic viewing conditions by focus cues and accommodation, strongly linked to the synkinesis between accommodation and vergence. The present experiment is thus an extension of previous studies that showed an effect of accommodation on perceived depth from stereopsis (Hoffman et al., 2008; Watt et al., 2005). The results of this experiment also emphasize the fact that the accommodation distance in stereoscopic viewing conditions is not bounded to the screen plane position and, therefore the effect of accommodation distance is strongly dependent on the depth-of-focus of the eyes. Egocentric distance perception may be based on vergence signal, specifically in absence of retinal information (Mon-Williams & Tresilian, 1999). In that, the effect of accommodation could be obtained on the basis of accommodative vergence (Mon-Williams & Tresilian, 2000). By increasing the depth-of-focus of the eyes, the accommodative vergence contribution to the final vergence response is weakened and, therefore the accommodation weight to distance perception is diminished.

♦

Accommodation and vergence signals are two potent sources of information in stereopsis when the viewing distance does not exceed a few meters. These two sources of information seem to affect the final perception of object 3D shape specified by binocular disparity in such a way that the distance used to scale these disparities is overestimated or underestimated. It has thus been shown that displaying object with disparity-induced vergence distance could lead observers to overestimate object depth displayed with crossed disparity (closer distances) and to underestimate object depth with uncrossed disparity (larger distances). Accommodation could also influence depth perception if the estimation of scaling distance results from a compromise between the vergence distance and the accommodation (screen-constrained) distance. The contribution of this study about the depth-of-focus was to show that the role of focus cues in the final estimate of 3D shape is reduced by increasing the depth-of-focus of the eyes. With an increased depth-of-focus, depth constancy also appeared reduced because the available visual information was also reduced. As there are no blur cues contained in stereograms, the influence of accommodation on the perception of 3D shape could thus be the results of the information from the state of oculomotor muscles of the accommodation system.

The results of the present study have strong practical implications on possible way to improve stereoscopic experience. The first is that the wearing of pinholes is not adapted in order to correct the misestimations obtained in 3D shape estimation task. Indeed, even though pinholes removed the blurred circle that is used by the visual system and affects both the accommodation response and the vergence response, they may also impact the tonic state of accommodation and decrease the light entering in the eyes. For this reason, the accommodation system will continue to impact the vergence response so as to yield persisting biased perception. The second implication concerns the highlight this study provided concerning the dependence of the oculomotor system on the depth-of-focus of the eye. Although it is difficult to create an efficient system that could increase the depth-of-focus without corollary effects, it is worth noting that the lighting conditions may play a potent role when watching stereoscopic displays (e.g. the loss of light due to the glasses). This assertion is true for short viewing distance because the depth-of-focus increases proportionally with the accommodation distance.

What can be derived from these observations is that depth perception from stereopsis shows strong perceptual distortions that are mediated when the visual scene contains different induced vergence distances other than the screen plane region. This effect can be explained by the tendency for the visual system to use a scaling factor (i.e. depth-to-width ratio) when interpreting the stereoscopic visual scene. This scaling factor seems to be linked to the state of accommodation because it can be observed in stereoscopic displays that do not mediate defocus information. As a consequence, the following chapter proposes the study of a non-comprehensive body of arguments that are either in favor or not of extra-retinal contribution to distance and depth perception. The idea is that this scaling factor could be due to the shape of the eye-lens that transforms the retinal projection or to the sensorimotor feedback from the ciliary muscles.

Chapter 6. Perceptual bias in stereopsis & Optical models of the eye

6.1 Introduction

Binocular disparity provides an ambiguous cue to depth perception because the interpretation of depth from stereopsis is strongly dependent on viewing distance: any pair of points on the retina gives rise to an infinite possibility of depth relations (Kaufman, 1974). A pioneer study has shown that perceived depth from stereopsis is consistently overestimated at near distance and underestimated at far distance (Johnston, 1991). This contraction bias has been interpreted as the consequence of using impoverished visual stimulation, but also as the use of an incorrect estimate of scaling distance (Johnston, Cumming & Landy, 1994; Johnston, 1991). A consistent body of studies provided experimental evidences that when the information from the state of ocular muscles (i.e. vergence and accommodation) is the lone cue to estimate a spatial extent, this contraction bias is amplified (e.g. Foley, 1980). Moreover, when objects are presented in isolation, observers tended to estimate objects at an intermediate distance, this has been labeled the "specific distance tendency" (Gogel & Tietz, 1973; Gogel, 1969). This phenomenon has thus been suggested as a basis for the poor depth constancy in 3D shape perception from stereopsis.

By comparing estimations of 3D shape from stereopsis in conditions where accommodation and vergence distances were matched or mismatched, our study and two previous others (Hoffman et al., 2008; Watt et al., 2005) concluded that focus cues may play a potent role to establish the egocentric distance presumably used to scale binocular disparities. Indeed, by converting perceptual thresholds into estimated scaling distances, the inferred distance was shown to be overestimated when the vergence distance was smaller than the accommodation distance and underestimated when the vergence distance was larger than the accommodation distance. In other words and in the case of stereoscopic displays, objects displayed in front of the screen plane are overestimated in depth whereas objects displayed behind the screen plane are underestimated in depth. The contraction bias observed by varying together the accommodation and vergence distance was thus magnified by increasing the stereoscopic demand and, therefore the discrepancy between the accommodation and vergence distances.

Related to the present effect of accommodation is the misjudgment often observed of apparent size and distance of object subtending a fixed angle as measured from the eye (Holway & Boring, 1941). The example, in stereopsis can be observed by simply increasing or decreasing the shift between the images/objects displayed for the left eye and for the right eye, inducing a change in vergence angle. The image or object is perceived smaller with increasing convergence and larger with decreasing convergence angle. As a consequence, larger image may indicate the object was nearer than its physical distance, while a smaller image may suggest the object was further away (Charman & Koh, 1995). Similar to the effect of accommodation inducing the use of an incorrect scaling distance for stereopsis, the change in apparent size with the change in the shift between the two screen images could also involve accommodation. In order to account for the contribution of accommodation in misestimating size and distance, two main hypotheses on the role of accommodation in providing significant visual information that are not related to the blur circle hypothesis have been put forward. In the next sub-section, these two hypotheses are described and confronted.

6.2 Origin of the scaling factor

Potential sources of distance information can be found in the retinal image however, the most are ordinal or relative and, therefore cannot indicate where an object is located in the environment. The uses of extra-retinal information from vergence and accommodation can in principle provide the basis for an egocentric distance perception, and thus allow for judgments of depth relations. Accommodation can provide ordinal information for close objects to the observer thanks to the presence of blur defocus gradient that helps interpreting stereopsis (e.g. Hoffman & Banks, 2010). But, when these blur cues are removed from the visual scene, accommodation can still play a role (Frisby, Buckley & Horsman, 1995) potentially thanks to its extra-retinal contribution. Two main hypotheses thus account for the contribution of accommodation in perceived size and distance: the first is labeled the “zoom lens” hypothesis (Roscoe, 1989) and the second is the “accommodative effort” hypothesis (Lockhead and Wolbarsht, 1989).

6.2.1 Optical effect of accommodation and the “zoom lens” hypothesis

The first hypothesis remains a retinal contribution and concerns a category of effects that are directly related to the change in the shape of the eye-lens modifying the light beam cast onto the manner light is cast onto the retina. Indeed, the basic idea is that accommodation change may produce a significant shift in the nodal points of the eye that affects the retinal image size. To this effect, could be added the fact that the optical axis does not coincide with the visual axis because the fovea is slightly toward the temporal side as compared to the optical axis symmetry axis of the eye, passing through the center of the pupil entrance.

The optical effects that have been considered as a potential source of apparent size and distance misestimation with varying viewing distances can be regrouped under the label “optical effect of accommodation”. Mainly three aspects have been discussed. The first is that focusing on a near object can provide slightly larger retinal image size than when focusing further away because the nodal points move forward with accommodation (see Figure 6.1, panel A). Optical simulations have been performed to unravel the possible involvement of the variations in eye-lens shape (e.g. Smith, Meehan & Day, 1992). Even though the eye model predicting accommodation dependent changes on retinal image may be subject to limitations due to huge inter-individual variability, it is worth noting that such optical simulations did not succeed in explaining the large effects observed on perceived size and distance. This type of optical simulation can be performed using an eye-model (e.g. the Gullstrand’s eye) and a retinal image formation model (e.g. using a six cardinal points eye model); the optical data used there, are represented in Table 6.1 (see *Appendices*) and the six cardinal points model is represented in Figure 6.1.

Another matter concerning the optical effects is the possibility that inappropriate levels of accommodation may lead to changes in the retinal image size (Smith, Meehan & Day, 1992). This possibility is directly related to the effect we observed in stereoscopic conditions with different virtual distances (i.e. distances induced by disparity). Indeed, while vergence is induced by binocular disparity mediated by two independent images for the left and right eyes, the accommodation system has to keep a focus around the screen plane region and depending on the depth-of-field. This viewing condition entailing an accommodation error relative to the vergence response may impact the retinal image size.

Hence, the idea here was to explain the effect obtained in stereoscopic viewing conditions for matched accommodation and vergence distances (DA=DV) and mismatched accommodation and vergence distances for a stereoscopic demand of about -0.27 D (i.e. DA = 96 and DV = 201 cm) as it was observed in the previous experiment. In the match condition (DA = DV = 96 cm), the scaling distance estimated was about 121 cm whereas in the mismatch condition (DA = 201, DV = 96 cm) the scaling distance was on average equal to 151 cm. The scaling distance was thus overestimated by 24.7 % and the question is then to know whether an optical effect of accommodation could explain this large shift.

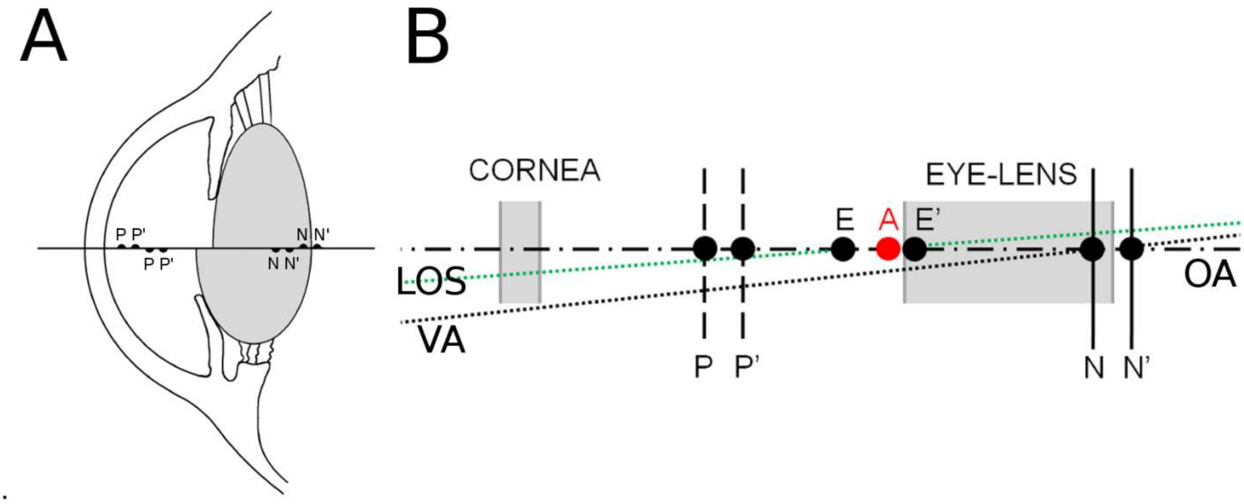

Figure 6.1: Schematic representation of an optical image formation model using six cardinal points. Panel A represents a schematic eye with its upper part corresponding to a relaxed accommodative state (distance vision) while its lower part represents an accommodated state (near vision). P and P' are the position of the principal points for object side and image side respectively. N and N' are the position of the nodal points for object side and image side respectively. Panel B is a schematic representation of the axes and lines passing through the schematic eye. The visual axis (VA) and the line of sight (LOS) pass are cast onto the foveal center. The visual axis (dotted black line) passes through the fovea and nodal points (respectively: N and N'). The line of sight (dotted green line) passes through the fovea and the entrance and exit pupil centers (respectively: E and E'). *A* represents the physical aperture stop center of the optical system, imaged on E and E' by ocular elements on each side.

As for the effect coming from the accommodation change, again the effects derived from optical simulation showed very small effects (0.3 % of change in retinal image size for the -0.27D stereoscopic demand that supposes no accommodative error) that cannot account for the measured effect of misestimations of depth. This result replicated the optical simulation results of others that showed that the optical effect of accommodation is insufficient to explain misjudgments of size and distance (Smith, Meehan & Day, 1992).

A third issue in the optical effects potentially impacting size and distance perception concerns the wearing of corrective spectacles in observer with ametropic eyes. Indeed, the corrective glasses presenting convergent power (correcting hyperopia) or divergent power (correcting myopia) creates a prismatic effect depending on this power. Hence, the wearing of correction glasses with negative power lenses would induce the minification of the retinal image size while correction with positive lenses would provoke magnification of the retinal image size. In natural vision, observers generally adapt to this change. However, wearing correction glasses while watching stereoscopic visual scene may affect the stereoscopic demand in such a way that intended displayed depth will not be viewed in the same way if the correction spectacles induce a change in the vergence angle. Moreover, a supplementary prismatic effect happens when the eyes converge as the chief is deviated for each eye depending on the glasses refractive index and surface curvature.

6.2.2 Accommodation contribution from sensorimotor feedback

The second hypothesis to explain how accommodation state may affect size and distance perception, concerns the influence of the extra-retinal signal, namely the sensorimotor feedback from the activity of the muscle organization of the accommodation system. It has been proposed that the perceived changes in size or distance could be related to the accommodative effort, i.e. the neural signals innervating the ciliary body (Lockhead & Wolbarsht, 1989). This assertion is supported by some experimental studies on the use of pinhole because it generates accommodation open-loop (Hogervorst & Eagle, 2000; Frisby, Buckley & Horsman, 1995; Charman & Koh, 1995, Holway & Boring, 1941). For example, Charman & Koh (1995) have shown that the wearing of pinholes reduced the perceptual errors in size matching. It is interesting to note that pinhole viewing also showed to increase the range of errors in accommodation response and also reduces the propensity to accommodate as object distance varies (Ward & Charman, 1987). Indeed, the use of pinhole viewing may provoke the accommodation system to adopt a level of accommodation remaining close to the tonic/resting level as it can be observed in the dark (Ward & Charman, 1987; Hennessy et al., 1976). Other experimental manipulation, such as the field-of-view restriction has also shown to affect accommodation (Gu & Legge, 1987; Hennessy & Leibowitz, 1971) and to reduce perceptual errors in size perception with varying distance (Charman & Koh, 1995; Holway & Boring, 1941). Even though it was sometimes argued that accommodative effort is not reliable enough to be a distance cue (Kunnapas, 1968; Heinemann et al., 1959), other investigators supported the contribution of accommodation together with the activity of convergence (Mon-Williams & Tresilian, 2000; Harvey and Leibowitz, 1967; Leibowitz and Moore, 1966).

It has thus been shown that accommodation contribution to distance estimate may depend on vergence (i.e. accommodative vergence) rather than accommodation (Mon-Williams & Tresilian, 2000). Indeed, considering the neural interactions of accommodation and vergence, it is possible that the visual system uses the vergence signal stemming from the accommodative response as a cue to distance. As a matter of fact, vergence was shown to be effective as distance cue, at least for short viewing distance (Tresilian & Mon-Williams, 1999; Foley, 1980). Moreover, it was also shown that the vergence signal could be used to scale binocular disparity (Bradshaw et al., 1996). In contrast, Mon-Williams & Tresilian (2000) have shown that observers had strong difficulty to perform a pointing task when accommodation was the only distance cue available. In addition, they showed accommodation signal was not reliable because observers failed to make the task precisely when an additional ambiguous cue was added in the visual scene. They thus concluded that accommodation could be used as a distance cue in full-cue environment and through the contribution of the vergence signal as a distance cue.

6.3 Conclusion

From these theoretical considerations, it appears that the accommodation state can participate to size and distance perception. However, it is unlikely that the indirect effect of accommodation on depth perception originates from the lens shape-dependent changes. Rather, this phenomenon is more possibly expected to arise from the contribution of extra-retinal information even if, some effects could stem from the forces exerted on the retina by the ciliary body for accommodating. This effect could move the retina forward to the lens and, therefore reduce the perceived image size because less receptors would be stimulated (Miles, 1975). However, this scaling factor appears more probably due to psychological factors rather than optical factors.

The interactions of vergence and accommodation shows that distance perception can be indirectly influenced by accommodation in such a way that the activity of accommodation can dramatically impact distance and depth perception despite the fact that accommodation does not directly contribute to the distance percept (Mon-Williams & Tresilian, 2000). The design of stereoscopic media that is expected to provide perceptual veridical information must consider these interactions and indirect effects in order to avoid perceptual distortions.

♦

In the previous sub-part, some arguments have been put forward that apparently suggest that the effects of accommodation in stereoscopic viewing conditions are related to the extra-retinal information from muscles activity of the accommodation system. The contribution of accommodation as a cue to distance perception (i.e. for scaling disparity) could be indirect, in the sense that most of the effect would be mediated by the accommodative vergence neural cross-link of the oculomotor system. This assertion is very important because it subtends that the mismatch is rather in the motor than in the sensorial component.

Considering these perceptual distortions arising in stereoscopic viewing conditions, the following part of this thesis was thus devoted to the study of the possible personalization of stereoscopic media based on estimated perceptual thresholds. The major focus of the next chapter is to investigate perceptual distortions of moving-in-depth stereoscopic objects and their relations to perceived naturalness.

Chapter 7. Distortion from stereopsis in motion-in-depth with linear perspective

7.1 Introduction

Perception of 3D shape from stereopsis is known to be inaccurate (Watt et al., 2005; Johnston, 1991). These misperceptions can therefore alter the realism of a rendered visual content (Vienne et al., 2013). Generally, when observers are asked to match the depth of an object to its height, they report systematic depth overestimation errors at near distances and underestimation at far distances while estimation is nearly accurate for a 80 cm observation distance (Johnston, 1991). It has also been shown that these systematic biases could arise with vergence distances being induced by binocular disparities (Watt et al., 2005). This phenomenon strongly impacts the quality of experience in stereopsis in such a way that observers perceive stereoscopic objects as distorted and, moreover, that perceived shapes of solid objects moving in depth may change considerably while the physical displayed shape remains the same (Scarfe & Hillis, 2006).

When motion is applied to visual objects (e.g. rotation), systematic distortions tend to be reduced (Brenner & Landy, 1999; Johnston, Cumming & Landy, 1994), presumably because apparent motion helps disambiguate disparity information (Johnston, Cumming & Landy, 1994; Richards, 1985). In Brenner & Landy (1999), participants had to adjust the size and shape of randomly textured ellipsoids so as to match a tennis ball and could adjust the simulated distance of the reference stimulus. They provided estimations that were largely biased – far from veridicality – but, rotation of the reference stimulus helped the matched shape to be perceived more circular. Indeed, Richards (1985) suggested that motion and stereopsis combined together should allow observers to recover correct 3D structure of the visual object. Specifically, stereo cues are insufficient to provide an accurate estimate of depth as the retinal projections can represent various angular disparities depending on fixation distance estimates. In the same way, structure-from-motion information cannot provide unambiguous perception of shape since each pattern of motion contains at least two possible interpretations, one being the reflected image of the other. However, when combined, each of these two signals may eliminate the ambiguity of the other: disparity provides the sign to suppress the reflection ambiguity and motion yields the relation between points to weaken the scaling distance dependency. Experimental evidences seemed to support these considerations (i.e. Johnston, Cumming & Landy, 1994). In a set of four experiments, participants had to match the depth to the height of hemi-cylinders presented with stereo cues only, motion cues only or with both signals together. The main result of this study is that stereo and motion cues could interact together so as to provide greater depth constancy, but the combination rule was not balanced, the weight attributed to motion was overall more important than the one of disparity (Johnston, Cumming & Landy, 1994).

A majority of reports do not seem in agreement with the opinion that stereo-motion combination offers veridicality. Hence, a large group of studies indicates that improving in shape from stereo with motion cues confers only slight improvement and that veridical perception is not necessarily achieved (Tittle & Braunstein, 1993; Bradshaw, Parton & Glennester, 2000; Champion, Brenner, Mamassian & Simmons, 2004), indicating that stereo with the addition of motion cues may not be sufficient. For instance, Bradshaw et al. (2000) observed no effect of adding motion parallax to disparity signal in a depth nulling task, a depth matching task and an absolute 3D shape judgment task; performances were overall worse in the latter condition that required observers to recover the fixation distance used to estimate absolute depth. Tittle et al. (1995) also provided evidence that the combination of disparity and motion cues did not noticeably change 3D shape perception or, at least did not allow accurate depth judgment of various stimuli such as the hemi-cylinder, the fronto-parallel dihedral angle and the coplanar random dot stereograms. Among a set of four experiments in their study, Norman et al. (1996) performed an experiment where they presented stimuli with disparity cues, motion cues (rotation) or both cues combined and stated that performances on three observers were similarly equal for all three conditions with no systematic effect indicating facilitation in performance due to one or the other cue type. Furthermore, it is interesting to note that two other studies also observed that disparity tended to be predominant over motion in judging 3D shape when disparity and motion specify different depth structures (Rogers & Collett, 1989; Tittle & Braunstein, 1993). Lastly, Champion et al. (2004) could not find any evidence of transfer of viewing distance information from motion to binocular cues across two objects presented sequentially. Overall, these findings suggest that disparity and motion remain relatively independent for the perception of 3D shape (Domini, Caudek & Tassinari, 2006).

Disparity and motion combination in judging 3D shapes thus remains unclear. As a matter of fact, motions such as rotation and motion parallax can help the recovery of 3D structure only under some circumstances (Brenner & Landy, 1999). Another instance was also investigated concerning the case of motion in depth that was studied in relation to participants' ability in depth constancy (Scarfe & Hibbard, 2006). Scarfe & Hibbard (2006) studied depth constancy for random-dot hemi-cylinders that were either static or moving in depth and defined using uncrossed disparities (i.e. located behind the screen plane). Their main objective was to compare performances in stereo-matching from the change in information in a single moving object with those in matching depth for objects located at different disparity-induced distances. They first asked participants to perform matching of elliptical cylinders and then used disparity patterns corresponding to circular cylinders for participants to evaluate the part of the stimuli used in the motion task. The moving stimuli required eye version movements because the motion path was shifted relative to observers' midline (28 degrees). Their results indicated that motion-in-depth was not sufficient to reach depth constancy because approaching objects appeared to expand in depth and receding objects tended to contract with motion as expected from the results for static stimuli by Johnston (1991). In all cases, this phenomenon created a potent perspective distortion that may be perceived even for slightly different distances specified relative to the screen plane (Vienne et al., 2013).

Here the major focus is put on the ability for a given observer to reach depth constancy for static virtual distances and for stimuli moving toward the observer. Crossed disparity-induced motion-in-depth stimuli may constitute very potent factors of immersion in imaging displays and, it is therefore worth investigating to what extent such stimuli may provide a vivid experience of depth perception. Our first objective was thus to study the benefit of correcting the disparities of the simulated scene according to individual thresholds of perceived depth-to-width ratio for transparent rotating cylinders. We asked participants to estimate whether cylinders were perceived as expanding or contracting over a range of four virtual disparity-induced distances. In a second condition, we asked the same participants to judge the expansion/contraction of the same stimuli approaching toward them from the screen plane position to one pre-determined fixation distance in order to study whether observers can reach the same perception of circular cylinders in motion as for static stimuli. In agreement with Scarfe & Hibbard (2006), results show that the changing information available in a moving object did not suffice to reach depth constancy. As we observed, stimuli needed to be modulated to provide a better perception of object shape, we investigated the perceived naturalness of approaching cubes in two stimulus conditions, linear perspective and disparity being the provided cues. Given that stimuli combining conflicting disparity-defined depth and perspective can be expected to generate rather larger distortions, we designed a last experiment to evaluate the propensity for observers to be disturbed by such depth distortions.

7.2. Experiment 1

7.2.1 Method

7.2.1.1 Stimuli

Stimuli were random dot stereograms depicting transparent elliptical cylinder, so that dots in the front elliptical part and in the back elliptical part were both visible (see Figure 7.1). Stimuli were rendered using OpenGL so that they could be carefully adjusted as a function of simulated distance and cylinder depth. For creating stereograms, the measured inter-ocular distance of each participant was used. Stereograms were composed of 200 small white dots that were previously defined randomly on a 2D gray texture itself used to texture-map the cylinder. The cylinder axis was oriented in the vertical direction. In order to disambiguate the disparity signal of the transparent cylinder, a rotation about the vertical axis was applied, with a mean velocity of about 45 deg.s-1 and a random rotation direction for each trial. The cylinders could be located at four possible disparity-induced distances chosen such that all participants could fuse the stimuli. These distances corresponded to fixation points being nearer than the screen plane to one of four angular disparities: 0, 0.5, 1 and 1.5 degrees, that is, 130, 110, 96 and 85 cm away from the participant’s eyes, considering a standard pinhole projection model. The radius and height of the cylinder located in the screen plane were set to 5 cm and 10 cm; these lengths were adjusted with the induced vergence distance so as to keep the same size-distance ratio. Cylinder sizes were displayed according to an adaptive double staircase procedure in order to estimate the point of subjective equality (PSE) and slope at this point for each condition.

Figure 7.1: The rotating transparent cylinders used in experiment 1 arranged for cross fusion. It is worth noting that this stimulus was easier to fuse when the cylinder was in rotation.

7.2.1.2 Apparatus

Stimuli were viewed on a 46 in. monitor (Hyundai S465D 46" HDTV LCD Polarized monitor) through polarized glasses. The monitor was located 1.3 meters away from the participants. The display resolution was 1920 x 1080 pixels. We minimized head movement by placing participants' head in a chin and headrest. The center of the display screen was carefully aligned along the subject's midline; this setting together with the gray background and white dots allowed minimizing spatial crosstalk.

7.2.1.3 Participants

A total of five observers participated in the study. They were on average 29.6 years old (SD = 11.1). All had normal or corrected vision and presented stereoacuity thresholds less than a chosen criterion of 30 arc minutes as assessed by the Randot Stereo Test. They gave their informed consent before beginning the experiment.

7.2.1.4 General procedure

Participants were required to complete first the static cylinder task before running the motion task given that thresholds in the first task were used to compute the stimuli that moved along the line of sight in the second task. The estimation in the static cylinder task was performed by randomly counterbalancing the distance levels. The stimuli were always presented on the subject's midline. An adaptive double staircase procedure was used both for the static and moving-in-depth stimuli.

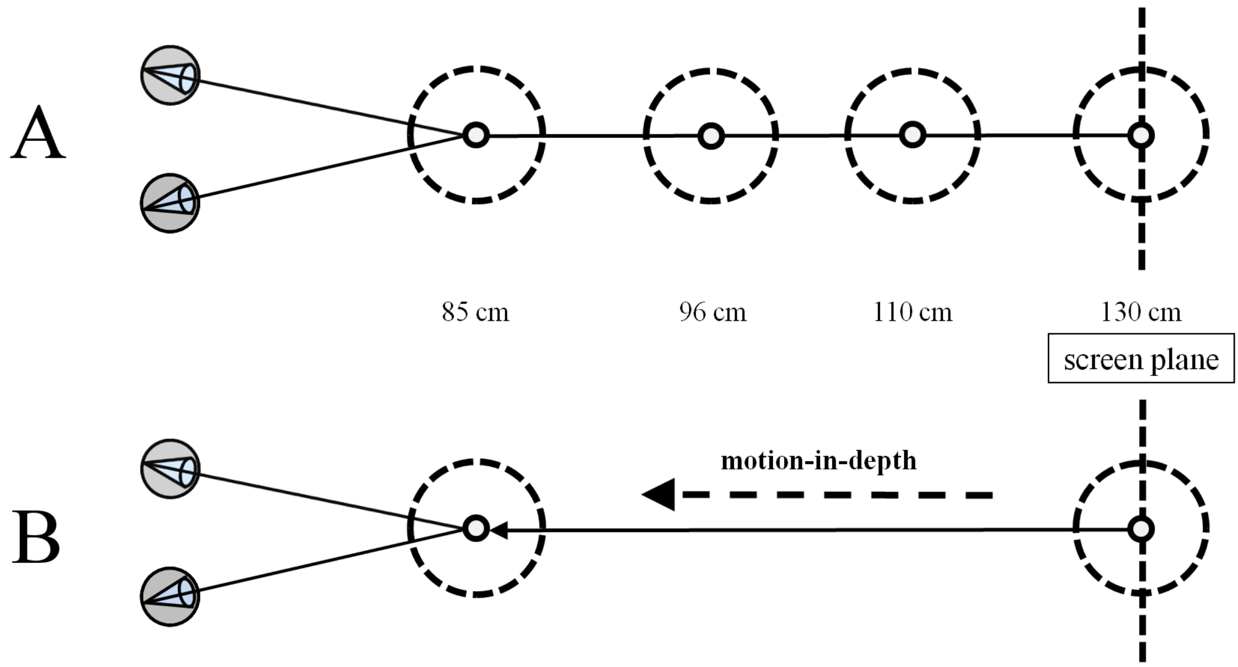

Figure 7.2: Top view of the experimental layout used in the static task (A) and in the motion task (B). The upper panel (A) represents the virtual viewing distances used in the static task (130, 110, 96 and 85 cm). These vergence distances were induced by binocular disparities. The lower panel represents the motion-in-depth along the line of sight that starts from the screen plane toward the fourth virtual distance (85 cm). The cylinders were always rotating on the vertical axis. Drawing is not to scale.

7.2.1.5 The static cylinder task

Observers had to judge whether the depth diameter of the rotating cylinder was smaller or larger than its horizontal diameter. In other words, they had to judge whether the cylinder was contracted or expanded in depth. The cylinders were displayed within an adaptive double staircase procedure consisting of two ascending and two descending staircases each containing fifteen steps (a total of 60 trials were used to compute the PSE for each distance). The initial value of the ascending staircase was set to a scaling of 0.7 whereas that of the descending staircase was 1.3, meaning that the displayed cylinder had a depth diameter that varied from 70% to 130% of the horizontal diameter respective of the staircase direction. Participants had to look at a fixation cross and then press a button to initiate the sequence. The presentation time of the cylinder was 3 seconds but participants had to respond as quickly as possible.

7.2.1.6 The motion task

Observers had to judge whether the depth diameter of the approaching cylinder was contracting or expanding during its motion-in-depth (motion along the line of sight). The motion always started at screen plane distance (130 cm) and ran at constant velocity of about $11.3 cm.s^{-1}$. Motion duration was 4 seconds. The stimulus also rotated about the vertical axis. At initial position, the scaling applied on the cylinder depth was the one estimated in the first task for the screen plane position. The cylinder size change during the motion along the line of sight was displayed according to an adaptive double staircase procedure consisting of two ascending and two descending staircases each containing fifteen steps. The expansion/contraction change of cylinder depth (diameter in depth) during the motion-in-depth along the line of sight within a trial was realized by performing a regression of the perceptual thresholds onto distance for each observer's results estimated in the previous task. Rather than a linear regression, we used a ratio of polynomials as curve fitting as the fit was better specifically for increasing distances. The ratio of polynomials used has the following form: $S(x) = (p1 \cdot x + p2)/(x + q1)$ where x is the distance between the center of the object and the screen plane (in mm) and S(x) is the corresponding depth-to-width ratio. The change in coefficient *p1* will change the shape of the curve while keeping the same starting point; an increase in coefficient *p1* produces an expansion of object depth while the decrease produces a contraction in the motion along the line of sight. Before using the regression value estimated from individual results during the first task, we checked that this fitting was correct and provided sufficient quality of statistical adjustment (using adjusted R-square). Moreover, it is important to note that this adjustment remained consistent when the distribution was rather linear. Hence, initial values of the ascending and descending staircases were based on this regression and this procedure was quite similar to varying the object depth at final positions between 0.3 and -0.3 as scaling factor.

7.2.2 Results

7.2.2.1 The static cylinder task

The individual PSEs across each disparity-induced vergence distances are presented in Figure 7.3. Exemplars of psychometric curves of each distance are also represented for one observer. A Friedman test realized on the PSEs across each distance revealed a significant effect of distance on observers' judgment of circularity ($\chi 2=14.44$, df=3, $p<0.01$) even though the magnitude of the effect was weak, as expected, given the slight dioptrical demand on the observers. This aspect of the design was to ascertain that all observers could fuse the stimuli. It should be noted that the effect of distance is not homogeneously distributed across observers, indicating clear individual differences. Indeed, two observers showed a marked effect (o3 and o5) whereas the two others provided smaller effects, especially one participant (o4).

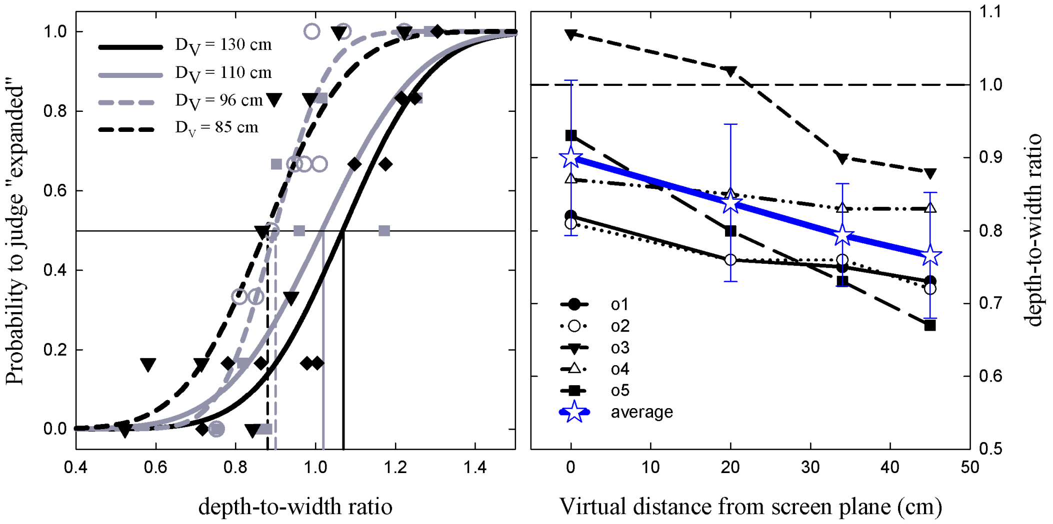

Figure 7.3: Individual and group results obtained in the static cylinder task. The left panel represents the probability to judge the transparent rotating cylinder as being expanded in depth as a function of the displayed depth-to-width ratio for one observer (o3); the curves are fitted with cumulative Gaussians. From this figure, the effect of the disparity-induced distances on the ability to perceive the rotating cylinder as expanded in depth can be observed. The graphical estimation of the PSEs can be obtained thanks to the four drop lines (thin vertical lines). The right panel represents the PSEs for all participants as a function of the virtual distance between the cylinder and the screen plane. The averaged PSE is represented in blue (error bars denote standard deviation); note that the required depth to perceive a circular cylinder is on average 13.4% smaller for the largest distance to screen

position (viewing distance of 85 cm) as compared to the screen plane position (viewing distance of 130 cm). The horizontal dashed line represents the depth-to-width ratio of non modulated cylinders.

The analysis performed on the slopes at the PSEs for the five participants revealed a significant effect of the disparity-induced distance ($\chi2=8.14$, df=3, $p<0.05$); the slope significantly decreased with increasing virtual distance from the screen plane.

7.2.2.2 The motion task

Psychometric functions (cumulative Gaussians) were fitted to the proportion of "expanded" responses for the approaching cylinder, and PSEs were extracted for each observer. Figure 7.4 provides a comprehensive summary of the results in the motion task. From these results, it can be seen that the presented cylinder depth must be contracted when approaching so as to provide the perception of a circular cylinder in motion-in-depth, therein replicated some of the results by Scarfe & Hibbard (2006).

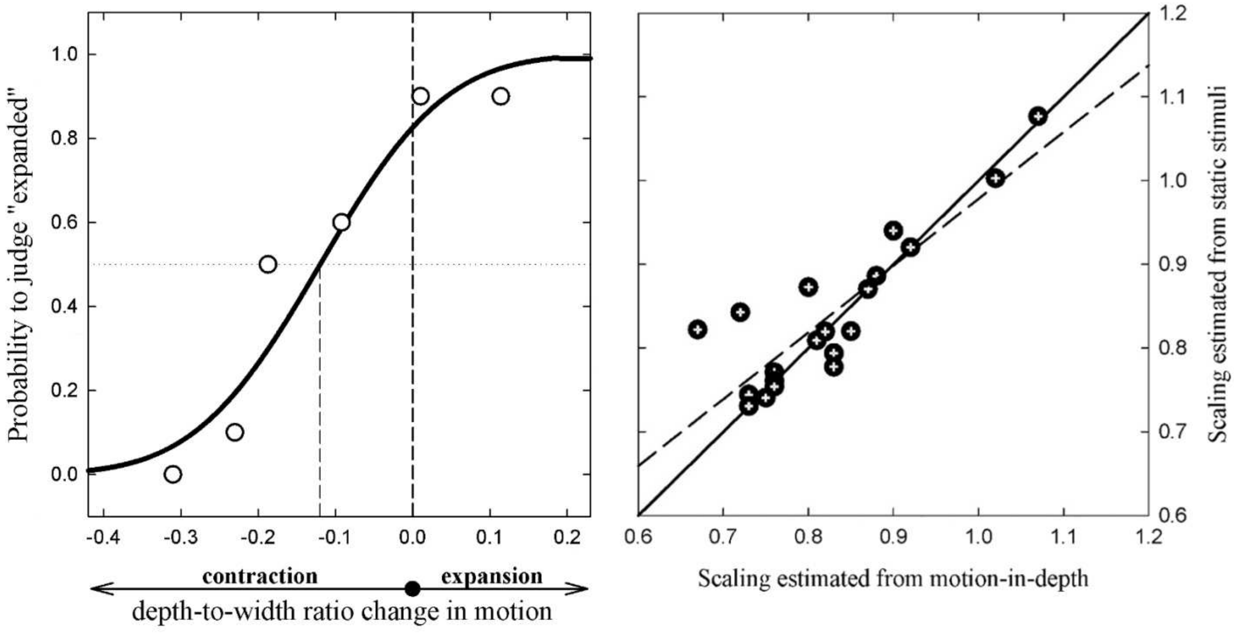

Figure7.4: This figure provides a representative psychometric function of the results of one observer (o1) (left panel) and the relationship between the depth-to-width ratio estimated in the first task and those retrieved from the regression obtained in the motion task (right panel). In the left panel, the data are fitted with a cumulative Gaussian: the grey drop dashed line allows retrieving the PSE, the black dashed line represents the object shape that remains physically identical during motion. On abscissa, the change in depth-to-width ratio indicates how the shape of the object is perceived contracted or expanded during motion-in-depth. In

the right panel, dots are pairs of scaling estimated from the static task and those retrieved on the basis of the regression from the motion task. Each pair of dots corresponds to one specific disparity-induced distance (M=4) for each observer (N=5). The solid black line represents the "perfect" correspondence between the estimated PSE in the static task and those estimated in the motion task. The dashed line corresponds to the linear regression line performed on the data.

We also extracted the scaling factors from the estimation made in the motion task (on the basis of the modulated object shape changes in motion) and plotted them as a function of those estimated in the first task where cylinders were displayed static at fixated distances. The correlation between the estimated scaling factors in the static and the moving task was 0.87 (Spearman coefficient, $p<0.0001$). The slope of the linear regression performed on the relationship between the scaling estimated from motion-in-depth and those estimated for static stimuli was 0.80. This could indicate a slight propensity for some observers to be more sensitive to the distance effect for motion-in-depth stimuli than for static ones.

7.2.3 Discussion

In the present experiment, participants were asked first to judge the circularity (contraction/expansion in depth relative to horizontal dimension) of transparent rotating cylinders defined by random-dot stereograms for four possible virtual distances in front of the screen plane. We expected to identify a significant effect of the disparity-induced distances on the ability to judge 3D shape as it was previously reported (e.g. Hoffman et al., 2008). A second task aimed at investigating whether observers' stereoscopic distortions were similar for static and moving-in-depth transparent rotating cylinders. Such a confirmative result is very useful for knowing how to adjust content for each stereoscopic user, because this indicates that (1) motion-in-depth is not sufficient in itself to provide depth constancy and therefore, that (2) adjustment through motion may be realized according to static stimuli estimations.

The disparity-induced distance had a significant effect on perceiver's ability to judge rotating cylinders as circular, and also showed increased uncertainty in judgment given the effect observed on the slopes at the PSE for each distance. This effect was rather small in this experiment, possibly because the maximum vergence demand was low compared to studies in which the observation distance was smaller (below 60 cm; e.g. Watt et al., 2005; Hoffman et al., 2008) and where there was a stronger stereoscopic demand (beyond 1D). Hence, the effect of distance was significant but unfortunately cannot be differentiated between an effect of accommodation distance and an effect of vergence distance, since we did not test the viewing conditions with the corresponding accommodation distance. The effect on the slopes at the PSE could either support an increase in task difficulty because of the increased vergence demand with stereoscopic distances or, an increase in judgment confidence due to the increased stereoscopic distance providing a somewhat uncommon viewing condition (e.g. mismatch between vergence and accommodation cues). It could however be noted that these results strongly support the possibility that stereoscopic object shapes can be adjusted according to observer's characteristics.

As compared to a disparity matching task – where two disparity-defined objects at different distances have to be matched (e.g. Glennerster, Rogers & Bradshaw, 1996) – the 3D shape judgment task concerns only one visual object of which at least two dimensions are used to provide the judgment. When asked to judge the expansion/contraction of such a cylinder, participants had to compare the horizontal dimension (width) with the depth information, and both of these dimensions need proper information about viewing distance in order to be estimated and compared (Johnston, 1991). There are experimental evidences to support that retinal size and disparities can be scaled according to the same viewing distance (van Damme & Brenner, 1996). As a consequence, the origin of systematic distortions in judging 3D shape could be due to a bias in perceiving correctly the viewing distance; this ability would be dependent on the actual vergence and accommodative states (Watt et al. 2005; MonWilliams & Tresilian, 2000), on the size ratios of horizontal and vertical disparities (Rogers & Bradshaw, 1995), and on cognitive pictorial cues such as familiar size or perspective (O'Leary & Wallach, 1980). The overestimation of cylinders depth presented in this study could therefore be the outcome of interacting vergence (stimulated by disparities), accommodation (constrained by the screen display distance and resolution) and vertical disparities. An indirect effect of accommodation cue is also possible given the mismatch with the vergence cue (Watt et al., 2005).

In a second task, observers had to judge how object shape could change during motion-in-depth. They were asked to judge whether an approaching rotating cylinder expanded or contracted in depth. They showed relatively good performances in doing the task precisely when compared to previous static estimations, as shown in Figure7. 4. The results provided here are thus replicating part of those obtained by Scarfe & Hibbard (2006) supporting the conclusion that shape information from motion-in-depth does not improve the perception of

3D shape from stereopsis. In addition, an approaching object must be contracted to be perceived as being constantly circular in depth and, therefore it is thus possible that even though information about object change is available, the same distance could be used all along the motion to scale object disparity. It is very important to note that contrary to Scarfe & Hibbard (2006), the stimuli used in this study present congruent angular size variation with distance. The information contained in the object size change, also called changing size or size looming, can provide metric information, e.g., for interceptive task (Montagne et al., 2000) and, therefore could be a source of information for relative depth perception (Wann, Mon-Williams, McIntosh, Smyth & Milner, 2001). Looming is thus considered as a powerful cue to motion-in-depth (Erkelens & Regan, 1986; Regan & Beverley, 1978) and its absence in the visual scene can reduce motion-in-depth perception. The change in retinal size also provided strong information about motion-in-depth but, however, this was insufficient to provide neither a correct estimate of distance nor a contribution to the estimate of viewing distance. In that, the present experiment showed that perceptual estimation of 3D shape is affected even for a small range of disparities and 3D shape perception seems similar for static and for moving stimuli.

7.3. Experiment 2

The role of binocular disparities in the perception of depth has been considerably studied since the invention of the random-dot stereograms (RDS); it provides a way to study the relation of disparity to perceived depth by displaying disparity in isolation (Julesz, 1971). It is thus predominantly used in the study of 3D shape perception whereas it is difficult to generalize all results from reduced-cue viewing conditions to natural (or cues-composed) viewing conditions. The nature of stimulus is thus rather underestimated in laboratory tests; random-dot stereograms are used to infer 3D shape with depth relation based on binocular disparity. However, visual depth relation can also be derived from monocular perspective cues such as linear perspective (Howard, 2002), e.g., in slant perception (Olson, 1974; Gillam, 1968; Freeman, 1966; Flock, 1965) and depth perception from monocular image (Saunders & Backus, 2006). Linear perspective or, the convergence of lines, is a potent information that recedes in the visual world but, in order to be used, this cue requires one fundamental assumption of lines parallelism for the same surface. Yang and Kubovy (1999) presented cube-like line drawings in a monocular image. They varied the relationships between the size and the amount of perspective and asked participants to identify the best cube. They observed a preference for images approaching the geometric solution, but a wide range of stimuli were identified as good cubes. The monocular information from perspective can thus help observers to perceive depth relation even though cues are depicted on flat images. Stevens and Brookes (1988) performed a study in which the interaction between linear perspective and stereopsis was investigated, showing that depth perception was dominated by linear perspective rather than disparity when there were contradictory. On the other hand, it was shown that adding texture information to random-dot stereograms increased perceived depth (Johnston, Cumming & Parker, 1993). It thus seems that in case of conflict between these two sources of information, a compromise can be reached depending on the reliability of the cues alone, even though in some conditions observers can experience bi-stability (van Ee, van Dam & Erkelens, 2002). As a consequence, it remains to be known how cues can interact when binocular disparities are incorrectly interpreted by increasing the disparity-induced distance of the stimulus.

The main hypothesis of the next experiment was thus the possibility that perceived distortion from disparity in motion-in-depth could be affected by the available cues in the stimulus. Indeed, stimuli with perspective monocular information may reinforce the distortion as compared to stimuli including only disparity information (i.e. random-dot stereogram). The ability for observers to discriminate object shape changes in depth for two types of stimuli was thus compared for stimuli that were either a random-dot stereogram defined cube where faces were defined by a texture of dots or an outlined cube defined by its edges intersecting at each vertex. Each cube was transparent and constantly rotating about its vertical axis. The striking difference between these two stimuli is that the edges-defined one contained an explicit linear perspective information with the vanishing lines from the vertices strongly informing the viewer of perspective distortion, whereas the random-dots textured one contained only disparity signals, a weak texture information and an implicit perspective from the shape defined by stereopsis. In that, it was expected that the two stimuli would not convey distortion from the disparity-induced distance with the same strength.

7.3.1. Method

7.3.1.1 Stimuli

In this experiment, the stimuli could be of two types: they were either random dot stereograms depicting transparent spinning cube (with dots in the front and in rear faces all visible) or edges-defined transparent cube (that contained strong coarse perspective and disparity cues, see Figure 7.5). Again, stimuli were rendered using OpenGL: they could be adjusted as a function of simulated distance and object depth. Random-dot cubes were composed of 100 small white dots per face, defined by four pre-generated 2D textures used to texture-map the cube. The cube was oriented with a constant tilt of 15 degrees relative to the horizontal axis in such a way that when approaching it did not appear as a fronto-parallel surface. Moreover, a vertical rotation was applied, with a mean velocity of about 45 deg.s-1 and a random rotation direction from one trial to the next. The cube was always moving in depth and the estimated scaling on the cylinder depth was chosen at starting position using the estimate computed in the first task of experiment 1 for the screen plane position. The motion ended when it reached the maximal fixation distance induced by 1.5 degrees of angular disparity. The constant velocity of the motion along the line of sight was about 11.3cm.s-1. In opposition to the previous task, here we used a constant stimuli procedure because we noticed some hysteresis behavior. The task was to judge the naturalness of the object displacement in motion-in-depth. We decided to measure separately two points of subjective equality: unnatural-to-natural and natural-to-unnatural. Specifically, stimuli were defined according to variations of the coefficient of the regression done on the results of experiment 1 for each observer. The possible values were -0.45, -0.35, -0.25, -0.15, -0.05, 0.05, 0.15, 0.25, 0.35 and 0.45 for the scaling factors. As a consequence, there were ten possible levels of object depth that were repeated ten times, resulting in a total of 100 trials for each stimulus type.

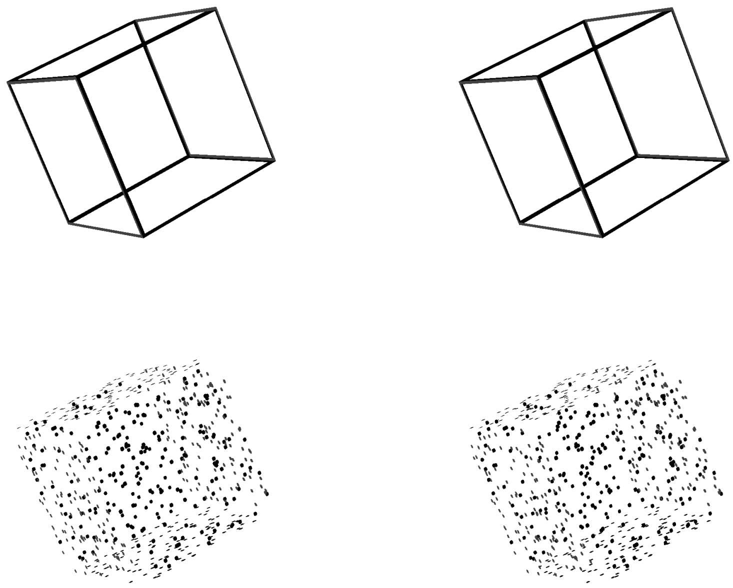

Figure 7.5: Pairs of stereoscopic views used in experiment 2 and arranged for cross fusion. The upper stereoscopic pair represents an outlined cube whose 3D shape can be perceived from apparent edges and from binocular disparity when fused. The lower stereoscopic pair represents a cube defined by a random-dot texture on each face; its 3D shape is then based on the rather weak texture information and on binocular disparity. Note that in each case, cubes are transparent to increase the depth range used in the scaling judgment. When the outlined and RDS cubes, depicting the same shape according to disparity, are compared, one can see that the first contain strong perspective information due to the presence of vanishing lines.

7.3.1.2 Apparatus & Participants

The same five observers participated in this second experiment. We also used the same experimental apparatus.

7.3.1.3 Procedure

From the first thresholds estimation where a rotating cylinder was shown for four fixed simulated distances and from the constant stimuli procedure used here, a total of 200 trials was presented in random order. In this experiment the task was to judge the naturalness of the video sequence where the spinning cube approached from the screen plane toward the individual. Objects were judged as natural when presenting no distortion and smooth zooming associated with motion. Otherwise, the cube was considered unnatural. Participants had to wait for the end of the motion before responding, but then they had to provide their response as quickly as possible.

7.3.2 Results

The probability of response “natural” was fitted for each participant and stimulus condition with Gaussians and the maximum was extracted from each fit (see Figure 7.6). The analysis performed on the maximal probability of response “natural” did not provide significant difference between the two types of stimulus over the five participants ($t=-0.99$, $df=8$, $p>0.05$). Given this first analysis and the well known inter-individual variability in judging 3D shape, we investigated individual observers’ data. Figure 7.6 and 7.7 provide a comprehensive description of the results for this motion task. From Figure 7.6, observer o3 provided somewhat marginal and poor performances, especially for the outlined cube judgments (below the chance level), as compared to the four other participants. Also, it can be observed that for four out of five observers, the maximal probability to provide a response “natural” was larger for the outline cube than for the RDS cube. This observation was supported by the fact that, when discarding observer o3, the statistical difference between the maximal probability to provide a response “natural” was larger for the outlined cube than for the RDS one ($t=-2.90$, $df=6$, $p<0.05$). When considering the data of all participants together, it can however be seen that the position of the maximal probability of response “natural” is quite well aligned with the threshold that was used to construct the shape changes during motion-in-depth in experiment 1 (the main black and dashed vertical line in Figures 7.6 and 7.7). This observation is true for the two stimulus types. Given that not all the PSEs could be computed for participant o3 and due to his atypical results, we discarded him for the next analysis; one possible interpretation of these results is provided in the discussion section. The PSEs of each psychometric function were also extracted (two per Gaussian fit) but they did not provide any significant difference, both for the unnatural-to-natural PSEs ($t=2.02$, $df=6$, $p>0.05$) and for the natural-to-unnatural PSEs ($t=0.26$, $df=6$, $p>0.05$). Figure 7.7 provides averaged data curve fittings with participant o3 omitted.

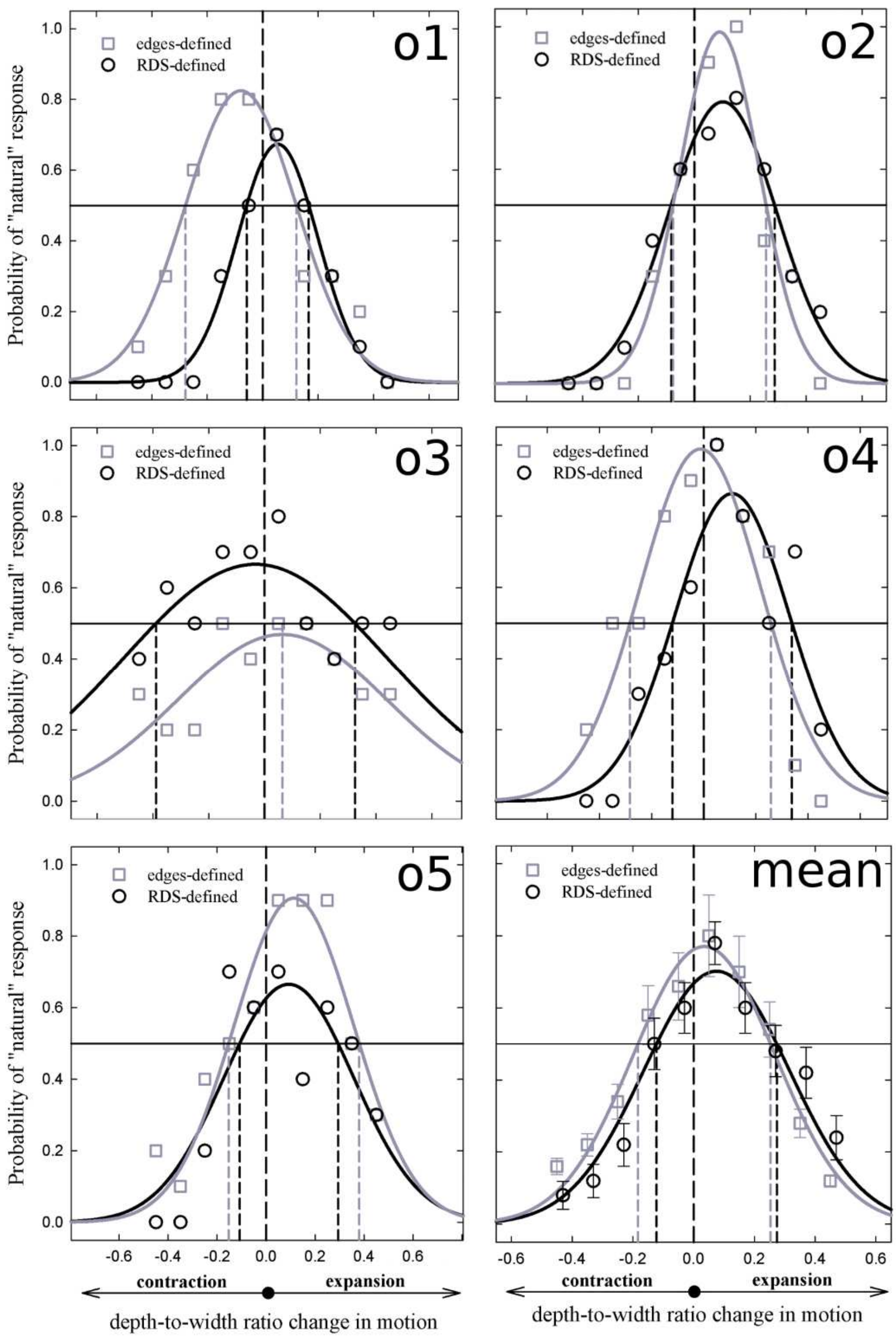

Figure 7.6: This figure represents the results in the naturalness judgment task for the two stimulus conditions, for each observer and the mean across all observers. Each panel represents the probability to judge the cube as “natural” as a function of its expansion/contraction in depth relative to the previous “best fitting” in the cylinder task (0 contraction/expansion) during motion-in-depth. The data are fitted with scaled Gaussians. The gray symbols (circles) and curves represent data for the edges-defined cube condition and the black symbols (squares) and curves represent data for the RDS-defined cube condition. Error bars denote standard error of the mean. Here the main dashed and black lines represent the thresholds estimated in the previous experiment for computing motion-in-depth. Note that

corresponding drop lines are drawn to retrieve the PSE (two per stimulus conditions) and show the extent to which observers were able to discriminate the shape changes as natural or unnatural.

A supplementary analysis is shown in Figure 7.7 in which reaction times are shown as function of the object shape changes in motion (depth-to-width ratio). When reaction times are compared to the probability of response, the distribution of reaction times is in general asymmetrical relative to the PSE of the probability distribution and displays a mode at the location of each PSE. Figure 7.7 shows that the distribution for the two stimulus conditions can be fitted by different polynomials: on average, the reaction time distribution seemed bimodal for the outline cube (four coefficients polynomial fitting), while it appeared unimodal for the RDS cube (two coefficients polynomial fitting).

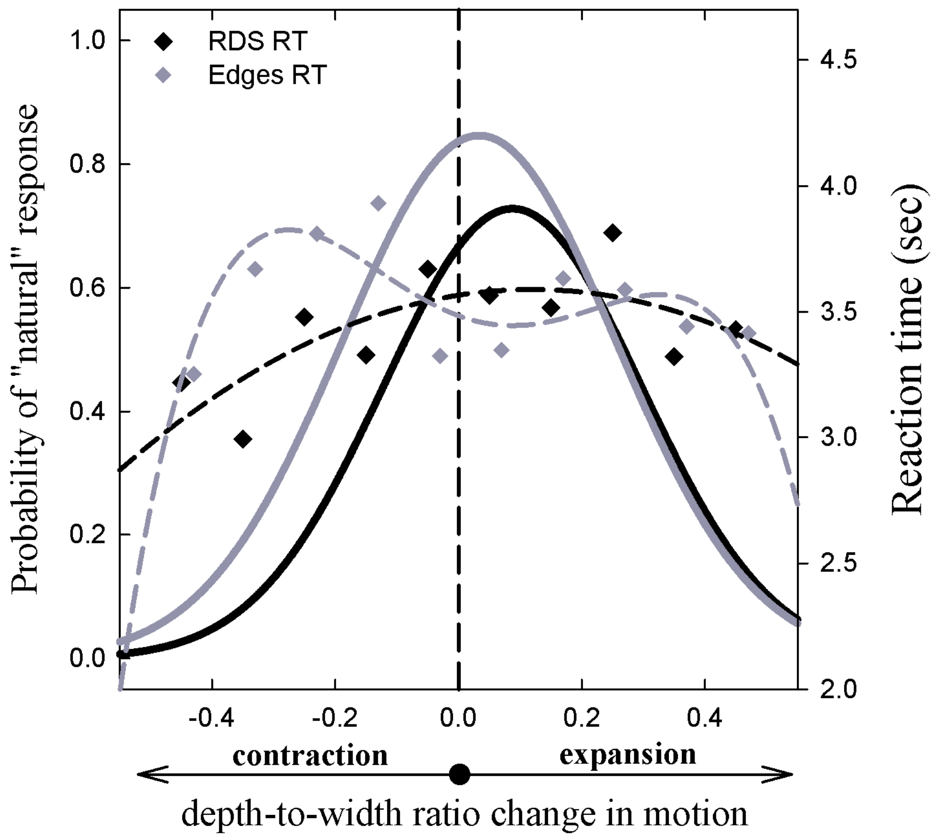

Figure 7.7: This figure represents the averaged results for judgment and reaction times in the naturalness task for the two stimulus conditions and for observers o1, o2, o4 and o5. The figure represents the fitted Gaussians to the probability to judge the cube as “natural” as a function of its expansion/contraction in depth relative to the previous “best fitting” in the cylinder task (0 contraction/expansion) during motion-in-depth. Reaction times are fitted with Polynomials. The gray symbols and curves represent data for the edges-defined cube

condition and the black symbols and curves represent data for the RDS-defined cube condition. The main dashed and black line at x=0 represents the threshold estimated in the previous experiment for computing motion-in-depth.

7.3.3 Discussion

The two main objectives of this second experiment were to study the effect of motion-in-depth on 3D shape perception for a stimulus different than the one used in experiment 1 and to compare the perceptual estimation of naturalness between stimuli with weak monocular perspective (random-dot stereogram) and strong monocular perspective (edges-defined stereogram). Two major results could be highlighted from this experiment. The first was that the maximum probability to report an object displacement as natural was greater for the outlined cube than for the RDS cube. Hence, stimuli providing more important perspective information (with the edges-defined cube), provided stronger propensity to discriminate the object shape changes in depth as natural compared to stimulus defined by random-dot stereograms (except for one observer). Yang and Kubovy (1999) presented cubes of various perspective changes on two-dimensional support to participants who were asked to select from two pictures the one that was perceived better as a cube; participants could only rely on perspective information. They found that participants were rather tolerant regarding what represented the best cube. Conversely, in our study, the cubes were more rarely judged as natural. As a matter of fact, only two participants showed an expansion/contraction value that was always judged as natural, and only for the edges-defined stimulus. It is thus possible that either the task difficulty was too important or that observers were really not convinced of the naturalness of such stimuli. The second main result was that the maximum of probability to provide the response “natural” was roughly aligned with the threshold estimated for cylinder’s motion-in-depth in Experiment 1, and there was no striking difference for the two types of stimuli. This observation suggested that object shape changes in depth were perceived as natural or unnatural as a function of the perceived expansion/contraction of object depth in motion and, that the cues present for judging shape did not change the perception of such a distortion. Moreover, this observation was important in the sense that the distortion from stereopsis due to increasing virtual distance from the display was not changed when compared to the one in Experiment 1 using a rotating cylinder. As a consequence, apart from one observer, the distortion could be obtained with a different and somewhat more complex stimulus. The poor performances of participant o3 could be explained by the fact that this observer may have become tired following the sessions due to the stereoscopic demand (e.g. Hoffman et al., 2008) and that his ability to extract disparity diminished to a point he may have experienced cube reversibility (e.g. Robert & Arger, 1968) that one can observe in absence of disparity information. Hence, an approaching cylinder perceived as bi-stable could certainly be judged less natural. Moreover, the difference between the performances for this observer for the two stimulus conditions could result from the fact that edges-defined cubes induce more easily depth reversal than RDS-defined cubes, leading to a greater probability to judge as unnatural an edges-defined cube.

7.4. General Discussion

Previous studies on 3D shape perception provided evidence that the viewing distance affected the way observers perceive the depth of objects (Hoffman et al., 2008; Watt et al., 2005; Johnston, 1991). In addition, misestimations were also observed when the viewing distance of objects was induced by content disparity that changed the vergence distance (Vienne et al., 2013; Hoffman et al., 2008; Scarfe & Hibbard, 2006; Watt et al., 2005). Hence, perceptual estimation of 3D shape is overestimated at near distances and underestimated at far distances (Johnston, 1991). This observation has been interpreted as a consequence of the use of an incorrect scaling distance. When the vergence distance is varied stereoscopically, objects displayed in front of the reference surface are perceived overestimated (elongated) and objects displayed behind the reference surface are perceived underestimated (contracted) as compared to the same object displayed in the screen plane (Hoffman et al., 2008; Watt et al., 2005). Motion-in-depth is insufficient to cancel this misestimation; approaching objects must be contracted and receding objects must be expanded to provide the object depth matching with the width or height (Scarfe & Hibbard, 2006). In comparison to the present investigation, it is worth noting that these cited studies presented stimuli only with disparity and weak texture information (i.e. random-dot stereograms). Here, the stimuli were designed by taking into account the size-distance ratio and objects were constantly in rotation, thereby providing multiple cues for proper judgment of shape. Moreover, stimuli consisted of transparent objects in order to provide a larger range of depth for the judgment.

Observers were first asked to judge static cylinders as being contracted or expanded in depth and showed that static cylinders displayed in front of the screen must be more and more contracted as they become closer to the observer. In a second task, they had to judge the expansion/contraction of approaching rotating cylinders and reported that these objects must be more and more contracted as they were approaching the observer to appear of equal width/depth. In addition, when PSEs were compared across tasks, they were actually very close for static and moving stimuli. In a second experiment, observers had to judge the naturalness of more complex objects (i.e. rotating cubes) that were either defined with disparity and weak texture (random-dot stereogram) or with disparity and strong (linear) perspective (contour-defined stereogram). The resulting performances showed that naturalness was judged as a function of the perceived expansion/contraction of object in motion-in-depth and that this estimation was very close from that of the previous task based on cylinders. A second result was that the same distortion (scaling) was observed in the two stimulus conditions, thereby indicating that the linear perspective information did not improve depth constancy. On the contrary, a greater probability to judge the object displacement as natural was obtained for the edges-defined cube as compared to the RDS-defined cube. This rather showed that perspective helps to discriminate naturalness, thereby contributing to strengthen the perceived distortion from stereopsis for instance via a cue interaction mechanism. This observation supports the fact that, to provide depth constancy in stereopsis, the single solution consists in adjusting binocular disparity according to perceptual thresholds. Overall, these results highlight the importance of appropriate binocular disparity information in combination to monocular signal for depth perception.

♦

The main objective of this study was to reinforce the knowledge about stereoscopic perceptual distortions, with the particular intention to establish a relationship between what is perceived distorted in depth (i.e. expanded or contracted) and what appears natural or unnatural. The idea behind is to study the capacity of observers in discriminating small variation of depth metric and to gather additional data concerning the perceived distortions so as to derive the possible content adaptation processing. Knowing the perceptual error in depth judgments and the viewing distance of the observer, it is thus possible to transform the depth mediated by the differences between the left and the right images on the basis of perceptual thresholds.

Chapter 8. Conclusion

Stereoscopic three-dimensional technology has recently received new considerations for two main reasons: the growing attractiveness in rendering depth for media entertainment, and the advance in 3D display technologies which now broadly dominate the use of anaglyphs. In fact, 3D displays may provide a more informative and more immersive visual experience because observers can extract the binocular disparities displayed between the left and the right views, more efficiently retrieve the depth of the observed visual scene, and thus, give visual content another dimension. The counterpart of this interest is the list of to-be-solved problems that interfere with the spectators' quality of experience. It has been reported that visual discomfort accompanied by a variety of visual symptoms may result from the use of stereoscopic displays (e.g. Shibata et al., 2011; Kim et al., 2011; Hoffman et al., 2008; Ukai & Howarth, 2008; Emoto et al., 2005; Wann et Mon-Williams, 2002; Rushton & Riddell, 1999; Wann, Rushton, & Mon-Williams, 1995). One of the main reason considered as the origin of these visual symptoms is the stereoscopic demand, providing an imbalance between the inputs that dramatically impact the oculomotor system. The stereoscopic viewing conditions have also shown to provide inaccurate perception of depth (e.g. Hoffman et al., 2008; Watt et al., 2005), the accommodation signal could influence the perceived depth in stereopsis.

The major objective of this thesis work was to investigate the main factors affecting QoE in stereopsis in order to provide better understanding and potential guidelines towards the improvement and further use of stereoscopic systems. As it was presented, three key aspects of QoE in stereoscopic viewing conditions were addressed: (1) the sources and causes of visual fatigue and discomfort, (2) the perceptual distortions due to the use of virtual distance induced by disparity and, (3) the improvement of QoE by adapting the disparity of the visual scene to previously measured perceptual thresholds. Vergence eye movements were measured both in stereo-display and in dual-screen display enabling the presentation of congruent disparity and blur-defocus information when studying visual fatigue. The effect of stereopsis on vergence movements was studied in order to test whether vergence tracking can specifically be used as indicator of visual fatigue. The next topic investigated the consistency in stereoscopic 3D shape perception as a function of vergence distance induced by binocular disparity and the screen-constrained accommodation distance. The role of the depth-of-focus in stereopsis was evaluated by manipulating the pupil aperture size with two controlled peripheral luminance conditions. Finally, the improvement of 3D shape perception was addressed through content adaptation according to individual perception thresholds measurement for motion-in-depth stimuli.

In that, this thesis has provided a more detailed analysis about vergence dynamics and 3D shape perception in stereoscopic viewing conditions as compared to real-world viewing conditions. In fact, the manipulation of vergence and accommodation distances allowed to unravel the contribution of each signal that is accompanied by a potential cross-link interaction to vergence eye movements and 3D perception. In addition, were investigated an indicator of stereoscopically-induced visual fatigue, the possibility to improve visual comfort by increasing the depth-of-focus and the extent to which it is feasible to personalized the stereoscopic visual scene on the basis of individual perception thresholds.

8.1 Visual fatigue and discomfort in stereoscopic displays

Visual Fatigue may emerge from the prolonged observation of stereoscopic contents which contain too large binocular parallax. The experience of visual symptoms in these conditions has been attributed to difficulties for observers to decouple accommodation and convergence (i.e. Hoffman et al. 2008), both systems being naturally coupled. The vergence system is at stake in stereopsis because it is supposed to be the main visual system component involved when observers explored stereoscopic visual scene. In natural vision, ocular vergence responses can be broken down into several components. The vergence system reacts mainly to two stimuli: the disparity between the two retinal images and the defocus information. The comparison of vergence responses in stereoscopic display condition and “real-world” condition showed that vergence dynamics is slower in stereoscopic viewing conditions; this observation could stem from the ambiguity that the visual system has to solve. Effects of stereoscopic exposure on disparity-step vergence response have also been observed when using stereoscopic display. The visual fatigue simulator including several crossed/uncrossed targets binocular disparities for which participants had to perform relative depth judgments showed to strongly stimulate the oculomotor system. The range of disparities was small for one condition for which no major significant effects were observed but, when his range was larger it provided an overall increase in vergence velocity. Hence, depending on the displayed range of disparity, the temporal dynamics of vergence became faster in the post-test as compared to the pre-test. This result could be due to the adaptation of the accommodative crosslink but, when the visual system is in fact unable to cope with the stereoscopic demand and, because of larger binocular parallax, the intensity of visual symptoms and the vergence dynamics changes are expected to be negative (Goss, 1995). Hence, it was seen that the vergence system can be altered in stereoscopic viewing systems and that its prolonged used can lead to changes in vergence responses accompanied by subjective discomfort. The present question is thus to know how this can be used to help improving the QoE in stereoscopic displays.

From the above discussed results, the first solution coming to mind to improve visual discomfort is to increase the depth-of-focus of the eyes in order to minimize the propensity for the visual system to perform accommodation responses. In line with that, in order to reduce visual fatigue in stereoscopic displays, the literature proposed mainly two solutions. The first is to reduce the range of disparities. Indeed, researchers observed with subjective tests that when the displayed content did not exceed a value of about one degree of angular disparity then viewers did not complain about visual discomfort or fatigue (e.g. Wöpking, 1995). However, here the problem is that reducing disparity will reduce the 3D effect. Hence, an alternative solution is to process parts of the content which disparities exceed the recommended limits, say by blurring (Leroy, Fuchs & Moreau, 2010). But, again, there are now two main problems arising from this solution. Indeed, the blurred parts may concern artistic choice of the content creator and, it could be annoying for stereoscopic viewer to try focusing on blurred parts.

Significant changes in the vergence system were observed following the observation of 3D content that were dependent on the presented range of disparities. The changes in vergence can thus be used as indicator of fatigue in the viewer visual system. As compared to other visual fatigue sensors (e.g. the eye-blink rate), vergence responses are directly related to stereoscopic depth perception and give an approach from a performance perspective. An advantage is also that vergence step responses can be easily implemented using the same imaging method as the stereoscopic content and, therefore there is no need for a supplementary processing system.

Two direct applications are thus possible. The first is to evaluate potential changes in the visual system of controllers of stereoscopic media in quality control rooms. This way, it would be possible to rate any stereoscopic media in its propensity to produce visual fatigue and, therefore to adapt the content part that could be annoying. The second solution is to evaluate visual fatigue in stereoscopic viewer and to provide a fatigue alert that informs the viewer to reduce the disparity through a depth adjustment processing or to make a break before continuing watching the content. This solution could be realized at home with the development of relatively low-cost eyetracking system for entertainment now appearing on the market.

Using an indicator of visual fatigue in stereoscopic viewing conditions, it is possible to detect when the visual system of viewers becomes tired in order to perform a content adaptation processing that could be useful to reduce discomfort at this stage. The use of an eye-tracking system during visualization of stereoscopic media is not disadvantageous, because it could also help to compute the parameters of the visual scene that could be fitted or personalized in the case of a single user (e.g. computing the gaze-dependent image blur depth-of-field, adjusting disparities as a function of the inter-ocular distance).

8.2 Perceptual distortions in stereopsis with virtual vergence distances

Together with the problems related to visual fatigue, numerous perceptual distortions have been reported in stereoscopic viewing conditions (Woods et al. 1993). While two classes of distortion are generally identified with respect to the possible visual annoyance – generation-related distortions vs. display-related distortions (Lambooij et al. 2009) –, one class often omitted is the perception-related distortions – or human-factor-related. This category of distortions is particularly salient in stereo estimation of 3D shape, specifically when disparity cues dominate perspective cues (Buckle, & Frisby, 1993). For instance, when observers are asked to match the depth of an object according to its width, they show systematic biases that depend on the location of the plane of reference that viewers use to estimate of depth (Johnston, 1991). More specifically, S3D viewers tend to underestimate the depth of stereoscopic objects when the viewing distance increased while they overestimate depth when the viewing distance is reduced. This phenomenon has been related to the viewers' propensity to scale binocular disparity as a function of an established egocentric distance. This distance could be based on several signals, such as the actual vergence and accommodative states (Watt et al., 2005; MonWilliams & Tresilian, 2000), as well as the size ratios of horizontal and vertical disparities (Rogers & Bradshaw, 1995).

Hence, a major problem of stereoscopic systems is the fact they provide binocular disparity with a screen-constrained accommodation distance. While performing vergence eye movements, a stereoscopic observer thus has to keep an accommodative state that matches with the depth-of-focus of the eyes given a screen-constrained distance. Here the impact of increasing the depth-of-focus of the eye on 3D shape estimations has been investigated using different disparity-induced vergence distances for a large pool of participants. The light-modulated depth-of-focus showed that observers provided better 3D shape estimate of a dihedral angle when the depth-of-focus of the eye was increased (with small pupil size). By varying both the accommodation distance and the vergence distances, it was shown that the weight assigned to the accommodation signal decreased when the depth-of-focus was increased, leading observers to provide estimations on the basis of the vergence signal for scaling distance. These results have been compared to previous that used pinhole viewing conditions; we argued that contrary to pinhole systems that may render accommodation in its resting state and increase the pupil size behind the hole, our method was more ecological because this is the depth-of-focus of the eyes *per se* that was manipulated. Considering that the signal from the state of accommodation contributes to distance perception and stereopsis, these results emphasize the importance of appropriate accommodation information in combination to binocular disparities.

8.3 On the possibility to adjust 3D content on the basis of perceptual thresholds

The fact that observers may experience perceptual distortions creates a serious problem in that veridicality in the perception of 3D shape cannot be reached when one attempts to render the metrics of a captured 3D world. One of the main objectives of this thesis work was to investigate whether observers may have a more vivid experience of depth when the stereoscopic displayed scene is modulated according to perceptual thresholds of object depth-to-width ratio. The results show that participants could precisely retrieve the best modulation between presented depth and width. As this effect could be amplified with stimuli containing stronger perspective cues (i.e. linear perspective), we asked participants to judge the naturalness of spinning cubes that were either edges-defined or defined by randomly textured surfaces (dots). The results show that linear perspective did not improve depth constancy for motion-in-depth object and, therefore appropriate binocular disparity information in combination to monocular signal is needed for depth perception in stereopsis.

Observers of stereoscopic visual scene can thus experience visual distortions, specifically when the visual scene presents objects at different virtual distances. In this thesis work, the possibility for observers to estimate the correct depth that should be displayed to provide the stereoscopic visual scene as undistorted was studied. Another observation was that observers can also perceive distortions for approaching object. In this case, the displayed depth should also be corrected and this can be done on the basis of perceptual thresholds estimated for static objects. The nature of the stimulus also plays a potent role because when the stereoscopic object, which depth is corrected, contains strong monocular perspective information, observers were shown better to perceive the object as natural as compared to stimuli defined by random-dot stereograms. This observation suggests that, it is worth adjusting stereoscopic depth of displayed visual scene, specifically for naturalistic visual scene.

It has been shown that object depth can be perceived expanded or contracted in depth as a function of the virtual distance from the screen plane position. This effect is strong enough to subsist even with motion-in-depth and linear perspective. From these observations, we can thus find out some solutions to adjust the depth of the displayed visual scene. Hence, we observed significant changes in the perception of object depth that were dependent on the virtual distances used to display the object. Here the first solution that comes to mind is to change the pattern of depth in the visual scene presented for a given virtual distance.

However, there is a possible solution that can be applied at the content creation step. As we know that: (1) depth is overestimated at near and underestimated at far physical (screen plane) distances and, that (2) objects in front of the screen are overestimated and objects behind the screen are underestimated, in order to weaken the perceptual distortions it is possible to provide in advance the virtual distance that will be displayed as a function of the observation distance. For example, for short viewing distance (around 50 cm) the best virtual distance will be behind the screen plane and for viewing distance around 3 m, the best virtual distance will be in front of the screen plane.

Now, concerning the content creation with computer generated imagery camera, it is possible to change the rendering specification at the content creation step in adapting the depth of object and visual scene. In so doing, it is thus required to modify the camera parameters to change the depth scaling of objects in the visual scene. Otherwise, the last solution arises at the post-production step and required the modification of the content disparity based on local disparity and therefore needs the use of a dense disparity map (Doyen, Blondé & Borel, 2011).

To conclude, these results on perceptual distortions can be used to provide a new rendering model of stereoscopic depth perception that can be applied at different steps of the content creation or the post-production to better render the scene depth and improve QoE in stereoscopic displays. This rendering model can be applied for an individual personalized 3D content or for a group of individuals on the basis of the mean effect.

8.4 Understanding & Improving QoE in stereoscopic displays

The studies reported in this thesis, furthermore support that stereopsis is strongly dependent on inter-individual differences and that limits of the visual system are idiosyncratic to the individual. In that, future improvement in stereoscopic QoE should attain a best entertainment quality based on personalized stereoscopic experience for individual user.

One parallel that can be made between the studies on visual fatigue and discomfort and those on perceptual distortions is the fact that distorted stereoscopic objects could as well become annoying for the visual system. Indeed, it would not be surprising that an approaching "popping-out" stereoscopic object could provide discomfort and difficulty to fuse if the visual system itself detects at the same time the object depth as distorted (or expanding/contracting when this object moves in depth). As a matter of fact, it is known that people may have difficulty to fuse disparity gradient that exceeds the value of one (Trivedi & Lloyd, 1985). This value has been determined for matched accommodation and vergence distances and, therefore it could be interesting to know whether this assertion holds for perceived or physical distances.

A last interesting question is about the necessity of metric information about depth and distance for stereoscopic visualizations. In the case of stereopsis, there is a large body of studies that suggested that depth constancy from stereopsis does not yield precise information because in natural vision, perceptual biases may be surmounted by the combination with other depth cues (e.g. motion parallax). Moreover, performances in task requiring the participants to perform disparity matching are generally better as compared to those for absolute estimation (e.g. the method used in this thesis) (see Glennerster, Rogers & Bradshaw, 1996). The differences between these two tasks mainly reside in the visual strategy used by the visual system to perform the task. In most everyday actions, the visual system may use binocular disparity in a depth matching task, both in rich visual scene and when the observer can act upon the environment. There is a kind of perception-action dissociation that I experienced at the beginning of this thesis. Imagine, you are looking a large 3D display placed at one meter and half away from you and, on which is displayed in its center a small circle. This small circle is shifted in opposite direction on the left and right images so that it should specify a depth location in front of the display. When presented in isolation and because it is sufficiently away from the display frame, very few people can perceive the correct geometrical distance that is intended (i.e. vergence alone is not sufficient). But, when they stretched one arm in the direction of the circle, it suddenly pops-out. The same effect is produced by using a rod, so the embodied spatial information in stretching the arm is not the solution to this question. Rather, when acting on our world, the visual system may use a simple and advantageous strategy of disparity matching more than a complex computation of depth relation based on the computation of an egocentric distance. However, what happens in the theaters is that people are just contemplative visual systems that should make several on-line computations – if they have the time – to retrieve depth from stereopsis in combination to other depth cues. Again, the design of stereoscopic media designed to provide veridical information must consider perceptual effects in order to better render the quality of experience in stereoscopic perception.

References

Adams, O.S. (1955). Stereogram Decentration and Stereo-Base as Factors Influencing the Apparent Size of Stereoscopic Pictures. The American Journal of Psychology. 68(1), 54-68.
Adams W. J., Mamassian P. (2004). Bayesian combination of ambiguous shape cues. Journal of Vision, 4, 921–929.
Akeley, K., Watt, S. J., Girshick, A. R., & Banks, M. S. (2004). A stereo display prototype with multiple focal distances. ACM transactions on graphics (TOG), 23(3), 804-813.
Allison, R. S. (2004). The camera convergence problem revisited. In Proceedings of SPIE, Vol. 5291, pp. 167-178.
Alpern, M. (1958). Variability of accommodation during steady fixation at various levels of illuminance. JOSA, 48(3), 193-197.
Alpern, M. (1946). The after effect of lateral duction testing on subsequent phoria measurements. American journal of optometry and archives of American Academy of Optometry, 23(9), 442.
Alvarez, T. L., Semmlow, J. L., & Pedrono, C. (2005). Divergence eye movements are dependent on initial stimulus position. Vision research, 45(14), 1847-1855.
Alvarez, T. L., Semmlow, J. L., Yuan, W., & Munoz, P. (2002). Comparison of disparity vergence system responses to predictable and non-predictable stimulations. Current Psychology of Cognition, 21(2/3), 243-262.
Anderson, H. A., Glasser, A., Manny, R. E. et Stuebing, K. K. (2010). Age-related changes in accommodative dynamics from preschool to adulthood. Investigative Ophthalmology & Visual Science, 51, 614-620.
Andreassi, J.L, (2000). Pupillary response and behavior. In: Psychophysiology: Human Behavior & Physiological Response, 4th ed. Lawrence Erlbaum Assoc., Mahwah, N.J., pp. 218-233.
Aslin R., Jackson R. (1979). Accommodative-convergence in young infants: development of a synergistic sensory–motor system. Canadian Journal of Psychology, 33, 222–231.
Atchison, D. A., Charman, W. N., & Woods, R. L. (1997). Subjective depth-of-focus of the eye. Optometry and vision science: official publication of the American Academy of Optometry, 74(7), 511.
Bahill, A. T., Clark, M. R., & Stark, L. (1975). The main sequence, a tool for studying human eye movements. Mathematical Biosciences, 24(3), 191-204.
Bharadwaj, S. R., Wang, J., & Candy, T. R. (2011). Pupil responses to near visual demand during human visual development. Journal of Vision, 11(4).
Blakemore, C., & Julesz, B. (1971). Stereoscopic depth after-effect produced without monocular cues. Science, 171, 286–288.
Blehm C., Vishnu S., Khattak A., Mitra S., Yee R. W. (2005). Computer Vision Syndrome: A Review. Survey of Ophthalmology. 50, 253–262.
Bradshaw, M. F., Parton, A. D., & Glennerster, A. (2000). The task-dependent use of binocular disparity and motion parallax information. Vision research, 40(27), 3725-3734.
Bradshaw, M. F., Glennerster, A., & Rogers, B. J. (1996). The effect of display size on disparity scaling from differential perspective and vergence cues. Vision research, 36(9), 1255-1264.

Brainard, D. H. (1997). The psychophysics toolbox. Spatial vision, 10(4), 433-436.
Brenner, E., & Landy, M. S. (1999). Interaction between the perceived shape of two objects. Vision research, 39(23), 3834-3848.
Brenner, E., & Van Damme, W. J. (1998). Judging distance from ocular convergence. Vision research, 38(4), 493-498.
Brenner, E., Van Den Berg, A. V., & Van Damme, W. J. (1996). Perceived motion in depth. Vision Research, 36, 699–706.
Brunswik, E. (1956). Perception and the representative design of psychological experiments. Berkeley: University of California Press.
Buckley, D., & Frisby, J. P. (1993). Interaction of stereo, texture and outline cues in the shape perception of three-dimensional ridges. Vision research, 33(7), 919-933.
Burian, H. M., & Von Noorden, G. K. (1974). Binocular vision and ocular motility: Theory and management of strabismus (pp. 183-191). CV Mosby.
Busettini, C., Fitzgibbon, E. J., & Miles, F. A. (2001). Short-latency disparity vergence in humans. Journal of Neurophysiology, 85(3), 1129-1152.
Busettini, C., Miles, F. A., Schwarz, U., & Carl, J. R. (1994). Human ocular responses to translation of the observer and of the scene: dependence on viewing distance. Experimental Brain Research, 79(2), 484-494.
Bülthoff, H. H., & Mallot, H. A. (1988). Integration of depth modules: stereo and shading. JOSA A, 5(10), 1749-1758.
Campbell, F. W., & Westheimer, G. (1960). Dynamics of accommodation responses of the human eye. The Journal of Physiology, 151(2), 285-295.
Campbell, F., Robson, J. et Westheimer, G. (1959). Fluctuations of accommodation under steady viewing conditions. The Journal of Physiology, 145, 579-594.
Campbell, F. W. (1957). The depth of field of the human eye. Journal of Modern Optics, 4(4), 157-164.
Champion, R. A., Brenner, E., Mamassian, P., & Simmons, D. R. (2004). No evidence for sequential effects of the interaction of stereo and motion cues in judgments of perceived shape. Vision Research, 44, 483-487.
Chapanis, A., & McCleary, R. A. (1955). Interposition as a cue for the perception of relative distance. American Journal of Psychology, 48, 113-132.
Charman, W. N. (1995). Optics of the eye. Handbook of optics, 1, 24-18.
Charman, W. N., & Koh, L. H. (1995). Effects of Accommodation, Vergence and Pupil Diameter on Size Estimation When Viewing Displays. University of Manchester Institut of Science and Technology (UK) Department of Optometry and Vision Sciences.
Charman, W. N., & Whitefoot, H. (1977). Pupil diameter and the depth-of-field of the human eye as measured by laser speckle. Journal of Modern Optics, 24(12), 1211-1216.
Ciuffreda, K. J. (2006). Accommodation, the pupil, and presbyopia, in: Borish's Clinical Refraction, Benjamin, W. J. ed., 2nd ed., Elsevier, Birmingham, UK, pp. 93-144.
Ciuffreda, K. J., & Kruger, P. B. (1988). Dynamics of human voluntary accommodation. Am J Optom Physiol Opt, 65(5), 365-370.
Clark, J. J., & Yuille, A. L. (1990). Data fusion for sensory information processing systems. Kluwer Academic Publishers.
Cooper, J., Selenow, A., Ciuffreda, K., Feldman, J., Faverty, J., Hokoda, S., & Silver, J. (1983). Reduction of asthenopia in patients with convergence insufficiency

after fusional vergence training. American journal of optometry and physiological optics, 60(12), 982-989.

Corbett A., Maples W.C. (2004). Test-retest reliability of the Saladin card. Optometry - Journal of the American Optometric Association. 75(10), 629–639.

Cutting, J. E. (1998). Information from the world around us. Perception and Cognition at Century's End: History, philosophy, and theory, 69-93.

Cutting, J. E. (1997). How the eye measures reality and virtual reality. Behavior Research Methods, 29(1), 27-36.

Cutting, J. E., & Vishton, P. M. (1995). Perceiving Layout and Knowing Distances: The Integration, Relative Potency, and Contextual Use of Different Information about Depth, Perception of Space and Motion, (pp. 69-117). Perception of Space and Motion, Academic Press, Inc.

Cutting, J. E. (1991). Four ways to reject directed perception. Ecological Psychology, 3(1), 25-34.

Daum, K. M., & McCormack, G. L. (2006). Fusion and binocularity. Teoksessa Benjamin, William J.(toim.): Borish’s clinical refraction, 2, 145-191.

Domini F., Caudek C., Tassinari H. (2006). Stereo and motion information are not independently processed by the visual system. Vision Research, 46, 1707–1723.

Domini, F., Adams, W., Banks, M.S., (2001). 3D after-effects are due to shape and not disparity adaptation. Vision Res. 41, 2733–2739.

Dowley, D. A. V. I. D. (1989). Fixation disparity. Optometry and vision science: official publication of the American Academy of Optometry, 66(2), 98.

Didier, D., Laurent, B., & Thierry, B. (2011). The Use of a Dense Disparity Map to Enhance Quality of Experience in Stereoscopic 3D Content. In SMPTE Conferences (Vol. 2011, No. 6, pp. 1-13). Society of Motion Picture and Television Engineers.

Dunn, B. E., Gray, G. C., & Thompson, D. (1965). Relative height on the picture-plane and depth perception. Perceptual & Motor Skills, 21, 227-236.

Eadie, A. S., Gray, L. S., Carlin, P., & Mon-Williams, M. (2000). Modelling adaptation effects in vergence and accommodation after exposure to a simulated virtual reality stimulus. Ophthalmic and Physiological optics, 20(3), 242-251.

Eadie, A. S., & Carlin, P. J. (1995). Evolution of control system models of ocular accommodation, vergence and their interaction. Medical and Biological Engineering and Computing, 33(4), 517-524.

Emoto, M., Nojiri, Y., & Okano, F. (2004). Changes in fusional vergence limit and its hysteresis after viewing stereoscopic TV. Displays, 25(2), 67-76.

Emoto, M.; Niida, T.; Okano, F. (2005). Repeated vergence adaptation causes the decline of visual functions in watching stereoscopic television. Journal of Display Technology, 1(2), 328- 340.

Engbert, R. (2005). Microsaccades: A microcosm for research on oculomotor control, attention, and visual perception. Progress in brain research, 154, 177.

Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. Vision research, 43(9), 1035-1045.

Erkelens, C. J., & Regan, D. (1986). Human ocular vergence movements induced by changing size and disparity. The Journal of physiology, 379(1), 145-169.

Erkelens, C. J., & Collewijn, H. (1985). Eye movements and stereopsis during dichoptic viewing of moving random-dot stereograms. Vision research, 25(11), 1689-700.

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. Trends in cognitive sciences, 8(4), 162-169.
Fincham, E. F., & Walton, J. (1957). The reciprocal actions of accommodation and convergence. The Journal of Physiology, 137(3), 488-508.
Fisher, S. K. & Ciuffreda, K. J. (1990). Adaptation to optically-increased interocular separation under naturalistic viewing conditions. Perception, 19, 171-180.
Fisher, S. K., Ciuffreda, K. J. (1988). Accommodation and apparent distance. Perception 17, 609-621.
Fisher, S. K., & Ebenholtz, S. M. (1986). Does perceptual adaptation to telestereoscopically enhanced depth depend on the recalibration of binocular disparity?. Attention, Perception, & Psychophysics, 40(2), 101-109.
Fiorentini, A. & Maffei, L. (1970). Electrophysiological evidence for binocular disparity detectors in human visual system. Science 169:208–209.
Flock, H. R. (1965). Optical texture and linear perspective as stimuli for slant perception. Psychological review, 72(6), 505-514.
Foley, J. M. (1980). Binocular distance perception. Psychological review, 87(5), 411.
Freeman, R. B. (1966). Effect of size on visual slant. Journal of Experimental Psychology, 71(2), 170-176.
Frisby J. P., Buckley D., Horsman J. M. (1995). Integration of stereo, texture and outline cues during pinhole viewing of real ridge-shaped objects and stereograms of ridges. Perception, 24, 181–198.
Frisby, J. P., Buckley, D., Wishart, K. A., Porrill, J., Gårding, J., & Mayhew, J. E. (1995). Interaction of stereo and texture cues in the perception of three-dimensional steps. Vision research, 35(10), 1463-1472.
Fukushima, T., Torii, M., Ukai, K., Wolffsohn, J. S., Gilmartin, B. (2009). The relationship between CA/C ratio and individual differences in dynamic accommodative responses while viewing stereoscopic images. Journal of Vision, 9(13), 1–13.
Garding J., Porrill J., Mayhew J. E., Frisby J. P. (1995). Stereopsis, vertical disparity and relief transformations. Vision Research, 35, 703–722.
Gibson, J. (1979). J. The ecological approach to visual perception. Original 1979 edition, Houghton-Mifflin; 1986 edition, Erlbaum.
Gibson, J. J. (1950). The perception of the visual world.
Gillam, B. J. (1968). Perception of slant when perspective and stereopsis conflict: experiments with aniseikonic lenses. Journal of Experimental Psychology; Journal of experimental Psychology, 78(2), 299-305.
Glennerster, A., Rogers, B. J., & Bradshaw, M. F. (1996). Stereoscopic depth constancy depends on the subject's task. Vision research, 36(21), 3441-3456.
Gogel, W. C., & Tietz, J. D. (1973). Absolute motion parallax and the specific distance tendency. Attention, Perception, & Psychophysics, 13(2), 284-292.
Gogel, W. C. (1972). Depth adjacency and cue effectiveness. Journal of Experimental Psychology, 92(2), 176.
Gogel, W. C. (1969). The sensing of retinal size. Vision research, 9(9), 1079-1094.
Gogel, W. C. (1961). Convergence as a cue to absolute distance. The Journal of Psychology, 52(2), 287-301.
Goss, D. A. (1995). Ocular Accommodation, Convergence, and Fixation Disparity: A Manual of Clinical Analysis, 2nd ed., Butterworth-Heinemann, Boston, MA.
Green, D. G., Powers, M. K., & Banks, M. S. (1980). Depth of focus, eye size and visual acuity. Vision research, 20(10), 827-835.

Gu, Y., & Legge, G. E. (1987). Accommodation to stimuli in peripheral vision. JOSA A, 4(8), 1681-1687.
Harvey, L.O. and Leibowitz, H.W. (1967) Effects of exposure duration, cue reduction, and temporary monocularity on size matching at short distances. J. Opt.Soc.Am. 57, 249-253.
Häkkinen, J., Kawai, T., Takatalo, J., Leisti, T., Radun, J., Hirsaho, A., & Nyman, G. (2008). Measuring stereoscopic image quality experience with interpretation based quality methodology. In Proc. SPIE (Vol. 6808, p. 68081B).
Hasebe H., Oyamada H., Ukai K., Toda H., Bando T. (1996). Changes in oculomotor functions before and after loading of a 3-D visually guided task by using a head-mounted display. Ergonomics, 39(11), 1330–1343.
Heinemann, E.G., Tulving, E. & Nachmias, J. (1959). The Effect of Oculomotor Adjustments on Apparent Size. The American Journal of Psychology, 72(1), 32-45.
Hennessy, R. T., & Leibowitz, H. W. (1971). The effect of a peripheral stimulus on accommodation. Attention, Perception, & Psychophysics, 10(3), 129-132.
Hering, E. (1879). Spatial sense and mo_ements of the eye. Baltimore, MD: American Academy of Optometry translation by C.A. Rahde.
Hermans, T.G. (1937). Visual size constancy as a function of convergence. Journal of experimental Psychology., 21, 307-324.
Heron, G., Charman, W. N., & Schor, C. (2001). Dynamics of the accommodation response to abrupt changes in target vergence as a function of age. Vision research, 41(4), 507-519.
Hibbard, P. B., Goutcher, R., O'Kane, L. M., & Scarfe, P. (2012). Misperception of aspect ratio in binocularly viewed surfaces. Vision Research, 70(1), 34-43.
Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture and disparity cues: Optimal cue combination. Journal of Vision, 4(12).
Hillis, J. M., & Banks, M. S. (2001). Are corresponding points fixed?. Vision research, 41(19), 2457-2473.
Hochberg, C. B., & Hochberg, J. E. (1952). Familiar size and the perception of depth. The Journal of Psychology, 34(1), 107-114.
Hoffman, D. M., & Banks, M. S. (2010). Focus information is used to interpret binocular images. Journal of vision, 10(5):13.
Hoffman, D. M., Girshick, A. R., Akeley, K., & Banks, M. S. (2008). Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. Journal of Vision, 8(3):33, 1–30.
Hofstetter, H. W. (1945). The zone of clear single binocular vision. American Journal of Optometry and Physiological Optics, 22, 301-333.
Hogervorst, M. A., & Eagle, R. A. (1998). Biases in three–dimensional structure–from–motion arise from noise in the early visual system. Proceedings of the Royal Society of London. Series B: Biological Sciences, 265(1406), 1587-1593.
Holway, A.H. and Boring, E.G. (1941) Determination of apparent visual size with distance variant. Am.J.Psychol.54, 21-37.
Howard, I. P. (2002). Depth perception. In Steven's Handbook of Experimental Psychology. John Wiley & Sons, Inc.
Howard, I. P., & Rogers, B. J. (1995). Binocular vision and stereopsis. Oxford: Oxford University Press.
Howarth, P. A. (2011). Potential hazards of viewing 3D stereoscopic television, cinema and computer games: a review. Ophthalmic and Physiological Optics, 31(2), 111-122.

Hung, G. K. (1998). Sensitivity analysis of the stimulus-response function of a static nonlinear accommodation model. Biomedical Engineering, IEEE Transactions on, 45(3), 335-341.
Hung, G. K., Zhu, H., & Ciuffreda, K. J. (1997). Convergence and divergence exhibit different response characteristics to symmetric stimuli. Vision Research, 37(9), 1197-1205.
Hung, G. K., Ciuffreda, K. J. et Rosenfield, M. (1996). Proximal contribution to a linear static model of accommodation and vergence. Ophthalmic and Physiological Optics, 16, 31-41.
Hung G. K., Ciuffreda K. J., Semmlow J. L., Horng J. L. (1994). Vergence eye movements under natural viewing conditions. Investigative Ophthalmology & Vision Science. 35, 3486–3492.
Hung, G. K. et Ciuffreda, K. J. (1988). Dual-mode behaviour in the human accommodation system. Ophthalmic and Physiological Optics, 8, 327-332.
Hung, G. K., Semmlow, J. L., & Ciuffreda, K. J. (1983). Identification of accommodative vergence contribution to the near response using response variance. Investigative Ophthalmology & Visual Science, 24(6), 772-777.
Hung, G. K. et Semmlow, J. L. (1980). Static behavior of accommodation and vergence: Computer simulation of an interactive dual-feedback system. IEEE Transactions on Biomedical Engineering, 27, 439-447.
Inoue, T., Ohzu, H. (1997). Accommodative responses to stereoscopic three-dimensional display. Applied Optometry. 36, 4509–4515.
Ittelson, W. H. (1951). Size as a cue to distance: Static localization. The American journal of psychology, 64(1), 54-67.
Ittelson, W. H., & Ames, A. (1950). Accommodation, convergence, and their relation to apparent distance. The Journal of Psychology, 30(1), 43-62.
Johnston, E. B., Cumming, B. G., & Landy, M. S. (1994). Integration of stereopsis and motion shape cues. Vision research, 34(17), 2259-2275.
Johnston, E. B., Cumming, B. G., & Parker, A. J. (1993). Integration of depth modules: Stereopsis and texture. Vision research, 33(5), 813-826.
Johnston, E. B. (1991). Systematic distortions of shape from stereopsis. Vision Research, 31, 1351–1360.
Jones, R. K., & Lee, D. N. (1981). Why two eyes are better than one: the two views of binocular vision. Journal of Experimental Psychology: Human Perception and Performance, 7(1), 30.
Judge, S. J., & Miles, F. A. (1985). Changes in the coupling between accommodation and vergence eye movements induced in human subjects by altering the effective interocular separation. Perception, 14(5), 617-629.
Julesz, B. (1971). Foundations of cyclopean perception. Chicago: University of Chicago Press.
Kasthurirangan, S., & Glasser, A. (2005). Characteristics of pupil responses during far-to-near and near-to-far accommodation. Ophthalmic and Physiological Optics, 25(4), 328-339.
Kasthurirangan, S., Vilupuru, A. S., & Glasser, A. (2003). Amplitude dependent accommodative dynamics in humans. Vision research, 43(27), 2945.
Kaufman, L. (1974). Sight and mind and introduction to visual perception. New York: Oxford University Press.
Khaustova, D. Y., Blondé, L., Huynh-Thu, Q., Vienne, C., & Doyen, D. (2012, February). Method and simulation to study 3D crosstalk perception. In

IS&T/SPIE Electronic Imaging (pp. 82880X-82880X). International Society for Optics and Photonics.
Kawahata, H. (1954). A study on pupillary reflex. Acta Soc. ophthal. jap. 58, 841-850.
Heinemann, E.G., Tulving, E. and Nachmias, J. (1959) The effect of oculomotor adjustments on apparent size. Am.J.Psychol. 72, 32-45.
Kennedy, R.S., Lane, N.E., Berbaum, K.S. & Lilienthal, M.G. (1993). Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness. International Journal of Aviation Psychology, Vol 3(3), 1993, 203-220.
Kesten H. Accelerated Stochastic Approximation (1958). The Annals of Mathematical Statistics. 29(1), 41-59.
Khosroyani, M., & Hung, G. K. (2002). A dual-mode dynamic model of the human accommodation system. Bulletin of mathematical biology, 64(2), 285-299.
Kim, J., Shibata, T., Hoffman, D., & Banks, M. (2011). Assessing vergence-accommodation conflict as a source of discomfort in stereo displays. Journal of Vision, 11(11), 324-324.
Koffka, K. (1935). Principles of Gestalt psychology. New York: Harcourt, Brace & World.
Koh, L. H., Charman, W. N. (1998). Accommodative responses to anisoaccommodative targets. Ophthalmic and Physiological Optics, 18(3), 254–262.
Kommerell, G., Gerling, J., Ball, M., de Paz, H., & Bach, M. (2000). Heterophoria and fixation disparity: A review. Strabismus, 8(2), 127-134.
Kooi, F. L., & Toet, A. (2004). Visual comfort of binocular and 3D displays. Displays, 25(2), 99-108.
Kotulak, J. & Schor, C.M. (1987) The effects of optical vergence contrast and luminance on the accommodative response to spatially bandpass filtered targets. Vision Res. 27, 1797-1806
Krishnan, V. V., Farazian, F., & Stark, L. (1973). An analysis of latencies and prediction in the fusional vergence system. American journal of optometry and archives of American Academy of Optometry, 50(12), 933.
Kruger, P. B., & Pola, J. (1986). Stimuli for accommodation: blur, chromatic aberration and size. Vision research, 26(6), 957-971.
Künnapas, T. (1968). Distance perception as a function of available cues. Journal of experimental psychology, 77(4), 523.
Lambooij, M., IJsselsteijn, W., Fortuin, M., Heynderickx, I. (2009). Visual discomfort and visual fatigue of stereoscopic displays: A review. Journal of Imaging Science and Technology, 53, 1–14.
Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. Vision research, 35(3), 389-412.
Legge, G. E., Mullen, K. T., Woo, G. C., & Campbell, F. W. (1987). Tolerance to visual defocus. JOSA A, 4(5), 851-863.
Leibowitz, H. and Moore, D. (1966) Role of changes in accommodation and convergence in the perception of size. J. Opt.Soc.Am. 56, 1120-1123.
Leibowitz, H.W. and Owens, D.A. (1978) New evidence for the intermediate position of relaxed accommodation. Documenta Ophthalmologica 46, 133-147.
Leroy, L., Fuchs, P., & Moreau, G. (2010). Real time comfort enhancement in stereoscopic displays by disparity and content-adapted blur. In Industrial Electronics (ISIE), 2010 IEEE International Symposium on (pp. 3595-3600).

Lockhead, G.R. and Wolbarsht, M.L. (1989) The moon and other toys. In The Moon Illusion, edited by M.Hershenson, Lawrence Erlbaum, Hillsdale, N.J., pp.259-266.
London, R., & Crelier, R. S. (2006). Fixation disparity analysis: Sensory and motor approaches. Optometry-Journal of the American Optometric Association, 77(12), 590-608.
Long, N. R., & Over, R. (1973). Stereoscopic depth aftereffects with random-dot patterns. Vision Research, 13, 1283–1287.
Loomis, J. M., Da Silva, J. A., Fujita, N., & Fukusima, S. S. (1992). Visual space perception and visually directed action. Journal of Experimental Psychology: Human Perception and Performance, 18(4), 906-921.
Maddox, E. E. (1893). The clinical use of prisms; and the decentering of lenses. J. Wright.
Mamassian, P. (2006). Bayesian inference of form and shape. Martinez-Conde, Macknik, Martinez, Alonso & Tse (Eds.). Progress in Brain Research, Vol. 154.
Massaro, D.W. & Friedman, D. (1990). Models of integration given multiple sources of information. Psychological Review, 97, 225-252.
Mather, G. (1996). Image blur as a pictorial depth cue. Proceedings of the Royal Society of London. Series B: Biological Sciences, 263(1367), 169-172.
Mayhew, J. E. W., & Longuet-Higgins, H. C. (1982). A computational model of binocular depth perception. Nature, 297(5865), 376-378.
Mays, L. E., & Gamlin, P. D. (1995). Neuronal circuitry controlling the near response. Current opinion in neurobiology, 5(6), 763-768.
Meesters, L. M., IJsselsteijn, W. A., & Seuntiëns, P. J. (2004). A survey of perceptual evaluations and requirements of three-dimensional TV. Circuits and Systems for Video Technology, IEEE Transactions on, 14(3), 381-391.
Miles, P.W. (1975) Errors in space perception due to accommodative retinal advance. AmJ.Optom.Physiol.Opt. 52, 600-603.
Mon-Williams, M., Tresilian, J.R. (2000). Ordinal depth information from accommodation? Ergonomics 43, 391–404.
Mon-Williams, M., & Wann, J. P. (1998). Binocular virtual reality displays: When problems do and don't occur. Human Factors: The Journal of the Human Factors and Ergonomics Society, 40(1), 42-49.
Mon-Williams, M., Warm, J. P. and Rushton, S. (1993), Binocular vision in a virtual world: visual deficits following the wearing of a head-mounted display. Ophthalmic and Physiological Optics, 13: 387–391.
Montagne, G., Fraisse, F., Ripoll, H., & Laurent, M. (2000). Perception–action coupling in an interceptive task: First-order time-to-contact as an input variable. Human Movement Science, 19(1), 59-72.
Neveu, P., Philippe, M., Priot, A. E., Fuchs, P., Roumes, C. (2012). Vergence tracking: a tool to assess oculomotor performance in stereoscopic displays. Journal of Eye Movement Research, 5(2):1, 1-8.
Nojiri, Y., Yamanoue, H., Ide, S., Yano, S., & Okana, F. (2006). Parallax distribution and visual comfort on stereoscopic HDTV. In Proc. IBC (pp. 373-380).
Norman, J. F., Todd, J. T., Perotti, V. J., & Tittle, J. S. (1996). The visual perception of three-dimensional length. Journal of Experimental Psychology: Human Perception and Performance, 22(1), 173-186.
O'leary, A., & Wallach, H. (1980). Familiar size and linear perspective as distance cues in stereoscopic depth constancy. Attention, Perception, & Psychophysics, 27(2), 131-135.

Ohzawa, I., DeAngelis, G.C. & Freeman, R.D. (1990) Stereoscopic depth discrimination in the visual cortex: neurons ideally suited as disparity detectors. Science 249:1037–1041.

Okada, Y., Ukai, K., Wolffsohn, J. S., Gilmartin, B., Iijima, A., Bando, T. (2006). Target spatial frequency determines the response to conflicting defocus and convergence-driven accommodative stimuli. Vision Research. 46, 475–484.

Ogle K. N. (1964). Researches in binocular vision. London: Hafner.

Ogle, K. N. (1954). Fixation disparity. The American orthoptic journal, 4, 35.

Olson, R. K. (1974). Slant judgments from static and rotating trapezoids correspond to rules of perspective geometry. Attention, Perception, & Psychophysics, 15(3), 509-516.

Owens, D. A., & Leibowitz, H. W. (1980). Accommodation, convergence, and distance perception in low illumination. American Journal of Optometry and Physiological Optics, 57, 540-550.

Owens, D. A., & Leibowitz, H. W. (1976). Oculomotor adjustments in darkness and the specific distance tendency. Attention, Perception, & Psychophysics, 20(1), 2-9.

Panum, P. L. (1858). Physiologische Untersuchungen uber das Sehen mit zwei Augen. Schwerssche Buchhandlung, Kiel.

Pastoor, S. (1991). 3D-television: A survey of recent research results on subjective requirements. Signal Processing: Image Communication, 4(1), 21–32.

Patel, N., & FIRTH, A. Y. (2003). Vertical vergence adaptation does improve with practice. Optometry & Vision Science, 80(4), 316-319.

Patterson, R., & Martin, W. L. (1992). Human stereopsis. Human Factors, 34(6):669-692.

Peli, E. (1998). The visual effects of head-mounted display (HMD) are not distinguishable from those of desk-top computer display. Vision Research, 38(13), 2053.

Percival, A. S. (1892). The relation of convergence to accommodation and its practical bearing. Ophthal. Rev, 11, 313-328.

Phillips, N. J., Winn, B., & Gilmartin, B. (1992). Absence of pupil response to blur-driven accommodation. Vision research, 32(9), 1775-1779.

Regan, D., Giaschi, D., Sharpe, J. A., & Hong, X. H. (1992). Visual processing of motion-defined form: selective failure in patients with parietotemporal lesions. The Journal of neuroscience, 12(6), 2198-2210.

Regan, D., & Beverley, K. I. (1978). Looming detectors in the human visual pathway. Vision research, 18(4), 415-421.

Regan D. and Beverley, K. I. (1973) Disparity detectors in human depth perception: evidence for directional selectivity. Science, 181, 877-879.

Richards, W. (1985). Structure from stereo and motion. JOSA A, 2(2), 343-349.

Richards, W. (1971). Anomalous stereoscopic depth perception. JOSA, 61(3), 410-414.

Richards, W. (1970). Stereopsis and stereoblindness. Experimental Brain Research, 10(4), 380-388.

Riccio, G. E., & Stoffregen, T. A. (1991). An ecological theory of motion sickness and postural instability. Ecological psychology, 3(3), 195-240.

Ritter, M. (1977). Effect of disparity and viewing distance on perceived depth. Attention, Perception, & Psychophysics, 22(4), 400-407.

Robert, H. C., & Arger, R. (1968). Necker cube perspective dominance as a function of retinal disparity. Perceptual and Motor Skills, 26(2), 367-370.

Rogers, B.J., Bradshaw, M.F. (1995) Disparity scaling and the perception of frontoparallel surfaces. Perception, 24, 155–179.
Rogers, B. J., & Collett, T. S. (1989). The appearance of surfaces specified by motion parallax and binocular disparity. The Quarterly Journal of Experimental Psychology, 41(4), 697-717.
Roscoe, S.N. (1989) The zoom lens hypothesis. In The Moon Illusion, Erlbaum, Hillsdale, N.J., pp.31-57
Rosenfield, M., Ciuffreda, K. J., & Hung, G. K. (1991). The linearity of proximally induced accommodation and vergence. Investigative ophthalmology & visual science, 32(11), 2985-2991.
Rushton, S. K. et Riddell, P. M. (1999). Developing visual systems and exposure to virtual reality and stereo displays: Some concerns and speculations about the demands on accommodation and vergence. Applied Ergonomics, 30, 69-78.
Saladin, J. J. (2005). Stereopsis from a performance perspective. Optometry & Vision Science, 82(3), 186-205.
Saunders, J. A., & Backus, B. T. (2006). The accuracy and reliability of perceived depth from linear perspective as a function of image size. Journal of Vision, 6(9):3, 882-897.
Scarfe, P., & Hibbard, P. B. (2006). Disparity-defined objects moving in depth do not elicit three-dimensional shape constancy. Vision research, 46(10), 1599-1610.
Schaeffel, F., Wilhelm, H., & Zrenner, E. (1993). Inter-individual variability in the dynamics of natural accommodation in humans: relation to age and refractive errors. The Journal of Physiology, 461(1), 301-320.
Schor, C. M. (2009). Neuromuscular plasticity and rehabilitation of the ocular near response. Optometry and Vision Science, 86, 788-802.
Schor, C. M., Maxwell, J. S., McCandless, J., & Graf, E. (2002). Adaptive control of vergence in humans. Annals of the New York Academy of Sciences, 956(1), 297-305.
Schor, C. M. (1992). A Dynamic Model of Cross-Coupling Between Accommodation and Convergence: Simulations of Step and Frequency Responses. Optometry and Vision Science. 69(4), 258–269.
Schor, C.M. (1986) Adaptation of accommodative vergence and vergence accommodation: Fry Lecture. Am. J. Optom. and Physiol. Optics 63, 587-609.
Schor, C., Wood, I. & Ogawa, J. (1984). Binocular sensory fusion is limited by spatial resolution. Vision Research, 24, 661-665.
Schor, C. M. (1983). Fixation disparity and vergence adaptation. Vergence eye movements: Basic and clinical aspects, 465-516.
Semmlow, J. L., Yuan, W., & Alvarez, T. (2002). Short-term adaptive control processes in vergence eye movement. Current Psychology of Cognition, 21(4/5), 343-376.
Semmlow, J., Hansmann, D., & Stark, L. (1975). Variation in pupillomotor responsiveness with mean pupil size. Vision research, 15(1), 85-90.
Semmlow, J. L., Hung, G. K., & Ciuffreda, K. J. (1986). Quantitative assessment of disparity vergence components. Investigative ophthalmology & visual science, 27(4), 558-564.
Semmlow, J. L., & Hung, G. K. (1983). The near response: theories of control. Vergence eye movements: Basic and clinical aspects, 175-195.
Semmlow, J. L., Hung, G. K. (1981). Experimental evidence for separate mechanisms mediating accommodative vergence and vergence accommodation. Documenta Ophthalmologica. 51, 209–24.

Semmlow, J., & Wetzel, P. (1979). Dynamic contributions of the components of binocular vergence. JOSA, 69(5), 639-645.

Semmlow, J., & Venkiteswaran, N. (1976). Dynamic accommodative vergence components in binocular vision. Vision research, 16(4), 403-410.

Sheedy, J. E., Hayes, J., Engle, J. (2003). Is all asthenopia the same? Optometry and Vision Science. 80, 732–739.

Sheedy, J. E., & Saladin, J. J. (1978). Association of symptoms with measures of oculomotor deficiencies. American journal of optometry and physiological optics, 55(10), 670.

Shibata, T., Kim, J., Hoffman, D. M., Banks, M. S. (2011). The zone of comfort: Predicting visual discomfort with stereo displays. Journal of Vision, 11(8).

Smith, G., Meehan, J.W. and Day, R.H. (1992) The effect of accommodation on retinal image size. Human Factor's 343, 289-301.

Speranza, F., Tam, W. J., Renaud, R., & Hur, N. (2006). Effect of disparity and motion on visual comfort of stereoscopic images. In Proc. SPIE (Vol. 6055, No. 60550B, pp. 94-103).

Stakenburg, M. (1991). Accommodation without pupillary constriction. Vision research, 31(2), 267-273.

Stevens, K. A., & Brookes, A. (1988). Integrating stereopsis with monocular interpretations of planar surfaces. Vision research, 28(3), 371-386.

Stiles, W. S., & Crawford, B. H. (1933). The luminous efficiency of rays entering the eye pupil at different points. Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character, 428-450.

Suryakumar, R., Meyers, J. P., Irving, E. L., & Bobier, W. R. (2007). Vergence accommodation and monocular closed loop blur accommodation have similar dynamic characteristics. Vision research, 47(3), 327-337.

Tittle, J. S., Todd, J. T., Perotti, V. J., & Norman, J. F. (1995). Systematic distortion of perceived three-dimensional structure from motion and binocular stereopsis. Journal of Experimental Psychology: Human Perception and Performance, 21(3), 663-678.

Tittle, J. S., & Braunstein, M. L. (1993). Recovery of 3-D shape from binocular disparity and structure from motion. Attention, Perception, & Psychophysics, 54(2), 157-169.

Toates, F. M. (1975). Control theory in biology and experimental psychology. Hutehison, London.

Todd, J. T. (2004). The visual perception of 3D shap. Trends in cognitive sciences, 8(3): 115-121.

Todd, J. T., & Norman, J. F. (2003). The visual perception of 3-D shape from multiple cues: Are observers capable of perceiving metric structure?. Attention, Perception, & Psychophysics, 65(1), 31-47.

Torii, M., Okada, Y., Ukai, K., Wolffsohn, J. S., Gilmartin, B. (2008). Dynamic measurement of accommodative responses while viewing stereoscopic images. Journal of Modern Optics, 55.

Tresilian, J. R., Mon-Williams, M., & 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.

Trivedi, H. P., & Lloyd, S. A. (1985). The role of disparity gradient in stereo vision. Perception, 14(6), 685-690.

Tsuetaki, T. & Schor, C.M. (1987) Clinical method for measuring adaptation of tonic accommodation and vergence accommodation. Am. J. Optom. & Physiol. Optics 64, 437-449

Turner, J. E., Horwood, A. M., Houston, S. M., Riddell, P. M. (2002). Development of the response AC/A ratio over the first year of life. Vision Research. 42(22), 2521-2532.

Tyler, C. W. (1990). A stereoscopic view of visual processing streams. Vision research, 30(11), 1877-1895.

Ukai, K., & Howarth, P. A. (2008). Visual fatigue caused by viewing stereoscopic motion images: Background, theories, and observations. Displays, 29(2), 106-116.

Ukwade, M. T., Bedell, H. E., & Harwerth, R. S. (2003). Stereopsis is perturbed by vergence error. Vision research, 43(2), 181-193.

Van Damme, W., & Brenner, E. (1997). The distance used for scaling disparities is the same as the one used for scaling retinal size. Vision research, 37(6), 757-764.

Van Ee, R., van Dam, L. C., & Erkelens, C. J. (2002). Bi-stability in perceived slant when binocular disparity and monocular perspective specify different slants. Journal of Vision, 2(9): 597-607.

Viénot, F., Bailacq, S., & Rohellec, J. L. (2010). The effect of controlled photopigment excitations on pupil aperture. Ophthalmic and Physiological Optics, 30(5), 484-491.

Vienne, C., Mamassian, P., Doyen, D. & Blondé, L. (2013). Perception of stereo at different vergence distances: implications for realism. International Conference on 3D Imaging (IC3D) (2012). In Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series (2013).

Vienne, C., Blondé, L., & Doyen, D. (2012, February). Visual fatigue versus eye-movements. In IS&T/SPIE Electronic Imaging (pp. 828812-828812). International Society for Optics and Photonics.

Vishwanath, D., & Blaser, E. (2010). Retinal blur and the perception of egocentric distance. Journal of Vision, 10(10):26, 1-16.

Vishwanath D., Girshick A. R., Banks M. S. (2005). Why pictures look right when viewed from the wrong place. Nature Neuroscience, 8, 1401–1410.

Wallach, H., & Halperin, P. (1977). Eye muscle potentiation does not account for adaptation in distance perception based on oculomotor cues. Attention, Perception, & Psychophysics, 22(5), 427-430.

Wallach, H. & Frey, K.J. (1972b). The nature of adaptation in distance perception based on oculomotor cues. Perception & Psychophysics, 11, 110-116.

Wallach, H., Moore, M. E., & Davidson, L. (1963). Modification of stereoscopic depth-perception. The American journal of psychology, 76(2), 191-204.

Wallach, H., & Zuckerman, C. (1963). The constancy of stereoscopic depth. The American journal of psychology, 76(3), 404-412.

Wann J. P., Mon-Williams M. (2002). Measurement of visual aftereffects following virtual environment exposure. In Stanney K. M. (Ed.), Handbook of virtual environments: Design, implementation, and applications. London, UK: Lawrence Erlbaum Associates.

Wann, J. P., Mon-Williams, M., McIntosh, R. D., Smyth, M., & Milner, D. A. (2001). The role of size and binocular information in guiding reaching: insights from virtual reality and visual form agnosia III (of III). Experimental brain research, 139(2), 143-150.

Wann, J. P., Rushton, S., & Mon-Williams, M. (1995). Natural problems for stereoscopic depth perception in virtual environments. Vision research, 35(19), 2731-2736.
Wang B, Ciuffreda KJ. Depth-of-Focus of the Human Eye: Theory and Clinical Implications. Survey of Ophthalmology. 2006;51(1):75–85.
Ward, P. A., & Charman, W. N. (1987). On the use of small artificial pupils to open-loop the accommodation system. Ophthalmic and Physiological Optics, 7(2), 191-193.
Watt, S. J., Akeley, K., Ernst, M. O., & Banks, M. S. (2005). Focus cues affect perceived depth. Journal of Vision, 5(10):7, 834–862.
Wheatstone, C. (1838). Contributions to the Physiology of Vision.--Part the First. On Some Remarkable, and Hitherto Unobserved, Phenomena of Binocular Vision. Philosophical Transactions of the Royal Society of London , 128, 371-394.
Wismeijer, D. A., & Erkelens, C. J. (2009). The effect of changing size on vergence is mediated by changing disparity. Journal of Vision, 9, 1-10.
Woods, A. (2010). Understanding crosstalk in stereoscopic displays. In Keynote Presentation at the Three-Dimensional Systems and Applications Conference, Tokyo, Japan (pp. 19-21).
Woods, A., Docherty, T., Koch, R. (1993). Image distortions in stereoscopic video systems. Proc. SPIE, 1915, 36–48.
Woodworth, R. S. (1938). Experimental Psychology. New York: Holt.
Wöpking, M. (1995), Viewing comfort with stereoscopic pictures: An experimental study on the subjective effects of disparity magnitude and depth of focus. Journal of the Society for Information Display, 3: 101–103.
Xing, L., You, J., Ebrahimi, T., & Perkis, A. (2011). Factors impacting quality of experience in stereoscopic images. In IS&T/SPIE Electronic Imaging (pp. 786304-786304). International Society for Optics and Photonics.
Yao, P., Lin, H., Huang, J., Chu, R., & Jiang, B. C. (2010). Objective depth-of-focus is different from subjective depth-of-focus and correlated with accommodative microfluctuations. Vision research, 50(13), 1266-1273.
Yang, T., & Kubovy, M. (1999). Weakening the robustness of perspective: Evidence for a modified theory of compensation in picture perception. Attention, Perception, & Psychophysics, 61(3), 456-467.
Yano S., Emoto M., Mitsuhashi T. (2004). Two factors in visual fatigue caused by stereoscopic HDTV images. Displays, 25, 141–150.
Yeh, Y. Y., & Silverstein, L. D. (1990). Limits of fusion and depth judgment in stereoscopic color displays. Human Factors: The Journal of the Human Factors and Ergonomics Society, 32(1), 45-60.
Young, R. S., Han, B. C., & Wu, P. Y. (1993). Transient and sustained components of the pupillary responses evoked by luminance and color. Vision research, 33(4), 437-446.
Zaroff, C. M., Knutelska, M., & Frumkes, T. E. (2003). Variation in stereoacuity: normative description, fixation disparity, and the roles of aging and gender. Investigative ophthalmology & visual science, 44(2), 891-900.
Zee, D. S., Fitzgibbon, E. J., & Optican, L. M. (1992). Saccade-vergence interactions in humans. Journal of Neurophysiology, 68(5), 1624-1641.

Appendices

The estimation of depth from stereopsis is likely dependent on the scaling distance the viewer must consider to retrieve depth. For the dihedron vertical angle used in our experiments, the scaling distance can be defined such as:

$$SD = \frac{\theta.IOD}{\eta \pm \eta.\theta}, \qquad \text{(formula 2.1)}$$

Where θ is the half width of the dihedron, IOD is the inter-ocular distance and η is the disparity between the peak and through of the dihedral angle.

Parameter	Notation	Relaxed value	Accommodated value
Focal length object-sided	f [mm]	17.055	14.169
Focal length image-sided	f' [mm]	22.785	18.930
Refractive power	F [dpt]	58.636	70.57
Location entrance pupil	E [mm]	3.045	2.667
Location exit pupil	E' [mm]	3.664	3.211
Principal point object-sided	P [mm]	1.348	1.772
Principal point image-sided	P' [mm]	1.602	2.086
Nodal point object-sided	N [mm]	7.078	6.533
Nodal point image-sided	N' [mm]	7.332	6.847
Length	L [mm]	24.387	

Table 6.1: Optical data of the Gullstrand eye. The locations of each parameter are referenced to the anterior surface of the cornea (mm = millimeter, dpt = diopter).