Stereoscopic 3D Technologies for Accurate Depth Tasks: A Theoretical and Empirical Study

FRONER, BARBARA

Stereoscopic 3D Technologies for Accurate Depth Tasks: A Theoretical and Empirical Study

by

Barbara Froner

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy

School of Engineering and Computing Sciences
Durham University
United Kingdom

Abstract

**Stereoscopic 3D Technologies for Accurate Depth Tasks:
A Theoretical and Empirical Study**

Barbara Froner

In the last decade an increasing number of application fields, including medicine, geoscience and bio-chemistry, have expressed a need to visualise and interact with data that are inherently three-dimensional. Stereoscopic 3D technologies can offer a valid support for these operations thanks to the enhanced depth representation they can provide. However, there is still little understanding of how such technologies can be used effectively to support the performance of visual tasks based on accurate depth judgements. Existing studies do not provide a sound and complete explanation of the impact of different visual and technical factors on depth perception in stereoscopic 3D environments.

This thesis presents a new interpretative and contextualised analysis of the vision science literature to clarify the role of different visual cues on human depth perception in such environments. The analysis identifies luminance contrast, spatial frequency, colour, blur, transparency and depth constancies as influential visual factors for depth perception and provides the theoretical foundation for guidelines to support the performance of accurate stereoscopic depth tasks.

A novel assessment framework is proposed and used to conduct an empirical study to evaluate the performance of four distinct classes of 3D display technologies. The results suggest that 3D displays are not interchangeable and that the depth representation provided can vary even between displays belonging to the same class. The study also shows that interleaved displays may suffer from a number of aliasing artifacts, which in turn may affect the amount of perceived depth.

The outcomes of the analysis of the influential visual factors for depth perception and the empirical comparartive study are used to propose a novel universal 3D cursor prototype suitable to support depth-based tasks in stereoscopic 3D environments. The contribution includes a number of both qualitative and quantitative guidelines that aim to guarantee a correct perception of depth in stereoscopic 3D environments and that should be observed when designing a stereoscopic 3D cursor.

Declaration

The material contained within this thesis has not previously been submitted for a degree at Durham University or any other university. The research reported within this thesis has been conducted by the author unless indicated otherwise. This research has been documented, in part, within the following publications:

- B. Froner, N.S. Holliman and S.P. Liversedge. A comparative study of fine depth perception on two-view 3D displays. Displays, 29(5):440-450, December 2008.

- N.S. Holliman, B. Froner and S.P. Liversedge. An application driven comparison of depth perception on desktop 3D displays. In *Stereoscopic Displays and Virtual Reality Systems XIV, Proceedings of SPIE-IS&T Electronic Imaging*, Volume 6490A, January 2007.

Further relevant research conducted by the author has been documented in the following publications:

- N.S. Holliman, C. Baugh, C. Frenk, A. Jenkins, B. Froner, D. Hassaine, J. Helly, N. Metcalfe and T. Okamoto. Cosmic cookery: making a stereoscopic 3D animated movie. In *Stereoscopic Displays and Virtual Reality Systems XVII, Proceedings of SPIE-IS&T Electronic Imaging*, Volume 6055A, San Jose (CA-USA), January 2006.

- B. Froner and N.S. Holliman. Implementing an Improved Stereoscopic Camera Model, in Eurographics Theory and Practice of Computer Graphics 2005, Canterbury (UK), June 2005.

Copyright Notice

Acknowledgements

I would like to express my gratitude to my academic supervisors Dr Nick Holliman and Prof Malcolm Munro, for providing me with the opportunity to pursue a PhD and for their advice and guidance during my research.

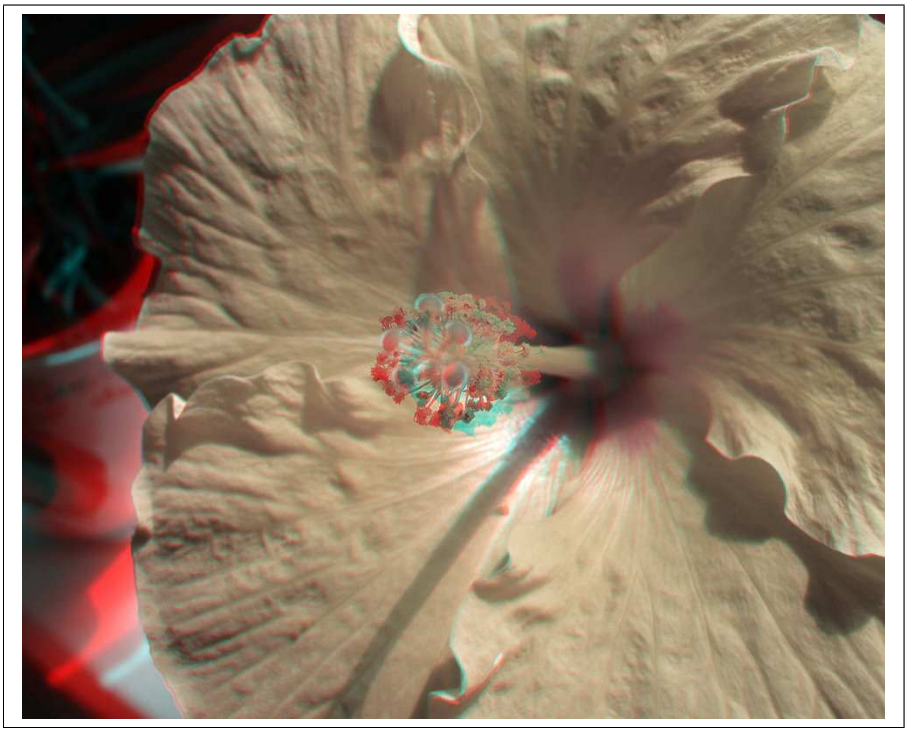
A heartfelt thank you and love go to my husband Dr Nikolaos Galiatsatos for standing by me throughout these six years of part time studies and for putting up with my moods; I do not know how I would have been able to accomplish this without your continuous support, encouragement and patience. An equally heartfelt thanks goes to my parents Maria Grazia and Mario, for believing in me and supporting me throughout my entire education. Many thanks also to the rest of my family for their care and understanding.

I am grateful to Prof Simon Liversedge, Dr Gustav Kuhn and Prof John Findley for their invaluable advice on statistics and perceptual psychology, and to Prof Luisa Mich for her encouragement and counsel. I also wish to thank my current employer, Foster Findley Associate, for their flexibility and for allowing me the time to complete my studies.

Special thanks go to Dale Norton, Dr Oliver Vogt, Tamzin Cloke and Simon McConway for their excellent job in proof reading my draft and their suggestions, and to Dr James Lowell for the brilliant time we had when we were sharing the office at uni. For their contribution to this work thanks also go to my research group members Dr Djamel Hassaine, Paul Gorley and Geng Sun.

I would like to express a very special thank you to Stefano Pellegrini for his invaluable support and friendship, and for standing by me in difficult times. Finally, I wish to thank Roberta Cristelli, Dr Adam Eckersley, Prof Keith Gallagher, Dr John Bailey, Dr Pim van 't Hof, Andrea Bonfá and all my other friends for their encouragement and continual belief.

Dedicated to my parents, Maria Grazia and Mario



Hibiscus, Copyright © by anaglyph
(original title "Another flower, anaglyph", from anaglyph.deviantart.com)

Contents

List of Figures

List of Tables

List of Abbreviations

AR Augmented Reality

ANOVA ANalysis Of VAriance

CAD Computer-Aided Design

CAVE Cave Automatic Virtual Environment

CSF Contrast Sensitivity Function

DOF Depth Of Field or Depth Of Focus

FOV Field Of View

HCI Human-Computer Interaction

HMD Head Mounted Display

GUI Graphical User Interface

IPD InterPupillary Distance

LIDAR LIght Detection And Ranging

LCD Liquid Crystal Display

M Mean

PC Personal Computer

PPI Pixels Per Inch

RDS Random Dot Stereogram

SD Standard Deviation

TMS Transcranial Magnetic Simulation

VR Virtual Reality

XGA eXtended Graphics Array

Chapter 1

Introduction

"A painting, though conducted with the greatest art and finished to the last perfection, both with regard to its contours, its lights, its shadows and its colours, can never show a relievo equal to that of the natural objects, unless these be viewed at a distance and with a single eye."

Leonardo Da Vinci (1452–1519) in Wheatstone [1838] (pg. 1)

In this sentence Leonardo Da Vinci summarises the essence of stereo vision and the unique sensation of depth that one can gain when looking at objects binocularly, i.e. with the input of both eyes. Besides this compelling perceptual effect, binocular 3D vision provides a number of advantages over "flat" 2D vision, including detailed information on the spatial relationship between objects in the observed scene and, therefore, more accurate relative depth judgements [Howard, 1919].

This thesis explores stereoscopic 3D technologies to support application fields that can benefit from the enhanced perception of depth offered by binocular vision and whose results critically depend on accurate depth judgements.

1.1 Context of Work

The general context of this thesis is Human-Computer Interaction (HCI), i.e. the knowledge domain that deals with all different relationships between humans and machines. From an historical point of view, a fundamental milestone in the area of HCI goes back to 1963 and is represented by Sutherland's Sketchpad [1963b] [1963a]. The Sketchpad represents one of the first interactive Graphical User Interfaces (GUIs) and the ideas presented in that work still influence the way we interact with computers nowadays. Since then the area of computer graphics and HCI has continued to evolve. The introduction of new technologies, both hardware and software, made it possible to visualise objects in a more realistic way (e.g. 3D Computer-Aided Design or CAD) and to manipulate them using ever more sophisticated devices (e.g. 3D mice, data gloves, etc.). Stereoscopic 3D displays represent one recent step of this evolution process. By exploiting the principle of stereopsis (literally "solid seeing"), stereoscopic 3D displays give the user the sensation of seeing the scene in a true three-dimensional space and provide a realistic perception of depth. Realistic perception of depth represents a key advantage against standard 2D displays, especially when the user is required to perform tasks based on depth judgements in a 3D space. The first stereoscope was invented by Sir Charles Wheatstone in 1833, but the first electronic 3D displays were only introduced into the market during the 1990s. Since then considerable advances have been made both in stereoscopic 3D technologies and Virtual Reality (VR) environments in order to enhance the level of immersion of the user and the degree of interactivity offered by the system. However, there is still little understanding of how stereoscopic 3D technologies can be used effectively to assist operators in performing depth-based tasks in real application fields. This thesis investigates desktop stereoscopic 3D solutions to support such tasks.

1.2 Statement of the Problem

In the past ten years an increasing number of application fields have expressed the need to visualise and interact with data sets that are intrinsically three-dimensional. In such domains operators often have the necessity to understand the spatial relationships of the objects in the 3D scene and carefully interact with these objects on the basis of their characteristics and depth position. These tasks are normally carried out in a common desktop working environment, using standard technologies such as a 2D display and a 2D cursor driven by a common mouse device. However, there is evidence that such tasks can benefit from using stereoscopic visualisation and 3D technologies thanks to the enhanced depth representation they can provide [Hassaine et al., 2010] [Parkin et al., 2001] [Goodsitt et al., 2000] [Higashida et al., 1988]. Given this scenario, a stereoscopic 3D paradigm of the current 2D working environment seems like a natural progression toward supporting accurate depth judgement tasks.

If we were to substitute a normal desktop 2D display with a stereoscopic 3D counterpart and the classic 2D mouse cursor we use everyday with a stereoscopic 3D version of it, we would obtain what, for the purpose of this thesis, is defined as a *desktop stereoscopic 3D environment*. But what are the characteristics that such a working environment should have in order to effectively support the operator in performing visual tasks that require accurate depth judgements? From this research question a number of sub-problems follow:

- What are the different factors that affect human depth perception in stereoscopic 3D environments?

- How should a stereoscopic 3D cursor for desktop stereoscopic 3D environments be designed in order to account for these factors and effectively support tasks based on accurate depth judgements?

- Does the type of 3D display used have an impact on the perceived depth of the user?

If yes, how can a stereoscopic 3D cursor counterbalance this effect?

Previous studies of human depth perception in stereoscopic 3D environments do not provide a clear answer to these questions.

1.3 Aim and Objectives

The aim of this research is to investigate desktop stereoscopic 3D technologies in the context of visual tasks that require accurate depth judgements. To fulfill this aim, a number of objectives have been identified as specified below.

1. **Investigate influential factors to the perception of depth for depth-based tasks.**

   In order to address the research problems outlined in section 1.2, it is important to develop a thorough understanding of the visual and technical factors that can affect the perception of depth in stereoscopic 3D environments. The vision science literature on human depth perception is copious, however there is little understanding of how different visual factors may impact the performance of tasks based on accurate depth judgement in stereoscopic 3D environments. Furthermore, there is a lack of scientific literature on the effect that the technical properties of the 3D display system used to perform the task could have on the perception of depth of the user. One of the objectives of this thesis is to present a contextualised and complete review of all such factors.

2. **Develop a quantitative assessment framework for 3D display performance.**

   As explained in the previous point, the knowledge regarding the effect of the 3D display technical properties on depth perception is still incomplete and fragmented. This thesis will propose a new method that can be used to assess the performance of different 3D display systems and clarify the role of their technical properties with regard to human depth perception in stereoscopic 3D environments.

3. **Perform an empirical comparative study on a set of representative 3D displays.**

   Previous investigations on human depth perception on 3D displays have almost exclusively studied single displays or compared 3D displays with 2D displays. This thesis will present a study performed to quantify human depth perception threshold levels across a range of representative categories of desktop 3D displays. Empirically investigating and quantifying depth perception is critically important for the aim of this project. It is believed that the different hardware, software and optical characteristics of the 3D display system may have an effect on the quality of the depth representation in the final stereoscopic 3D image and, consequently, on the performance of tasks that require accurate depth judgements.

4. **Formulate a number of guidelines for the design of a stereoscopic 3D cursor.**

   Despite the large amount of literature dedicated to interaction techniques for 3D environments, little work is available regarding the design of a stereoscopic 3D cursor suitable for supporting depth-based tasks in desktop stereoscopic 3D environments. This thesis will present a number of guidelines to address this issue based on existing literature and the novel research carried out as part of this project.

It is important to note that the focus of this research is on task performance and human depth perception and not on the assessment of different interaction techniques or input devices that could be used to drive the stereoscopic 3D cursor. Furthermore, the main interest is the investigation of desktop technologies that do not require the user to change their desktop working habits, so large scale stereoscopic 3D environments, e.g.

Cave Automatic Virtual Environments (CAVEs), are beyond the scope of this thesis. Finally, particular attention has been paid to guarantee that the adopted research method is scientifically sound. In this regard, note that in this thesis the term *methodology* denotes the collection of tools used during the scientific investigation and is not intended in the strict philosophical meaning of the word, i.e. the science of methods and procedures.

1.4 Thesis Overview and Research Contributions

This thesis is organised into a number of chapters and appendices, each focusing on a different aspect of the stated problem, as summarised in the following paragraphs.

- Chapter 2 - Depth Perception and Stereoscopic Environments. Begins with an overview of the mechanisms that enable depth perception in humans and a high level description of the main depth cues and their integration. This is followed by an introductory discussion on the principles behind stereoscopic viewing on 3D display systems and a detailed taxonomy of the different optic design that these electronic devices can adopt. The chapter concludes with a critical review of previous works on stereoscopic 3D cursors and depth-based tasks in 3D environments.
- Chapter 3 - Design of a Stereoscopic 3D Cursor. Investigates the design of a novel 3D cursor to support depth-based tasks in stereoscopic 3D environments. The chapter presents a contextualised, in-depth review of the visual, technical and synthetic factors that can affect depth perception and depth-based task performance in both real scenes and synthetic scenes presented via stereoscopic 3D displays. It concludes with a discussion of the main ideas that emerged from the review and a number of guidelines for designing a stereoscopic 3D cursor.
- Chapter 4 - Assessment of 3D Technologies: Methodology. Presents a robust and sensitive methodology for the investigation of human depth perception on different

3D displays, which has been used in this research project to assess and compare the performance of a set of desktop 3D displays. This chapter also illustrates the rationale behind the hypothesis and predictions for the 3D display comparative study and describes the pilot experiment performed to assess and refine the methodology.

- Chapter 5 - Assessment of 3D Technologies: Results. Describes the experimental details of the main 3D display comparative study and reports the empirical results. The most salient aspects of the collected data and the statistical analysis are thoroughly reported in this chapter.

- Chapter 6 - Discussion. Presents an interpretative discussion of the main findings that emerged from the analysis of the collected data and examines in detail the reasons why in some instances the theoretical predictions were not supported by the empirical results. The chapter concludes with a summary of the main outcomes of the comparative study and the implications that these have on the design of a stereoscopic 3D cursor.

- Chapter 7 - Conclusions and Future Work. The first part of the chapter reviews the research work presented in this thesis and discusses its novel contributions. The second part outlines a number of directions for future work and concludes the chapter and the thesis.

- Appendices. Appendix A contains a glossary of technical terms, of which first usage is marked in bold in the main text. The questionnaires and the forms used during the empirical study are presented in Appendix B, whilst the raw empirical data are reported in Appendix C. Finally, Appendix D illustrates the geometry of perceived depth on planar 3D displays.

The research presented in this thesis offers a number of academic contributions. First, it provides a novel interpretation of the vision science literature in the context of 3D scenes in stereoscopic 3D environments. In this way it contributes an improved and comprehensive understanding of the role that different visual cues play in the perception of depth in such environments. Second, it presents a new assessment framework that classifies 3D display technologies based on their optical design and technical specifications, and uses human perceived depth as a means of evaluating 3D displays' performance and fidelity in reproducing the depth embedded in the input 3D image. Third, it presents a 3D display comparative study that successfully uses this assessment framework to identify the tradeoffs between different displays in supporting visual tasks based on depth judgements. This trial represents the first empirical study to investigate human depth perception on a set of representative electronic 3D displays. Finally, the outcomes of the analysis of the influential factors for depth perception and the empirical study are used to propose a novel universal 3D cursor prototype suitable to support depth-based tasks in stereoscopic 3D environments. Regarding the latter, the research contribution includes a number of both qualitative and quantitative guidelines that should be observed when designing such a cursor.

Chapter 2

Depth Perception and Stereoscopic 3D Environments

2.1 Introduction

This chapter presents the theoretical framework that forms the foundation of the research work presented in this thesis and provides a critical review of previous work related to depth-based tasks performed in stereoscopic 3D environments.

Section 2.2 describes the phenomenon of depth perception in humans and consists of a brief introductory paragraph on visual perception (subsection 2.2.1), an overview of the different depth information sources used by the visual system to infer the third dimension (subsections 2.2.1, 2.2.3, 2.2.4 and 2.2.5) and a discussion of the existing theory regarding the integration of these depth information sources (subsection 2.2.6). Section 2.3 gives an explanation of the principles behind stereoscopic viewing on electronic devices (subsections 2.3.1 and 2.3.2), followed by a survey of the different 3D display technologies available on the market (subsection 2.3.3) and a discussion on the perceptual issues of stereoscopic viewing on 3D displays (subsection 2.3.4). Section 2.4 presents an in-depth review of previous work on depth-based tasks in 3D environments and stereoscopic 3D cursors (subsections 2.3.1 and 2.3.2). The chapter concludes in section 2.5 with a summary of the key points that emerged from the literature review presented in the previous sections.

2.2 Human Vision and Depth Perception

2.2.1 Introduction

Vision is the process that allows individuals to assimilate and interpret visual information from the surrounding environment. It consists of a sequence of steps that begins when light enters the eyeball and reaches the retina. Although located at the back of the eye, the retina can be considered an extension of the brain that acts as an interpreter responsible for converting the light focused by the eye lens into neural signals [Sekuler and Blake, 1994]. These signals, under the form of electrical impulses, are then processed by other parts of the brain and are eventually distributed to different sets of neurons in the **visual cortex**. Each set of cortical cells is sensitive to particular features of the visual information and is responsible for a specific aspect of the visual perception experience of the individual. For example, the wavelength of light is one of the features that our visual system processes in order to enable us to see colour. Colour perception, along with pattern perception, are neural mechanisms that help to define the nature of an object, i.e. *what* we are looking at [Sekuler and Blake, 1994]. In many circumstances though, we need to understand *where* objects are located in the 3D space and how far they are from us. Some animals, e.g. bats, can extract spatial information via auditory processing [Griffin, 1959], but humans and other primates primarily rely on the elaboration of visual information through a process called *depth perception*.

2.2.2 Seeing in 3D

Depth perception is the faculty of perceiving the world in three dimensions in order to identify the spatial layout of objects and surfaces that constitute our surroundings [Fleming and Anderson, 2003]. But how can we perceive the world in 3D given the fact that the inputs to our visual system, i.e. the images projected onto our two retinas, are 2D? This apparent paradox has fascinated philosophers and researchers for centuries, but it was only with the English empiricists, and in particular with George Berkley and his "New Theory of Vision" [Berkeley, 1709], that the modern theory of depth perception started to take shape. Today, it is widely accepted in literature that the human visual system infers the third dimension by elaborating and combining a number of clues and signs contained in the retinal images and extracted from the movements of the eyes. These perceptual units are what psychologists call **depth cues**.

Depth cues are generally classified into two broad categories: *visual cues* and *oculomotor cues* [Sekuler and Blake, 1994]. As their name suggests, oculomotor cues derive from the movements of the parts that constitute the eye, e.g. the contraction of the ocular muscles. On the other hand, visual cues are extracted directly from the images projected onto the retinas by the lens and can be further classified into *monocular cues* and *binocular cues*. In turn, monocular cues can be *static* or *dynamic*, depending if they are available to a stationary observer looking at a motionless scene or they require movements of the observer and/or the object the observer is viewing [Palmer, 1999] [Sekuler and Blake, 1994]. Figure 2.1 illustrates this taxonomy.

The following sections give an overview of the main depth cues and their role in human depth perception. For clarity they are organised into three main groups: oculomotor, monocular and binocular depth cues.

2.2.3 Oculomotor Depth Cues

Oculomotor cues, also sometimes referred to as ocular, optical or physiological cues, are associated with the physical state of the eyes themselves and their components; for example the angle of inclination of the eyes or the level of strain in the ocular muscles. They can

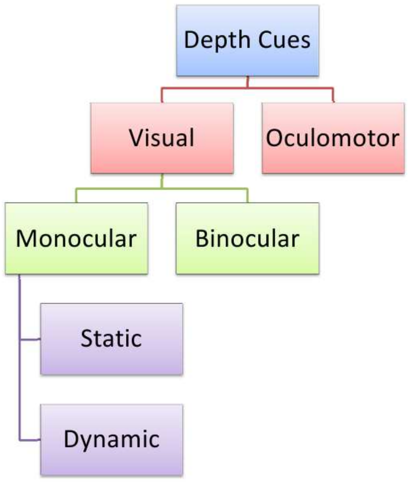


Figure 2.1: Depth cue taxonomy.

provide unambiguous information about the absolute depth of the object fixated by the observer [Sekuler and Blake, 1994], however they can only provide information about a single object at a time [Bruce et al., 2003]. Oculomotor cues to depth include vergence and accommodation.

Vergence Ocular vergence (or convergence) is the angle between the optical axes of the two eyes [Palmer, 1999]. When the observer fixates on an object, the eye balls rotate and in doing so they converge. The degree to which they converge is dictated by the distance of the fixated object; if this is placed at infinity, the optical axes of the observer's eyes are parallel. On a stereoscopic system, stereo images that are not well calibrated can force the observer to diverge their eyes and cause outward rotation. This phenomenon is called *wall-eyed vision* and if too pronounced it can introduce eye strain and headache [McAllister, 1993].

Accommodation Ocular accommodation is the process through which the eye changes its optical power to focus on objects at different viewing distances [Bossong et al., 2009]. As the observer focuses on distant objects, the ciliary muscles stretch the eye lens making it thinner and allowing it to bring the objects into sharp focus on the retina. Conversely, when the observer focuses on near objects the muscles contract making the lens thicker and more powerful. In this way these kinesthetic movements control the amount of accommodation of the eyes. An interesting fact is that accommodation in humans declines with age: most people start having difficulties focussing on near objects (*presbyopia*) during their 40s and by mid-50s the eye's ability to accommodate is lost [Atchison et al., 2008] [Sun et al., 1988]. The stimulus that drives accommodation and triggers an accommodative response by the eye is image blur or **defocus blur**, which can itself be an effective depth cue as explained in detail in section 3.2.5 of the next chapter.

Vergence and accommodation are considered to be relatively weak depth cues and become ineffective at signaling distances beyond 6 to 8 feet (i.e. about 2 metres) [Allison et al., 2009] [Shirley and Marschner, 2005] [Bruce et al., 2003] [Palmer, 1999]; at this distance the ciliary muscles are already at their most relaxed state and vergence angles change very little beyond that [Palmer, 1999]. However, experimental results showed that for shorter distances people rely on them as effective sources of depth information and to compute the perceived size of objects [Palmer, 1999]. Since a typical viewing distance for a desktop *3D display* falls below the threshold of two metres, it is fair to assume that the observer of the artificial 3D scene will make use of vergence and accommodation as effective cues.

In this regard there is another important consideration to be made. In real life accommodation and vergence are usually coupled in the human visual system [Palmer, 1999] [Ware, 1995], so a change in vergence triggers a change in the accommodative state of the eye and vice versa, but **planar 3D displays** introduce a discrepancy between these two cues [Bowman et al., 2005] [Ware, 1995], as further discussed in section 2.3.4.

2.2.4 Monocular Depth Cues

Monocular cues are 2D sources of information that can be perceived through the input of one eye only, hence their name, and are usually experiential [Holliman, 2006]. They are often called *pictorial cues* because several of them were discovered by the Renaissance artists in the attempt to reproduce depth in their paintings [Sekuler and Blake, 1994]. Raphael made superb use of them in his famous work *The School of Athens* (Figure 2.2); remarkably, this painting provides a good impression of the third dimension despite being drawn on a flat surface.

Figure 2.2: *The School of Athens* by Raffaello Sanzio, 1509-1511.
[From http://en.wikipedia.org/wiki/The_School_of_Athens, last accessed: 10.07.2010]

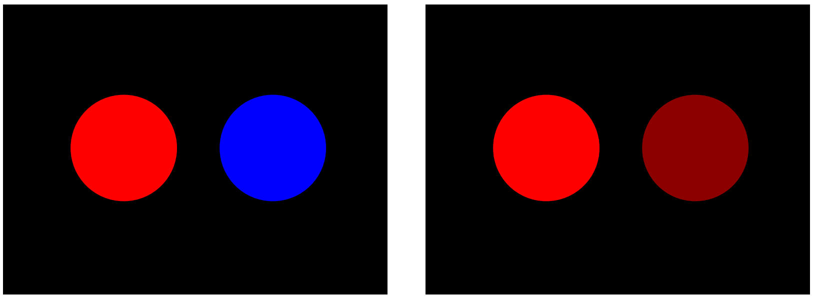
Monocular cues to depth differ in their nature: some of them are based on the geometry of the scene, while others relate to illumination or atmospheric conditions [Sekuler and Blake, 1994]. Besides this, they can be static or dynamic. A background overview of the most common monocular cues is given here based on [Ware, 2008] [Bowman et al., 2005]

[May and Badcock, 2002] [Palmer, 1999], while a more contextualised discussion on depth cues can be found in Chapter 3.

Motion parallax It is the only monocular dynamic cue and it occurs, for instance, when looking out of a car window whilst the car is moving: near objects appear to pass by rapidly while far objects appear to move at a much slower speed but in the same direction as the car. This apparent relative motion of the observed objects whilst the observer is moving is defined as motion parallax and it can give a strong hint about the relative depth of objects.

Colour The perceived colour of an object derives from the physical characteristics of the light it scatters and the way this interacts with the photoreceptors in the human eye; however it also highly depends on context, i.e. the physical properties occurring elsewhere in the visual field [Knoblauch and Shevell, 2003]. Cues to depth deriving from colour can be strong and it has been demonstrated that under some circumstances they can even override geometric cues [Guibal and Dresp, 2004]. In this regard, there are two interesting concepts concerning the perception of colour. First, objects that have the same shape and size and are placed at the same relative distance can be perceived by the observer as at different distances because of their difference in *hue*; this phenomenon is called *colour stereoscopic effect* [Vos, 1966] or simply *chromostereopsis* [Allen and Rubin, 1981] and it is particularly vivid when, for example, red and blue objects are placed adjacent to each other against a black background. Second, bright objects usually appear to be closer to the observer than dark objects, as an effect of **luminance contrast** simulating the principles of **atmospheric perspective** [Guibal and Dresp, 2004]. These two visual phenomena are shown in Figure 2.3 (a) and (b) respectively. A more in depth review of the visual properties of colour is given in section 3.2.3 of the next chapter.

Occlusion When two objects overlap, the principle of occlusion suggests that the oc-



(a) (b)

Figure 2.3: Properties of colour as depth cue. In (a) the red and blue circles appear to be at different depths, while in (b) the bright red circle appears to be closer than the dark red circle.

cluded object lies behind the occluding one. Occlusion is a very strong depth cue and it can even prevail over **retinal disparity** in situations where the two cues are in conflict [Kaufman, 1974]. It is important to note that occlusion only allows understanding of relative depth of objects but does not provide information about their absolute distance to the observer. Occlusion and its effect on depth perception is thoroughly discussed in section 3.2.6.

Size The size of the image of an object on the retina is inversely proportional to the object's relative distance to the observer. Hence, as an object moves away from the observer, its retinal image becomes smaller. If the observer is familiar with the size of the observed object then they can judge its distance by its retinal image. Size can be a good cue to distance in absence of other cues [Ittelson, 1951].

Perspective The term perspective refers to the change of the appearance of the objects as their distance to the observer varies. From a perceptual point of view there are at least three ways in which perspective can be used as a cue to depth.

- **Linear perspective** As objects recede from the observer, their apparent size becomes smaller and their position with respect to the foreground and the background are established by parallel lines converging on the horizon at infinity; this principle can be seen in the way railway tracks appear to converge at a distant point. Linear perspective is a strong depth cue and in some cases it can even contradict the depth information provided by retinal disparity [Stevens and Brookes, 1988]. The geometry of linear perspective was originally elaborated by Leonardo Da Vinci and other Renaissance artists during the fifteenth century [Sekuler and Blake, 1994].

- **Texture gradient** When a uniformly textured surface is projected onto a plane, for example when looking at a field covered by gravel, the texture elements become smaller, denser and undefined as the distance to the observer increases.
- **Size gradient** More distant objects to the observer appear to be smaller than nearer objects with a similar physical size.

Atmospheric perspective Because of light scattering due to vapor and particles in the atmosphere, near objects are sharp and clear, while far objects appear to be hazy and bluish in colour. This phenomenon goes under the name of atmospheric perspective or aerial perspective and can lead to different perceptual experiences depending on the atmospheric conditions. The use of aerial perspective as a cue to depth is further discussed in Chapter 3, sections 3.2.1 and 3.2.5.

Depth from focus and Depth of focus The lens of the eye can adapt its shape to bring the image of objects at different distances into sharp focus on the *fovea*, i.e. the region of the retina associated with the most acute vision. As a result, objects that are at a different distance than the fixated object become blurry, even though the **depth of focus** varies with distance [Ware, 2004]. Blur and depth of focus are thoroughly discussed in section 3.2.5 of the next chapter.

Lighting and Shadowing Shadows can be a strong visual cue. The way in which an object is shaded provides information about its shape and position relative to the light source. Cast shadows give cues about the height of an object and also about the relative position and size between objects in a scene [Ware, 2004].

2.2.5 Binocular Depth Cues and Stereopsis

Stereopsis can be defined as the perception of relative depth that derives from binocular vision [Sekuler and Blake, 1994] and it is by definition a binocular cue, i.e. it requires the input from both eyes to be perceived. In humans, stereopsis is possible because the eyes are laterally separated and their fields of view overlap. When looking at objects in a three-dimensional space, the eyes perceive two slightly different views of the same scene. As a result, the two retinal images of the objects in the overlapping portion of the scene have a lateral displacement; in technical terms this displacement is called **retinal disparity** [Palmer, 1999]. The human brain has the ability to fuse the two images into a single 3D image and process retinal disparity in order to extract three-dimensional information about the scene, giving a sensation of depth [Poggio and Poggio, 1984]. This concept is illustrated in Figure 2.4.

Stereopsis is a particularly strong depth cue and, according to Poggio and Poggio [1984], one of the most important and accurate. By processing retinal disparity information humans can make extremely fine depth judgements that in some circumstances would be impossible when using only one eye [Sekuler and Blake, 1994]. Previous studies have demonstrated that humans can detect depth difference induced by image disparities of only a few seconds of arc [Howard, 1919], in some subjects as small as 2 seconds of arc [Berry, 1948]. In a recent study Allison et al. [2009] have shown that stereopsis can support depth discriminations beyond typical laboratory distances (i.e. 1 to 6 metres) and significantly improve the precision and the accuracy of depth estimates up to 18 metres. However,

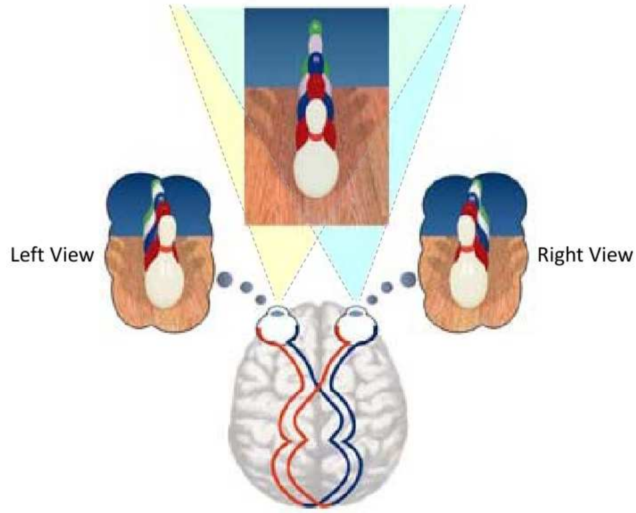


Figure 2.4: The concept of stereopsis. The human brain merges the two 2D images perceived by the eyes into one 3D image. By processing their relative binocular disparity it extracts three-dimensional information about the scene, enabling judgement of the depth of objects and their relative distance.
[Adapted from http://www.vision3d.com/stereo.html, last accessed: 13.07.2010]

according to Richards [1970], nearly 2.7% of the human population posses no stereopsis in one of the two hemisphere (i.e. portion of the visual field to the left or right of the fixation point) and as much as 20% presents a degree of stereo-anomaly. As explained in section 2.3.1 of this chapter, stereopsis represents the key concept upon which conventional 3D displays base their design.

In binocular vision science, there are a number of terms and concepts that are used in relation to stereopsis and stereoscopic performance. Stereoscopic acuity or **Stereoacuity** is an indication of how well a person can perceive depth based on disparity information and is defined as the smallest depth difference that can be detected binocularly. Another important concept is the one of *fusion limits*, which relates to the largest binocular disparity for which a binocular visual stimulus can be fused and perceived as a single 3D image (single vision). In the literature, a distinction is often made between *horizontal fu-*

sion limits, sometimes referred to as *maximum detectable depth*, and *vertical fusion limits*, depending if the focus is on horizontal disparities or vertical disparities.

Fusion of images, and hence single vision, occurs in a limited portion of the visual space, around where the eyes are fixating. This region is defined as *Panum's fusional area* and is centered on the *horopter*, i.e. the set of points in space whose image physically falls on corresponding points in the two retinas. Within the Panum's fusional area, objects that are located nearer or further away than the horopter are perceived as a single 3D image; outside of this region double vision occurs.

Stereoacuity and fusion limits (i.e. maximum detectable depth) are often measured in terms of lower and upper *disparity thresholds. Lower disparity threshold* is defined as the the minimum amount of disparity required so as for the observer to be able to perceive depth, while *upper disparity threshold* refers to the maximum amount of disparity that can be reliably perceived as depth before single vision beaks down and double vision occurs. Lower and upper disparity thresholds are sometimes referred to as *lower* and *upper disparity limits* respectively.

2.2.6 Depth Cues Integration

As explained in the previous sections, the human visual system uses various information sources, i.e. depth cues, to extract depth information and compute the distance and layout of objects in the surrounding space. Individual depth cues can be ambiguous, but when they are combined a stable, unambiguous depth map of the observed scene can be created [Bruce et al., 2003] [Sekuler and Blake, 1994].

In the last forty years numerous studies have been conducted to investigate how our visual system combines this variety of information sources in order to achieve an unambiguous solution. Marr [1982] initially proposed that we represent the world around us using what he defined as *the 2.5D sketch*, an observer-centred view of the scene that combines the information provided by the different depth cues. The vast majority of the studies in the area support this theory. Some of them suggest that depth cues are combined in an additive fashion following a linear model, where the perceived depth is given by a weighted average of the depth information contained in every single cue [Johnston et al., 1994] [Johnston et al., 1993] [Young et al., 1993] [Bruno and Cutting, 1988]. Others provide some evidence that the combination of depth cues does not always follow a linear, additional model and that cues can sometimes interact with each other in a non-linear fashion [Bradshaw and Rogers, 1996] [Reinhart, 1991]. A more recent study suggests that the cue integration system in humans presents a degree of plasticity, that is the weighting of individual cues may vary depending on the context defined by the visual stimuli and the relevance of the specific cue in that context [Harwerth et al., 1998]. Regardless of what the exact combination rules are, the point that all these studies seem to imply is that the presence of multiple depth cues helps to disambiguate the visual stimuli and the more depth cues, the better the depth perception of the observer.

Another important fact that emerges from the reviewed literature is that cue conflict can degrade depth perception, which in turn suggests that no single source of information is predominant. Natural scenes do not usually present strong cue discrepancies. However, when generating content for stereoscopic 3D systems care should be taken to ensure that the depth cues contained in the **stereoscopic images** are concordant with each other in order to avoid ambiguous stimuli and a poor 3D experience for the user.

2.3 Electronic 3D Display Systems

2.3.1 Introduction to 3D Technologies

In order to deliver the 3D effect, conventional 3D displays exploit the principle of stereopsis. This means that when showing a stereoscopic image, these displays must provide a mechanism that allows the left eye to see only the left view and the right eye to see only the right view at a certain moment in time. There are several ways to achieve this and since the invention of the first **stereoscope** by Sir Charles Wheatstone in 1838 many solutions have been proposed resulting in a range of different 3D display systems [Holliman, 2006]. The following sections give an overview of the main optical designs for 3D display technologies.

2.3.2 Stereoscopic versus Autostereoscopic Systems

In the literature, 3D displays are often divided into two main categories: stereoscopic and autostereoscopic. Stereoscopic systems require the user to wear a device, usually headgear or glasses, that ensures the left and right views are perceived correctly by the eyes. Examples of this 3D displays family are the RealD StereoGraphics CrystalEyes® Workstation, which uses Liquid Crystal (LC) shutter glasses (Figure 2.5 a), and the Miracube (ex Pavonine-Dimen) S-series displays that instead require the user to wear polarised glasses (Figure 2.5 b).

On the other hand, autostereoscopic displays automatically direct the left and right views to the correct eye so the user is not required to wear any special device. The Dimension Technologies Inc (DTI) Virtual Window® 19 (Figure 2.6 a) and the SeeReal C-series displays (Figure 2.6 b) are examples of autostereoscopic 3D displays.

2.3.3 3D Display Taxonomy

There are several technical approaches that can be adopted in order to provide the user with a binocular image. The following sections give an overview of the 3D display technologies available on the market at the time of writing this thesis, based on the different types of 3D experience they provide to the user. Given the scope of this research, the survey mainly focuses on desktop 3D displays with two views.

a **b**

Figure 2.5: Examples of stereoscopic 3D displays: RealD StereoGraphics CrystalEyes4 LC shutter glasses (a) and Miracube G170S display with required polarised glasses (b).
[From http://www.inition.co.uk, last accessed: 11.07.2011]

a **b**

Figure 2.6: Examples of autostereoscopic 3D displays: DTI Virtual Window 19 (a) and SeeReal C-i display with head tracker (b).
[From http://www.inition.co.uk, last accessed: 11.07.2011].

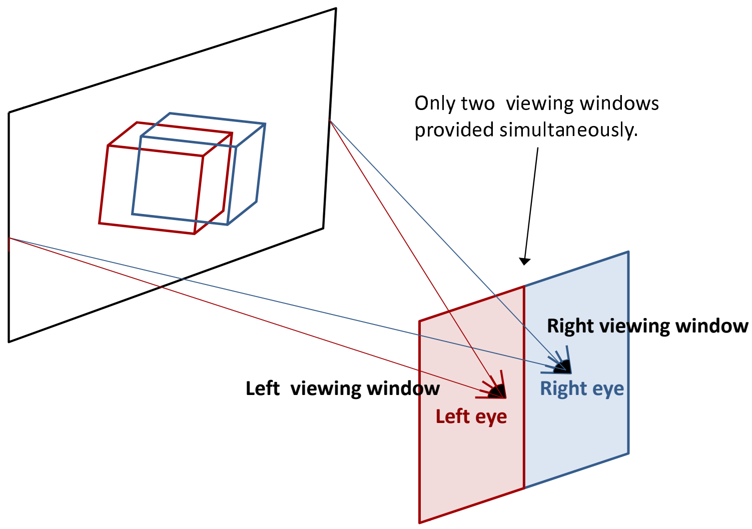


Figure 2.7: Two-view 3D display. The two views are shown in two separate viewing windows simultaneously so the user can perceive the stereo effect and see the image in 3D. [Adapted from [Holliman, 2006]]

Two-View Displays

Two-view 3D displays can generate only two views (i.e. images) at a time, one for the left eye and one for the right eye. These views are displayed in two different sets of pixels on the display screen and are visible through two separate **viewing windows** or **viewing zones** [Holliman, 2006], as shown in Figure 2.7. As long as the eyes of the user stay within the relative viewing window, the left and right views can be seen and fused, and a 3D representation of the scene can be perceived.

The width of the viewing windows, and hence the freedom of lateral movement, depends on the design of the display. On autostereoscopic two-view displays this is usually quite limited (about 2 to 3 cm) and in order to experience the 3D effect the user has to maintain a central viewing position. To counter this limitation, some autostereoscopic two-view displays incorporate a *tracking & steering mechanism* that can detect the position of the head of the user and adjust the viewing windows accordingly. Alternatively, the two views can be displayed using a stereoscopic solution, which allows more freedom of movement to the user but requires the use of spectacles or head gear. A description of the main optical designs adopted for two-view displays currently available on the market is presented below.

Twin-LCD Autostereoscopic 3D Displays This type of display generates two images using two separate **Liquid Crystal Display (LCD)** elements, one LCD per image, and then directs one image to the left eye and the other image to the right eye. This approach has two main advantages. First, the user is provided with a full resolution 3D image. Second, twin-LCD systems present little or no **crosstalk**, i.e. there is no interference between the left and right view, which means that the output 3D image is usually crisp and free from disturbing artifacts. On the other hand, this type of display is suitable for single users only and can have quite high implementation costs due to the twin-LCD nature and the often sophisticated optical design. The Iris-3D [McKay et al., 2007] [McKay, 2005] [McKay et al., 1999] and the Kodak 3D Stereo Display [Cobb, 2005] are two examples of two-view, twin-LCD autostereoscopic displays.

Single-LCD Autostereoscopic 3D Displays Similar to the previous category, two-view single-LCD displays generate two images simultaneously but in this instance they only employ one LCD element. In order to correctly direct the two views to the eyes, these displays can use a range of micro-optical elements like parallax barriers, lenticular optics or micropolarisers [Holliman, 2006]. The micro-optics, in combination with the LCD unit, enable the user to see both views simultaneously and experience the 3D effect without the need to wear any special device. Compared to twin-LCD displays, single-LCD displays benefit from considerably lower implementation costs, however they also deliver inferior image quality [Holliman, 2006]. In particular, the quality of the final 3D image can be compromised by visual artifacts introduced by the micro-optics (e.g. crosstalk and ghosting) and by the fact that the left and right images only have half of the resolution, either vertically or horizontally, of the original stereo pair. Finally, single-LCD autostereoscopic displays are usually suited for a single user at a time due to the limited viewing freedom. Many of the 3D displays that have been released to the market belong to this category. The Sharp LL-151-3D (desktop solution) and the Sharp Actius AL3DU (laptop solution) are examples of single-LCD autostereoscopic displays implemented using parallax barrier technology [Montgomery et al., 2001a] [Woodgate et al., 2000a] [Woodgate et al., 2000b]. Lenticular elements are the key to the optical design of the SeeReal [Schwerdtner and Heidrich, 1998b] [Schwerdtner and Heidrich, 1998a] and the DTI Virtual Window® autostereoscopic display ranges [Eichenlaub and Gruhlke, 1999] [Eichenlaub, 1997], while a micropolariser-based technology was invented and successfully employed by VRex [McAllister, 2005] [Faris, 2001] [Faris, 1994]. Notwithstanding, the majority of the VRex commercial displays are stereoscopic solutions (see below).

Single-LCD Stereoscopic 3D Displays This class of displays is similar in principle to the single-LCD autostereoscopic class in that it uses a single LCD display in combination with micro-optical elements in order to generate two views and direct them to the user's eyes correctly. Unlike the former, these are stereoscopic systems and as such they require the user to wear a special device. Usually this consists of polarised glasses [Holliman, 2006], either linearly polarised or circularly polarised, which enable the user to see the 3D image by working in conjunction with a method of polarising the two views in the display optics. These 3D displays have the advantage of being suitable for multiple users, but like other stereoscopic solutions they suffer from the drawback that the user is required to wear spectacles. In addition, they have a similar limitation to their autostereoscopic counterpart, i.e. visual artifacts introduced by the micro-optics, such as crosstalk and ghosting, as well as left and right images that only have half of the resolution of the original stereo pair. The

Miracube G170 and the VRex VR-FPD3O are two commercial examples of single-LCD stereoscopic displays that use micropolariser optics.

Time-Multiplexed Stereoscopic 3D Displays Time-multiplexed displays, also known as field-sequential displays, present the left and right views on the screen in sequence and use optical techniques to occlude the right eye when the left view is shown and vice versa [McAllister, 1993]. These displays are stereoscopic systems and, as such, they usually require the user to wear glasses, either active shutter glasses or passive polarised glasses [Bos and Haven, 1989]. They provide full resolution images for both eyes, but given their field-sequential nature the effective video frame rate per eye is half of the screen nominal one. For this reason, they are usually driven by screens with high refresh frequency, in order to avoid disturbing flickering effects. RealD StereoGraphics CrystalEyes® are an example of LC active shutter glasses that can be used to visualize 3D images on a normal desktop station display [Lipton and Ackerman, 1990].

Multi-View Displays

Unlike two-view displays, multi-view displays generate multiple views (i.e. more than two) simultaneously [Holliman, 2006], as shown in Figure 2.8, but, depending on their viewing position, the user can see only two out of these many views at a certain moment in time. The views are presented so that the user can always perceive a valid stereo pair irrespective of the position and their number can vary according to the system; some multi-view displays have a fixed amount of views while with others the number of views is programmable. If enough horizontal viewing freedom is available, multi-view systems can support multiple users.

Multi-view display solutions can be implemented using either a temporally multiplexing approach or a spatially multiplexing approach. Mixed designs that use spatiotemporal

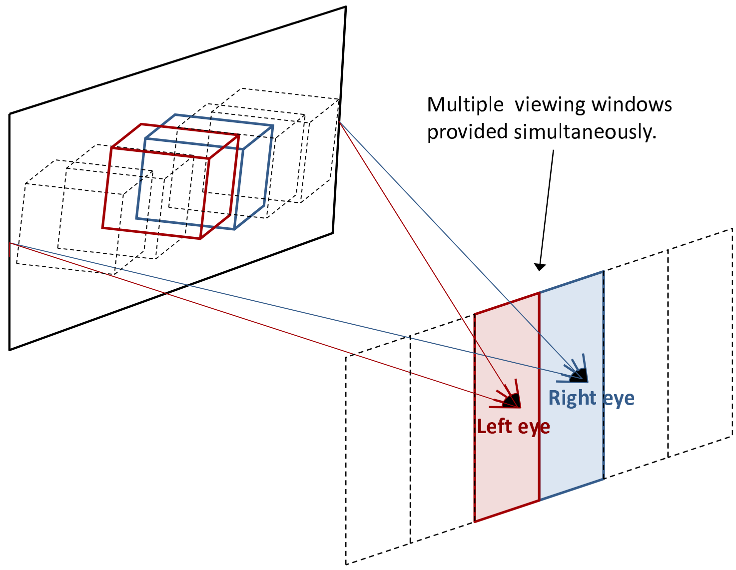


Figure 2.8: Multi-view 3D display. Multiple views are shown simultaneously in multiple viewing windows, providing a range of viewing positions to one or more users at a time. [Adapted from [Holliman, 2006]]

multiplexing have been adopted as a solution for large scale 3D display projectors [Jang et al., 2004], but these systems are out of the scope of this thesis.

Temporally multiplexed multi-view displays use the same principle as field-sequential two-view displays (see section 2.3.3, page 27); that is they show the multiple viewing window images sequentially in time but at a faster refresh frequency than the human eye *flicker fusion threshold*, so they appear as they were shown simultaneously. This type of display has the advantage that each view can be shown with full screen resolution, but suffers the downside of requiring extremely fast display elements and broadband circuits [Holliman, 2006], and therefore can be expensive to manufacture. On the other hand, space-multiplexed systems divide the display screen into several pixel subsets that are directed to different viewing windows and can be used to show the different images truly simultaneously. The main advantages of these displays are the great viewing freedom they can provide and the fact that they can be manufactured at relatively low cost, using standard display and optics elements. However, compared to two-view displays and to their time-multiplexed counterpart, they have considerably lower resolution per view with a consequent loss of quality and detail in the final 3D image.

Most of the commercial multi-view displays are designed to use spatial multiplexing, mainly due to the lower implementation costs. The Spatial View SVI MU series and the Magnetic Enabl3D^{TM} series are only two examples of the new generation of autostereoscopic multi-view displays available on the market. An example of time-multiplexed autostereoscopic multi-view display is the Cambridge display [Moore et al., 1996].

Volumetric Displays

The type of displays reviewed until now all have the common characteristics of being planar, i.e. they present the output 3D image on a static flat surface. There is a category of 3D displays that adopts a totally different approach and instead renders the 3D image inside a physical 3D volume; these display devices are defined as volumetric or multiplanar displays [McAllister, 2005] [Anderson, 2005]. Volumetric displays employ moving mirrors in combination with other optical and physical devices in order to project or reflect light dots in a defined 3D space above the display device itself. Compared to planar 3D displays they have the advantage that they produce solid 3D images, where accommodation and vergence are coupled, as in the real world [Sullivan, 2005]. Furthermore, they are *full-parallax* systems and they allow multiple users at the same time. However, the current state of the technology only supports a limited number of colours and limited image resolution, which both represent a disadvantage.

The Sony Volumetric 3D and the Actuality Systems Perspecta Spatial 3-D Display 1.9 [Chun et al., 2005] [Favalora, 2002] [Favalora et al., 2002] are two examples of volumetric displays.

Others

Under the umbrella of 3D display technologies there are at least two more categories that are worth mentioning in this review: Head Mounted Displays (HMDs), also known as helmet mounted displays, and holographic displays. HMDs are wearable optic displays that are mounted in front of the operator's eyes and can be monocular or binocular. They are often combined in a single helmet that incorporates earphones for 3D sound and head tracking sensors and if binocular they can support stereoscopic viewing. The first HMD was invented by MIT American professor Ivan Sutherland during the 1960s, but was only used for the first time in real applications during the 1970s in military operations. Since then their use has extended progressively into a variety of fields, such as medicine, education and, more recently, entertainment (e.g. 3D video games) [Wu and Ho, 2005]. Binocular HMDs can be considered a special category of two-view 3D displays, but their technical characteristics and the viewing experience that they offer to the user are totally different to desktop 3D displays.

Holographic displays are optic devices that are capable of reproducing a 3D image from a flat 2D screen by generating a light field identical to the one that would be emitted by the original 3D scene. The result is a hologram, that is a 3D image that appears to "pop out" from the 2D screen surface and contains full horizontal and vertical disparity information across a range of viewing angles. In order to achieve this, these displays employ diffraction-based techniques that exploit the physical properties of coherent light (e.g. laser) [Reichelt et al., 2010]. Similar to volumetric displays, holographic displays do not suffer from the accommodation-convergence decoupling issues that characterise planar 3D displays and can reproduce very high quality 3D images. Nevertheless, their high implementation costs and the fact that they require a large number of mathematical calculations to operate represent a barrier to their commercialisation.

The Sensics HDSight^{TM} [Sensics Inc., 2008] and the SeeReal Sub-Hologram and Viewing Window Technology prototype [Reichelt et al., 2010] are examples of HMDs and holographic displays respectively.

Summary

The previous subsections presented an overview of the different 3D display designs and their different characteristics, both in terms of user experience and underlying technologies used to reproduce the 3D effect. One way to classify 3D displays is to use these two key factors, i.e. user perceptual experience and technology, and categorise the systems based on the amount of parallax the display can reproduce at a specific instant in time and its technical characteristics [Benton, 2001] [N.S. Holliman and Pockett, 2011]. Table 2.1 categorises the 3D display systems reviewed above based on these two criteria. *Two-view* displays offer the lowest degree of parallax as they can produce only two views simultaneously. Intermediate systems, i.e. *horizontal-parallax* displays, can reproduce multiple degrees of parallax simultaneously even though only in the horizontal direction. Finally, *full-parallax* displays can reproduce multiple degrees of parallax both horizontally and vertically and for this reason they are the closest to the real world parallax experience.

Table 2.1: 3D display taxonomy based on the amount of parallax that the display can reproduce simultaneously and its technical characteristics.
[Adapted from [Benton, 2001]]

Two-View			Horizontal-Parallax		Full- Parallax
Stereoscopic	Autostereoscopic	Stereoscopes	Stereoscopic	Autostereoscopic	Autostereoscopic
Single-LCD space-multiplexed with polarised glasses (linear or circular)	Single-LCD with parallax barriers	Wheatstone stereoscope	Multi-view time-multiplexed	Multi-View space-multiplexed	Volumetric
Single-LCD time-multiplexed with active shutter glasses	Single-LCD with lenticular lenses	HMD	Multi-View mixed design		Holographic
Single-LCD time-multiplexed with passive polarised glasses	Twin-LCD (single user)				

2.3.4 Depth Perception on 3D Displays

The main advantage of using a 3D display, compared to a normal 2D display, is the ability to see an image in depth. The perception of depth on a 3D display is largely based on the depth cues theory presented in section 2.2, but it differs from the way we experience depth in the natural world on a number of points.

A first important difference is that planar 3D displays introduce a mismatch between accommodation and vergence. In particular, on such devices the objects that populate the 3D scene always physically lie on the screen plane and therefore are all in focus at the same focal depth [Cruz-Neira et al., 1993]. This means that in order to get a crisp image of the objects, the observer always has to focus (accommodate) at the depth of the screen plane. At the same time though, their eyes are forced to converge according to the virtual depth of the object they are fixating, so as to be able to fuse the two views and see the 3D scene in stereo. When the virtual depth of the object and the physical distance to the screen do not coincide, i.e. when the object is drawn so as to be perceived in front or behind the screen plane, the accommodation-convergence mismatch occurs. If exacerbated this mismatch can lead to visual fatigue and serious discomfort, such as double vision and image blurring, and can even hinder visual performance [Hoffman et al., 2008] [Ware, 1995] [Yeh, 1993].

Visual fatigue may also derive from interocular crosstalk, i.e. the interference of the information contained in the left view channel with the one contained in the right view channel and vice versa. On a planar 3D display there are two situations in which crosstalk may occur [Lipton, 1993]. The first is when the mechanism used to separate the left and right views fails to completely isolate the two channels, so information from one channel leaks into the other. 3D displays based on lenticular technology or other type of micro-optics inherently produce a certain amount of this type of crosstalk and so do displays based on polarisation techniques. Also, when attempting to see a stereoscopic image on an autostereoscopic two-view display from an unsuitable viewing position, leakage between the two eye views occurs and crosstalk is perceived by the user. The other cause of crosstalk is related to the afterglow effect due to the slow phosphor decay rate, which causes residual information of one view to be still visible when the other view is shown on the screen [Lipton, 1987]. This last factor only affects depth perception on field-sequential 3D display systems based on CRT monitors, e.g. first generation shutter glasses, which despite being a mature technology are still widely used.

Crosstalk largely depends on the optical design of the display and it results in perceived ghost images around objects; it can cause visual discomfort and headache and it is known to have a detrimental effect on depth perception [Pastoor, 1995] [Yeh, 1993]. Also, it is well recognised in literature that the ghosting effect due to crosstalk becomes more prominent with increasing contrast and increasing image disparity [Engle, 2008] [Seuntiëns et al., 2005] [Pastoor, 1995] [Lipton, 1993]. This means that under some viewing conditions crosstalk can pose a limit to the amount of depth that can be reproduced and comfortably perceived by users on a 3D display.

Finally, there are a number of other factors that may influence the perception of depth on a 3D display, which under real-world conditions are not generally experienced. These include inter-channel variations in image attributes such as brightness and contrast [Lambooij et al., 2009, Holliman, 2006], and display-related variables, e.g. misalignments between the left and right view and **aliasing**. Aliasing is a very broad term that can be used to describe different artifacts. In signal processing and computer graphics, the term refers to the visual distortions that occur when a sampled signal is badly reconstructed so that the reconstructed image differs from the original one [Harris, 2004], and it can be caused by a number of factors. In the specific context of stereoscopic environments a major cause of aliasing artifacts is the 3D display system used to visualise the 3D scene and its technical characteristics, as it has emerged from the literature reviewed in section

2.4.2. An in-depth discussion of aliasing and other technical factors is presented in the next chapter, section 3.3.

2.4 Depth-Based Tasks and Stereoscopic 3D Technologies

2.4.1 Application Fields

In the last decade the use of stereoscopic 3D technologies have become increasingly popular in several professional fields, mainly due to the benefits that stereoscopic visualisation can provide. Application domains such as medicine, geoscience and bio-chemistry have registered an increasing need to visualise their intrinsically 3D data sets in a real 3D space and interact with them in an efficient and reliable way. In this context, critical decisions are often made by identifying the position in depth of specific objects in the 3D space in order to extract important features from the image. Such tasks are often performed with the aid of a 3D cursor and require accurate and reliable depth judgements; failure to do so can cause data misinterpretation and lead to undesirable outcomes, like an inaccurate structural model (geoscience) or an incorrect diagnosis (medicine).

In ophthalmology, standard techniques for the diagnosis of the diabetic retinopathy and the treatment of glaucoma require the operator to judge the depth and pick anomalous elements in the retinal *fundus images* of patients [Habib et al., 2008] [Habib et al., 2005]. Similar techniques are used in diagnostic radiology for the early detection of occult carcinomas. For example, stereomammography and 3D technologies can offer accurate support to radiologists in locating atypical masses in the breast, thus improving breast cancer detection [Mello-Thoms et al., 2003] [Mello-Thoms et al., 2002] [Goodsitt et al., 2000]. Stereoscopic angiograms have also been employed successfully to extract depth information about blood vessels for the study of vascular disease or the evaluation of other physical entities [Higashida et al., 1988]. Similarly, 3D technologies have proved to be beneficial in assisting other fields of medical imaging, such as bone image segmentation [Falcão and Udupa, 2000] and image-guided Transcranial Magnetic Simulation (TMS) [Neggers et al., 2004].

In structural biology, 3D technologies have been successfully employed to segment tomographic 3D data and improve the interpretation of complex 3D structures. In this context, 3D tools, and in particular a 3D cursor, are often employed to identify and isolate regions of interest in 3D images, which can be subsequently visualised and analysed in detail [Li et al., 1997].

Further applications for stereoscopic 3D technologies can be found in the domain of geoscience. Here, the ability to judge the location in depth of rock features is particularly important and it directly affects the building process of the geological structural model. This is the case for geoscientific studies based on LIght Detection And Ranging (LIDAR) technology, where the interpretation of geological structures, such as faults, is often done by identifying and selecting small elements within the rock outcrop 3D data that result from the LIDAR laser scanning [McCaffrey et al., 2005].

2.4.2 Previous Work on Stereoscopic 3D Cursors

In spite of their disparate background, the tasks carried out in the application fields cited above all have one requirement in common: the need to be able to accurately judge the position in depth of objects in the 3D data set and the ability to interact with them in the 3D space. If the task is to be performed in a desktop stereoscopic 3D environment, it is essential that the cursor used for interaction is also rendered stereoscopically, as pointed out in [Habib et al., 2008] [Jin et al., 2007] [Habib et al., 2005] [Neggers et al., 2004] [Stein and Coquillart, 2000] [Basdogan et al., 2000] [Li et al., 1997].

Interaction in 3D environments is a vast research topic that has received widespread attention in the past twenty years. Numerous studies have investigated different aspects of it, but little work can be considered directly relevant to depth-based tasks performed in stereoscopic environments using a 3D display and a 3D cursor. The aim of the following sections is to present a critical review of the most relevant works under this umbrella, divided into categories of pertinence.

Medical Imaging

Goodsitt et al. [2000] investigated whether the improved depth perception associated with stereomammography could be further enhanced with the use of a virtual 3D cursor. Similarly, Habib et al. [2008, 2005] performed a study where participants were asked to use a stereoscopic 3D cursor to assess the morphology of the optic disc in stereoscopic fundus images. In both studies a stereoscopic 3D cursor was developed by the authors and used by participants in order to interact with the stereoscopic images used during the experiments. The results of these studies show that stereoscopic technologies can be used effectively to support depth-based tasks and interact with the stereo images in order to highlight features of interest. However, they also indicate that under some conditions the characteristics of the cursor used during the experiments might have had a detrimental effect on task accuracy.

These studies do not provide a satisfactory answer to the question of how to design a stereoscopic 3D cursor to support depth-based tasks, but they suggest [Habib et al., 2008] [Habib et al., 2005] and acknowledge [Goodsitt et al., 2000] the fact that more research is needed to investigate the effect of characteristics such as cursor shape and contrast on performance of tasks of this kind.

Computer Graphics and 3D Aided Design

Slater [1992] developed an algorithm to support 3D interaction on a standard workstation without the need of advanced hardware, the focus of the study being on the efficiency of 3D rendering techniques. Despite being motivated by the need to provide a software environment for a new 3D input device used to drive a 3D cursor, the Desktop Bat, this study does not offer any insight on the design of the 3D cursor itself.

Zhai et al. studied the effect of semi-transparency, or partial occlusion, and **binocular disparity** on depth perception in 3D interaction, and introduced a new technique for dynamic 3D target acquisition [Zhai et al., 1996] [Zhai et al., 1994]. The empirical results of their study showed that the semi-transparent cursor, the *Silk Cursor*, led to better accuracy and shorter task completion time for both monoscopic and stereoscopic display modes, and suggest that in terms of depth cues partial occlusion can be as effective as binocular disparity. The novel idea of using semi-transparency as a cue to depth is an important contribution towards a better understanding of how to design a 3D cursor to support depth-based tasks. However Zhai et al. only explored the effect of one of the many factors that could affect performance of such tasks in desktop stereoscopic 3D environments.

Ware and Lowther [1997] investigated the use of a one-eyed 2D cursor for target acquisition and selection in a stereoscopic environment using shutter glasses. For this purpose they performed a *Fitts's Law* [Fitts, 1954] experiment where they compared task performance using the one-eyed 2D cursor, i.e. a cursor presented in only one eye, against a 3D cursor. Their results showed that error rates and movement times were lower for the one-eyed 2D cursor. The authors theorised that this would be the case because, for the specific task tested in their experiment, 2D screen movements are shorter than 3D movements and even where the distance is the same, selecting a 2D target generally requires less precision in fewer dimensions than selecting a 3D target. The study of Ware and Lowther introduces an interesting approach to support interaction tasks in stereoscopic 3D environments and it suggests that a one-eyed 2D cursor can successfully be used to perform a 3D task. Nevertheless, the use of such a cursor in a stereoscopic environment raises a number of perceptual issues, especially in terms of discomfort due to prolonged use and **binocular rivalry** [Howard, 2002b]. That is, when the eyes are presented with two conflicting stimuli, as in the study of Ware and Lowther, one stimulus is suppressed at any one time and therefore it can not be seen by the observer. Furthermore, users with pronounced ocular dominance might experience discomfort if the cursor is presented in their non-dominant eye. For these reasons it is believed that a one-eyed 2D cursor might not be an effective and suitable technique to support 3D tasks in stereoscopic environments for a prolonged time.

Stein and Coquillart [2000] investigated the problem of positioning 3D points within a geometric modeling system and proposed the *Metric Cursor*, a 3D cursor designed to support positioning tasks in such environments. In their work the authors briefly discuss various aspects of the problem, including cursor movement (how to move the cursor within the 3D space and which input device to use to drive it), cursor behaviour (how the cursor can be used in order to complete the task) and cursor appearance (cursor shape and other visual characteristics that can be used in order to enhance depth perception). However, they do not give an in-depth analysis of depth cues and visual factors that can affect depth perception and do not justify their choice of using a crosshair as a cursor shape. Furthermore their focus is on a 3D working environment shown on a conventional 2D screen, not via a stereoscopic 3D screen. Similarly, van Overveld [1989] designed and presented a flexible 3D cursor that could be employed as a 3D locator device or as a geometric construction tool in the context of 3D geometrical design on conventional 2D displays. In this study, the author focused on the implementation aspects of the 3D cursor and did not consider the implications of the visual and technical factors related to working in a stereoscopic 3D environment.

Butts and McAllister [1988] developed a real-time stereoscopic cursor prototype and investigated its use for interaction with 3D models in a stereoscopic environment using shutter glasses. The goal of their study was to evaluate the use of this cursor as a stereoscopic depth measuring device and as a means to draw and manipulate 3D objects, and to examine the perceptual issues related to such operations. Based on their observations, the authors identify a number of interesting perceptual issues that should be considered when designing and implementing a stereoscopic 3D cursor. These include: the possibility of using a reference cube around the cursor to enhance depth judgement (i.e. association of stereopsis with other depth cues), the picket fence problem due to aliasing, the *perceptual zooming* effect experienced when size constancy breaks down (see section 3.2.4 for details) and a number of perceptual issues due to the limitations of the hardware used at the time, which no longer constitute a problem.

Barham and McAllister [1991] presented an interactive stereoscopic system to enable users to draw and modify B-Splines curves, and investigated different types of stereoscopic cursors in order to support such tasks. During the comparative experiment the number of errors in selecting a target curve were counted and participants were asked to give a subjective ranking of the cursors used on a scale of 1 (worst) to 5 (best). The results showed that a shaded circle had the highest rank combined with a 0% task performance error rate. The crosshair (a two-dimensional shape given by two intersecting lines forming a "+" shape) and the jack (a three-dimensional shape given by three intersecting lines forming a six-pointed star) both ranked highly, mainly for the extra sense of depth they generated; however they also presented some drawbacks, namely fusion problems and distorted perception of the depth of the hotpoint of the cursor due to the reduced thickness of the lines that form them. The authors suggest that a solution to this problem could be the combination of the jack or crosshair with a circle placed at the hotpoint of the cursor. The triangle and pyramid were ranked average and appeared to be efficient only when their size was big enough to obscure the target points. The square could be used effectively when it was scaled to the size of the target point, while the tri-axis had a low rank and caused the highest error rate. Generally, participants showed a preference for cursors characterised by a size similar to one of the target points and a colour that would provide a good level of contrast against the target points; cursor shapes that ranked better generally had good interposition properties. The main outcomes of the study of Barham and McAllister, which is probably the most relevant to the objectives of this thesis, are that the nature of the cursor and its visual characteristics are critical factors that can severely affect task performance, and that the choice of the cursor depends on the specific task to be performed.

Carver and McAllister [1991] investigated human interface design issues for interactive stereo drawing applications, including stereo cursor development and presented *3-D Draw*, a 3D drawing application designed specifically to work on a DTI autostereoscopic display that includes a 3D cursor functionality. In a similar context, Wright and McAllister [1993] presented an application to create and manipulate Bezier surfaces in stereo using a stereoscopic field-sequential display, where the stereo manipulation was instead accomplished via a three-button mouse and a monoscopic 2D cursor.

The studies presented by McAllister et al. provide a number of valuable qualitative inputs to the question of how a 3D cursor should be designed in order to support depth-based tasks in desktop stereoscopic 3D environments. However, the conclusions are mainly based on what the authors observed and on the subjective feedback and personal preference of participants; no empirical data and statistical analysis was presented in the above reviewed papers. Another drawback in their work is the fact that the perceptual issues observed during the experiments and the suggested potential solutions are specific to the particular 3D display used for the study. One final important point is that these studies highlight the limitations of using a 2D cursor for tasks such as picking and selecting objects in a stereoscopic 3D environment and further reinforce the need to use a stereoscopic 3D cursor instead.

In a recent study, Azari et al. [2010] investigated the implementation of a stereoscopic 3D cursor as the natural extension of the conventional 2D cursor and performed two evaluative experiments . Even though this idea is valuable and in line with the objectives of this thesis, the study of Azari et al. analyses the problem from a different perspective and does not consider the effect that the characteristics of the different 3D display technologies and the visual attributes of the cursor could have on the performance of depth-based visual tasks. Besides, the experiments performed by the authors focus more on the benefit of using stereoscopic 3D technologies as compared to 2D technologies and are characterised by non-repeatability and weak scientific methodology.

Virtual Reality, Augmented Reality and Interaction Techniques

During the last decade new approaches have been presented in order to improve interaction in VR, Augmented Reality (AR) and immersive 3D visualisation environments. The majority of the studies in this field focus primarily on the functionalities of the virtual environment, the level of immersion of the user while performing the task and the degree of interactivity offered by the system; this is the case for the interactive desktop environments presented by Alpaslan and Sawchuk [2006, 2004] and Turner et al. [1996] and for the X-Rooms™ system proposed by Isakovic et al. [2002]. In these studies, the authors analyse the problem of supporting interaction tasks in a stereoscopic environment from a totally different perspective than the one adopted for this thesis and do not offer any insight into how a 3D cursor for such environments should be designed.

Similarly, Siegl et al. [Siegl et al., 2007] [Siegl and Pinz, 2004] and Biocca et al. [2006a,b] investigated novel ways to interact with objects in a stereoscopic AR environment using a 3D cursor. However, the focus of their studies is on the usability and augmentation aspect of the system itself [Siegl et al., 2007] [Siegl and Pinz, 2004] and the interaction technique interface [Biocca et al., 2006a] [Biocca et al., 2006b] rather than the nature of the cursor used to perform the task. The most recent study of Siegl et al. [2007] once again suggests that the characteristics of the 3D cursor could have had an effect on the performance of the object localisation task, which consisted of determining the position and size of objects in the augmented 3D scene used during their system evaluation experiment. The authors briefly discuss and acknowledge this fact, but they do not offer any scientific explanation for how these factors could affect task performance, nor how the design of the 3D cursor should be improved.

Takaki et al. [2005] looked at direct manipulation applied to virtual environments and developed a new interface for 3D object manipulation controlled by the user's gaze and posture. In a similar context, Yoshimura et al. [1994] and Venolia [1993] studied 3D direct manipulation techniques for object manipulation and interaction through a 3D cursor. None of these studies utilises stereoscopic images to visualise the 3D scene. Besides, even when a 3D cursor is employed to interact with the 3D scene, the focus of the research is not on the cursor's visual attributes and the effect that these might have on task performance but rather on the evaluation of the interactive environment as a whole, which the specific 3D cursor design adopted by the authors is part of. The only contribution towards the aim of this thesis is offered by Yoshimura et al. who suggest that for the specific selection task adopted in their experiment a **beam cursor** [Yin and Ren, 2006] is more effective than a beam-less cursor; however the methodology adopted to perform the comparison lacks a robust experimental design and therefore the results have limited scientific validity.

Grossman and Balakrishnan [2004] investigated pointing tasks on a 3D volumetric display where the target size varies in the three spatial dimensions. In particular, they studied the effect of target size, movement amplitude and movement angle necessary to perform the task, and the interaction between the three on task performance. Furthermore, they proposed and evaluated a predictive theoretical model to characterise this 3D pointing behavior - a modified version of Fitts's law. However interesting, the results presented by Grossman and Balakrishnan are only valid for volumetric displays and can not be extended to planar 3D displays. Furthermore, their validity might be limited due to the weak scientific methodology adopted during the experiment, i.e. potentially insufficient repetitions for the high number of experimental conditions tested and statistical analysis performed on an arbitrary subset of the collected empirical data. In a separate study Grossman and Balakrishnan [2006b] explored different 3D selection techniques for volumetric displays in a sparse target environment. Based on their experimental results the authors suggested a number of improvements that could be applied to the *ray casting* technique [Roth, 1982] in order to successfully support object selection on this type of display. Similarly to their previous work, this study by Grossman and Balakrishnan suffers from the methodological drawback that the statistical analysis is applied to an arbitrary subset of the collected data. Besides, the authors' suggested techniques are only valid for volumetric 3D displays and can not be directly extended to any category of planar 3D displays, nor used as a foundation to base the design of a stereoscopic 3D cursor to support depth-based tasks.

In a more recent study Vanacken et al. [2009] presented a number of guidelines to improve existing interaction techniques in order to support object selection in stereoscopic 3D environments characterised by high target density and restricted target visibility. Based on these, the authors proposed and evaluated new forms of the bubble cursor and ray casting technique augmented by sensory feedback. The empirical results of their evaluative experiments suggest that visual feedback had a beneficial effect on the selection task performance, while both auditory and haptic feedback seemed to have no effect. Interestingly, participants expressed a liking for the auditory feedback and, on the other hand, a dislike for haptic feedback. Generally, the study presented by Vanacken et al. provides some valuable input for the aim of this thesis, as discussed in section 3.4. Nevertheless it also has two major drawbacks. Firstly, the complexity of the experimental design used for the evaluative experiments could have potentially compromised the reliability of the empirical results. Secondly, and most importantly, the results are limited to shutter glasses 3D displays, the technology used for stereo viewing during the experiments. In fact the authors did not consider the 3D cursor design issues related to other types of 3D display technologies nor did they discuss how different visual factors and depth cues could affect the perception of depth, and therefore task performance, in stereoscopic 3D environments.

In the context of VR and AR, the most relevant studies for the purpose of this thesis are perhaps the ones conducted by Hou [2001] and Ty [2001], the latter being a refined version of the first, and by Sands et al. [2004]. Hou and Ty studied 3D cursor alignment tasks in stereoscopic AR and investigated the effect of target surface texture, target position, 3D cursor shape and binocular disparity on task performance using shutter glasses. The results show that the the target surface texture density and the target position had an effect on placement accuracy, with participants performing better with highly textured surfaces and when the target point was placed in the centre of the stimuli. Statistically, binocular disparity had also an effect on accuracy, but in practice the effect was caused by other factors, which were not very well controlled during the experiment. Regarding cursor shape, the statistical analysis shows that this factor had no significant influence on task accuracy. However, the subjective results indicate that candidates preferred a 3D-volumetric cursor to a 2D-areal or a 1D-linear one, mainly because the volumetric cursor offers more features along X, Y and Z axes, and low texture surfaces, as they are easier to fuse. Overall the volume cursor was the most favourable one, while the line cursor was the least favourable.

Sands et al. [2004] studied selection tasks in 3D AR using a 3D cursor and investigated the potential benefit of adding synthetic sources of depth information to the cursor and the AR environment. In particular, the authors performed two experiments where they assessed the effect of a number of cues (shadowing, texture, scalar and coordinate information), cursor shape (cone versus crosshair) and type of display (stereoscopic HMD versus conventional monoscopic display) on performance of a target alignment task. Their results are in line with previous studies and suggest that consistent cues have a positive, additive effect on performance, with certain cues having a stronger effect than others and the combination of all cues being the most effective. On the other hand, conflicting cues had a detrimental effect on performance. Coordinate information and shadowing appeared to have the most significant effect individually and seemed to outweigh the benefits of stereoscopic viewing. Regarding the latter, the stereoscopic condition gave better performance than the monoscopic one, however participants expressed a preference for the monoscopic 2D display. Although very interesting and relevant, the study of Sands et al. does not seem to follow a rigorous scientific methodology and does not consider the effect of other relevant visual cues (e.g. size and occlusion) and the more technical factors related to the type of 3D display used. Furthermore, the authors' main focus is on AR environments and the issue of guaranteeing that the synthetic depth cues associated with the cursor are in agreement with the natural image of the AR. Nevertheless, this study still represents a valid starting point for the research presented in this thesis.

Perception and Other Relevant Studies

A number of authors have presented valuable material on dealing with stereoscopic systems, even though their studies do not directly relate to the design of a stereoscopic 3D cursor.

Hsu et al. [1996] discussed important issues that should be considered when conducting studies that aim at evaluating stereoscopic technologies. Among these: viewing conditions, subject stereoacuity, visual artifacts like flickering and ghosting, stereoscopic image generation, task nature and difficulty. Under the same umbrella, Hodges [1991] presented an overview of salient factors that might affect the generation and the display of stereoscopic images, and that should be considered when developing stereoscopic software. These are: depth geometry and mapping, parallax and camera separation, crosstalk, display refresh rate in case field-sequential technologies are employed to display the stereoscopic image, interaction device, image scaling perceived depth and viewing position.

Reinhart et al. looked at the effect of both monocular and binocular depth cues on performance for a number of depth-judgement tasks using a field-sequential stereoscopic display [Reinhart, 1991] [Reinhart et al., 1990] [Reinhart, 1990]. With specific regard to the cursor positioning task, the results suggested that depth cues have an interactive, rather than additive, effect on task performance [Reinhart, 1991]. Positioning errors on the Z-axis were significantly smaller in the presence of each of three tested depth cues (i.e. **luminance**, size and binocular disparity), but the largest performance benefit occurred from the addition of the first depth cue, with diminishing returns associated with the addition of the second and third cues.

Drascic and Milgram [1991] investigated cursor positioning performance in a stereoscopic remote environment using field-sequential technologies. The aim of the study was to combine a stereoscopic video with a computer generated stereoscopic cursor as a tool for measuring distances in a real remote world, and to assess this idea via an empirical experiment. Their results showed that the virtual pointer could be accurately aligned with the real target displayed in the stereoscopic video. The procedure followed in the experiment of Drascic and Milgram is complex and takes into account several variables, hence it is possible that the results are not completely unbiased. Nevertheless, this study showed the potential of using 3D virtual pointers for operations that require depth-judgements in remote environments. In addition, the authors identified a number of salient points that should be considered when implementing a 3D cursor. Among these: effects of combining monoscopic and stereoscopic depth cues in carrying out such tasks, effect of luminance on depth perception and visualisation issues introduced by computer graphics such as aliasing.

Jää-Aro and Kjelldahl [1997] conducted a series of experiments on both a high resolution work station and a low resolution flight helmet HMD in order to study the effect of image resolution, stereopsis and aliasing on depth perception. The results suggest that both low resolution and aliasing artifacts have a detrimental effect on task performance and that anti-aliasing is more important than binocular disparity for accurate distance estimates. Interestingly, the authors speculate that the low levels of image contrast that characterise the HMD seemed to have caused a blurring effect similar to anti-aliasing, which in turn seemed to have had a positive effect on depth perception and distance judgements during the experiments. Note that this consideration simply reinforces the positive effect of antialiasing on the tested distance judgements task performance and does not contradict the well-known fact in literature that low level of image contrast have a detrimental effect on stereoscopic performance (see section 3.2.1). In accordance with the findings of Carver and McAllister [1991], Butts and McAllister [1988], and Drascic and Milgram [1991], the work of Jää-Aro and Kjelldahl highlights the negative effect of visual aliasing artifacts on the performance of tasks that require accurate depth judgements.

2.5 Conclusions

In this chapter, a survey of a number of topics that relate to depth-based tasks in 3D environments was presented. The first part of the chapter gave an overview of human depth perception and visual cues to depth. The second part focussed on electronic 3D display technologies and discussed the perceptual issues of stereoscopic viewing on 3D displays as compared to natural scenes. Finally, the third part presented a critical review of previous works on stereoscopic 3D cursors as a means to support depth-based tasks.

From the review it has emerged that in the past twenty years considerable effort has been dedicated to explore new interaction techniques for 3D environments, mainly due to the advantages that 3D visualisation can offer and the increasing usage of 3D displays in a variety of application fields. Notwithstanding, a limited number of studies can be considered directly relevant to the design of a stereoscopic 3D cursor to support depth-based tasks in desktop stereoscopic 3D environments. Under the umbrella of VR, the majority of the reviewed works adopted a weak scientific methodology and focussed on aspects of the 3D experience of the user that are not relevant to the aim of this thesis, e.g. their level of immersion or the input device to be used for interaction. In other instances the authors presented and evaluated a specific 3D cursor design, often as part of a whole 3D interactive environment, without justifying the design choices with a thorough review of the visual perception literature. In the research areas of computer graphics, 3D-aided design and perception, the approach adopted by the authors was at times more scientific and the results more pertinent to this project. The most relevant works in this context are probably the ones of Vanacken et al. [2009] and Sands et al. [2004], together with the older study on 3D cursors presented by Barham and McAllister [1991], which provided the first inspiration for the research presented in this thesis. However, they only represent a starting point to answer the question of how a stereoscopic 3D cursor should be designed in order to effectively support depth-based tasks in a desktop stereoscopic 3D environment. In fact, none of the studies reviewed in this chapter provide a sound and complete explanation of how different visual and technical factors may impact depth perception in such environments, and subsequently how they may affect the design of a stereoscopic 3D cursor. The aim of the following chapters is to fill these knowledge gaps by adopting a rigorous and scientific approach.

Chapter 3

Design of a Stereoscopic 3D Cursor

3.1 Introduction

Chapter 2 highlighted the benefits of stereoscopic visualisation for application fields dealing with visual data that are intrinsically three-dimensional in nature, like medical imaging, structural biology or geoscience. In such domains, the ability to accurately locate objects in depth and perform target selection or delineation is crucial to a correct interpretation of the data. This chapter investigates the design of a novel 3D cursor to support such tasks in stereoscopic 3D environments. In order to fulfill this objective it was necessary to perform an in-depth analysis of the existing literature that deals with the factors that can affect depth perception in both real and synthetic scenes presented via stereoscopic 3D displays. For clarity, these factors are discussed and grouped into three main categories as follows.

Section 3.2 focuses on the visual aspect of the factors that affect the perception of depth. The factors that are more intrinsically connected to the technical properties of the 3D display used to perform the task are discussed in section 3.3. Section 3.4 describes the synthetic factors that can be used in order to provide the user with feedback. The main ideas emerging from these three sections are discussed in section 3.5 and concluded in section 3.6.

3.2 Visual Factors

Visual perception is an extremely complex process and the way different visual characteristics of the stimuli can interact with each other in order to form the final visual **percept** in the observer can be very intricate. This section of the chapter focuses on the visual factors that can interact with stereopsis and affect the perception of depth when observing a 3D scene. For an explanation of the binocular visions science terminology used in the following subsections, refer to section 2.2.5.

3.2.1 Luminance Contrast

The term contrast generally refers to the difference in brightness between a specific object and the other objects within the same *Field Of View* (FOV). The human sensory systems are particularly responsive to relative temporal and spatial changes in stimuli intensities, rather than to their absolute intensities, and this is especially true for vision. The visual system in fact is more sensitive to contrast than absolute **luminance** (L), so as to enable us to perceive the world in a coherent manner regardless of the difference in light intensity between different places and different times of the day. In accordance with this behaviour scientists developed the concept of **luminance contrast**, which generally describes the ratio between the luminance change to which the eye is exposed and the average luminance to which the eye is adapted. A more formal definition that is often used in physiology is the one of *Michelson contrast* [Michelson, 1927], which defines contrast as:

$$\frac{L_{max} - L_{min}}{L_{max} + L_{min}} \tag{3.1}$$

where L_{max} and L_{min} represent the two luminances in question, e.g. the luminance of the target in a stimulus and that of its background respectively. Michelson contrast assumes that the eye of the observer is adapted to the sum of the two luminances, expressed by the denominator, and it is suitable for measuring the contrast in periodic stimuli that do not have a clear distinction between background and foreground (e.g. grids), which are commonly employed in vision science studies. The mathematical definition of Michelson contrast yields a value between 0 and 1 and therefore it can be easily expressed as a percentage.

Another common metric used in vision science is the one that expresses contrast as a simple ratio, as specified in the following equation:

$$\frac{L_{max}}{L_{min}} \tag{3.2}$$

The ratio between the brighter luminance, L_{max}, and darker luminance, L_{min}, can also be expressed in the form of an $N : 1$ ratio; this simply indicates that L_{max} is greater than L_{min} by a factor of N. Figure 3.1 shows an example of high and low contrast ratio images of the same subject.

The level of luminance contrast in the visual stimuli is an important factor in binocular vision. Its effect on depth perception has been the focus of several studies and it is now widely recognised in literature. There are several interesting aspects regarding the way contrast can affect stereopsis, the most salient of which can be summarised as follows.

A number of studies demonstrated a significant effect of binocular variations of contrast on stereoacuity, with stereoscopic thresholds dropping gradually as target contrast in the stimuli raises [Halpern and Blake, 1988] [Legge and Gu, 1989] [Cormack et al., 1991] [Smallman and McKee, 1995] [Rohaly and Wilson, 1999]. In other words, authors registered better stereoscopic performance at higher contrast levels in the stimuli. Furthermore, most of the afore-mentioned studies suggest that stereoacuity varies as a power-law

a **b**

Figure 3.1: Examples of high (a) and low (b) contrast images.

function of contrast [Halpern and Blake, 1988] [Legge and Gu, 1989] [Rohaly and Wilson, 1999]. Finally, Halpern and Blake [1988] found that for stimuli characterised by an already high level of contrast (i.e. equal or higher than 21 **dB** relative to the contrast threshold for detection of the stimuli used in the study) a further increase in binocular contrast did not result in better stereoacuity, which suggests that the effect of contrast on stereoacuity eventually saturates.

An interesting common finding is that reducing the target contrast monocularly (i.e. only in one eye) had higher detrimental effect on stereoacuity than reducing the target contrast binocularly (i.e. in both eyes) by an equivalent amount [Halpern and Blake, 1988] [Legge and Gu, 1989] [Schor and Heckmann, 1989] [Cormack et al., 1997] [Schor et al., 1998]. This means that differences in contrast between the left and right view of a stereo pair do have a disruptive effect on stereoacuity, and stereoscopic performance in general, and links to the fact that people with marked differences in contrast sensitivity in the two eyes or that suffer from **amblyopia** typically show diminished stereoacuity [Goodwin and Romano, 1985]. Interestingly, results of physiological investigations that focused on the neural aspects of binocular vision suggest that the visual system has a compensation mechanism at synaptic level so that if different contrast values are presented in each eye, the signal from the weak eye is boosted [Truchard et al., 2000].

A similar behaviour is observed also when **stereo matching** is considered. In this regard Smallman and McKee [1995] showed that features in the left and right eye match successfully only if the ratio between their contrasts lies in a specific range. This interesting finding suggests that image features that are visible monocularly in both eyes only match and result in proper stereo vision if their contrasts are similar; if this condition is not met they remain unmatched and a perceived depth near the fixation plane is assumed. The authors found that the critical contrast ratio, i.e. the ratio between the high contrast feature in one eye and the contrast of the same feature required in the other eye for stereo matching to occur, varied greatly among subjects. In order to guarantee that all subjects would be able to perform stereo matching successfully the minimum contrast ratio required was 2.3% : 1%.

On the other hand, vertical fusion limits seem to be independent of contrast [Schor et al., 1989] [Schor and Heckmann, 1989]. In particular, the study of Schor and Heckmann [1989] showed that interocular contrast differences had no effect on vertical fusion limits, even when the contrast ratio between the two eyes was as high as 10% : 40%.

Interestingly, Hess et al. [2003] found that for *fractal* stereo pair images, reducing the contrast of the stimuli only in one eye had no more disruptive effect on both stereoacuity and the maximum detectable depth (i.e. fusion limits) than reducing the contrast by a comparable amount in both eyes. These results suggest that changes in contrast did cause a loss of stereoscopic performance, however performance for the changed monocular contrast conditions were not significantly different than performance for the changed binocular contrast conditions.

Finally, a rather interesting fact is that contrast plays a role in the process of depth perception not only as a stereoscopic depth cue but also as pictorial, monocular cue to depth by simulating the effects of aerial perspective. O'Sea et al. [1994] investigated this phenomenon and showed that contrast as a pictorial cue has a significant effect on perceived depth of objects in the stimuli, with low-contrast objects appearing further away than high-contrast objects; this theory is also supported by Mallot [1997]. Notably, the authors registered this trend not only when subjects observed the stimuli binocularly but also in the case of monocular observations, i.e. in the absence of stereopsis.

Even though beyond the scope of this thesis, from a neurological point of view the studies mentioned so far suggest that the effect of contrast on stereopsis has both monocular and binocular components and that contrast information is probably elaborated by both monocular and binocular neurons at different stages of vision [Halpern and Blake, 1988] [Rohaly and Wilson, 1999]. Truchard et al. [2000] specifically investigated the neural aspects of the monocular and binocular mechanisms underlying contrast information encoding; the authors concluded that, from a neurological point of view, the **contrast gain control** effect was strong at the monocular stage, before the information from the two eyes is merged, but not at the binocular stage, suggesting that contrast gain control is mainly a monocular process. Based on their empirical studies Truchard et al. also suggest that the visual system has a compensation mechanism at the synaptic level so that if different contrast values are presented in each eye, the signal from the weak eye is boosted. However, for the purpose of this thesis the focus is on the perceptual effect of contrast on depth perception and not on its physiological or neurological aspects. In this regard, it is worthy to note that the empirical study performed by Habib et al. [2008, 2005], where subjects used stereoscopic 3D displays and a stereoscopic 3D cursor to assess optic disc morphology, contrast mismatches of the cursor shape between the left and right view indeed had a detrimental effect on task performance and in some conditions prevented subjects from successfully fusing the two views, leading to *diplopia*.

On the basis of the literature reviewed so far, it is recommended when designing a 3D cursor to support depth-based selection tasks in stereoscopic environments that the following points are taken into consideration. First, the contrast levels between the actual cursor shape and the background should be maximised in order to boost stereoacuity. However, it is also important to bear in mind that the effect of contrast on stereoacuity seems to saturate [Halpern and Blake, 1988], which suggests that further increasing the contrast between cursor shape and background beyond a certain limit will not lead to better task performance. In the study of Halpern and Blake [1988] this limit was set to be 21 dB relative to the contrast threshold of detection of the stimuli used for the experiment. In the absence of further quantitative studies on this matter, it is therefore advised to adopt a cursor-background contrast level that is no higher than 21 dB relative to the contrast detection threshold, i.e. the minimum contrast level required for the cursor shape to be detectable against the background.

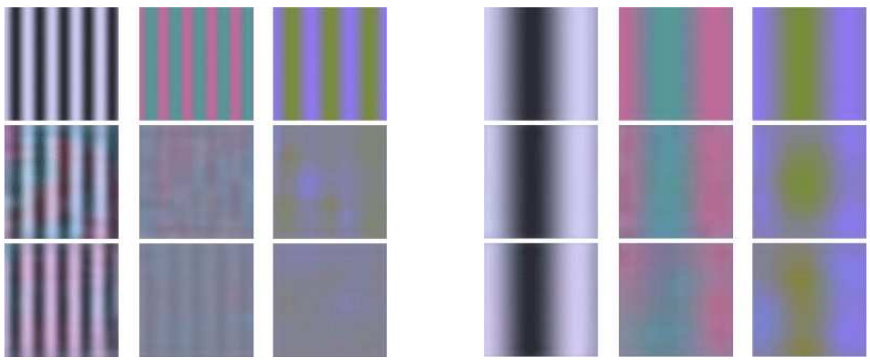
Second, no interocular differences in contrast should be shown for both cursor shape and scene, for this may have a detrimental effect on task performance. For working environments like the ones adopted in geoscientific applications, this is easily achieved because the left and right views used to render the data stereoscopically intrinsically present no differences in contrast, colour or other visual factors. On the other hand, in application fields such as medical imaging, the left and right view are often acquired via the use of two cameras and therefore interocular differences in such visual factors may be unavoidable when the data are presented to the operator stereoscopically. In order to minimise the disruption caused by contrast differences between the left and right view, it is recommended that data undergo a "smoothing" process before being stereoscopically presented to the operator for analysis.

This said, the studies reviewed above suggest that the human visual system is tolerant to some interocular contrast differences. So in situations where contrast differences between the left and right view can not be removed, some interocular contrast difference in the cursor shape could be allowed in order to improve the cursor-background contrast level. But what is the best tradeoff between cursor-background contrast level and interocular contrast differences in the cursor shape in order to maximise stereoscopic performance? From the literature review it has emerged that there is no agreement on what the contrast ratio limits between left and right view should be in order not to affect stereoscopic performance; different studies yielded different results with large differences among subjects. It is fair to assume that the high variance that characterises the results is due to the different nature of the stimuli used in the various experiments. This implies that in order to be effective the best tradeoff should be carefully determined on application-by-application basis, according to the characteristics of the stereoscopic images used to visualise the data. In practical terms though, this approach may be extremely difficult and time consuming and therefore it is not recommended. Therefore, the final recommendation is that interocular contrast differences in the cursor shape should be avoided even for applications where the left and right view used to visualise the data stereoscopically are intrinsically characterised by interocular differences. The study performed by Habib et al. [2008, 2005] supports this point as it clearly showed that interocular contrast differences in the cursor shape do have a detrimental effect on stereoscopic task performance.

Finally, it is suggested that operators who are undertaking critical depth judgement tasks are screened for amblyopia, for this may be accompanied by low sensitivity to contrast and hence poor stereoacuity. Amblyopia is a visual condition that affects 1 to 5% of the adult population [Weber and Wood, 2005] and that can be difficult to diagnose because of the lack of any obvious underlying ocular pathology. It usually occurs only in one eye and it can seriously impair stereoacuity [Weber and Wood, 2005] [Goodwin and Romano, 1985], one of the reasons being the interocular differences that results from the unclear visual image of the the amblyopic eye compared to the higher quality image of other eye.

3.2.2 Spatial Frequency

In general scientific terms, *spatial frequency* can be defined as the measure of how rapidly a signal changes in space. In the specific context of visual perception, spatial frequency is usually expressed in number of cycles per degree (c/deg) of **visual angle** and is an indicator of the degree of "smoothness" or "homogeneity" of the visual stimuli: high spatial frequency areas are characterised by frequent and abrupt changes in the spatial content of the image (like in the presence of edges or choppy patterns), while low spatial frequency areas are generally smooth, homogenous and register low changes in spatial content. The images of Figure 3.2 clearly depict the difference between high and low spatial frequency stimuli.



a **b**

Figure 3.2: Examples of high (a) and low (b) spatial frequency images.
[From [Brainard et al., 2008]]

Spatial frequency plays a crucial role in binocular vision and its effect on depth perception and stereoscopic performance has been widely studied. Because of its close relation to contrast, most studies in this research area investigated the joined effect of these two variables together rather than their individual effect; it is therefore difficult to treat them as purely individual factors. However, for clarity, the first part of this section attempts to isolate and summarise the effects of spatial frequency alone while the second part discusses how binocular vision gets affected by both the spatial frequency content and the contrast characteristics of a stereoscopic image.

A first important fact is that interocular differences in the spatial frequency of the target stimuli generally have a detrimental effect on stereoscopic performance [Schor et al., 1998]. In this regard, the studies performed by Hess et al. [2003, 2002] yielded some interesting findings, which can be summarised as follows. First, minimum detectable depth (i.e. stereoacuity) was negatively affected by low-pass filtering but remained unaltered when high-pass filtering was applied to the stimuli; furthermore, monocular low-pass filtering had a more disruptive effect than the equivalent binocular low-pass filtering, while no significant difference was registered between monocular and binocular high-pass filtering. Second, maximum detectable depth (i.e. binocular fusion limits) was also negatively affected by monocular low-pass filtering, even though to a much lesser extent than stereoacuity, while there was no loss in stereo performance in the case of high-pass filtering; it is worth noting here that these trends were similar for the monocular and binocular filtering conditions.

According to the study of Schor and Wood [1983], both lower and upper disparity limit thresholds decrease as spatial frequency increases, with the lower disparity limit decreasing (i.e. stereoacuity improving) at a faster rate than the upper disparity limit, but remain almost constant for target stimuli with spatial frequencies above 2.4 c/deg. This finding is in accordance with the results of Legge and Gu [1989] and suggests that stereoacuity performance are better at higher spatial frequencies and that, similarly to contrast, the effect of spatial frequency on stereoscopic performance eventually saturates. In particular, in this regard Legge and Gu [1989] found that disparity thresholds were inversely proportional to spatial frequency for low spatial frequencies (0.25 - 3.0 c/deg), were lowest at medium spatial frequencies (around 3.0 c/deg) and were marked by an irregular behaviour and individual differences for high spatial frequencies (5.0 - 20.0 c/deg).

At low spatial frequencies (0.8 and 1.6 c/deg) a similar behavior was also observed by Schor and Heckmann [1989] for vertical fusion limits, which seem to decrease with increasing spatial frequency.

Based on the results of these studies it can generally be concluded that stereo performance improves with increasing spatial frequencies. However this positive effect seems to saturate at medium spatial frequencies (around 2.4 - 3.0 c/deg) and can potentially become disruptive at high and very high spatial frequencies (5.0 - 20.0 c/deg or above).

With regard to the interlinked effect of spatial frequency and contrast, it is widely recognised in literature that interocular differences in contrast have a detrimental effect on stereoacuity performance (see section 3.2.1). This phenomenon goes under the name of *stereo contrast paradox* [Cormack et al., 1997] and has been observed over a vast range of target spatial frequencies. However, the negative effect of interocular contrast differences is stronger when the spatial frequency of the stimuli is low (1.2 c/deg or less) [Halpern and Blake, 1988] [Schor and Heckmann, 1989] [Cormack et al., 1997], tends to disappear at relatively high spatial frequency (5.0 c/deg) [Cormack et al., 1997] and, in line with this last point, is absent in random-element stereograms because of the high-frequency nature of the stimuli [Cormack et al., 1997]. In other words, stereoacuity seems to be less affected by interocular changes in contrast at higher spatial frequencies [Halpern and Blake, 1988]. This interesting behavior is probably linked to the fact that stimuli characterised by low spatial frequency have a lower apparent contrast than high spatial frequency stimuli [Howard and Rogers, 2002b]. Furthermore, Schor et al. [1998] found that for matched interocular spatial frequencies **transient stereoscopic performance** was worse in the case of unmatched interocular luminance contrast compared to the case with matched interocular contrast; on the other hand, for unmatched spatial frequencies stereoscopic performance was best when the contrast of the stimuli was also unmatched.

These findings lead to the general conclusion that the human visual system is tolerant to interocular contrast mismatches when the stimuli's spatial frequency content is high, while it is remarkably intolerant to equivalent interocular contrast mismatches in the presence of low spatial frequency stimuli.

In the light of this review, the recommended guidelines in terms of 3D cursor design include the following points. A first general point that seems straightforward but is worth mentioning is that the cursor shape should not present any difference in spatial frequency content between the left and right view. The logical choice of using the same shape to represent the cursor in the two views is sufficient to guarantee that this requirement is satisfied. As it can be easily anticipated, the adoption of different shapes will inevitably lead to problems at the very fundamental levels of the stereoscopic depth perception process, like for example the difficulty in matching corresponding features of the cursor shape between the two views.

Furthermore, it is recommended that the shape chosen to represent the cursor in the working scene is characterised by a moderate degree of spatial frequency, while very high and low spatial frequency shapes should be avoided. One interesting fact that has emerged from the studies analysed above is that stereoscopic performance is generally better at high rather than low spatial frequencies, with the best results at spatial frequencies in the range of 2.4 up to 3.0 c/deg [Schor and Wood, 1983] [Legge and Gu, 1989], depending on the nature of the stimuli. However, similarly to contrast, the effect of spatial frequency on stereoscopic performance seems to saturate and can potentially become disruptive when visual stimuli are characterised by particularly high frequency content; according to the study of Legge and Gu [1989] this was the case for spatial frequencies in the range of 5.0 to 20.0 c/deg. From a pragmatic point of view these results suggest that extremely detailed 3D shapes, e.g. a 3D star, and completely smooth shapes, e.g. a sphere, should be avoided.

But what cursor shape is the most suitable to support accurate depth judgment tasks? In order to answer this question, it would be necessary to perform an experiment where the spatial frequency of different cursor shapes is measured and the impact that this has on stereoscopic task performance quantified. However, in the absence of detailed empirical data, a box-shaped cursor seems to represent a sensible choice and is to be preferred to a spherical cursor because its higher spatial frequency content should lead to better depth perception than a totally smooth shape such as a sphere.

Finally, this section can be concluded with a consideration that is not directly related to the design of a 3D cursor but is certainly worth noting. One of the salient points that has emerged from the literature review is that interocular contrast mismatches and low contrast levels have a particularly detrimental effect on stereoscopic performance in the presence of stimuli with low spatial frequency content. Therefore, with regard to application fields that may make use of stereoscopic images characterised by relatively low spatial frequency, e.g. the fundus images used in ophthalmology for the detection of retinal diseases, it is strongly recommended that data undergo a uniforming process in terms of image contrast between the left and right views before being utilised for diagnosis. However, it is important that this uniforming process is edge-preserving and does not cancel important features present the images, for this would have severe consequences on the outcome of the interpretive process performed by the operator and potentially lead to the wrong diagnosis.

3.2.3 Colour

Colour is the property that describes the appearance of an object in terms of one's perception of its **hue**, **saturation** and **brightness**. In order to experience colour an organism must be able to distinguish objects based on the variations in spectral reflectance in the surrounding environment, a capability that goes under the name of colour vision.

Colour vision is widespread in vertebrates, while in mammals it is less developed with the exception of primates [Pokorny and Smith, 2003]. In humans, the nervous system derives colour by elaborating the responses of the photoreceptors in the retina of the eye: the *cones* and the *rodes*. Cones work best in bright light conditions and are sensitive to different portions of the spectrum. Rodes are more sensitive to light than cones; they can function at low light levels and are mainly responsible for night vision. Although their signals can interact with cone signals, rodes are characterised by only one type of light-sensitive pigment and therefore play a limited role in colour vision [Pokorny and Smith, 2003]. This explains why at night objects usually appear to be all greyish and no colour is perceived. Cones on the other hand contain one of three pigments sensitive to a particular range of wavelengths and provide the basis for colour vision. Based on their spectral sensitivity they are usually divided into three categories: S or *blue cones*, maximally sensitive to short wavelengths and to what we perceive as blue-violet light, M or *green cones*, highly responsive to medium wavelengths and mainly sensitive to light perceived as green, and L or *red cones*, with peak sensitivity at longer wavelengths in correspondence of light perceived as yellow-green but also responsible for the perception of red, as Figure 3.3 illustrates. The RGB colour model used in electronic media is based on the physiology of the human retina and the way we perceive colour.

The range of wavelengths that humans can perceive through the cone cells goes approximately from 400 nm to 700 nm [Palmer, 1999] and defines what is known as the *visible light spectrum*, shown in Figure 3.4. When white light (e.g. sun light) reaches an object perceived, for example, as yellow, the light components outside the wavelength range correspondent to yellow in the visible spectrum are captured or absorbed by the object, while the yellow wavelength components, around 500 nm, are reflected back. When the latter reach the human eye, they stimulate the M and L cones allowing us to perceive the object as yellow. Based on this theory, objects that in sun light appear as black absorb

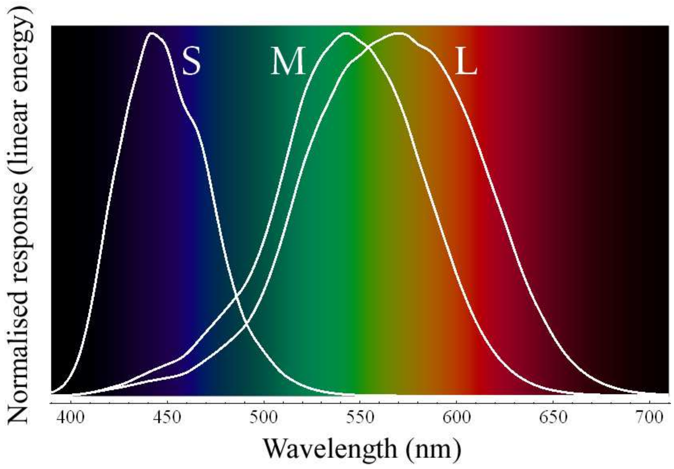


Figure 3.3: Normalised response spectra of human cones to monochromatic spectra stimuli. The spectral sensitivity of S cones picks at around 430 nm, the one of M cones at 545 nm and the one of L cones at 575 nm.
[Adapted from http://en.wikipedia.org/wiki/Color_vision, last accessed: 05.01.2011]

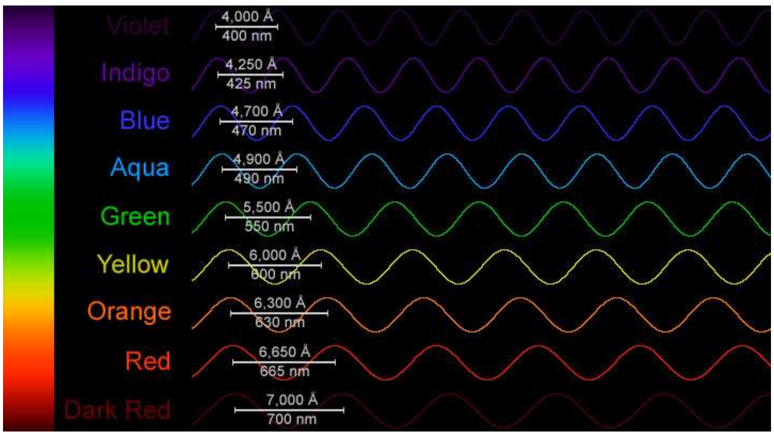


Figure 3.4: The visible light spectrum.
[From http://www.windows2universe.org, last accessed: 05.01.2011]

light at all wavelengths and scatter little light. Vice versa, white objects absorb a minimal quantity of sun light and instead reflect back most of it.

Humans are usually classed as *trichromats*, i.e. they possess three different types of cone in the retina, but a number of molecular genetics studies suggest that a substantial percentage of women have a fourth type of colour receptor, which can greatly enhance their colour discrimination [Jameson et al., 2001]. However, even in the presence of "normal" trichromacy, according to Judd and Kelly [1939]:

"*[...] there are about ten millions surface-colours distinguishable in daylight by the trained human eye.*" (pg. 359)

Interestingly, there is also evidence that from a linguistic and anthropological point of view people typically use only about eleven basic names to classify colours of objects seen in everyday life: white, black, red, green, yellow, blue, brown, purple, pink, orange and grey [Boynton and Olson, 1990] [Berlin and Kay, 1969].

The role of colour vision in binocular depth perception has been the object of research for several decades but results are mixed and not always in agreement; given their tight relationship, colour and luminance cues are often studied together. Lu and Fender [1972] studied the contribution of visual information based on differences in colour (*colour contrast*) and difference in luminance (*luminance contrast*) to the perception of depth in **Random Dot Stereograms** (RDSs) and concluded that colour contrast alone "does not give rise to the percept of depth". Furthermore, they suggest that the human visual system relies on independent coding of colour and luminance information and that binocular vision principally relies on the latter, with only mild secondary effect of colour. Livingstone and Hubel [1987] performed a series of psychophysical experiments on different aspects of visual perception and reached similar conclusions. Likewise, Krauskopf and Forte [2002] found that stereoacuity was poorer for chromatic stimuli than for achromatic stimuli, suggesting that colour information does not enhance stereopsis; based on their results the authors also concluded that luminance and colour channels are not independent in the human disparity interpretation mechanism but they rather interact.

On the other hand, a number of studies indicate that colour information does support stereopsis and therefore plays an important role in the detection of depth in stereoscopic stimuli. The experiments of de Weert and Sadza [1983] showed that RDSs containing chromatic information can be fused both at different luminance contrast levels and at **isoluminance**. This result was confirmed by Scharff and Geisler [1992], who concluded that for the conditions tested in their study depth could be reliably detected for chromatic RDSs at isoluminance. Similarly, Kingdom and Simmons and their collaborators reported that depth could be perceived in chromatic stimuli characterised by isoluminant RDSs [Kingdom and Simmons, 1996] [Simmons and Kingdom, 1997] [Kingdom et al., 1999]; regarding the interaction between colour and luminance information, they suggest that there are two independent stereopsis mechanisms in humans, one sensitive to chromatic contrast and the other to luminance contrast [Simmons and Kingdom, 2002] [Simmons and Kingdom, 1997], and that "chromatic stereoscopic processing is less precise than luminance processing" [Kingdom and Simmons, 1996]. In a more recent study den Ouden et al. [2005] investigated the role of colour information in the perception of depth in complex stereoscopic stimuli consisting of a number of solid shapes and concluded that chromatic information has a positive effect on stereopsis and is used by the visual system "to solve the **binocular matching** problem in complex images". More than a decade before Stuart et al. [1992] described a similar finding for RDSs and concluded that colour contrast provides valuable input to stereopsis and can be used "at least as effectively as luminance contrast" in solving stereo matching. Based on their results they also suggest that the human stereopsis mechanism seems to involve a channel that is sensitive to both colour contrast and luminance contrast.

Summarizing, the studies reviewed so far seem to suggest that colour information, and in particular colour contrast, does contribute to the percept of depth in stereoscopic stimuli and that in the human stereopsis mechanism colour and luminance contrast are closely related and seem to interact. Generally, depth perception in the presence of colour contrast stimuli at isoluminance is possible but is weaker than when both colour contrast and luminance contrast are simultaneously present. Whether colour and luminance information are conveyed and processed individually and then joined or are instead elaborated together is not clear, but this aspect of human stereo vision is beyond the scope of this thesis.

An interesting aspect of the human visual system is the way the response of the photoreceptors in the retina co-varies with the spatial frequency content of stimuli. There is evidence that the **Contrast Sensitivity Function** (CSF) for luminance of monochromatic sinusoidal grating (i.e. achromatic stimuli with constant colour information and varying luminance contrast levels) is a band-pass function of spatial frequency, with sensitivity being highest in the range between 2.0 and 5.0 c/deg [De Valois, 2003]. On the other hand, the CSF for isoluminant grating with chromatic information (i.e. chromatic stimuli with varying colour contrast and constant luminance contrast levels) is a low-pass function of spatial frequency. This behaviour is illustrated in Figure 3.5, where CSFs are reported for luminance varying stimuli (full diamonds) and for isoluminant colour-varying stimuli for the L and M cones (squares) and the S cones (triangles) respectively. Interestingly, the S cone CSF loses sensitivity faster than the L-M cone CSF as spatial frequency increases [De Valois, 2003], possibly due to the fact that the S cones constitute only 7% of the whole cone photoreceptor population of the human retina [Curcio et al., 1991].

These results show that at high spatial frequencies (5.0 c/deg and above) our sensory systems are not particularly good at detecting isoluminant colour-varying patterns, which in turn suggests that colour vision can only marginally contribute to the detection of fine spatial details. On the other hand the colour vision system becomes crucial in the presence

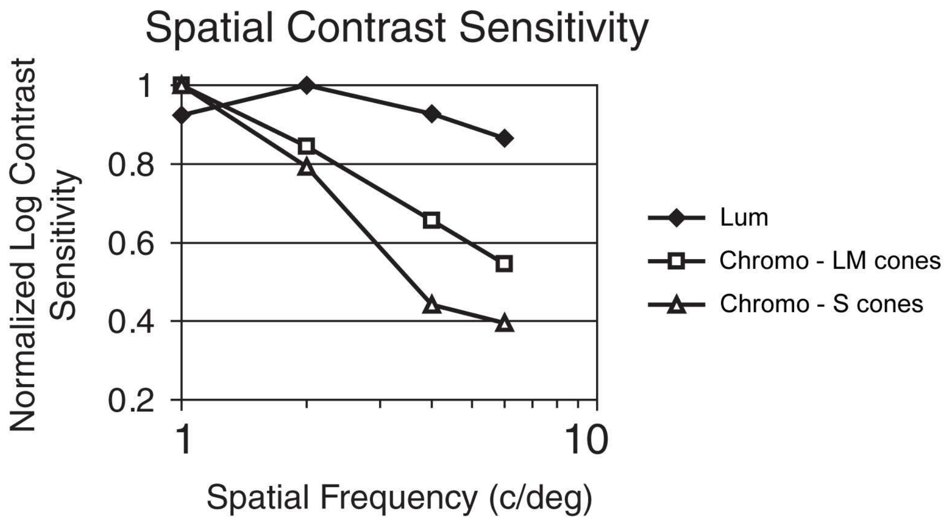


Figure 3.5: Normalised spatial CSFs for luminance-varying gratings (full diamonds), isoluminant chromatic gratings varying only in L and M cone absorption (squares) and isoluminant chromatic gratings varying only in S cone absorption (triangles).
[Adapted from [De Valois, 2003]]

of low spatial frequency information (2.0 c/deg and below), where the luminance signal is instead weak. The experiments of Chaparro et al. [1993] confirm this theory.

As mentioned in section 2.2.4, another interesting phenomenon regarding colour vision is what goes under the name of chromostereopsis. Chromostereopsis, sometimes also referred to as colour stereopsis or colour stereoscopic effect, was first described by Goethe [1810] and refers to the effect of colour contrast on perceived depth when stimuli are observed binocularly at isoluminance [Guibal and Dresp, 2004]. For example, when colour patches of comparable luminance values are placed side by side on a flat surface and are viewed stereoscopically, chromostereopsis is experienced as the illusion of the depth difference between the patches even though they are coplanar and no retinal disparity information is present in the stimuli [McClain et al., 1990]. This effect is well documented in literature and it is believed to be due to chromatic aberrations that occur in the eye, which induce different disparities in the images of the observed objects purely on the basis of the wavelength of their colours or hue [Howard and Rogers, 2002a] [McClain et al., 1990].

Interestingly, colour cues in the presence of chromostereopsis can interact with luminance information and cause depth reversal, depending if the chromatic objects are darker or lighter than the background [Guibal and Dresp, 2004] [Dengler and Nitschke, 1993] [Kishto, 1965]. Finally, given its optical nature, chromostereopsis can also be experienced by colour-blind people as it is unrelated to how colours are coded in the retina [Kishto, 1965].

In an interesting study, McClain et al. [1990] investigated the interaction and effect of chromostereopsis and stereopsis on depth perception on a shutter glasses 3D display; the authors concluded that chromostereoscopic effects derived by hue do affect the amount of perceived depth, especially when associated to relatively small disparity levels. In their experiments, short-wavelength colours (e.g. blue) on black background generally induced an illusory positive disparity effect, short-wavelength colours (e.g. red) produced a negative disparity effect and colours near the centre of the visible spectrum (e.g. green) had almost no effect. Besides, the atmospheric perspective effect of luminance was also observed, with brighter colours being perceived closer than darker colours. Based on these results the authors conclude that cautions should be taken when choosing the hue of adjacent objects with a disparity difference between them of 3.39 minutes of arc or lower; in such a situation, colours on the extreme ends of the visible spectrum should be avoided unless they are intended to enhance or negate the depth perception induced by the binocular disparity of the two objects.

A final important aspect is the effect of colour on **attention** and on the performance of high level tasks such as segregation and target identification. A number of studies have shown that target detectability directly depends on the chromatic differences between target and distractors/background, and that when this is significant segregation and detection can be readily accomplished [Webster et al., 1998] [Nothdurft, 1993] [D'Zmura, 1991]. In this regard, Kingdom et al. [2001] investigated the effect of chromatic and luminance contrast on stereoscopic depth judgements with RDS stimuli and concluded that target-distractors differences in both colour and luminance are beneficial and facilitate stereoscopic depth discrimination. However, the authors also clarify that this was the case only if the target elements were in minority compared to the distractors or if participants had prior knowledge of the target colour or luminance characteristics.

A general conclusion that emerged from the literature review is that human colour vision does convey spatial analysis capabilities and can contribute to stereoscopic depth judgements, especially when coupled with luminance information. Based on this, a number of guidelines in terms of stereoscopic 3D cursor design can be drawn up. First, it is recommended that the cursor shape is presented with a good level of both colour and luminance contrast against the background. With specific regard to hue, in applications where features are usually analysed against a black background, e.g. geosciences, this is easily achieved by assigning any chromatic property to the cursor shape. For applications that instead utilise pre-acquired chromatic images, e.g. the fundus images used in ophthalmology, this can be accomplished by picking a cursor colour that is **complementary** to the dominating colour of the working scene. In this regard, note that colour contrast is particularly important for working scenes characterised by low spatial frequency (e.g. ophthalmology), for in such conditions our colour system is more effective at spatial analysis than our luminance system. Therefore, in the presence of low spatial frequency content, having a cursor that stands out in colour against the relatively uniform background is particularly beneficial for task performance. Conversely, in working environments characterised by high spatial frequencies, cursor-background luminance contrast should be boosted as our colour visual system is less effective.

Differences in colour play a beneficial role also in segregation and detection, with better performance associated with high levels of target-distractors colour contrast. Even though not directly related to depth judgements, segregation and detection are crucial in working environments that are characterised by numerous sparse objects, as it is important for the user to readily be able to isolate the cursor from the rest of the objects in the scene and identify it in order to be able to use it to perform other tasks. Therefore it is recommended that in such working environments the cursor's colour is significantly different than that of the other objects in the scene.

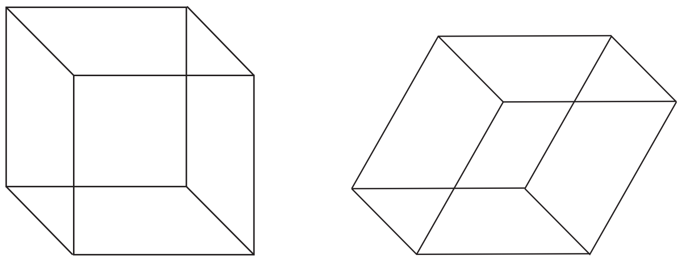
A third important point that emerged from the review is that chromostereopsis can affect the perception of depth in stereoscopic images shown on electronic 3D displays. This means that a difference in depth between the cursor and a target could potentially be perceived even when their retinal disparity is the same, which could compromise the performance of depth judgements tasks. In order to overcome this potential problem it is suggested that the user is provided with a form of feedback when one or more targets are located at the same stereoscopic depth of the cursor. An example of feedback would be to make the target(s) blink on and off the screen; a thorough discussion on different forms of feedback and synthetic cues is given in section 3.4.

Finally, it is worth mentioning that given the highly subjective nature of colour perception, there is no such a thing like the "best colour" to be used to represent a stereoscopic 3D cursor, or a cursor in general. For example, there is evidence that some colours are better than others in attracting attention, e.g. red [Guibal and Dresp, 2004]. However, red is simply a percept and it is unlikely that a particular colour uniquely defined by specific hue, brightness and saturation parameters will give rise to exactly the same perceptual experience in different people. Therefore, attention should be paid to relative colour differences between the cursor and the rest of the working scene rather than to colour as an absolute property of the cursor in itself.

3.2.4 Perspective, Distance and Depth Constancies

When observing a scene with no depth cues, objects may appear to reverse in depth [Howard and Rogers, 2002b]. This phenomenon goes under the name of *reversible per-*

spective and was first observed by the Swiss crystallographer Louis Albert Necker [Necker, 1832], the author of the famous homonymous optical illusion. The Necker cube, shown in Figure 3.6, consists of a wire-frame cube drawn in parallel perspective. Staring at the figure causes the cube to flip back and forth between two equally possible orientations, with one face sometimes appearing nearer and some other times appearing further away. Ambiguous stimuli that allow two possible interpretations like the Necker cube are referred to as bistable, or more generally as **multistable**.



a **b**

Figure 3.6: Necker cube as usually illustrated today in literature (a) and in the form of a rhomboid as originally presented by Necker (b) [Necker, 1832].

de Lussanet and Lappe [2008] investigated bistability in ambiguous stimuli and demonstrated that binocular disparity can disambiguate Necker cubes, while linear perspective does not help in making this stimulus less ambiguous. On the other hand, the authors found that perspective was a useful cue for disambiguating the depth of a point-light walker (Figure 3.7) while binocular disparity was of little help. These interesting findings suggest that three-dimensional objects can spontaneously reverse in depth even in the presence of depth cues. Back in 1838 Wheatstone observed this very phenomenon and described how a wire-frame cube that rotates about a diagonal axes can spontaneously reverse its depth and, simultaneously, direction of rotation [Wheatstone, 1838]. In this regard, Sinsteden [1860] noticed that when observing a windmill at an angle of approximately 30° from the plane of rotation and in dim light conditions, the vanes of the mill appear to spontaneously reverse direction of rotation.

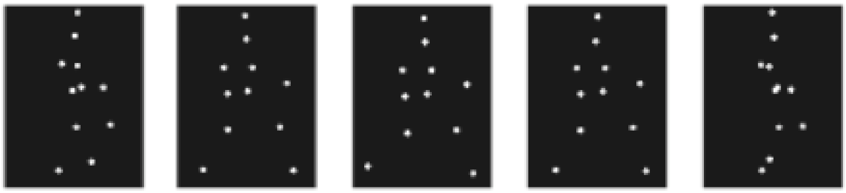


Figure 3.7: Example of point-light walker.

On the basis of these studies it is recommended that wire-frame shapes should be avoided when designing a stereoscopic 3D cursor to support depth-based selection tasks, as they could lead to ambiguous depth perception of the cursor itself, which in turn could have a detrimental effect on task performance. This is especially true for the task scenes typically used in geoscientific applications, where depth cues are frequently weak or absent.

An interesting aspect of the human visual system is the ability to perceive the visual features of a familiar object as intact or "constant" in spite of changes in viewing conditions, or to detect constant relationships between co-varying features of the object [Howard and Rogers, 2002b]. This phenomenon goes under the name of **perceptual constancy** and can concern different visual features of the observed object (colour, orientation, location, size, shape, direction). For instance, snow appears to be white regardless of whether it is observed in the dim illumination of the moonlight or in the bright light conditions of full sunshine. Given the scope of this thesis, the focus of this section is on **depth constancies**, i.e. constancies related to the visual perception of depth, and in particular on *size constancy*.

Size constancy is the ability to perceive the size of an object as remaining constant despite changes in its absolute distance along the line of sight and the consequent differences in its retinal image. Humans have a certain degree of size constancy [Howard and Rogers, 2002b] and this can be observed in everyday life; for example, when we look at a person walking away we do not perceive the person as shrinking in size but rather as moving away from us. Size constancy is known to be better when the object of interest is familiar to the observer [Marotta and Goodale, 2001] [Over, 1963] [Ittelson, 1951] and in an environment rich in depth cues [Koh and Charman, 1999] [Harvey and Leibowitz, 1967], but is not a directly measurable quantity itself. This notwithstanding, its underlying perceptual mechanism is formalised in the *size-distance invariance hypothesis*, a perceptual hypothesis that states that for an image of a given perceived size SR_P (correct or not correct), the ratio of the perceived object size S_P to the perceived object distance D_P is constant [Howard and Rogers, 2002b]:

$$SR_P = k\frac{S_P}{D_P} \tag{3.3}$$

This implies that the perceived size of the object of interest is proportional to its perceived distance and that misjudgements in object size lead to errors in judged distance, and vice versa.

The size-distance invariance hypothesis, also known as *size-distance scaling*, can be considered a perceptual interpretation of *Emmert's law*, which states that a visual afterimage projected onto a surface becomes smaller as the surface is brought nearer to the observer [Emmert, 1881]. Intuitively this makes sense and suggests that two objects that generate retinal images of the same size and are placed at different distances from the observer must have different physical size, i.e. the object placed further away must be bigger than the nearer object. The formal equivalence between size-distance invariance hypothesis and Emmert's law has also been proven scientifically [Mariko and Sachio, 2005]. However, it is worth noting that Emmert's law is not a perceptual hypothesis but an optical law based on geometrical rules that expresses the relationship of three objectively measurable quantities: the image size of an object projected onto the retina (or object angular size) SR, the physical size of the object (or object linear size) S and the physical distance of the object from the observer's eye along the line of sight D. The link between these quantities can be mathematically expressed as follows:

$$SR = k\frac{S}{D} \tag{3.4}$$

where k is a constant.

When at least two of the variables in the perceptual model represented by equation 3.3 (i.e. perceived image size, perceived object size and perceived distance) are correctly registered with the corresponding variables in the physical model of the scene represented by equation 3.4 (i.e. image size, physical object size and object distance) we are in the presence of "ideal perception"; that is, the way the visual stimulus is perceived by the observer matches exactly the physical nature of the stimulus in the real scene. On the other hand, if two or all three physical quantities are misjudged and their perception deviates considerably from their objective measures, the size-distance invariance breaks down and the observer could experience paradoxical effects or perceptual phenomena like the moon illusion (Figure 3.8).

Figure 3.8: The moon illusion effect. The observer's perception of the moon size varies greatly depending on the viewing conditions.
[From http://news.bbc.co.uk/1/hi/magazine/4619063.stm, last accessed: 17.01.2010]

Why does the moon at the horizon appear to be considerably bigger than the moon high up in the sky? This optical effect has fascinated people since ancient times and has been studied throughout the centuries. Even though no consensus has been reached yet on what is causing the effect, the latest studies on the matter suggest that the unfamiliarity of the object of interest (we do not have a concept of typical size for an object such as the moon) and, more generally, the lack of depth cues when the moon is observed at the zenith play an important role in it [Kaufman and Kaufman, 2000] [McCready, 2004] [Ross and Plug, 2002].

Cues to depth and perspective laws play a crucial role in the perception of stimuli even in 3D graphics. For example, the stereogram of Figure 3.9 contains two squares that subtend the same visual angle but that are characterised by different levels of binocular disparity; when the image is fused the presumed nearer square appears to be smaller than the further square. This occurs because the linear size of the two squares is scaled according to their difference in depth, which in turn is induced by the difference in disparity. If the assumption is made that the two squares are two identical objects (i.e. their physical size is the same), then the perceptual experience described above will be perceived as "wrong" and paradoxical by the observer. In fact, according to the principles of linear perspective, the near square would be expected to appear bigger than the far square, and not vice versa.

Based on the literature reviewed so far, it is recommended that care is taken when dealing with objects moving in a synthetic stereoscopic 3D space, especially when the 3D scene is not characterised by strong depth cues. In particular, it is important to respect linear perspective laws and Emmert's law otherwise the stimulus may lead to unwanted visual artifacts and to an unintended perspective experience. For the same reason conflicting depth cues must also be avoided [Harwerth et al., 1998].

With specific regard to 3D cursor design, it is important that the movement of the

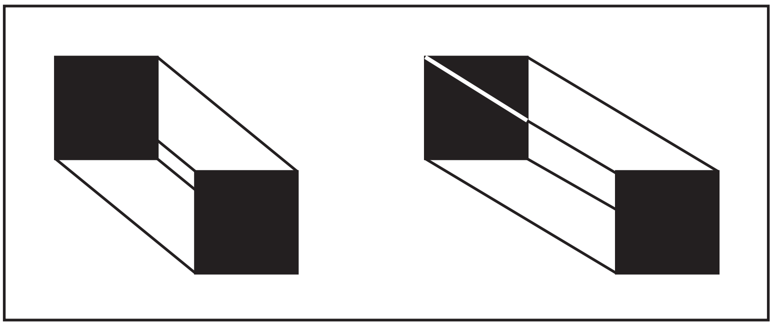


Figure 3.9: Perceived size and stereoscopic depth. The two squares are the same size but the stereoscopically nearer one appears smaller than the other and the parallel lines appear to converge.
[Redrawn from [Howard and Rogers, 2002b]]

cursor in depth respects the size-distance invariance and that the cursor shape appears to be perspective-correct. Failing to do so could confuse the operator and have a detrimental effect on task performance.

3.2.5 Blur and Depth of Focus

In everyday language the term blur refers to the condition of being not sharp in appearance and out of focus. In vision science there are at least three distinct ways in which blur can be used as a stimulus to the visual system in order to affect the perception of the depth of an object, as explained in this section.

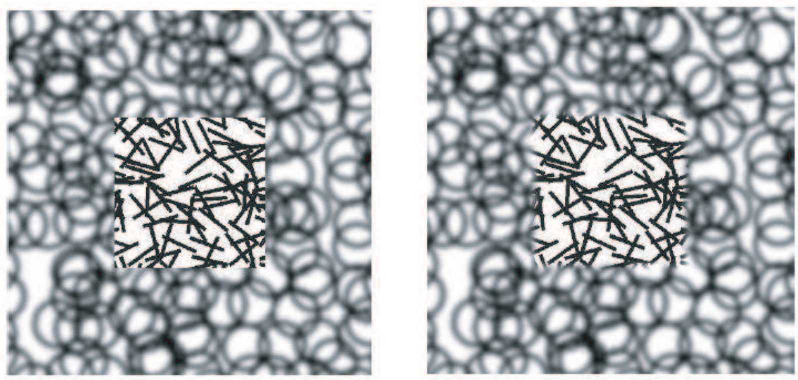
When a person takes the voluntary decision to focus on a particular object in a 3D scene, other objects that lie in the same depth plane as the fixated object appear to be sharp, while objects that are nearer or further away are perceived as out of focus [Howard, 2002a] and their image is characterised by a certain level of **image blur** or **defocus blur**. For a given plane of fixation, the range of distances along the eye's optical axis throughout which the objects in the scene are still perceived to be in focus and no image blur is detected by the observer is the *Depth Of Field*. In other words, the depth of field is the distance range in scene space within which the accommodative state of the eye can change without generating detectable blur. Its equivalent in image space is the *Depth Of Focus* [Howard, 2002a], but these two terms are often interchanged in common use and referred to with the same acronym, DOF. Similarly to geometric perceived depth (see Appendix D), blur varies as a function of object distance from the plane of fixation: the blurrier the object appears to be, the further it is from the plane of fixation. This concept can be expressed mathematically as follows [Mather and Smith, 2000]:

$$\tan\sigma = \frac{r|D-u|}{Du} \tag{3.5}$$

where, σ represents the image blur width (or radius), r is the observer eyes' pupil radius, D is the distance between the observer and the non-fixated object, and u is the distance between the observer and the fixated object.

It is well known that image blur and DOF can be used by the visual system as a cue to depth [Howard, 2002a] [Marshall et al., 1996] [Mather and Smith, 2000] [Pentland, 1987]. Artists and photographers often create an impression of depth by simulating the blurred appearance of objects that are not in the plane of fixation and by reducing the depth of focus so to include only the object of interest, while the rest of the image is left with different levels of blur [Held et al., 2010] [Howard and Rogers, 2002b]. In this way, the degree of image blur provides the information about the relative distance in depth between the object of interest and the other objects in the scene. However, when used in this way image blur is an ambiguous cue to depth [Howard and Rogers, 2002b] [Marshall et al., 1996] [Grossman, 1987] [Pentland, 1987] because it fails to provide information about the sign of the difference in depth between the two objects: an object that lies in front of the plane of fixation is characterised by the same amount of blur as an object that is at the same relative distance but behind the plane of fixation [Mather and Smith, 2000] [Marshall et al., 1996].

Notwithstanding, under specific viewing conditions this sign ambiguity can be solved and image blur can be used effectively. In particular, when in an image two regions with different degrees of image blur share a common boundary, the amount of blur of the boundary itself can be used to discern the depth ordering of the two regions. Figure 3.10 illustrates this concept.



a **b**

Figure 3.10: Occlusion edge blur. Depending on the degree of blur of the edge of the central square, the line-textured object appears in front (a) or behind (b) the circle-texture object. [From [Marshall et al., 1996]]

As it can be seen from the picture, when the edge of the central square is sharp (a) the inner sharp line textured object is perceived to be in front of a blurred circle texture object, while when the edge of the central square is blurred (b) the inner sharp line textured object is perceived to be behind the blurred circle-texture object, as if it was being observed through a hole. Marshall et al. [1996] called this effect *occlusion edge blur* and have demonstrated that it can act effectively as a cue to relative depth even when stereoscopic information is not present. In this regard Mather and Smith conducted a series of experiments and concluded that occlusion edge blur is a cue to depth but only a qualitative one, mainly because of the limited capability of the human visual system to detect small differences in blur [Mather and Smith, 2002] [Mather, 1996].

Mather and Smith [2000] investigated also the interaction between image blur and other depth cues and their main findings can be summarised as follows. When image blur and image disparity were the only depth cues present in the RDS stimuli, stereopsis was the dominant cue while the effect of blur was marginal and only present when the two depth cues were consistent with each other. On the other hand, when the interaction between image blur, interposition and contrast was investigated in all their possible combinations, depth judgement accuracy was best for the cue combinations that included image blur and worst for the combinations that included interposition, while response times were slowest for single cue combinations and fastest for three-cue combinations and combinations that included blur [Mather and Smith, 2004]. In this regard, Marshall et al. [1996] reported that the weighting of occlusion image blur in comparison to other depth cues can vary considerably across subjects.

On the basis of the discussion above it can be concluded that blur can serve as a coarse, qualitative cue to relative depth and can stimulate our sensory system in different ways. Image blur alone is intrinsically ambiguous because of its unsigned nature, but occlusion edge blur can effectively be used to solve this ambiguity and can be considered as a depth cue distinct from image blur. Finally, it is important to distinguish these two depth cues from pictorial blur (or haze blur). The latter operates at much larger distances, as a result of atmospheric perspective, and only affects the image of very far objects in a scene [Marshall et al., 1996] [Bruce et al., 2003], e.g. when we look at mountains from a distance as illustrated by Figure 3.11.

With regard to depth-based selection tasks, blur and DOF could only be used to make the working environment more realistic and enhance the depth perception of the 3D scene as a whole, but it can not be used effectively in order to improve task performance. In fact, applying image blur and DOF with the plane of fixation centered on to the 3D cursor could even have a detrimental effect on task performance, for the type of task considered in this thesis the position of the target is not known before showing the scene. This means

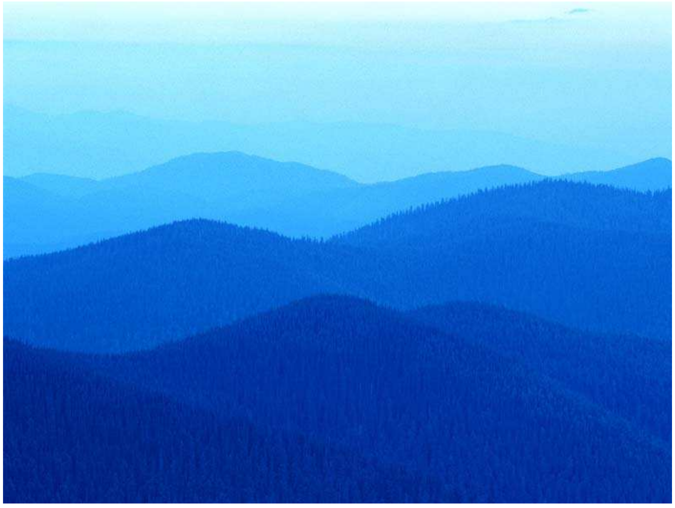


Figure 3.11: Pictorial blur as observed when looking at mountains on a hazy day.

that the blur in front and behind the cursor could obstruct the view of the operator and impair the localization of the target. Concerning this point, in the study performed by Sun and Holliman [2009] participants' score was lower for the two stereoscopic sequences that contained DOF in comparison to the other three stereoscopic sequences that did not contain DOF blur. These results suggest that image blur did not enhance the perceived depth quality of the stereoscopic videos but, on the contrary, worsened it. Furthermore, the majority of the participants that took part in this study explicitly stated that they disliked the DOF blur effect, especially when applied in the foreground in front of the target object, because it prevented them from gaining a clear view of the scene. These results suggest that the only way blur could be used effectively to improve depth perception in depth-based selection tasks is to track the eyes of the operator and update the image blur in the 3D scene in real time accordingly, on the basis of equation 3.5. However, this is not easy to achieve and there is a risk that the effort required to correctly incorporate the eye tracker into the working environment may not justify the potential gain in task performance. Furthermore, this solution is not suitable for applications where multiple operators are required to work on the same display.

3.2.6 Occlusion and Transparency

Occlusion, also known as *interposition* or *overlap*, is the condition in which an opaque object partially or totally obscures a more distant object [Fleming and Anderson, 2003]. This common circumstance can be observed very frequently in everyday scenes and plays a significant role in the perception of depth. The geometry of occlusion introduces some unequivocal qualitative rules about the relative distance of objects: nearer objects can occlude more distant objects but not vice versa. This principle is clearly illustrated in Figure 3.12, where the simple fact that the cathedral is occluded by the statue makes us infer that it must also be further away than the statue.

Figure 3.12: Occlusion. Nearer objects (statue) occlude further objects (cathedral).

In the presence of a 3D scene with stereoscopic information, the edge that separates the occluded object from the occluding one is the only feature that carries quantitative information about the local depth of the scene (i.e. disparity) [Anderson, 2003]. More precisely, the disparity information carried by the edge is used by the visual system to assign depth to the object the edge belongs to, while the only certain fact about the other object is that it must be either nearer or more distant. Assuming that the edge is associated with the occluding object, the disparity information carried by the edge is used to assign a depth measure to the occluding object, while the occluded object can only have the same or greater distance. In other words, it is impossible for the occluded object to be nearer than the occluding one. This simple geometrical fact is formalised in what Anderson et al. [2002] define as the *contrast depth asymmetry principle*, which states that:

"*The two sides of a luminance discontinuity with an associated depth value are constrained to either appear at the depth of the contour, or one side of the contour can appear more distant in depth*" (pg 163).

A visual condition that is strictly related to occlusion is **transparency**, sometimes also referred to as *semi-occlusion*. However, in the presence of transparency things get more complicated. When an object is visible through another object in the scene, the visual perceptual system not only has to organise depth perpendicularly to the line of sight, but also parallel to it, having to represent two or more distinct depths at the same image location. For example, when observing a scene similar to the one of Figure 3.13, our visual system has to segment the image "in depth" to extract the correct depth information and let us infer the different location of all the objects in the view, transparent (i.e. glass) and non-transparent (i.e. painting on the wall).

In order to achieve this, the visual system undergoes a perceptual process known as *scission* [Koffka, 1935] [Heider, 1933], through which a single perceptual property is separated into two or more components. In the case of transparency, a single perceived luminance value is decomposed into different contributing values, one for each "layer" in depth.

It is widely accepted in literature that in order for scission to occur and transparency to be perceived, there are several conditions that must be met [Singh and Anderson, 2002] [Westland et al., 2002]. According to Metelli [1970], who proposed the first quantitative model of the constraints for transparency perception and whose theory forms a common base to most modern theories in the field, these conditions can be divided into two distinct

Figure 3.13: Transparency. The painting on the wall is visible through the transparent glass.
[From http://www.thermoglance.com, last accessed: 12.01.2011]

categories. The first category of rules is usually referred to as *photometric conditions* and expresses the necessary relations between the different levels of luminance that characterize different areas in an image. Specifically, the photometric conditions for transparency state that the transparent surface can not increase the contrast nor alter the polarity of the contrast of the surface visible through it. In literature, these two photometric constraints are also referred to as the *magnitude constraint* and the *polarity constraint* respectively [Singh and Anderson, 2002].

The second set of rules for perceptual transparency goes under the name of geometrical or *figural conditions* [Metelli, 1974a] [Metelli, 1974b] [Kanizsa, 1979], which, as the name suggests, express the geometrical relations that must hold between different areas of an image. Similarly to the previous category of conditions, the figural conditions can be summarised with a number of constraints. In particular, for an area in the image to undergo scission it is necessary that the contour of the underlying layer does not present discontinuity where it meets the transparent layer in front of it. Likewise, the transparent layer in front must present good continuity at the location where it meets the underlying layer.

Failure to satisfy these two figural continuity constraints may reduce or eliminate the percept of transparency. Figure 3.14 summarises the conditions for perceptual transparency described so far.

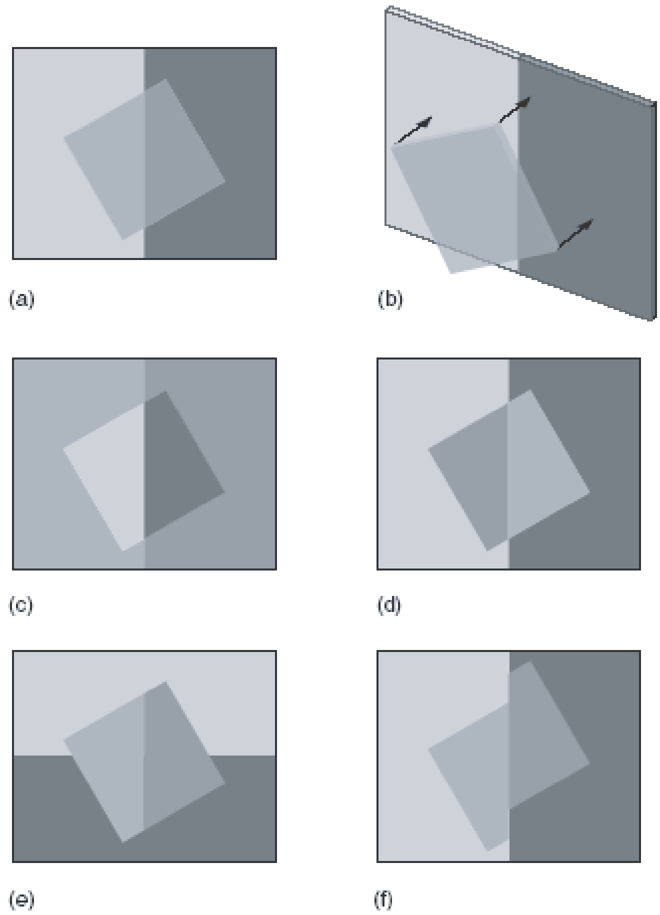
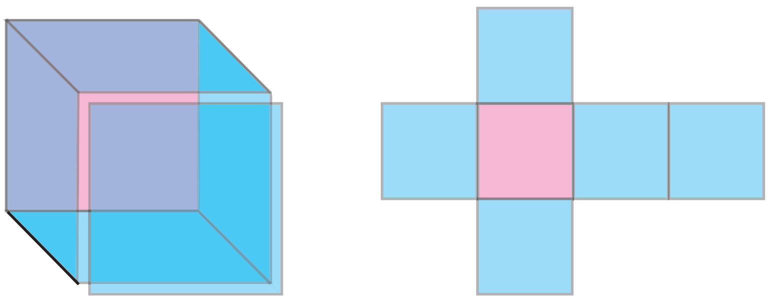


Figure 3.14: Conditions for perceptual transparency. All the conditions are met (a), perspective view of the scene (b), violation of the the photometric magnitude constraint (c) and of the the photometric polarity constraint (d), violation of the figural conditions for the underlying layer (e) and for the overlaying layer (f).
[Adapted from [Fleming and Anderson, 2003]]

An interesting fact is that in the presence of contrasting cues to transparency, figural conditions may override photometric conditions. In particular, in this regard Beck and Ivry [1988] found that when observing experimental stimuli that met the figural conditions to transparency but violated the photometric polarity constraint, subjects occasionally reported perceiving transparency anyway. This aspect of the perception of transparency is particularly important when dealing with abstract and synthetic stimuli that do not reassemble real life scenes. The model of Metelli is an attempt to describe how perceived transparency works through quantitative constraints based on how we see transparency in real life; if in a synthetically generated scene these constraints are infringed the final percept of transparency may be affected. However, in real scenes, or scenes that reassemble real conditions, the photometric and figural constraints are naturally met and in accordance with each other so scission readily occurs and transparency is perceived.

As far as transparency is concerned, the working environments in which the 3D task in question is performed are characterised by scenes that resemble real life viewing circumstances and do not present contrasting transparency cues. Therefore it is fair to assume that the photometric and figural conditions described above are met and transparency is readily perceived by the operator that performs the 3D selection task.

In computer-generated scenes, when the photometric and figural conditions are met, transparency can be used effectively as a cue to depth distinct from occlusion. Zhai et al. [1996, 1994] demonstrated this with the *Silk Cursor* experiment, where a semi-transparent (i.e. neither fully transparent nor fully opaque) box was successfully used to "capture" a target in a stereoscopic 3D environment. The advantage of using semi-transparency is that it does not completely block the view of objects that lie behind it while it still maintains the same strong depth information, formalised in the contrast depth asymmetry principle, that characterises total occlusion. However, transparency alone may introduce some issues. Specifically, when semi-transparency is applied to the faces of a box-shaped 3D cursor used to enclose a target like in the experiment of Zhai et al., the user has to "register" different levels of contrast in accordance to the scission principle described above, a process that for the sake of this discussion is defined as *contrast registration process.* In particular, when the target is characterised by the highest level of contrast it means that it lies in front of the cursor, while the middle and the lowest contrast conditions correspond to the target being inside and behind the cursor respectively. In order to achieve this the user has to interactively move the cursor through the target and correctly associate the three levels of contrast to the appropriate target position; failure to do so could have a detrimental effect on task performance. In their experiment Zhai et al. indirectly made the contrast registration process easier for the user by assigning the same level of transparency to all faces of the cursor box but the back face, which instead was characterised by a higher level of transparency. However, the ease of the contrast registration process could be further improved by assigning a colour to the rear face of the cursor box that differs from that of the other faces, as shown in Figure 3.15.



(a) **(b)**

Figure 3.15: Box-shape cursor with semi-transparent coloured faces (a) and relative face organisation (b).

3.2.7 Summary of Visual Factors

This first part of the chapter thoroughly analysed the most influential visual factors for depth perception in both real scenes and synthetic scenes presented via stereoscopic 3D displays; this formed the basis for a first set of recommendations for the design of a stereoscopic 3D cursor to support tasks that require accurate depth judgements in stereoscopic 3D environments. The outcomes of this analysis are summarised in table 3.1.

Table 3.1: Summary of influential visual factors for depth perception.

Factor	**Effect on Depth Perception and Recommendations**
Luminance Contrast (Section 3.2.1)	High cursor-background contrast to maximise stereoacuity (max 21 dB). Avoid interocular differences (screen operators for amblyopia).
Spatial Frequency (Section 3.2.2)	Moderate degree of spatial frequency in cursor shape, e.g. boxed shaped cursor. Avoid interocular differences. For working scenes with low spatial frequency content, equalise contrast between left and right view to improve stereo performance.
Colour (Section 3.2.3)	High cursor-rest of the scene colour contrast to ensure segregation (especially in the presence of scenes with low spatial frequency content). Provide visual feedback to mark objects at the same stereoscopic depth as cursor to avoid chromostereoscopic effects.
Perspective, Distance and Depth Constancies (Section 3.2.4)	Avoid wire-frame cursor shapes (especially in the presence of working scenes with weak depth cues). Respect linear perspective and Emmert's laws; ensure cursor movement respects size-distance invariance principle and cursor shape is perspective-correct. Avoid contrasting depth cues.
Blur and Depth of Focus (Section 3.2.5)	Avoid use of defocus blur.
Occlusion and Transparency (Section 3.2.6)	Use transparency to guide target interaction and adopt cursor shape with semitransparent faces. Adopt different colour for back cursor shape face to facilitate contrast registration process.

3.3 Technical Factors

The first part of this chapter presented an analysis of the influential visual factors for depth perception, both in natural scenes and in stereoscopic 3D environments. The aim of this second part of the chapter is to investigate the influential factors that are more intrinsically related to the properties of the 3D display system rather than to the way the human perceptual system works. As explained in the following sections, the technical characteristics of the 3D display system play a primary role in the reproduction of the depth embedded in the stereoscopic 3D images and can deeply affect the perception of depth of the observer. Therefore, when designing a 3D cursor to support depth-based selection tasks in stereoscopic environments, such technical factors must be thoroughly understood and accounted for.

3.3.1 Display Resolution

The term resolution refers to the level of detail that an image or a visual device (e.g. a computer monitor or a TV) can reproduce and it is usually measured in number of pixels per each dimension, i.e. horizontal and vertical. However, this common meaning of the word resolution is misleading because it does not provide any information about the size of the pixels, or **pixel pitch**, and the granularity of the visual information. One can easily infer that the smaller the pixel pitch, the closer the pixels are together, the higher the level of detail that can be reproduced. Therefore, a more accurate way to measure resolution is to calculate the pixel density of the device or image, expressed in Pixels Per Inch (PPI).

In a 3D display system the physical resolution of the screen is a factor that can influence the reproduction and subsequent perception of depth. Just as the resolution of a normal 2D display is determined by the number of pixels and their pitch, the resolution of a stereoscopic 3D display is determined by the number of **stereoscopic voxels** and their geometrical characteristics [Hodges and Davis, 1993]. Stereoscopic voxels, or simply voxels,

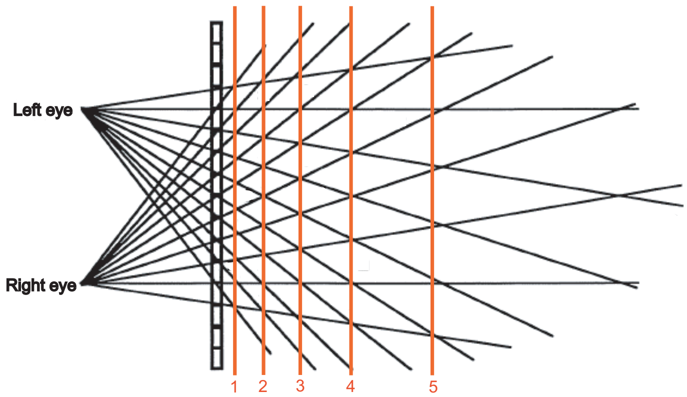


Figure 3.16: Stereoscopic voxels in a cross-section of a 3D display viewing volume. [Adapted from [Hodges and Davis, 1993]]

divide the viewing volume of a 3D display system into depth layers (Figure 3.16).

As the picture illustrates, the number of depth layers directly depends on the depth span of the voxels and defines the resolution of the 3D display in the third dimension. In turn, the depth span of a voxel for a given disparity between corresponding pixels in the left and right view depends on the horizontal pixel pitch. Therefore it can be concluded that the stereoscopic resolution of a 3D display system directly depends on the number of pixels per view or, in other words, on the horizontal screen resolution of the 2D views that form the stereoscopic image delivered to the user. In particular, by adopting the same logic used for 2D displays it can be easily understood that 3D display systems with a higher number of pixels per view have better stereoscopic resolution.

Horizontal screen resolution per view is a factor that can influence the level of accuracy of the 3D display system in reproducing the depth information contained in a stereoscopic image. Therefore, when a 3D display system is used in order to perform tasks based on depth judgements, factors such as its horizontal screen resolution per view become critical and must be taken into consideration as they could deeply affect performance and task accuracy.

3.3.2 Sampling and Aliasing Artifacts

In a 3D display system, or more generally in any electronic display system, the visual information contained in the source image is sampled by the framebuffer of the graphics card and then sent to the display device where it gets reconstructed and displayed to the viewer. In these terms, an image can be treated as a continuous signal that is transformed into a series of discrete values. This process goes under the name of *sampling* and its theoretical principles are formalised in the Nyquist-Shannon Sampling Theorem [Shannon, 1949] [Nyquist, 1928].

Static images are sampled only spatially, while animated sequences of images presented over time (e.g. a video clip or a movie) are sampled both spatially and temporally. When not performed adequately, spatial and temporal sampling may introduce visual artifacts in the reconstructed image. Such artifacts go under the name of *aliasing* and may appear in the reconstructed image as jagged edges, loss of details, disintegrating textures, discontinuity in long, thin objects and disappearance of small objects [Crow, 1977]. In the case of spatio-temporally sampled images further aliasing artifacts include objects moving with jerky motion, reversal of the direction of movement, small objects flashing on and off the screen intermittently and, for bigger objects, changes in size and shape [Bex et al., 1993] [Crow, 1977].

Finally, in the presence of stereoscopic images sampling errors can lead to a further level of aliasing artifacts, which can affect the disparity information contained in the image. These include horizontal and vertical disparity inconsistency between left and right view, size inconsistency, edge inconsistency and inconsistency between disparity and other depth cues, and can affect the perceived position, size, orientation and depth of objects and hinder the ability of the viewer to fuse the two views [Pfautz, 2002] [Castle, 1995]. Besides, sampling errors can also be the cause of *interperspective aliasing artifacts*, i.e. discontinuity between adjacent views in a stereoscopic display [Moller and Travis, 2005].

A detailed discussion of sampling artifacts in digital imaging is presented by Pfautz [2002], who proposes two methods to counterbalance the effects of aliasing in static images, animated images and stereoscopic images (both static and animated). Issues related to interperspective aliasing in stereoscopic images are discussed and tackled by Moller and Travis [2005], while Sung et al. [2002] propose a new method to diminish motion jerkiness in 2D animated images.

Even though interesting, a detailed analysis of sampling artifacts is not in the scope of this dissertation. However, from the brief review presented in this section it is easy to conclude that aliasing is a factor that can not be ignored when dealing with depth-based tasks in stereoscopic environments. With specific regard to the design of a 3D cursor to support such tasks, it is recommended that an anti-aliasing algorithm is to be used in order to avoid artifacts and inaccuracies in the reproduction of the cursor shape and depth.

3.3.3 Image Interleaving

In two-view 3D display systems that rely on a single LCD panel and do not adopt alternate-frame sequencing as a technique to deliver the stereoscopic image, the left and right views are spatially interleaved (see section 2.3.3 for details). That is, there must be a mechanism in place that eliminates every second column (or row) of pixels in the two views respectively and combines them into a single interleaved image, so to meet the pixel resolution limit of the LCD panel. There are different ways to achieve this, using hardware, software or both, and different interleaving algorithms can be adopted depending on the design of the 3D display system. Some of the graphics card drivers available on the market do support a number of different stereoscopic viewing modes and automatically apply an interleaving algorithm to the source image, without requiring the user to perform any manual image interleaving. At the time of writing this thesis, the NVIDIA Quadro drivers support the following interleaving modes: vertical interlaced stereo, color interleaved stereo and horizontal interlaced stereo.

Although very often the 3D display systems that are more accessible, both financially and in terms of ease in using the device, rely on image interleaving, the effect of the different interleaving techniques on the final depth percept of the user has not been thoroughly understood yet. The literature currently available in the field does not offer an in-depth analysis of the matter and it becomes apparent that there is a clear lack of understanding that could be addressed. On the basis of the work carried out by Pfautz [2002] it is easy to anticipate that if the interleaving technique used introduces sampling errors, then the reconstructed stereoscopic image could potentially present aliasing artifacts, which in turn would introduce inaccuracies in the reproduced depth of objects and hinder the depth perception of the observer. Accordingly, it is expected that such artifacts will lead to potential inaccuracies in the shape, size and location of the 3D cursor and therefore have a detrimental effect on task performance.

3.3.4 Summary of Technical Factors

The second part of this chapter investigated how depth perception in stereoscopic 3D environments can be affected by a number of technical factors, i.e. factors related to the technical characteristics of the 3D display system used. It can be anticipated that these play a primary role in the reproduction of the depth embedded in the stereoscopic 3D image and can deeply affect the amount of depth perceived by the observer. Notwithstanding, the scientific literature in the field is limited and was not sufficient to gain a clear understanding of how the effect of such factors could be counterbalance when designing a 3D cursor to support depth-based tasks. The outcomes of this analysis are summarised in Table 3.2.

Table 3.2: Summary of influential technical factors for depth perception.

Factor	**Effect on Depth Perception and Recommendations**
Display Resolution (Section 3.3.1)	The stereoscopic resolution of a 3D display system depends on the resolution of the underlying 2D screen(s). The horizontal screen resolution per view can affect the depth information contained in the displayed 3D image.
Sampling and Aliasing Artifact (Section 3.3.2)	Image sampling errors can lead to aliasing artifacts in the displayed 3D image and affect the amount of perceived depth. Adopt anti-aliasing techniques to avoid artifacts and inaccuracies in the reproduction of cursor shape and depth, and of objects in the working 3D scene.
Image Interleaving (Section 3.3.3)	Image interleaving may lead to aliasing artifacts due to image sampling errors, which can introduce inaccuracies in the reproduced depth of objects and in the shape, size and location of the cursor.

3.4 Synthetic Cues

The previous two sections of this chapter analysed and discussed the effect of different visual cues and technical factors on the perception of depth. The focus of this section is on a third category of factors that go under the name of *synthetic cues*.

Synthetic cues are particular stimuli that are added to the 3D environment on top of the natural cues in order to assist the user in performing a task. Depending on the nature of the stimuli and the sensory system they are targeting, synthetic cues can be further divided into *haptic cues*, which are related to the sense of touch, *aural cues*, if they target the sense of hearing, and *visual cues*, if instead they are related to the sense of sight.

Synthetic cueing has been used in a variety of applications in order to improve visual task performance. Bachelder and McRuer [2002] showed that synthetic cueing augmentation could be successfully employed in helicopter flying simulations in order to improve hover performance in night-vision operations using night vision devices. Similarly, Tannen et al. [2004] tested the effect of integrating the visual displays with spatial aural cueing in a simulated flight task where subjects were asked to maintain flight parameters while searching for air and ground targets; on the basis of the experimental results the authors concluded that coupled visual and spatial aural displays are an effective means of providing information about the target location. In the context of medical imaging, synthetic aural information can be used in addition to visual assessment of ultrasonic images as a tool to diagnose blood vessel damage or diseases, or as a supplement tool in the interpretation of radiological images [Chun, 2006].

With specific regard to selection tasks in stereoscopic 3D environments, in a recent study Vanacken et al. [2009] investigated the use of synthetic cues as a form of feedback to support the selection of objects in conditions of dense target environment or limited target visibility. In particular, this study has shown that visual feedback, when adequately provided, can significantly improve selection task performance, while on the other hand haptic and aural feedback did not have any significant effect on subjects' performance. Interestingly, some subjects reported that they did not like the presence of haptic feedback in some experimental conditions, whereas most subjects seemed to like the conditions where aural feedback was present and felt that it was useful despite its negligible effect.

In this regard, Bolia et al. [1999] showed the benefit of auditory feedback as an additional cue to support selection tasks. In this study, subjects were asked to perform a visual search task within a geodesic sphere (i.e. no use of stereoscopic devices) with and without spatial aural feedback; the empirical results indicate that the presence of the aural feedback significantly decreased task execution time, with no significant difference in term of task accuracy between the two conditions.

On the basis of the above discussion, it is recommended that a number of strategies are employed when designing a stereoscopic 3D cursor in order to support accurate stereoscopic depth tasks. First, it is suggested that the cursor makes use of visual feedback as a way to provide the user with information about the spatial location of target objects in the 3D scene. Furthermore, for target selection tasks based on depth judgements, an effective design strategy would be to highlight potential target objects within the volume of influence of the 3D cursor that are located at the same "real depth" as the cursor. It is important to clarify that in this context the term "real depth" indicates the depth that characterises the object in the source stereoscopic image, which in turn depends strictly on the amount of image disparity that characterises that object, and not the object's perceived depth, which instead can be biased by a number of factors, especially technical factors, as explained in section 3.3.

Second, it is advised that the cursor provides aural feedback when the user selects a specific target object within the 3D scene. Generally, people seem to like aural feedback and perceive it as helpful, even though the study of Vanacken et al. [2009] suggests that its effect on task performance is not significant. Furthermore, using aural feedback as a means of informing the user that selection has occurred could help in reducing the strain on the visual sensory system in conditions of overloaded visual information. Finally, one last advantage is that aural feedback does not require any specialised equipment and can be reproduced with no overhead costs [Vanacken et al., 2009]. Empirical results reported in literature are discordant about the effect of this type of feedback on task performance, however it is believed that for the purpose of this thesis it could be both beneficial and complementary to visual feedback.

Haptic feedback on the other hand does require specialised hardware, which can be difficult to use or require a considerable amount of training before it can be used effectively. Besides, haptic feedback seems to have no significant impact on 3D selection performance and it is not always appreciated by users [Vanacken et al., 2009]. Therefore, the last recommendation in terms of synthetic cueing is that a 3D cursor used in carrying out selection in a stereoscopic 3D environment is not required to provide haptic feedback. While in other application fields the use of haptic feedback could be very beneficial for the user or even essential, e.g. virtual motor rehabilitation systems [Kayyali and Shirmohammadi, 2009] and surgical training in 3D environments [Meier et al., 2001], for the type of tasks considered in this project it will not be studied further.

3.5 General Discussion

The factors that can affect depth perception are numerous and the way they interact with each other can be extremely complex and difficult to predict. Overall, when reproducing a synthetic stereoscopic scene on a 3D display system, conflicting depth cues should be avoided, and linear perspective laws and Emmert's law should be respected as much as possible.

In terms of 3D cursor design, it is recommended to adopt a simple 3D shape with limited amount of detail, for this will facilitate the fusion of the stereoscopic image and ensure that the visual artifacts introduced by potential aliasing errors are not exacerbated. On the contrary, completely smooth 3D shapes (e.g. a sphere) and extremely detailed 3D shapes (e.g. a 3D star) should be avoided as they could both hinder the user's stereoscopic performance.

The level of luminance contrast between the cursor shape and the scene background should be boosted in order to maximise stereoacuity, especially in the presence of high frequency content, i.e. 5.0 c/deg or more, in the working scene. However it is suggested that this is no higher than 21 dB of the luminance contrast detection threshold of the cursor shape. Where possible, the cursor colour should be complementary to the colour of the other objects in the 3D scene, in order to guarantee a significant level of colour contrast and ensure that the cursor is easily identified. This aspect is particularly important in working environments characterised by low spatial frequencies, i.e. 2.0 c/deg or less, (e.g. ophthalmology), where the spatial analysis capabilities of our luminance system are limited. Furthermore, no interocular differences in cursor shape, luminance and colour should be shown as this could again have a detrimental effect on the user's depth perception and task performance. For similar reasons, wireframe shapes should also be avoided, especially in the presence of working scenes where depth cues are weak or absent (e.g. some geoscientific applications).

With regard to blur and DOF, it is recommended to avoid the use of image blur as a cue to depth for the 3D tasks in question. On the other hand, transparency could be applied to the faces of the 3D cursor shape and serve as an effective cue for guiding the user during selection. Lastly, synthetic cueing could be used as a valid form of feedback to provide the user with target spatial location information.

In the light of the discussion above, a 3D box-shaped cursor with semi-transparent faces seems to represent a good candidate. The space inside the 3D box would represent the "volume of influence" of the cursor and the centre of the box the point at which its disparity is calculated. Objects located within this volume would be under the influence of the cursor and can be selected by the user; control over the size of the volume of influence could be granted to the user. In order to provide the user with target location information, visual feedback could be used to highlight any objects within the volume of influence that share the same depth as the cursor (i.e. are characterised by the same amount of image disparity as the cursor); a potential form of visual feedback would be to make the objects of interest blink on and off the screen alternatively. This measure would also counterbalance the negative effects of potential chromostereoscopic phenomena on depth perception, which could hinder task performance. The user could perform the actual target selection via a simple mouse click. When a click is performed, the nearest object to the centre of the cursor would be selected. Once the user has selected a target object, aural feedback could be provided and the colour of the selected object would change in order to communicate that selection has occurred. To assist the user in understanding which target objects are within the cursor volume of influence, it is suggested to render the rear face of the cursor box with a different colour than the other faces.

The prototype suggested above is best suited for tasks such as selection and target acquisition. However it can be used effectively even to delineate regions of interest in a stereoscopic image, a task often carried out in ophthalmology and more generally in medical imaging.

Finally, on a more general level it is recommended that operators required to perform the task in question are screened for amblyopia, as this can seriously impair stereoacuity and therefore task performance.

3.6 Conclusions

This chapter presented a review of the critical factors that can affect the perception of depth, with specific focus on depth perception in stereoscopic 3D environments. The vision science literature on the influential visual factors is copious and forms a solid basis to justify the suggested guidelines in terms of 3D cursor design. On the other hand, a lack of scientific literature was identified regarding the effects of the influential technical factors on the perception of depth when using a 3D display system; the knowledge is fragmented and there is no thorough understanding of how such factors can affect the design of a 3D cursor to support depth-based tasks in stereoscopic 3D environments. Synthetic cueing has also been discussed as a third category of factors that could be employed in order to support the user in performing such tasks.

On the basis of the reviewed literature, a number of guidelines for the design of a stereoscopic 3D cursor have been identified and the features of a cursor prototype suitable to support depth-based selection tasks in stereoscopic 3D environments have been described. It is believed that the work presented in this chapter is valuable for designing a 3D cursor for stereoscopic 3D environments and assist operators in performing accurate depth tasks. However, a better understanding of the technical factors is needed before the effects on the quality of the final 3D image can be thoroughly understood and the cursor design guidelines finalised. This knowledge gap forms the basis for the work presented in Chapters 4 and 5.

Chapter 4

Assessment of 3D Technologies: Methodology

4.1 Introduction

From Chapter 3 it emerged that the knowledge about the effect of different 3D display technologies on depth perception in stereoscopic 3D environments is still fragmented and incomplete. As 3D displays become increasingly available and start being used in applications where depth judgements are critical, a good understanding of such effects becomes fundamental. This chapter presents a robust and sensitive methodology for the investigation of human depth perception on different 3D displays, which can be used to assess how well different systems can reproduce the depth present in an input stereo image. It is anticipated that technical factors such as display resolution, image interleaving and aliasing may have a detrimental effect on the quality of the displayed 3D image and alter the shape, disparity and appearance of the objects it contains. With specific regard to 3D cursor design, these factors could pose a limit on the required minimum cursor size and affect its perceived position in depth. The work presented in this and the following two chapter will help to clarify and quantify these aspects.

This chapter is organised as follows. Section 4.2 summarises previous comparative studies on human depth perception in stereoscopic environments. The method adopted in order to compare and assess different 3D displays is presented in Section 4.3, while section 4.4 explains the rationale behind the predictions for this study. Section 4.5 reports the details of the pilot experiment performed to assess and refine the methodology. Lastly, section 4.6 concludes the chapter.

4.2 Background

Previous studies of human depth perception on 3D displays have almost exclusively studied single displays and have typically investigated fusion limits, i.e. the highest levels of image disparity a 3D display can support before fusion brakes down [Jones et al., 2001] [Holliman, 2006].

Yeh and Silverstein [1990] and Woods et al. [1993] studied fusion limits for stereoscopic desktop 3D displays. The results showed that the total range of depth comfortably viewable on a 3D display is limited. Similar results were demonstrated for autostereoscopic displays by Jones et al. [2001] who suggested a working perceived depth range of as little as 60 mm behind and 50 mm in-front of the display surface. Whatever the precise value of fusible range for desktop 3D displays, it is clearly limited and therefore it becomes increasingly important to understand how this limited range is represented on different 3D displays.

A useful discussion of experimental design for stereo imaging trials is presented by Hsu et al. [1996], however much of the discussion relates to LC shutter glasses, the predominant desktop 3D display technology at that time.

Rosen et al. [2004] conducted a study where the human perception of 3D spatial relations was used to compare a Perspecta Spatial 3D System [Favalora, 2002] against a normal 2D LCD display. The results indicate that under some specific experimental conditions performance for the Perspecta display was better, but the study does not seem to adopt a statistically robust approach. Besides, results and discussion are restricted to the Perspecta 3D system and are not generalisable. In a recent study Alpaslan et al. [2006] compared task performance on a 2D display, a LC shutter glasses stereoscopic display, a two-view autostereoscopic display and a multi-view autostereoscopic display. The authors investigated interaction performance in a trial where participants had to manipulate a 3D object to be in the same depth plane as a target object. The results suggest that better performance was obtained using the LC shutter glasses, but the study does not provide a hypothesis predicting this nor explain why this might be the case.

Grossman and Balakrishnan [2006a] also investigated 3D depth-judgement task performance using different 3D display techniques. During the experiment participants were asked to perform three different tasks (a depth-ranking task, a path-tracing task and a potential collision task) using a volumetric display, a LC shutter glasses stereoscopic display with head tracking, a LC shutter glasses stereoscopic display without head tracking and a 2D standard perspective display respectively. Generally, the results showed that the volumetric display had slightly better performance than the stereoscopic display with head tracking (even though this difference was not always significant), followed in order by the stereoscopic display without head tracking and, lastly, the 2D display. Similarly to the previous work though, the authors do not provide a valid hypothesis predicting these results.

None of the aforementioned studies attempted to quantify human depth perception threshold levels across a range of representative categories of desktop 3D displays, which is one of the objectives of this thesis. To empirically investigate and quantify depth perception is critically important, as it is believed that the different hardware, software and optical characteristics of the display system may have an effect on the quality of the depth representation and, consequently, on the design characteristics that a 3D cursor should have in order to effectively support the user in tasks that require accurate depth judgements.

In order to investigate perceived depth on different types of 3D displays, a new methodology was developed and an empirical study was performed. The ultimate aim of this study was to quantify inter-display human binocular depth perception differences and understand the role played by the different design characteristics of the tested systems on such perception differences. The methodology adopted during this study is described in detail in the following sections of this chapter, while the empirical research and relative results are presented in Chapter 5.

4.3 Research Method

4.3.1 Experimental Design

For the purpose of this comparative study, a **repeated-measures design** was adopted. Repeated-measures designs, as compared to **independent designs**, are statistically more robust and reduce unsystematic variability, providing greater statistical power to detect the effect of the different factors observed during the study [Field, 2005]. Data collected during the trials were analysed using a two-way repeated-measures ANalysis Of VAriance (**ANOVA**) and **Student's t-test** statistics [Student, 1908].

During the study two independent variables were manipulated: the type of display used to perform the task (*Display*) and the amount of disparity present in the input stereo image used as stimuli (*Disparity*). The dependent variables were the proportion of trials at which participants selected the correct target (*Score*) and the amount of time elapsed between when the stimuli was first displayed and when the subject selected their answer (*Response Time*); score and response time were used as measures of task accuracy and task difficulty respectively (high score = high task accuracy, high response time = high difficulty in performing the task). Each subject was asked to repeat the same task 28 times on each display for each level of disparity, i.e. for each experimental condition.

Image disparity, and hence perceived depth, was controlled in pixels and was randomly chosen within a range of possible disparity levels, each of which was distributed across the trials with equal probability. The order in which people performed the task on each display was counterbalanced and followed a **Latin square** design [Raghavarao and Padgett, 2005] in order to reduce noise and unsystematic variability in the collected data.

4.3.2 Task and Stimuli

The image used for the study consisted of two white squares on a black background, as shown in Figure 4.1. The squares were centered in the middle of the screen and were positioned horizontally one next to the other. Between the two squares there was a small square that marked the center of the screen and acted as the fixation point; participants were asked to maintain fixation on this point throughout each trial as they were performing the task. The square that acted as fixation point was 6 pixels wide, while the width of the other two squares was 64 pixels each. The size of the squares are integer multiples of 1 pixel so that the test image does not present artifacts due to sampling errors when the video signal is digitalised.

In each trial one square was always positioned on the display plane while the position in depth of the other square was randomly chosen among different crossed-disparity levels of depth in front of the screen. For half of the trials disparity was shown on the left square while for the other half disparity was shown on the right square. In this way the position of the square (i.e. left or right) that appeared to be closer to the participant was counterbalanced across trials. Stimuli were presented via the stereoscopic 3D displays and participants were asked to identify which square was the closest to them, using the keyboard for input. In particular, they were asked to press the letter C on the keyboard if the left square appeared to be closer to them or press the letter M if instead the right

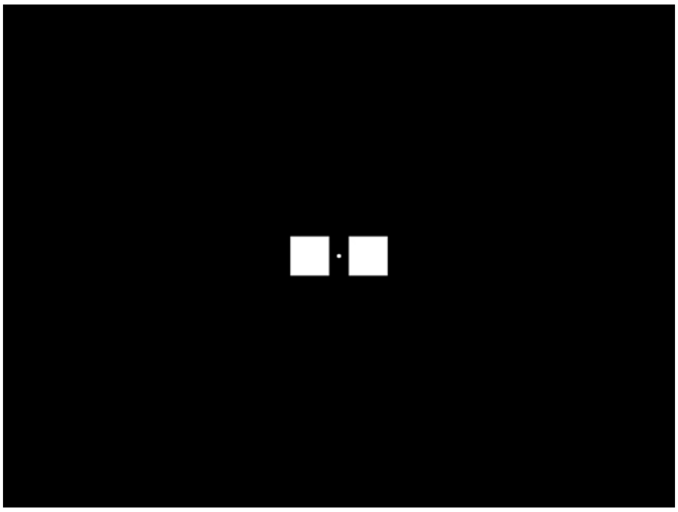


Figure 4.1: Stimuli used in the 3D display comparative study.

square appeared to be closer. In cases where they could not detect any difference between the two squares they were asked to guess (forced choice). Between two consecutive trials a black and white mask with no embedded depth information was shown on the 3D display for the duration of one second so as to remove any stimulus afterimage effect, which could introduce an unwanted bias in the results.

When defining task and stimuli for this study, a number of factors had to be considered carefully and choices had to be made accordingly and accurately so as to reduce potential biases in the results. A first important choice was about how to draw the stimuli, i.e. whether to use 3D graphics or 2D graphics. After a meticulous analysis it was decided to use simple 2D squares, as the use of 3D shapes (e.g. a box) would have introduced issues related to factors such as the type of perspective to be used, which in turn could have caused unwanted effects on the empirical results.

Another crucial choice concerned whether the stimuli should be drawn using a depth mapping algorithm [Froner and Holliman, 2005] with distances controlled in centimetres or by simply controlling the size and position of the shapes in pixels. Finally it was decided to control the size of the squares and the level of disparity applied to the stimuli in pixels rather than in centimetres, so as not to introduce artifacts due to 3D graphics rounding errors. In order to obviate this rounding problem the adoption of antialiasing was also considered, however with the danger of introducing further problems as it was soon realised. This is because in computer graphics antialiasing techniques are based on blurring and, as it was explained in section 3.2.5, blurred edges can deeply affect depth perception leading to potential unwanted effects on the empirical results. Besides, edge blurring would have introduced a certain amount of ambiguity when measuring the exact amount of disparity to be applied to the stimuli. On the other hand, quantifying disparity and square sizes in pixels is unambiguous and therefore it is believed that it was the right thing to do.

The choice of the task is also justified for several reasons. Firstly, the task used in this study is non-application specific and therefore requires no domain knowledge from the participants. Secondly, it is both unambiguous and simple to learn and perform. Finally, it forces the participant to use stereopsis as a unique cue to depth. In accordance with this last point, the stimuli used for the study did not present any of the additional depth cues that are normally associated with stereopsis in everyday scenes (see Chapter 3). This also justifies the choice of using a black and white image instead of a chromatic one that could instead contain some unwanted cue to depth due to chromostereoscopic hue effects, as explained in section 3.2.3. Furthermore, black and white stimuli also maximize the image contrast, which in turn is known to have a positive effect on stereoacuity, with better performance observed at higher contrasts (see section 3.2.1). These choices make the task suitable to measure depth perception only in relation to binocular disparity and to assess how well different 3D display systems support fine depth judgements at different image disparity levels, which is the purpose of the comparative study presented in this thesis.

4.3.3 Procedure

Volunteers were screened for stereo vision using the **Titmus** and the **TNO** [von Noorden and Campos, 2001] stereoacuity tests (Figure 4.2); candidates with suitable vision were then selected as participants in the study. Prior to the start of the experiment, participants reviewed the instructions and completed a number of practice trials in order to gain confidence with the 3D displays and correctly understand the task. Once ready, they completed the same number of trials on each display.

a

b

Figure 4.2: Titmus test (a) and TNO test (b).
[From http://www.haagstreituk.com, last accessed: 16.01.2010]

Trials started with an **orthoscopic test** on each display, where participants were shown a red box on black background in the left view and a blank screen in the right view. They were then asked if they could see the correct image in the correct eye and if necessary the left and right views were swapped prior the start of the trials. During the experiment head movements were minimized via the use of chin rests.

Participants were instructed to be as accurate as possible in their decision but not to spend too much time on each trial, even though no time limit was imposed. Answers could not be changed and both score and response time were recorded. In each trial, participants were assigned a score of 1 if they gave the correct answer and a score of 0 if they gave the wrong answer. In the trials where both squares had zero disparity, the key pressed was recorded and the score was calculated by always assigning right as the correct answer in the pilot experiment, and by using a randomization algorithm in the main experiment. This algorithm randomly assigned *left* as the correct answer for half of the trials and *right* as the correct answer for the remaining half of the trials. Note that both methods are acceptable as the number of repetitions was sufficiently high to ensure statistically reliable results in both cases. However, the randomization algorithm made it possible to obtain scores not different than chance (i.e. 50%), both statistically and numerically, even for participants that were systematically choosing left (or right) as an answer for all the trials where the difference in depth between the two squares was in fact not discernable. Response time was measured in milliseconds (ms).

At the end of each experimental session participants were asked to complete a questionnaire about the display they had just used (ease seeing in 3D, disturbing factors, level of discomfort, general comments about the display). They were also asked to fill in a more general survey at completion of the whole experiment and rank the displays from worst to best; to do so, participants were allowed to revisit the equipment at the end of all trials in order to remind themselves of the characteristics of the seven displays. Where possible, the questionnaires used a typical five-level **Likert scale** [Likert, 1932]. Finally, all participants were fully debriefed and were given the chance to ask questions. The questionnaires and forms used during the study are reported in Appendix B.

4.4 Geometric Predictions

This study focuses on the comparison of two-view desktop 3D displays. As the name suggests these systems generate two distinct views, one for the left eye and one for the right eye, each of which consists of a distinct set of pixels and is shown in a separate viewing window (see section 2.3.3). The design of a 3D display has an influence on how these two views are reproduced and therefore on the resolution and size of the 3D image when it is finally shown to the observer. For this reason, in order to predict the performance expected in terms of human depth perception on each system included in the study it was first necessary to carefully analyse the different typologies of two-view 3D displays based on their technical characteristics. Before proceeding with the analysis, a short section with the definition of the technical terms used to identify key features of the 3D displays and the image displayed on them is presented.

4.4.1 Terminology and Definitions

The terminology used in this thesis extends that used by Holliman [2006], in order to distinguish between the input stereo image that is to be fed to the 3D displays and the capabilities of the displays themselves.

For clarity, the definitions that are relative to the input stereo image are identified by the prefix *image*, while the definitions that relate to the display characteristics and the properties of its physical viewing windows are identified by the prefix *view*.

- *Image pixel* - Single pixel unit in the input stereo image that is to be displayed. It is considered to be the basic unit of addressable colour in one channel of the input stereo image. Note that an input stereo image for a two-view 3D display consists of two channels: a left and a right.

- *Image resolution* - Resolution of a single channel of the input image that is to be displayed. It is defined in image pixels and it is the same for both channels of the input image.

- *Image disparity* - Disparity between two corresponding points in the input stereo image to be displayed. It is measured in image pixels.

- *View pixel* - Basic addressable unit in a single view on a specific 3D display. Depending on the optics of the 3D display, view pixels can have different size and aspect ratio than the physical pixels of the underlying display screen.

- *View resolution* - Resolution, measured in view pixels, that can actually be displayed in each view on a specific 3D display and that is physically directed to each eye. Again, depending on the design of the 3D display this can be different from the resolution of the underlying display screen.
- *View disparity* - It is the physical disparity between two homologous points shown on a 3D display and it can be measured in view pixels or, being a physical quantity, in millimetres (mm). When compared among each other, it is expected that different 3D displays will yield different minimum values of view disparities, depending on their design.

If not specified otherwise, when the term disparity is used in this thesis it refers to image disparity, while the term display refers to a 3D display.

4.4.2 3D Display Classification

The two-view 3D displays available on the market at the time of the study can be classified into four different categories or *classes* based on the physical interleaving pattern used and, consequently, on the view resolution that they can support, as described below.

Full-Resolution Displays These displays show one full resolution view, both vertically and horizontally, for each eye for a total of two views, and provide double the number of pixels of an equivalent 2D display. This may be implemented by using two independent LCD panels, one for each view, or via temporal multiplexing and a single LCD panel.

Row-Interleaved Displays These displays provide two views, each with full resolution horizontally and half resolution vertically. This is achieved by spatially interleaving alternate rows of image pixels from the left and right channels of the input stereo image. The total number of pixels seen by the observer is unchanged as compared to an equivalent 2D display, with half of the pixels seen by the left eye and the other half seen by the right eye.

Column-Interleaved Displays These displays provide half horizontal and full vertical resolution in each of the two views by spatially interleaving alternate columns of image pixels from the left and right channels of the input stereo image. Again, compared to a 2D display of equivalent size the total number of pixels seen by the observer is unchanged, with half of the total seen by each eye respectively.

Colour-Column-Interleaved Displays Like the previous category, these displays show two views with half horizontal and full vertical resolution, and provide the same number of pixels of an equivalent 2D display with half of the total seen by each eye respectively. They spatially interleave left and right pixels from the input stereo image in alternate colour-columns at sub-pixel level.

Colour-Column-Interleaved displays could be considered a subcategory of the Column-Interleaved displays, but for the sake of this analysis and later discussion it is clearer to keep them in two distinct classes.

4.4.3 Theoretical Predictions and Hypotheses

presented via a 400x300 pixel window per image channel. This window is centered in the resolution of the underlying screen of the specific 3D display. hardware interface and the display

For the purpose of this study, the image resolution for the stimulus is defined as independent of the displays. During the trials, the stimulus is presented via a 400x300 pixel window per image channel, centered in the middle of the screen. The pixel size of this window is kept constant across displays, while the total pixel size of the stereo pair varies across displays and is equal to the native resolution of the specific display considered. The actual view resolution seen per eye is dependent on the software drivers, the hardware interface and the display optics.

The option of keeping the total pixel size of the stereo pair constant across the displays was also taken into account. For the purpose of this study though, the native display resolution is considered to be part of the technical characteristics of the display itself and therefore, in order to conduct a fair comparison, it should not be changed. For this reason it is believed that the choice of keeping the pixel size of the window in which the stimuli are presented steady and to use the native display resolution as total pixel size of the stereo pair was the correct design decision.

On the basis of these clarifications, on a Full-Resolution display the stimuli are shown in a 400x300 pixel window in each channel (image resolution) and each eye physically sees a stimulus of 400x300 view pixels (view resolution). On a Row-Interleaved display the aspect ratio of the window remains 4:3 but each eye physically sees 400x150 view pixels. Finally, on a Column-Interleaved display or a Colour-Column-Interleaved display the aspect ratio of the window remains 4:3 but each eye physically sees 200x300 view pixels. In the latter two cases the loss in view resolution (as compared to the original image resolution) does not mean a change in the shape of the displayed image as the aspect ratio of the physical screen image remains the same due to the effect of the interleaving.

Based on the above discussion, performance for any two-view 3D display can be predicted in terms of which levels of input image disparity can be reproduced as view disparity and hence perceived as depth by the observer, starting from the display published specifications. In particular, the theoretical predictions for each class of 3D display described in section 4.4.2 can be summarised as follows.

Full-Resolution Displays This class of displays is expected to have the capability to reproduce all input disparity levels as equivalent view disparity levels that will then be perceived by the observer at discrete depths.

Row-Interleaved Displays Because they support full horizontal view resolution, it is expected that these displays should be able to reproduce all input disparity levels as equivalent view disparity levels. Similarly to Full-Resolution displays, it should therefore be possible to show the entire input image disparity range. However, it is worth bearing in mind that these displays may have a built-in vertical offset of one view pixel that could potentially alter the amount of depth perceived by the observer.

Column-Interleaved Displays These displays only have half horizontal view resolution because the input image is subsampled horizontally by a factor of two as a result of the column interleaving. For this reason it is expected that each alternate increment in image disparity will be removed allowing a Column-Interleaved display to reproduce only half of the input image disparity range. This means that the observer should be able to perceive a difference in depth only at every second level of input image disparity, e.g. at 2- 4- 6- etc. image pixel disparity with the same perceived depth for 2- and 3- image pixel disparity, 4- and 5- and so on. Another consequence of this design is that the 0-pixel image disparity plane is not coincident with the physical screen plane but slightly in front of it by an offset of one physical screen-pixel [Holliman et al., 2007].

Colour-Column-Interleaved Displays Similarly to the Column-Interleaved displays, these displays also subsample the input image horizontally. The only difference is that the subsampling is performed at colour-column, sub-pixel level rather than at column-pixel level. This could potentially alter the perception of colour in the final display observed by the user but in terms of depth reproduction it is expected to have the same effect as the column interleaving approach. On the basis of these considerations, the prediction for the Colour-Column-Interleaved displays is that they will only have the capability to reproduce half of the values of the input image disparity range and that differences in depth will only be perceived at every second level of image disparity. As with the Column-Interleaved displays, even on these displays the 0-pixel disparity plane is not coincident with the physical screen plane but slightly in front of it, though in this case the offset is only one third of a physical screen-pixel [Holliman et al., 2007].

One of the objectives of this thesis is to understand how well different 3D display technologies can support visual tasks based on depth judgements. For this purpose, it is extremely important to investigate minimum threshold values of perceived depth in terms of the smallest number of units of input image disparity (i.e. image pixels) that a human with normal vision can perceive as depth. Overall, the prediction is that different 3D displays will reproduce different levels of input image disparity in a different way and that this will directly affect the minimum threshold values of perceived depth. In particular, it is predicted that Full-Resolution and Row-Interleaved displays will have a threshold level of 1-pixel image disparity, while Column-Interleaved and Colour-Column-Interleaved displays will have a threshold level of 2-pixel image disparity due the subsampling of the input image in the horizontal direction.

Detailed predictions in terms of score and response time are summarised in Table 4.1. Here, a score of 100% suggests an expectation that participants will be able to detect depth and correctly perceive the image disparity presented in the stimuli, while a score of 50% is no different than chance and indicates an expectation that participants will guess the answer and will be unable to perceive depth from the disparity presented. Furthermore, it is expected that in the conditions where participants will encounter difficulty to perform the task the response time recorded will be higher than the one associated with the conditions where participants will have no difficulty in performing the task and the

Table 4.1: General predictions for score, S, (%) and response time, RT, (qualitative).

Disparity	0		1		2		...	N	
	S	RT	S	RT	S	RT		S	RT
Full-Resolution	50	high	100	low	100	low		100	low
Row-Interleaved	50	high	100	low	100	low		100	low
Column-Interleaved	50	high	50	high	100	low		100	low
Colour-Column-Interleaved	50	high	50	high	100	low		100	low

depth induced by the amount of disparity applied to the stimuli will be readily perceived. Note that while it is easy to quantitatively predict score, it is not possible to do the same with response time.

Based on the above discussion, displays with full horizontal view resolution (i.e. Full-Resolution displays and Row-Interleaved display) are expected to yield a mean score of 100% and a low response time for the condition with an amount of image disparity bigger or equal to the 1-pixel threshold, and a mean score of 50% and a high response time for the cases where the stimuli are treated with no disparity. On the other hand, displays with half horizontal view resolution (i.e. Column- and Colour-Column-Interleaved displays) should lead to a mean score of 100% and a low response time for the trials treated with an image disparity of 2 pixels or higher, and a mean score of 50% associated to a high response time for image disparities smaller than the 2-pixel threshold level.

4.5 Pilot Experiment

Given the level of complexity of the proposed comparative study, it was important to refine the methodology before proceeding with the main experiment. For this reason, a number of informal experiments and a full pilot experiment were conducted in order to check the feasibility of the main research and the robustness of the experimental design. The following sections describe the details of the pilot experiment [Holliman et al., 2007].

4.5.1 Participants

A total of 14 participants, 11 male and three female, were recruited within the Durham University population. Their age varied between 20 and 34, with a mean age of 26 years. They were screened for stereoacuity using both a Titmus test and a TNO test and all met the minimum criteria for selection, i.e. stereoacuity at 40 seconds of arc. They were naive concerning the purpose of the experiment and received a nominal sum of five pounds per hour, for a total of 10 pounds each.

Participants were divided into two groups of seven people each. The experiment was carried out in two separate group sessions that followed the same experimental protocol, as described in section 4.3.3. The same experimental conditions were presented to all participants.

4.5.2 Experimental Conditions and Trial Blocks

During the pilot experiment, seven different two-view 3D displays were tested with four different levels of input image disparity, for a total of 28 experimental conditions. The four possible levels of image disparity to be applied to the stimuli were 0- 2- 4- and 6-image pixels respectively. The choice to increment the input image disparity by two image pixels ensures an input signal that all displays should be able to reproduce as perceivable depth (see Table 4.1).

Each subject was asked to repeat the task 28 times for each of the four levels of image disparity on each display (i.e. 28 repetitions for each experimental condition), giving a total of 112 experimental trials per display per subject. Both score and response time were recorded.

4.5.3 Apparatus

The chosen set of displays was representative of the four display classes described in 4.4.2 as specified in the following list.

- A time sequential stereoscopic CRT display using CrystalEyes LC shutter glasses [Lipton and Ackerman, 1990], an autostereoscopic Kodak display [Cobb, 2005] and an autostereoscopic IRIS-3D display [McKay et al., 1999] for the Full-Resolution category.
- A ColorLink linearly polarized stereoscopic display for the Row-Interleaved category.
- A DTI 2018 LCD display [Eichenlaub, 1993] and a SeeReal C-i display [Schwerdtner and Heidrich, 1998b] for the Column-Interleaved category.
- A Sharp LL-151-3D autostereoscopic display [Jacobs et al., 2003] for the Colour-Column-Interleaved category.

The displays were driven by seven independent machines that used the same kind of graphics card (NVIDIA Quadro FX family) and the same software driver (NVIDIA ForceWare Release 80). The experiment was conducted in a dark room, with minimal light levels and the equipment arranged as shown in Figure 4.3.

4.5.4 Stimuli Details

The trial stimulus consisted in two white neighboring squares on a black background with a small white square in the middle, as described in section 4.3.2. The distance between the two internal edges of the left and right square was 20 image pixels. Stimuli were presented to participants via the stereoscopic displays at the manufacturers' nominal viewing distance for the Kodak and Iris3D displays and at 650 mm for all other displays.

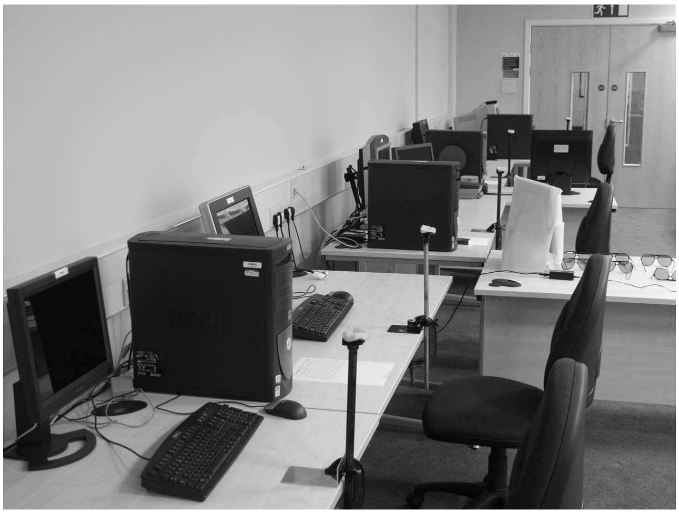


Figure 4.3: Environment used in the 3D display comparative study.

4.5.5 Results

Results are reported only for 12 of the 14 participants, as data from the other two were excluded because of poor score performance (average score of 49% and 52% respectively, against a minimum score of 74% of all other participants).

The empirical results obtained during the pilot experiment are presented in two separate sections: one for score and one for response time; subjective results are also briefly summarised in a third independent section. An interpretative discussion of these results is reported in Chapter 6.

Score

Overall, score performance was good. When participants could detect a depth difference between the two squares, an average score of 94% was achieved, which is close to the ideal score of 100%. Table 4.2 shows the mean value and the standard deviation of score for each experimental condition; mean scores are also shown graphically in Figure 4.4. Note that the data illustrated in this graph and the line graphs in the next subsection of this chapter (pg 126-133) are representative of discrete measurements; the lines that join score values belonging to the same display have illustrative purpose only and no interpolation should be applied between two adjacent score points in order to extract intermediate values.

Table 4.2: Score results: mean, M, (%) and standard deviation, SD.

Disparity	0		2		4		6	
	M	SD	M	SD	M	SD	M	SD
DTI	53.87	.31	95.83	.08	93.75	.13	95.83	.07
SeeReal	56.84	.23	52.98	.07	88.10	.22	88.69	.23
Colorlink	61.31	.22	88.39	.28	88.39	.28	89.29	.27
Sharp	63.10	.15	95.54	.07	95.24	.07	86.61	.21
Iris3D	72.02	.14	99.11	.02	100.00	.00	99.41	.01
Kodak	52.68	.24	96.43	.07	95.54	.08	95.24	.09
S/Glasses	60.71	.26	100.00	.00	98.51	.03	98.21	.05

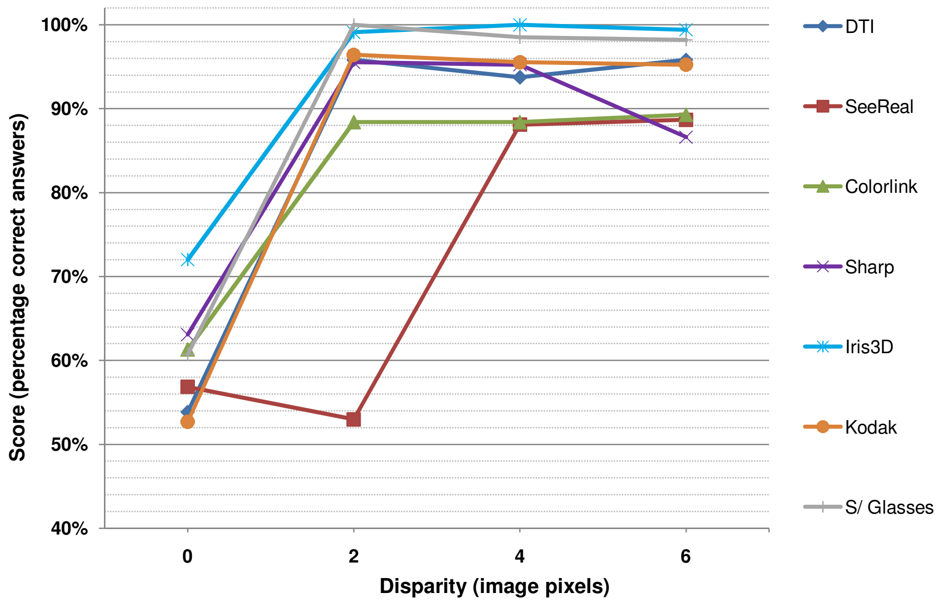


Figure 4.4: Score means for all experimental conditions.

Score data were first subjected to ANOVA, with *Disparity* and *Display* as within-subjects independent variables and *Score* (%) as the dependent variable. According to Mauchly's test [Mauchly, 1940] the assumption of sphericity had been violated for all main effects. Therefore degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity ($\varepsilon = .30$ for display, $\varepsilon = .38$ for disparity, $\varepsilon = .20$ for display*disparity).

The ANOVA statistics revealed that there was a significant effect of both display and disparity on performance (i.e. score), as well as a significant interaction between the two (all F values > 6.04 and all p values $<= .011$).

Performance at 0-pixel Image Disparity

Figure 4.5 shows the mean score and standard deviation for the conditions with 0-pixel disparity. When no disparity is applied to the input image, performance at chance (i.e. score = 50%) is expected for all displays. To evaluate this hypothesis a series of pairwise t-tests was conducted where the mean score associated with each display was compared against chance.

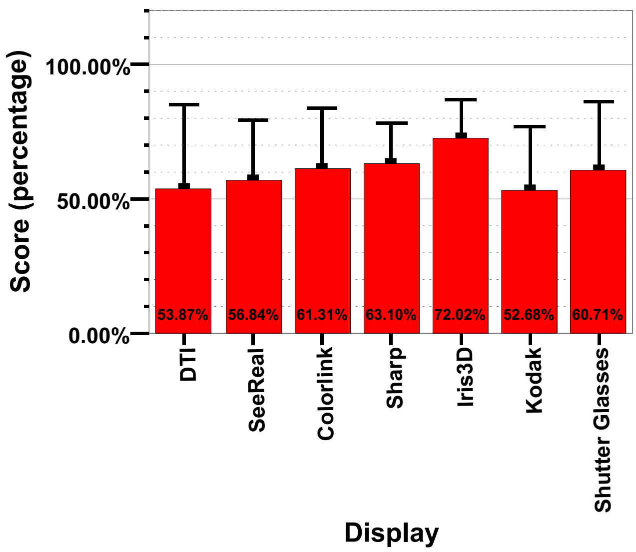


Figure 4.5: Mean score and standard deviation for the 0-pixel disparity condition.

Even though participants seemed to be slightly biased towards choosing the right square (all mean scores $> 50\%$), the tests showed that the mean scores for the DTI, Colorlink, SeeReal, Kodak and Shutter Glasses were not significantly different from chance (all $t(11)$ values < 1.77, all p values $> .11$). These results suggest that these five displays performed as predicted.

When using the Sharp display, participants were significantly biased towards selecting the right square even when no disparity was introduced at software level ($M = 63\%$, $t(11) = 3.04$, $p = .01$). However, a more detailed analysis showed that performance for the Sharp display was reliably better (i.e. closer to chance) than that observed for the Iris3D display ($t(11) = 2.44$, $p < .05$) and not reliably worse than performance for any of the other displays (all $t(11)$ values < 2.18, all p values $> .05$). These results suggest that there was no reliable difference in performance between the Sharp display and the other displays that performed at chance.

Concerning the Iris3D display, performance was significantly higher than chance ($M = 72\%$, $t(11) = 5.29$, $p < .001$) and that of all other displays, including the Sharp (all $t(11)$ values > 1.99, all p values $= .07$ or lower). This suggests that when using the Iris3D display, candidates were indeed not performing by chance, but they were systematically perceiving a difference in depth between the two squares (i.e. right square closer than left square). A detailed discussion of this point is presented in section 6.2.1.

Performance at 2-pixel Image Disparity

The next aspect of the score data that was considered in detail was performance at 2-pixel disparity. Mean scores and standard deviation for this condition are given in Figure 4.6. Note that the error bars in this and the following bar graphs in this section (pg 123 and 125) have the purpose to show the magnitude of the standard deviation and do not indicate that numerical score values above 100% were recorded during the experiment; 100% was in fact the highest achievable score in any experimental condition.

According to the predictions, all tested displays should have the capability to reproduce an image disparity of 2 pixels. In order to investigate this point, a series of paired t-tests was performed.

The tests revealed that overall the displays performed as expected with the only excep-

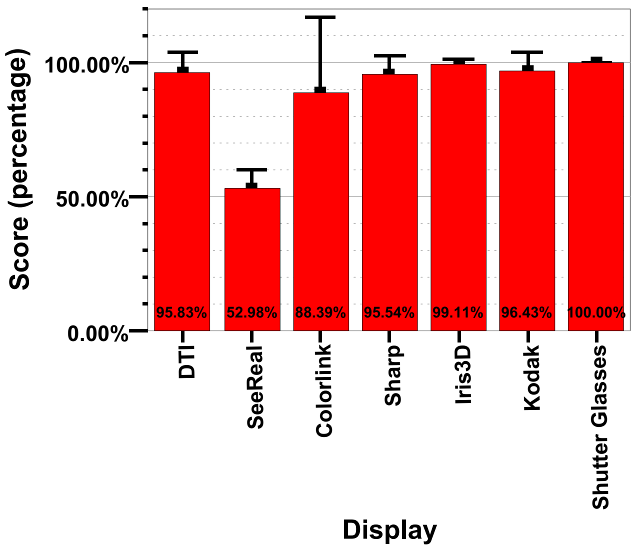


Figure 4.6: Mean score and standard deviation for the 2-pixel disparity condition.

tion of the SeeReal. In particular, the mean score for the SeeReal display ($M = 53\%$) was significantly lower than the mean score for all the other displays (all M values $> 88\%$; all $t(11)$ values > 4.28, all p values $= .001$ or lower) and not different than chance (SeeReal vs chance: $t(11) = 1.45$, $p > .1$). This suggests that when using this display participants were unable to detect any difference in depth between the two squares. By contrast, when using any of the other displays, candidates were clearly able to detect depth and were performing significantly better than chance (all $t(11)$ values > 4.71, all p values $= .001$ or lower).

In respect to the SeeReal display, pairwise comparisons also showed that at 2-pixel disparity participants performed significantly worse than at 4-pixel disparity ($t(11) = 5.16$ and $p < .001$) but not reliably differently than at 0-pixel disparity ($t(11) = 0.51$ and $p > .5$). This suggests that with the experimental conditions adopted for this study the SeeReal display does not have the predicted capability to reproduce 2-pixel image disparity.

Finally, a series of pairwise comparisons between displays revealed that performance for the Sharp display was reliably lower than performance for the Shutter Glasses display ($t(11) = 2.26$ and $p = .045$) and marginally lower than the one for the Iris3D display ($t(11) = 2.17$ and $p = .053$). The t-tests also revealed that when using the DTI display candidates' performance was marginally worse than when using the Shutter Glasses display ($t(11) = 1.90$ and $p = .084$). Regarding the other 3D displays, no other difference approached reliability (all $t(11)$ values < 1.69, all p values $> .1$).

On the basis of these results, it can be concluded that for the 2-pixel disparity condition performance was as expected or only marginally lower than expected for all displays with the only exception of the SeeReal display, which instead showed unexpectedly low performance.

Performance at 4-pixel Image Disparity

Mean scores and standard deviation for the condition with 4-pixel disparity are shown in Figure 4.7.

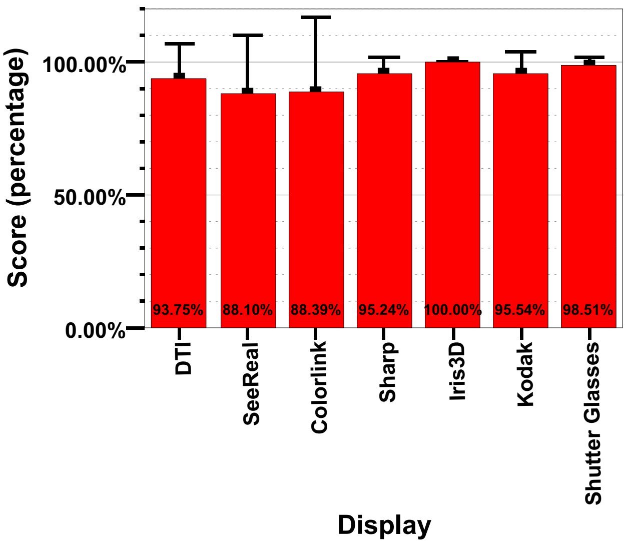


Figure 4.7: Mean score and standard deviation for the 4-pixel disparity condition.

The expected results for this experimental condition were that all displays should be able to reproduce an image disparity of 4 pixels. In line with this prediction, performance for all tested displays was overall high (all Mean values = 88% or higher).

As it can be seen from Figure 4.7, mean scores presented some numerical differences among the different displays. However, the only difference that was statistically reliable was the better performance observed for the Iris3D display as compared to the Sharp display ($t(11) = 2.46$ and $p < .05$).

There were also a number of marginal effects. In particular, performance for the Shutter Glasses display was marginally better than for the Sharp display ($t(11) = 2.03$ and $p < .1$). Furthermore, performance for the Iris3D display was marginally better than for the Kodak display ($t(11) = 1.99$ and $p < .1$). Finally, performance for both the Iris3D display and the Shutter Glasses display was marginally superior than for the SeeReal display (both $t(11)$ values > 1.82, both p values $< .1$). All other differences did not approach significance (all $t(11)$ values < 1.67, all p values $> .1$).

Overall, it can be concluded that when the stimuli were treated with 4-pixel image disparity, all displays performed as expected. Some numerical variability in performance was present, but the t-test analysis revealed that these differences in the mean score values were largely non-significant.

Performance at 6-pixel Image Disparity

When the input image was treated with a level of disparity of 6 pixels a situation similar to the one observed for the 4-pixel disparity arose. Figure 4.8 shows mean scores and standard deviation for the 6-pixel disparity condition.

A series of paired t-tests showed that performance for the Iris3D display was reliably better than for the DTI display ($t(11) = 2.25$ and $p < .05$). Furthermore, performance for both the Iris3D and Shutter Glasses was marginally better than for the Sharp display (both $t(11)$ value > 2.01 and both p values $< .1$). No other difference approached reliability (all $t(11)$ values < 1.72, all p values $> .1$).

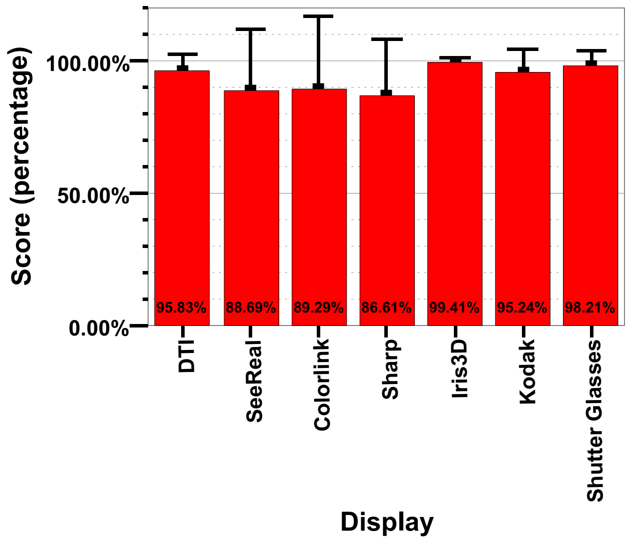


Figure 4.8: Mean score and standard deviation for the 6-pixel disparity condition.

On the basis of the numerical trends observed in the data, the Sharp display deserves particular attention. As the graph of Figure 4.4 illustrates, the Sharp is the only display that shows an apparent decrease in performance with increased disparity. To be specific, the mean score for this display drops from 95% at 2- and 4-pixel disparity to 87% at 6-pixel disparity. Nevertheless, pairwise comparisons showed that this decrement in performance was non-significant (all t values < 1.70, all p values $> .1$).

Overall, the data for the 6-pixel disparity condition are similar to those observed for the 4-pixel disparity condition in that few statistically reliable differences were obtained. Therefore it can be concluded that when the stimuli were treated with a 6-pixel image disparity all the displays performed as predicted.

Response Time

Table 4.3 shows the mean value and the standard deviation for response time for each experimental condition; response time means are also shown graphically in Figure 4.9.

Similar to score data, response time data were also subjected to ANOVA, with *Dispar-*

Table 4.3: Response time results: mean, M, (ms) and standard deviation, SD.

Disparity	0		2		4		6	
	M	SD	M	SD	M	SD	M	SD
DTI	3079	1556	1679	1034	1585	970	1953	1285
SeeReal	2862	1856	2600	1124	1650	1185	1817	999
Colorlink	4239	3443	2121	1366	2111	1516	2220	1729
Sharp	4121	4810	1983	1484	1945	1206	2437	1637
Iris3D	3540	1827	1524	1000	1379	1041	1446	808
Kodak	3505	1478	1478	893	1687	1049	1808	1088
S/Glasses	5238	7359	1314	730	1536	987	1718	1122

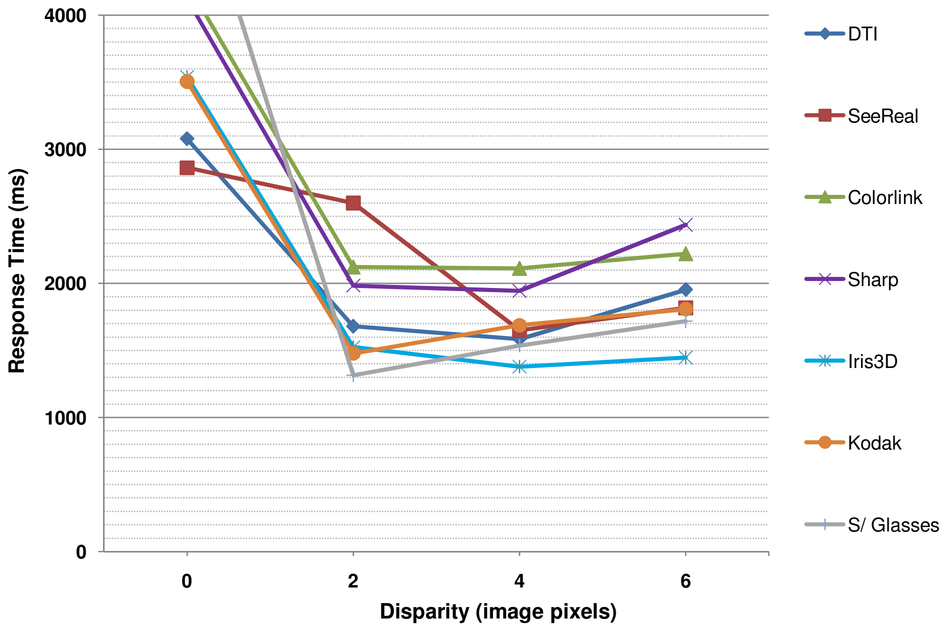


Figure 4.9: Response time means for all experimental conditions.

ity and *Display* as within-subjects independent variables and *Response Time* (ms) as the dependent variable. Again, Mauchly's test revealed that the assumption of sphericity was not met, so degrees of freedom were corrected for all main effects using the Greenhouse-Geisser estimates of sphericity ($\varepsilon = .45$ for display, $\varepsilon = .45$ for disparity, $\varepsilon = .10$ for display*disparity).

The ANOVA statistics revealed that there was a strong effect of disparity on response time ($F(1.35, 14.80) = 19.22$ and $p < .001$), while there was a non-significant main effect of display ($F(2.69, 29.61) = .57$ and $p = .620$) and a non-significant interaction between the type of display used and the amount of image disparity applied to the stimuli ($F(1.87, 20.58) = 1.02$ and $p = .372$).

Given the non-significance of the display variable main effect, it was not possible to proceed with paired comparisons across experimental conditions with the same disparity level and different display. For this reason, the detailed analysis of the time response data only consists of sets of t-tests where mean response times relative to the same display are compared across different disparity levels, comprising one set of paired comparisons for each display as specified below.

Performance for the DTI Display

Response time for this display showed the expected trend, i.e. a high response time at 0-pixel disparity ($M = 3079$ ms) followed by a drop at 2-pixel disparity ($M = 1679$ ms) and an almost constant behavior for higher disparities, even though with a slight increase at 6-pixel disparity ($M(4 - pixel) = 1585$ ms, $M(6 - pixel) = 1953$ ms). The t-tests confirmed that participants were significantly slower at 0-pixel disparity as compared to each of the other three disparity conditions (all $t(11)$ values > 2.52 and all p values $< .029$).

Mean differences between the 2-, 4- and 6-pixel conditions were all non-significant (all $t(11)$ values > 2.52 and all p values $<= .029$), which confirms that the numerical increase in response time registered at 6-pixel disparity is statistically non-meaningful.

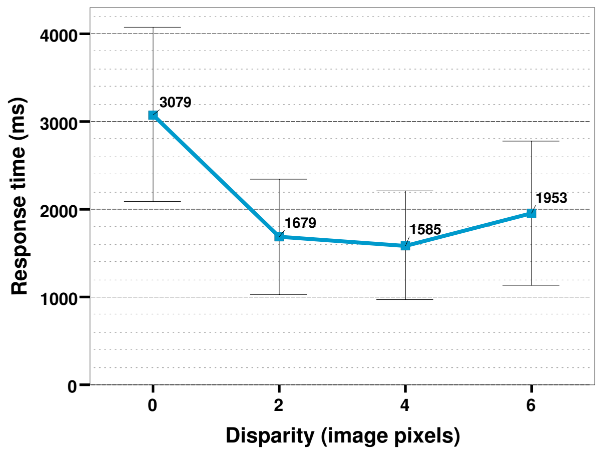


Figure 4.10: Mean response time and standard deviation for the DTI display.

The graph of Figure 4.10 shows mean response times and standard deviations for the DTI display.

Performance for the SeeReal Display

Response time means and standard deviations for the SeeReal display are shown in Figure 4.11.

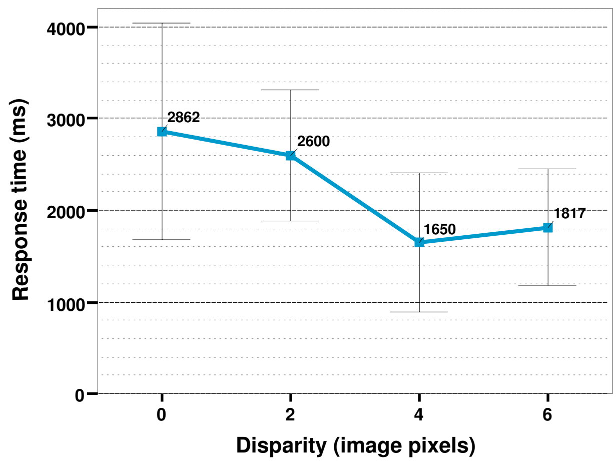


Figure 4.11: Mean response time and standard deviation for the SeeReal display.

As it can be seen from this graph, data for the SeeReal display yielded an unexpected trend, with mean response time at 2-pixel disparity ($M = 2600$ ms) only slightly lower than mean at 0-pixel disparity ($M = 2862$ ms) and notably higher than means at 4- ($M = 1650$ ms) and 6-pixel disparity ($M = 1817$ ms). This point was investigated in detail with a series of paired t-tests, which confirmed that participants were indeed slower at 2-pixel disparity as compared to 4- and 6-pixel disparity conditions ($t(11) = 3.24$, $p = .008$ and $t(11) = 2.81$, $p = .017$ respectively). Also, there was no significant difference between their response time at 2-pixel disparity and the one at 0-pixel disparity ($t(11) = .80$, $p = .003$). Even though in contrast with the predictions, these results are in line with the score results.

Interestingly, the t-tests also revealed that at 4- and 6- pixel disparity participants were only marginally faster than at 0-pixel disparity ($t(11) = 2.07$, $p = .063$ and $t(11) = 1.89$, $p = .086$ respectively).

Finally, the slight numerical differences in response time between the 4- and 6-pixel conditions were found to be non-significant ($t(11) = -1.70$, $p = .118$) as expected.

Performance for the Colorlink Display

Similar to the DTI display, the Colorlink display showed a trend in line with the predictions in that the response time was high at 0-pixel disparity ($M = 4239$ ms), registered a drop at 2-pixel disparity ($M = 2121$ ms) and from here it remained nearly constant ($M(4 - pixel) = 2111$ ms and $M(6 - pixel) = 2220$ ms), with only negligible numerical differences. This is clearly shown in the graph of Figure 4.12.

The pairwise t-test comparisons confirmed that when the stimuli were treated with 0-pixel disparity participants were significantly slower than when the stimuli were treated with 2-, 4- or 6-pixel disparity ($t(11) = 2.97$, $p = .013$; $t(11) = 3.11$, $p = .010$; $t(11) = 3.06$, $p = .011$ respectively). The small numerical differences between the 2-, 4- and 6-pixel conditions did not approach reliability (all $t(11)$ absolute values $< .52$ and all p values $> .61$).

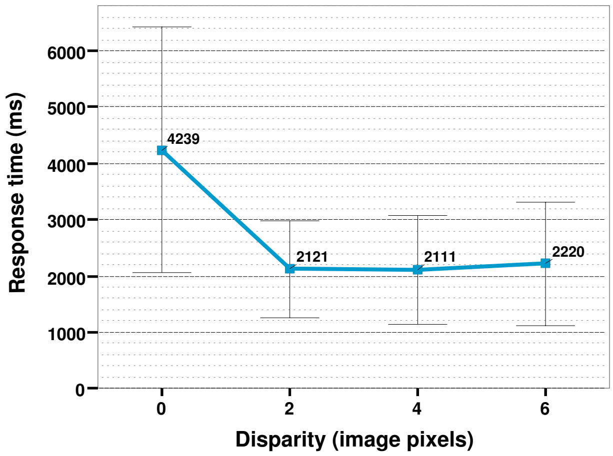


Figure 4.12: Mean response time and standard deviation for the Colorlink display.

Performance for the Sharp Display

Figure 4.13 shows response time means and standard deviations for the Sharp display.

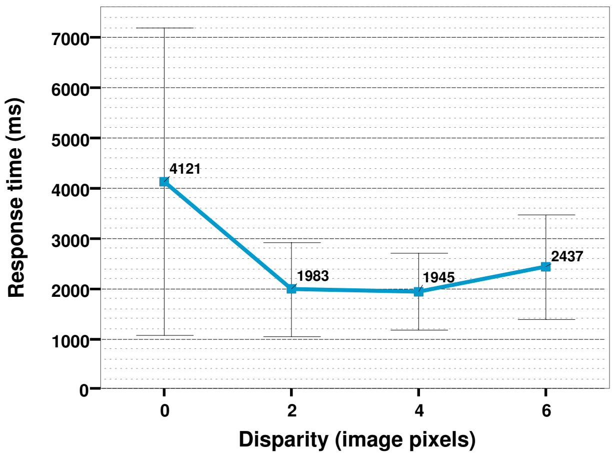


Figure 4.13: Mean response time and standard deviation for the Sharp display.

At first glance, the response time data for the Sharp display seem to follow the expected trend ($M(0-pixel) = 4121$ ms and $M(2-pixel) = 1983$ ms), with the only exception of an increase in mean value between the 4- ($M = 1945$ ms) and the 6-pixel disparity ($M = 2437$ ms) conditions. However, when analysed in detail the Sharp results yielded a number of interesting points. In particular, when the stimuli were shown with a disparity of 2 or 4 pixels participants were only marginally faster then when the stimuli were treated with 0-pixel disparity ($t(11) = 2.10$, $p = .060$ and $t(11) = 1.94$, $p = .078$ respectively). Furthermore, the difference between their mean response time at 0-pixel disparity and the one at 6-pixel disparity was non-significant ($t(11) = 1.59$, $p = .141$). Finally, at 4-pixel disparity participants were marginally faster than at 6-pixel disparity ($t(11) = -2.05$, $p = .065$).

On the other hand, the t-test statistics revealed also that there was no difference between participants' mean response times at 2-pixel disparity and the means at 4- ($t(11) = .17$, $p = .870$) and 6-pixel disparity ($t(11) = -1.28$, $p = .226$) respectively. These last results reflect what was expected according to the theoretical predictions.

Performance for the Iris3D Display

A graphical representation of response time means and standard deviations for the Iris3D display is given in Figure 4.14.

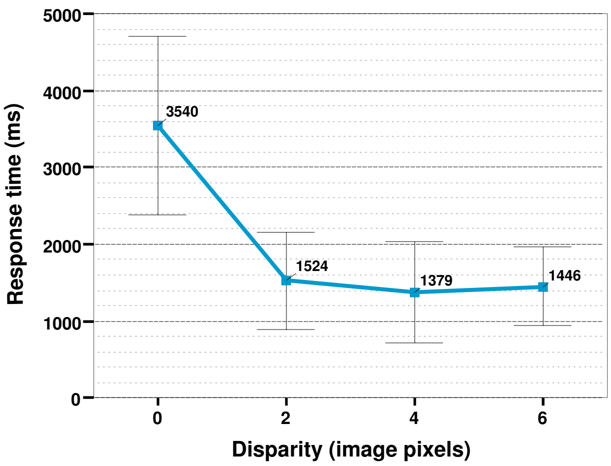


Figure 4.14: Mean response time and standard deviation for the Iris3D display.

Response time data for the Iris3D display are according to the expected trend and yielded the following mean values: $M(0 - pixel) = 3540$ ms, $M(2 - pixel) = 1524$ ms, $M(4 - pixel) = 1379$ ms and $M(6 - pixel) = 1446$ ms. Paired t-test statistics showed that when using this display participants were significantly slower at 0-pixel disparity than they were at 2-, 4- and 6-pixel disparity ($t(11) = 5.61$, $t(11) = 7.04$ and $t(11) = 6.46$ respectively; all p values $< .001$) and that the small differences in the participants' response time between the 2- and the 4-, the 2- and the 6-, and the 4- and 6-pixel disparity conditions were all non-significant ($t(11) = 1.53$, $p = .154$; $t(11) = .56$, $p = .586$; $t(11) = -.56$, $p = .590$ respectively).

Performance for the Kodak Display

Response time means and standard deviations for the Kodak display are shown in Figure 4.15.

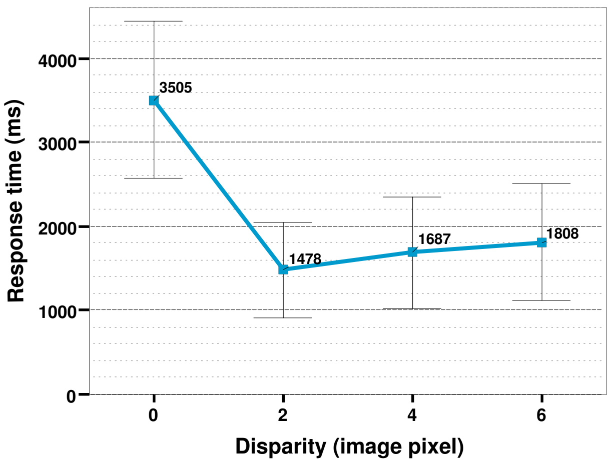


Figure 4.15: Mean response time and standard deviation for the Kodak display.

As it can be observed from the graph, response time data for the Kodak display follow the expected trend for the 0- ($M = 3505$ ms) and 2-pixel ($M = 1478$ ms) disparity conditions, but after that they show an apparent steady, moderate increase with increased disparity ($M(4 - pixel) = 1687$ ms and $M(6 - pixel) = 1808$ ms) that, based on the theoretical predictions presented in section 4.4.3, was not expected.

The paired t-test comparisons confirmed this pattern and showed that when performing the task at 2-pixel disparity participants were only marginally faster than when performing at 4-pixel disparity ($t(11) = -1.82$, $p = .096$) but significantly faster than what they were at 6-pixel disparity ($t(11) = -2.33$, $p = .040$). These results suggest that as disparity increased people may have found it harder to perform the assigned task.

On the other hand, differences in mean response time between the 0-disparity condition and the other three disparity conditions were all extremely significant (all $t(11)$ values > 5.99, all p values $< .001$) as expected.

Performance for the Shutter Glasses Display

The Shutter Glasses display yielded a response time data trend similar to the one of the Kodak display, in that the mean response time is high at 0-pixel disparity ($M = 5237$), it drops sharply at 2-pixel disparity ($M = 1314$) but it shows a steady though modest increase at higher disparities ($M(4 - pixel) = 1536$ ms and $M(6 - pixel) = 1718$ ms). This can readily be observed in the mean and standard deviation graph of Figure 4.16.

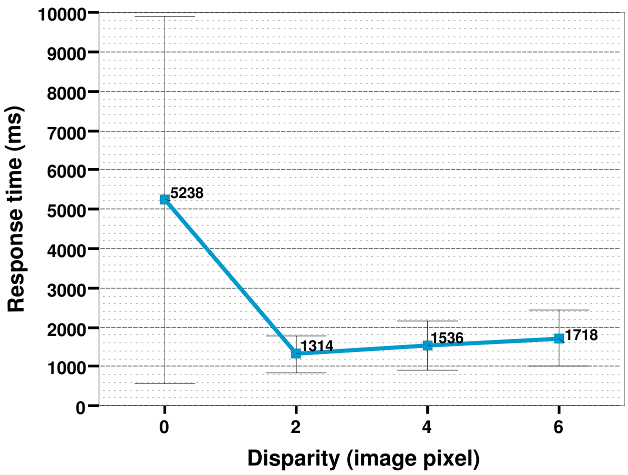


Figure 4.16: Mean response time and standard deviation for the Shutter Glasses display.

The t-test statistics partially confirmed this pattern and revealed that participants were only marginally faster at 2-pixel disparity as compared to the condition with 6-pixel disparity ($t(11) = -1.93$, $p = .079$), but there was no significant difference in their mean response times between 2- and 4-pixel disparity ($t(11) = -1.25$, $p = .237$) and between 4-

and 6-pixel disparity ($t(11) = -1.22$, $p = .247$).

Interestingly, the paired comparisons also revealed that the difference in mean response time between the 0-pixel disparity and the 2-pixel disparity conditions was only marginal ($t(11) = 1.84$, $p = .093$), and that the differences between the 0-pixel disparity condition and the 4- and 6-pixel disparity conditions respectively were both statistically non-significant ($t(11) = 1.76$, $p = .106$ and $t(11) = 1.66$, $p = .125$).

Subjective Results

At the end of each experimental session participants were asked to complete a detailed questionnaire concerning the display they had just used (level of ease seeing 3D, disturbing factors, level of discomfort, general comments about the display). They were also asked to fill in a more general survey at completion of the whole experiment.

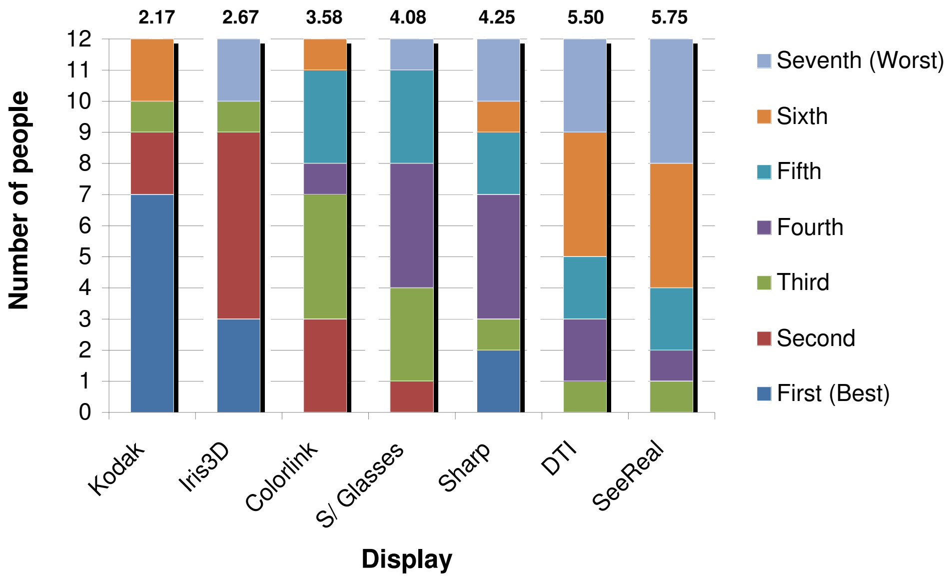


Figure 4.17: Subjective ranking of displays.

As part of the post-trial questionnaire, participants were asked to rank the seven displays from best to worst based on their personal experience in using the displays during the pilot experiment. Figure 4.17 shows a cumulative frequency histogram based on the participants' subjective ranking results from this questionnaire, with average ranking score for each display shown across the top of the graph. On average, the Kodak display was ranked as best (score = 2.17), while the SeeReal display was ranked as worst (score = 5.75).

At the end of each display session, participants were asked to express how comfortable it was to see the 3D image on the specific 3D display. They were also asked if they experienced any discomfort and, if yes, how strong the discomfort was on the specific display. Answers to the first and third questions were given via a 10-point rating scale.

Figure 4.18 shows a cumulative comparison of results for the first question, with average scores shown across the top of the graph. The scale used ranged from 1 = "not comfortable at all" to 10 = "extremely comfortable". Again, participants found that the Kodak was the best display in terms of the level of comfort in seeing the 3D image (score = 8.42), while the SeeReal display was the worst (score = 5.67).

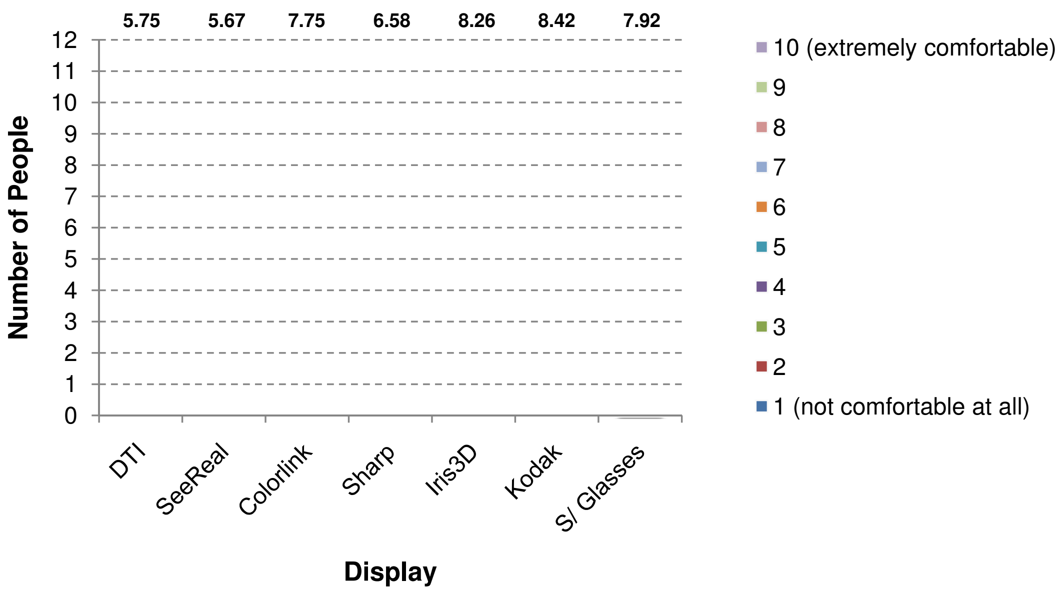


Figure 4.18: Subjective rating for level of comfort in seeing the 3D image on each display.

Results for the second and third question are shown in Figure 4.19 and 4.20 respectively.

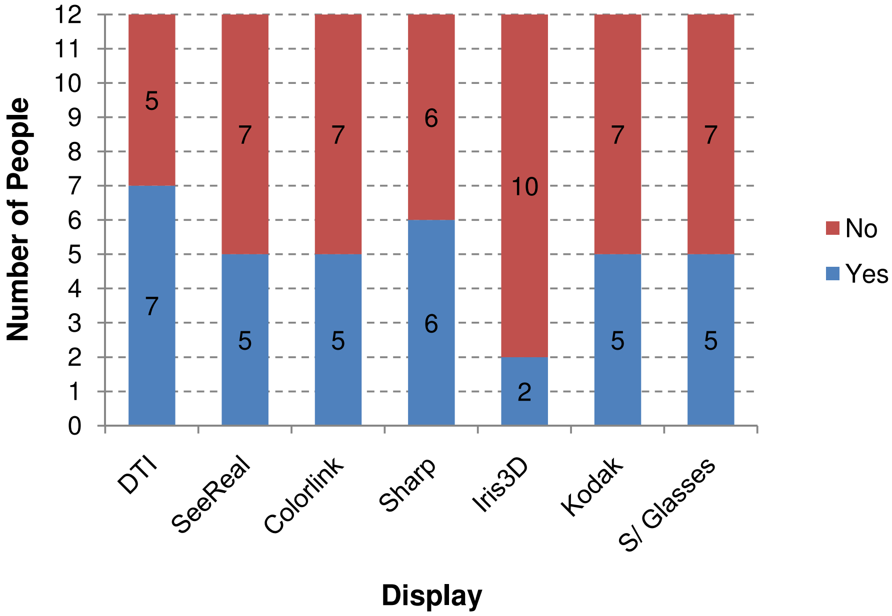


Figure 4.19: Subjective results for physical discomfort on each display.

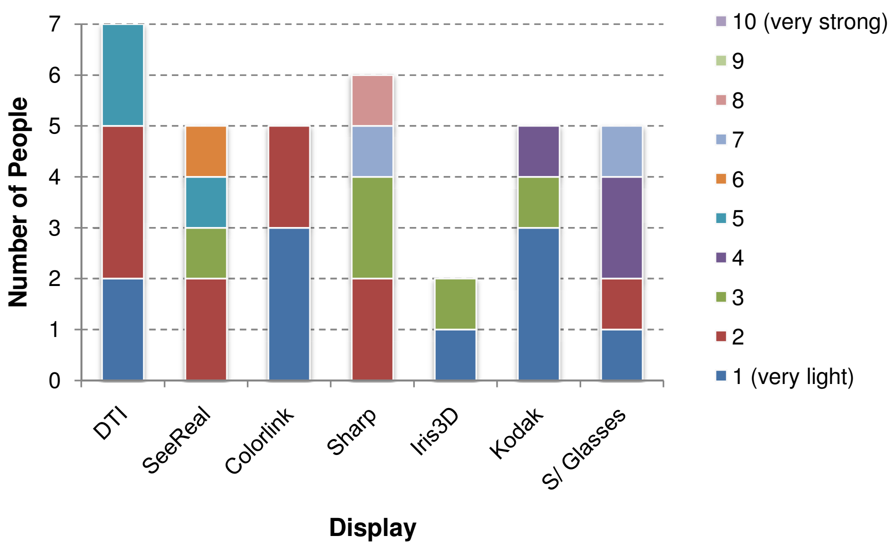


Figure 4.20: Subjective rating for level of discomfort on each display.

Overall, the percentage of people that experienced some form of discomfort at some point during the experiment was 48%, against the 52% that did not experience any discomfort. Looking at the results individually, when using the DTI display 7 participants out of 12 perceived discomfort, while with the Iris3D display these were only 2 out of 12. Furthermore, in the cases where discomfort was perceived, the Sharp display registered the highest average level of discomfort (score = 4.2), while the Colorlink registered the lowest (score = 1.4). These scores are expressed on the same scale used in the questionnaire, i.e. 1 = "very light discomfort" and 10 = "very strong discomfort".

4.6 Conclusions

This chapter presented a methodology that makes use of human depth perception in order to assess the performance of different stereoscopic 3D displays. This includes a general classification model for 3D display technologies based on their characteristics, and a detailed experimental design that can be used to carry out quantitative empirical research and perform a large scale comparative study of such technologies. Its robustness has been assessed via a pilot experiment during which seven desktop two-view 3D displays have been tested across four different levels of image disparity (0-, 2-, 4- and 6-pixel disparity). The details of the pilot study are reported in the second part of the chapter.

The data collected during the pilot experiment showed that the methodology is sound and that it provides a flexible and sensitive framework that can be used to assess and compare not only two-view desktop 3D displays but any kind of two-view 3D system. However, the data also provided a clear insight into aspects of the methodology's experimental design that could be improved and adjusted so as to reduce the amount of unsystematic variability in the results and, most importantly, to increase the amount of information that can be extracted from the empirical data themselves. In particular, the pilot experiment highlighted the need to run the full-scale experiment using single-pixel image disparity increments, rather than 2-pixel increments. This should identify performance differences between displays in representing disparity for fine depth-judgements and provide a better understanding of how the 3D displays' technical factors may affect the design of a 3D cursor to support depth-based tasks in stereoscopic 3D environments. The improved methodology and the main experiment are presented in the next chapter of this thesis.

Chapter 5

Assessment of 3D Technologies: Results

5.1 Introduction

Chapter 4 presented a new methodology that can be used to investigate and compare the performance of 3D displays. The pilot study performed in order to tune the experimental design revealed the need to run the main study on finer image disparity intervals. This chapter describes the main 3D display comparative study and reports the empirical results and the statistical analysis of the collected data. The methodology used for the main study has been improved as compared to the one adopted for the pilot study in order to better control for a number of factors that could bias the results and introduce noise in the data, as explained in section 5.2.1.

Section 5.2 reports the experimental details of the main study. The empirical results are then presented in three separate sections; section 5.3.1 looks at the score results, the most interesting ones for the purpose of this thesis, while response time and subjective results are reported in sections 5.3.2 and 5.3.3 respectively. The salient points that emerged from the statistical analysis of the collected data are concluded in section 5.4.

5.2 Experimental Details

5.2.1 Improvements in the Research Method

The methodology used for the main study largely corresponds to the one used for the pilot study. However, a number of improvements have been made in order to increase the robustness of the method and to better control the factors that could bias the results and introduce noise in the data. These improvements include:

- More comprehensive vision screening of the participants. In particular, for the main experiment participants were screened both for stereo vision and visual acuity as explained in section 5.2.2.

- Finer control of the image disparity levels used to treat the stimuli. By using single pixel image disparity increments it should be possible to identify performance differences between the displays in representing disparity for fine depth judgements. In addition, this should also provide a better understanding of the artifacts that the different technical characteristics of the displays can introduce in the output image shown to the observer.

- Amendment of the original stimuli image in order to ensure that the correct viewing position is maintained during the trials (see section 5.2.6 for details).

- Improved algorithm to compute score for the trials where the stimuli were treated with an image disparity of 0 pixels. As explained in section 4.3.3, the new algorithm allowed for scores not different than chance (i.e. at 50%), both statistically and numerically, even when participants were systematically choosing left (or right) as an answer for all the trials where both squares were presented with 0-pixel image disparity.

- More thorough practice sessions before the actual data collection so as to make sure that participants are fully familiar with the task and with the usage of the 3D displays tested during the experiment.

- Use of a more up-to-date set of displays to represent the different 3D display classes.
- Improvements of session and post-trial questionnaires.
- Stimuli presented at manufacturers' nominal viewing distance on each display. During the pilot experiment a viewing distance of 650 mm was adopted for all displays but the Iris3D and the Kodak, for which it was not possible due to their optical and mechanical design. For the main experiment it was decided to present the stimuli at manufacturers' nominal viewing distance on each display. The reason behind this decision is to allow for a fairer comparison, as viewing distance is considered to be an intrinsic characteristic of the 3D display itself.

5.2.2 Participants

A total of 14 participants, nine male and five female, were recruited within the Durham University population. Age varied between 20 and 51 while the mean age was 28 years. Participants were naive with regard to the purpose of the experiment and received a nominal sum of five pounds per hour, for a total of 10 pounds each.

Prior to the start of the experiment volunteers were screened for both stereo vision and visual acuity at 1 metre. For stereoacuity a Titmus test was used, while for visual acuity a Snellen Chart was reproduced on an LCD monitor using Thomson Software Solutions Test Chart 2000 Pro (Figure 5.1). Participants all achieved a minimum stereoacuity of 40 seconds of arc and had at least a (corrected to) binocular visual acuity of 20/20.

5.2.3 Procedure

Participants were divided into two groups of seven people each. The main experiment was carried out in two separate group sessions that followed the experimental protocol

Figure 5.1: Thomson Software Solutions Snellen Chart.

described in section 4.3.3.

5.2.4 Experimental Conditions and Trial Blocks

During the main experiment, seven different two-view 3D displays were tested with four different levels of input image disparity, for a total of 28 experimental conditions. The four possible levels of image disparity to be applied to the stimuli were 0- 1- 2- and 3- image pixels respectively; for this experiment single-pixel image disparity increments were used, as compared to the 2-pixel image disparity increments used during the pilot experiment. Image disparity, and hence perceived depth, was therefore controlled in pixels. The seven 3D displays used are described in detail in section 5.2.5.

Each subject was asked to repeat the task 28 times for each of the four levels of image disparity on each display, giving a total of 112 experimental trials per display per subject. Both score and response time were recorded.

5.2.5 Apparatus

The set of 3D displays tested during the main experiment was representative of the four display classes described in 4.4.2. Following is a detailed description of the technical properties and optic design that characterizes each tested display.

DTI Display

The DTI 2018XLQ Virtual Window (*DTI*) is a switchable 2D/3D autostereoscopic display. In order to generate the 3D effect it uses a particular kind of illumination system called *Parallax Illumination* [Eichenlaub, 1993] [Eichenlaub, 1992]. For the DTI 2018XLQ display, this consists of vertical stationary light lines, generated via fluorescent lamps combined with reflectors and lenses, situated behind a 1280x1024 LCD display. Alternate columns of pixels are seen by either the left or right eye of the viewer, with the left image placed only in the odd columns of pixels and the right image only in the even columns of pixels. The result is an output stereo image that consists of a left view and right view, each with a resolution of 640x1024 view pixels. The DTI display belongs to the Column-Interleaved display class.

SeeReal Display

The SeeReal C-n (*SeeReal*) is a newer display than the SeeReal C-i used during the pilot study. Similar to the DTI, the SeeReal C-n is an autostereoscopic display that uses a single 1600x1200 LCD panel to show two 800x1200 interleaved images in output. Unlike the former though, the SeeReal display employs a prism mask system [Schwerdtner and Heidrich, 1998b] [Schwerdtner and Heidrich, 1998a] in order to correctly direct the left and right view to the eyes of the viewer. Specifically, it uses an array of vertically oriented micro-prisms in front of the LCD screen that acts as a parallax element and alternately deflects the light coming from the screen into the viewer's left and right eye. In this way the two channels of the input stereo image are vertically interlaced, with the left view appearing in the odd numbered columns and the right view in the even numbered columns. Together with the DTI display, the SeeReal display is representative of the Column-Interleaved display class.

Sharp Display

The Sharp LL-151-3D (*Sharp*) is a single-LCD autostereoscopic display, like the DTI and the SeeReal. It is 2D/3D switchable and it uses parallax (slit) barrier technology [Jacobs et al., 2003] [Harrold et al., 2000] [Harrold et al., 1999] that for this display consists of a latent retarder barrier [Montgomery et al., 2001b] placed behind the single 1024x768 LCD panel. If the display is in 3D mode, this special retarder array operates in conjunction with a layer of micro polarizers in order to create a parallax barrier. The barrier is defined on a finer pitch than the DTI and SeeReal displays and it directs alternate colour-columns from pixels to the left and right eyes. A consequence of this approach is that the RGB triple seen as one pixel in each eye is physically generated using the colour components of two pixels of the LCD screen. This makes the task of interleaving the stereo pair slightly more difficult for the drivers as they have to combine views within a single pixel. On the eXtended Graphics Array (XGA) base panel used by the Sharp display this results in each eye having a view resolution of 512x768 pixels. The Sharp display is the only representative of the Colour-Column-Interleaved category.

Colorlink Display

The Pavonine-Colorlink G170A (*Colorlink*) is a passive stereoscopic display that uses polarized glasses and that can either be used as a 2D or 3D monitor. Unlike the linearly-polarized Colorlink display used for the pilot experiment, the Pavonine-Colorlink tested during the main experiment uses circular polarization. In particular, this display uses a single 1280x1024 LCD panel and a permanently attached film that creates different circular polarization states for alternate rows of pixels. Without glasses the display acts exactly as a normal LCD; with the appropriate glasses odd rows of pixels are only seen by the left eye while even rows are seen by the right eye. In stereoscopic mode the resulting screen resolution for each view is 1280x512 view pixels. The Pavonine-Colorlink display is the only Row-Interleaved display tested during this experiment.

Iris3D Display

The Iris3D (*Iris3D*) is an autostereoscopic display that uses two high resolution LCD panels (1600x1200 pixels each) and a type of dual channel projection optics [McKay, 2005] [McKay et al., 1999]. This system consists of two components: a projection system and a head unit. The projection system has two separate optical paths for the left and right images and uses lenses and mirrors to project the left and right panels onto a concave mirror in the head unit. The concave mirror acts as a directional screen keeping the left and right images separate and correctly directed to the left and right eye; the two views can be seen by the viewer through the two halves of a single circular exit pupil. This optical design allows the Iris3D to provide two 1600x1200 full resolution views and as the channels are optically isolated there is no crosstalk between the two images. The Iris3D display is one of the three Full-Resolution displays included in the main experiment.

Kodak Display

The Kodak Stereo 3D display (*Kodak*) is an autostereoscopic display that uses two 17" LCD panels as image sources and forms two virtual images of these at optical infinity [Cobb, 2005]. The optical design of the display employs two spherical combining mirrors, one for the left image and one for the right image, each centered on a monolithic ball lens; these components form the core of the two image engines. Each engine has its own distinctive optical path (see Cobb [2005] for details) and creates an image in output: one optical path generates the left image and the other optical path generates the right image.

The two output images are then correctly directed to the eyes, so when the user looks through the circular exit pupils a stereo image can be seen. Another characteristic of the Kodak display is that it has a wide field of view [Cobb, 2005]. As a result of this approach the display delivers a stereo image with a field of view of 45^ox36^o and a view resolution of 1280x1024 pixels. The Kodak display belongs to the Full-Resolution display class.

Shutter Glasses Display

The Stereographics CrystalEyes (*Shutter Glasses*) are LC shutter glasses that can be used to visualize 3D images using a high frequency CRT display. This system uses a field sequential approach [Lipton and Ackerman, 1990] where the left and right views are shown on the monitor alternately. This is synchronized using an infrared emitter to switch the LC shutters in the glasses on and off in time with the images on the screen. On high frequency monitors the visibility of the flicker is minimized and the viewer perceives the left and right images as if they were displayed simultaneously. In this way the user can fuse the two views and see the 3D image in stereo with full resolution in each eye. For the main experiment the Stereographics CrystalEyes were used with a 20" Hitachi Superscan 813 CRT display at a screen refresh rate of 120 Hz. The result is a view resolution of 1024x768 view pixels with a frequency of 60 Hz in each eye. The Shutter Glasses display is the third Full-Resolution display tested in this experiment.

Table 5.1 summarises the technical characteristics and the operating modes of the displays used in the main experiment. The displays were driven by seven independent Personal Computers (PCs) that used the same type of graphics card (NVIDIA Quadro FX family) and the same driver (NVIDIA ForceWare version 84.26). The screen brightness was adjusted to nominal level for each display and the experiment was conducted in a dark room, with minimal light conditions (see Figure 4.3). The stimuli were presented to participants via the 3D displays at the manufacturer nominal viewing distance. Specifically,

Table 5.1: 3D systems specifications and operating modes.

Display Model	Screen Size (diagonal)	Viewing Distance	Res in 2D (px/display)	Res in 3D (px/eye)	Graphics Card Mode in Stereo	Interface
DTI 2018XLC (*DTI*)	18.1”	800mm	1280x1024	640x1024	Shutter Glasses	Analogue
SeeReal C-n (*SeeReal*)	20”	760mm	1600x1200	800x1200	Vertical Interlaced	Digital
Pavonine Colorlink G170A (*Colorlink*)	17”	650mm	1280x1024	640x1024	Shutter Glasses	Digital
Sharp LL-151-3D (*Sharp*)	15”	650mm	1024x768	512x768	Colour Interleaved	Digital
Iris3D (*Iris3D*)	20.1”	920mm	1600x1200	1600x1200	Clone mode	Digital
Kodak Stereo 3D (*Kodak*)	17”	600mm	1280x1024	1280x1024	Clone mode	Digital
CrystalEyes (*Shutter Glasses*) with Hitachi Superscan 813	20”	650mm	1024x768	1024x768	Blue-line-code for Stereographics prod.	Analogue

the viewing distance was 800 mm for the DTI display, 760 mm for the SeeReal display, 920 mm for the Iris3D display, 600 mm for the Kodak display, and 650 mm for all the other displays.

5.2.6 Stimuli

The visual stimuli used for the main study are shown in Fig. 5.2.

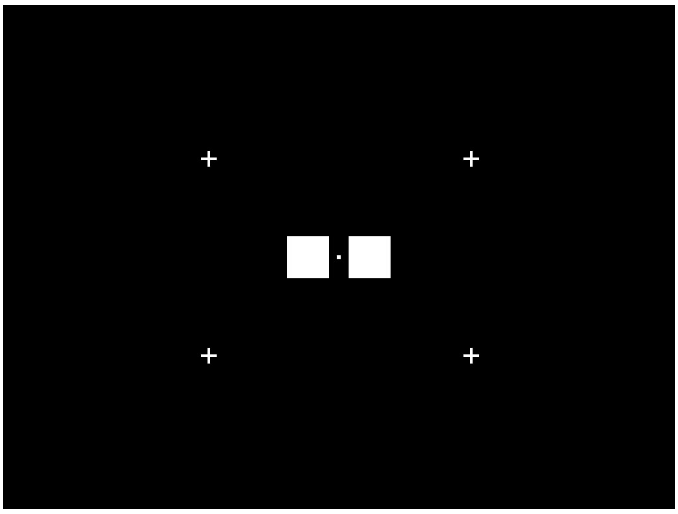


Figure 5.2: Stimuli used for the main experiment.

The stimuli consisted of two white neighboring squares on a black background with a small white square in the middle that acted as a fixation point, as described in section

4.3.2. The fixation point square was 6 image pixels wide while the width of the other two squares was 64 image pixels each. The distance between the two internal edges of the left and right square was 30 image pixels.

The stimuli also included four Greek crosses, placed peripherally around the two central squares and at a steady crossed-disparity of 4 image pixels; their only purpose was as a guide to help maintain the correct viewing position. Participants were instructed that when sitting in the correct viewing position, all four crosses should appear to be in front of the screen throughout the trials. The crosses had a size of 24x24 image pixels and had branches that were 4 image pixels thick.

5.2.7 Predictions and Hypotheses

Table 5.2 summarizes the predictions for the main experiment in terms of expected mean score and response time for every experimental condition. In the 0-pixel disparity condition a score of 50% and a high response time are expected for all displays. At image disparities of 2 and 3 pixels, all displays should show a score of 100% and a low response time. Finally, at 1-pixel image disparity Full-Resolution and Row-Interleaved displays are expected to yield a score of 100% and a low response time, while Column-Interleaved and Colour-Column-Interleaved displays should provide a score of 50% and a high response time.

Table 5.2: Predictions for score, S, (%) and response time, RT, (qualitative).

Disparity	0		1		2		3	
	S	RT	S	RT	S	RT	S	RT
DTI	50	high	50	high	100	low	100	low
SeeReal	50	high	50	high	100	low	100	low
Colorlink	50	high	100	low	100	low	100	low
Sharp	50	high	50	high	100	low	100	low
Iris3D	50	high	100	low	100	low	100	low
Kodak	50	high	100	low	100	low	100	low
Shutter Glasses	50	high	100	low	100	low	100	low

5.3 Results

Before proceeding with the presentation of the data and the statistical analysis it is worth remembering that in this study score, i.e. the proportion of trials at which participants selected the correct target, is used to assess the displays' performance intended as the capability of reproducing a certain amount of image disparity as view disparity and hence as perceivable depth. On the other hand, response time is defined as the amount of time elapsed between when the stimuli are first displayed and when the subject selects their answer, and is used as an indicator of the level of ease with which participants perform the assigned task on the different displays.

The results obtained during the main experiment are here presented in three separate sections: one for score, one for response time and one for subjective results. In each section, results are reported for 13 participants only, as data from a fourteenth participant were excluded from the analysis because of failure of the orthoscopic condition in one of the experimental sessions. An interpretative discussion of these results is reported in Chapter 6.

5.3.1 Score

Score results are summarised in Table 5.3, where mean value and standard deviation are reported for each experimental condition. Mean scores are also shown graphically in Figure 5.3. Note that the data illustrated in this graph and the line graphs in the next subsection of this chapter (pg 150-169) are representative of discrete measurements; the lines that join score values belonging to the same display have illustrative purpose only and no interpolation should be applied between two adjacent score points in order to extract intermediate values.

As it can be seen from the data, performance was generally good. In the conditions where there was a difference in disparity between the two squares (i.e. disparity level of

1 image pixel or bigger), participants' average score across displays was 84.18% or better, with an overall mean score of 91.50%.

Table 5.3: Score results: mean, M, (%) and standard deviation, SD.

Disparity	0		1		2		3	
	M	SD	M	SD	M	SD	M	SD
DTI	50.55	.09	58.51	.12	100.0	.00	100.0	.00
SeeReal	49.73	.04	50.00	.05	51.92	.09	99.46	.01
Colorlink	48.08	.05	94.78	.05	99.18	.02	99.18	.02
Sharp	52.75	.06	97.25	.06	99.73	.01	99.73	.01
Iris3D	51.92	.05	98.08	.03	98.35	.05	98.63	.03
Kodak	52.47	.10	91.76	.13	93.96	.14	92.03	.16
S/ Glasses	48.35	.07	98.90	.02	100.0	.00	100.0	.00

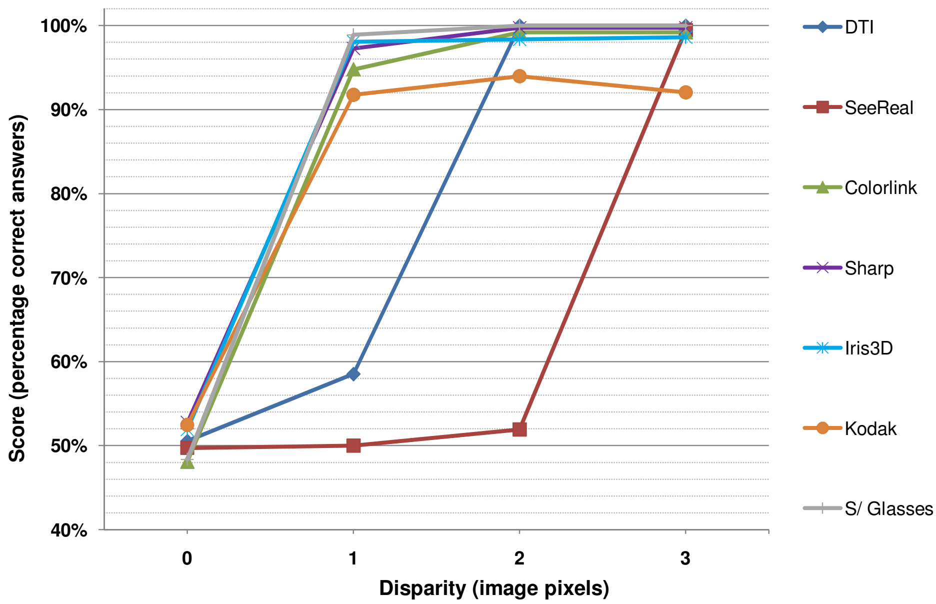


Figure 5.3: Score means for all experimental conditions.

Score data were first subjected to ANOVA, with *Disparity* and *Display* as within-subjects independent variables and *Score* as the dependent variable. Mauchly's test indicated that the assumption of sphericity had been violated for all main effects. For this reason, degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity ($\varepsilon = .31$ for the effect of display, $\varepsilon = .53$ for the effect of disparity, $\varepsilon = .28$ for the effect of display*disparity).

All effects are reported as extremely significant at $p < .001$. There was a significant effect of the type of display on participants' score ($F(1.85, 22.23) = 49.00$), as well as a significant effect of the level of disparity applied to the input image on score ($F(1.60, 19.18) = 853.95$). There was also a significant interaction effect between the type of display used and the level of image disparity applied to the input image ($F(5.06, 60.76) = 68.24$).

Performance at 0-pixel Image Disparity

The first aspect of the score data that was considered in detail in the analysis was performance at 0-pixel disparity, i.e. when no image disparity was applied to the stimuli. Fig. 5.4 shows mean scores of participants and standard deviation for this condition.

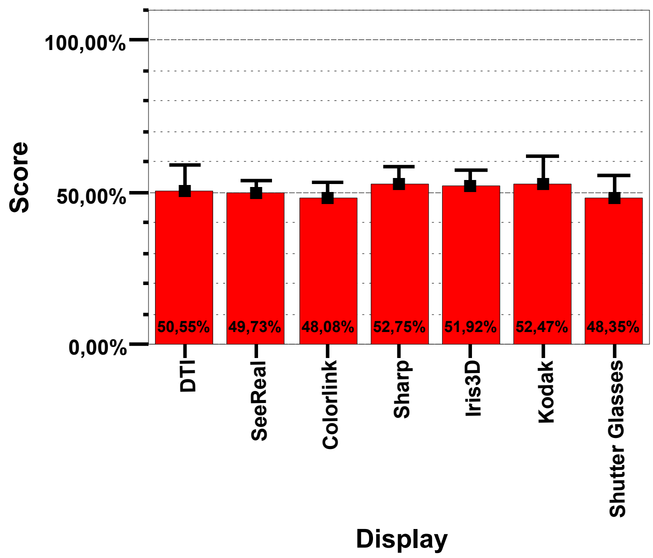


Figure 5.4: Mean score and standard deviation for the 0-pixel disparity condition.

There are two main reasons why it is important to examine performance at 0-pixel disparity. First, the analysis of these data gives a clear indication of how reliable the display is in conditions where no differences in depth should be shown; that is, if the display operates correctly, scores no different than chance should be recorded. Second, it helps to identify potential problems at both hardware level (e.g. optical, mechanical, electronic) and ergonomic level (e.g. viewing position, freedom of movement during usage), which can compromise performance and cause misperception of depth by the user.

In order to investigate if the seven displays performed as predicted, a series of one-sample comparisons was conducted where the mean score associated with each display was compared against chance (i.e. score = 50%). Despite slight numerical differences that ranged from 48.08% (Colorlink) to 52.75% (Sharp), the t-test revealed that none of the mean scores was significantly different from chance (all absolute $t(12)$ values $<= 1.69$, all p values $>= .117$).

Mean scores for the Sharp display and the Colorlink display were also compared against each other as they represent the upper and lower limits of the mean score data range. Even in this case the pairwise t-test revealed that the difference was non-significant ($t(12) = -1.817$, $p = .094$).

These results are in line with the predictions and suggest that, independently from the display used, participants were unable to detect any difference in depth between the two squares when no disparity was present in the test image.

Performance at 1-pixel Image Disparity

The second aspect of the data that was analysed in detail was performance at 1-pixel disparity. Mean scores and standard deviation for this condition are given in Figure 5.5. Note that the error bars in this and the following bar graphs in this section (pg 156 and 158) have the purpose to show the magnitude of the standard deviation and do not indicate that numerical score values above 100% were recorded during the experiment; 100% was in fact the highest achievable score in any experimental condition.

This aspect of the score data was particularly interesting for the following reason. In the case where the effects of antialiasing and 3D computer graphics are absent as in this

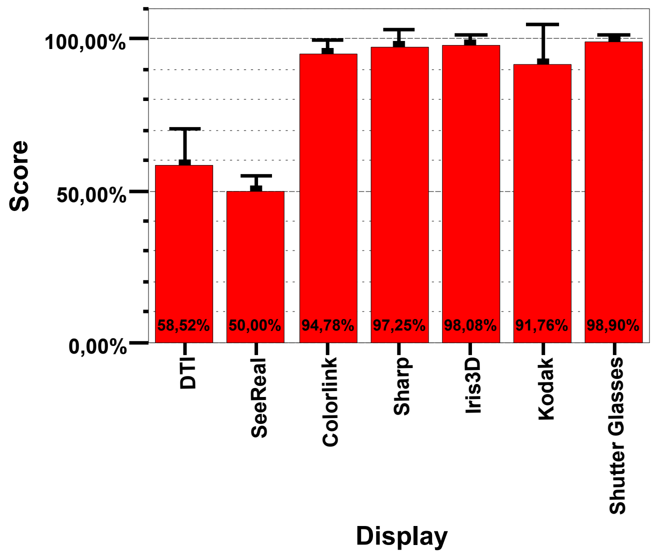


Figure 5.5: Mean score and standard deviation for the 1-pixel disparity condition.

study, 1-pixel image disparity represents the minimum reproducible disparity at software level on a generic stereoscopic 3D display. For the purpose of this thesis it was important to understand if the displays have the capability to reproduce this minimal amount of image disparity and if the differences in view resolution between the displays did have the predicted effect on depth reproduction when the stimuli is presented to the observer. The empirical data collected during this experiment can be used to clarify this point in a reliable way because the amount of depth induced by 1-pixel image disparity on any of the tested displays is greater than the minimum depth perceivable by a human with average stereoacuity (see Appendix D, section D.1). Therefore, if the display had the capacity to reproduce 1-pixel image disparity, participants should be able to detect it.

According to the theoretical predictions specified in 4.4.3, for the 1-pixel image disparity condition Column-Interleaved displays (DTI and SeeReal) and Colour-Column-Interleaved displays (Sharp) should lead to a score of 50%, while Row-Interleaved displays (Colorlink) and Full-Resolution displays (Iris3D, Kodak and Shutter Glasses) should lead to an optimum score of 100%. In order to investigate this hypothesis a series of t-test statistics were performed, which yielded a number of interesting points.

A first important finding is that the mean score for the SeeReal display ($M = 50.00\%$) was marginally lower than the mean score for the DTI display ($t(12) = 2.11$, $p = .056$) and significantly lower than the mean score for all the other displays (all $t(12)$ values > 11.05, all p values $< .001$). Besides, the t-tests also revealed that the mean score for the SeeReal display at 1-pixel disparity was no different than chance ($t(12) = .00$, $p = 1.0$) and no different than the one at 0-pixel disparity ($t(12) = .13$ and $p = .89$). These results are in line with the predictions and suggests that the SeeReal display does not have the capability to reproduce the depth induced by 1-pixel image disparity.

In this regard, it is interesting to note a surprising aspect of the data associated with the DTI and the Sharp displays. Specifically, performance for the DTI display ($M = 58.52\%$) was reliably higher than chance ($t(12) = 2.54$ and $p = .026$) but also significantly lower than optimum performance, i.e. score at 100% ($t(12) = 12.39$ and $p < .001$). Furthermore, performance for the DTI display was significantly lower than performance for the Colorlink, Sharp, Iris3D, Kodak and Shutter Glasses displays respectively (all $t(12)$ values $<= 7.23$, all p values $< .001$) and only marginally better than performance for the SeeReal display ($t(12) = 2.11$ and $p = .056$). Summarizing, these results mean that at 1-pixel disparity performance for the DTI display was better than expected, but still far from the optimum score of 100%. This suggests that to some extent the DTI display has the capability to reproduce 1-pixel image disparity as view disparity but participants still had substantial difficulty in detecting the difference in depth between the two squares.

Concerning the Sharp display ($M = 97.25\%$), performance was significantly better than chance ($t(12) = 28.21$ and $p < .001$) and not significantly different than optimum performance ($t(12) = 1.64$ and $p = .127$). Additionally, when using the Sharp display participants performed noticeably better than when using the DTI and the SeeReal displays (both $t(12)$ absolute values $>= 10.59$, both p values $< .001$) and not significantly differently than when using the Colorlink, Iris3D, Kodak or Shutter Glasses displays (all $t(12)$ absolute values $<= 1.52$, all p values $> .1$). These results are in contrast with the theoretical predictions and demonstrate that performance for the Sharp display was significantly better than the expected performance at chance and at the same level with performance of displays that have full horizontal view resolution (i.e. displays belonging to either the Row-Interleaved or the Full-Resolution classes).

The results for the remaining 3D displays, which all have full horizontal view resolution, can be summarised as follows.

Performance for the Shutter Glasses display ($M = 98.90\%$) was not significantly different than optimum performance ($t(12) = -1.76$ and $p = .104$), while performance for the Iris3D display ($M = 98.08\%$) was only marginally poorer than optimum performance ($t(12) = -2.01$ and $p = .068$). It can be concluded that performance for these two displays was at or close to the predicted 100%.

Performance for the Kodak display ($M = 91.76\%$) was significantly poorer than optimum performance (Kodak vs optimum: $t(12) = -2.33$ and $p = .038$) and than performance for the Shutter Glasses display ($t(12) = -2.26$ and $p = .043$), but only marginally worse than performance for the Iris3D display ($t(12) = 1.81$ and $p = .095$). No significant difference in performance was observed between the Kodak display and the ColorLink display ($t(12) = .96$ and $p = .357$). Concerning the Colorlink display ($M = 94.78\%$) performance was significantly worse than optimum performance ($t(12) = 4.16$ and $p = .001$) and also significantly worse than performance for the Shutter Glasses display ($t(12) = 4.21$ and $p = .001$) or the Iris3D display ($t(12) = 2.52$ and $p = .027$). In general terms, these figures suggest that performance for the Kodak and the ColorLink displays was below the predicted performance of 100%.

Summarizing the results for the 1-pixel disparity condition, not all displays performed as anticipated and there are examples of both half and full horizontal view resolution displays performing inconsistently with the performance predicted from their specifications. In particular, for the Column-Interleaved and Color-Column-Interleaved class the prediction was that no display should be able to show 1-pixel disparity; in these categories the SeeReal display performed as expected, the DTI display led to performance higher than predicted but far from optimum, while the Sharp display surprisingly demonstrated optimum performance. On the other hand, Full-Resolution and Row-Interleaved displays should all have the capability to show 1 pixel disparity; in these two categories the Shutter Glasses and Iris3D display essentially led to optimum performance, while performance for the Kodak and the ColorLink was below prediction. A more detailed discussion of these results can be found in Chapter 6.

Performance at 2-pixel Image Disparity

With a level of image disparity of 2 pixels all the displays were expected to be able to reproduce a detectable depth difference between the left and right square. The results for this experimental condition are shown in Figure 5.6.

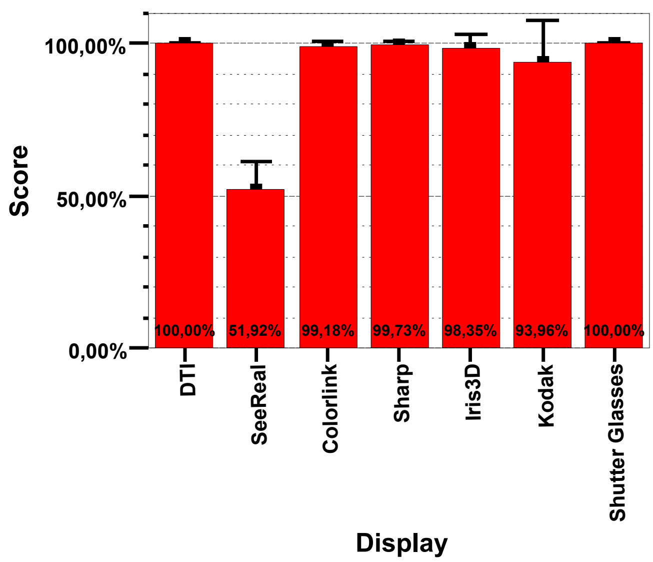


Figure 5.6: Mean score and standard deviation for the 2-pixel disparity condition.

At 2-pixel disparity performance was generally very high, with an overall mean score of 91.88%. Performance for the DTI display and the Shutter Glasses display was numerically as expected (both M = 100%, both s = .00), while performance for the Sharp display (M = 99.73%), the Iris3D display (M = 98.35%) and the Kodak display (M = 93.96%) was not significantly different than optimum despite the numerical differences in mean score (all $t(12)$ values < 1.57, all p values > .1). Performance for the Colorlink display was numerically high (M = 98.35%) and only marginally differed from optimal performance ($t(12) = 1.90$, $p = .082$); pairwise comparisons also revealed a marginal difference between performance for the Colorlink display and performance for the DTI and Shutter Glasses displays respectively (both $t(12)$ values = 1.90, both p values = .082).

The exception is represented by the SeeReal display, as the score for this display was surprisingly low (M = 51.92%) and remarkably worse than expected (M = 100%). Specifically, when using the SeeReal display participants' performance was significantly lower than when using any other display (all $t(12)$ values > 8.78, all p values < .001) and did not differ from performance at chance ($t(12) = .75$ and $p < .465$). This finding indicates that participants were unable to detect any difference in depth between the two squares and suggests that, in the conditions of this experiment, the SeeReal display did not have the capability to reproduce 2-pixel image disparity as perceivable depth.

Overall, in the 2-pixel disparity condition most displays performed reliably as expected, while the exception of the SeeReal display raised interesting questions about the reliability of 3D image presentation through the software, electronic and optical interfaces for this display.

Performance at 3-pixel Image Disparity

The results for the experimental condition with 3 pixels of image disparity are shown in Figure 5.7.

As it can be seen from the graph performance for this condition was generally very

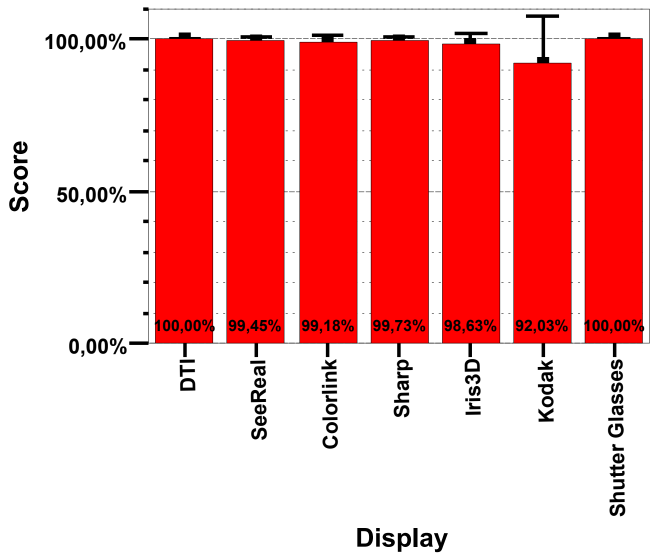


Figure 5.7: Mean score and standard deviation for the 3-pixel disparity condition.

high, with an overall mean score of 98.43%. Specifically, performance for the DTI display and the Shutter Glasses display was numerically as predicted (both M = 100%, both $s = .00$) and the same as for the condition with 2-pixel disparity, while performance for the SeeReal (M = 98.43%), Colorlink (M = 98.43%), Sharp (M = 98.43%) and Iris3D (M = 98.43%) displays was slightly lower in numerical terms but did not significantly differ from optimal performance (all $t(12)$ values = 1.59 or lower, all p values $> .1$). Performance for the Kodak display was also high (M = 92.03%) but marginally lower than optimum ($t(12) = 1.84$, $p = .091$) and also marginally poorer than performance for the DTI, Sharp and Shutter Glasses displays respectively (all $t(12)$ values = 1.81 or higher, all p values = .096 or lower). A series of pairwise comparisons among the displays' mean scores showed that no other numerical difference approached significance (all $t(12)$ values < 1.76, all p values $> .1$).

Overall, data for the 3-pixel disparity condition are similar to those observed for the 2-pixel disparity condition in that they revealed few significant differences in score. Notably, the SeeReal display now shows the expected optimal performance while there is a continuing wide variance in the Kodak display results.

Performance Trends for Individual Displays

To investigate the individual display related trends in the data, a number of paired t-test statistics were performed where the mean scores of the same display across the different disparity levels were compared against each other.

For the DTI display, the relatively low mean score value at 1-pixel disparity suggests that while participants could occasionally perceive a difference in depth between the left and the right square, the majority of times the two squares appeared to be in the same depth plane. In this regard the t-tests revealed that performance at 1-pixel disparity was reliably better than performance at 0-pixel disparity ($t(12) = -2.90$, $p = .013$), but significantly worse than performance at 2-pixel disparity ($t(12) = -12.39$, $p < .001$). This result reinforces the finding that with the DTI display the difference in perceived depth induced by 1 pixel image disparity was not reliably detectable.

Concerning the SeeReal display, the t-tests showed that there was no significant difference among mean scores at 0- 1- and 2-pixel disparity (all $t(12)$ absolute values $<= .79$, all p values $> .1$). This result supports the outcome of the analysis of data for 2-pixel disparity; that is, in the experimental conditions adopted for this study, the SeeReal display failed to show an image disparity of 2 pixels.

With respect to the Colorlink display, pairwise comparisons revealed that performance at 1-pixel disparity was significantly lower than performance at 2- and 3-pixel disparity (both $t(12)$ absolute values $>= 3.25$, both p values $< .01$). Despite this, the high numerical value of the mean score at 1-pixel disparity suggests that most of the time participants were correctly perceiving the depth difference between the two squares.

Finally, the analysis of the numerical trend of the Sharp, Iris3D, Kodak and Shutter Glasses displays revealed that for these displays there was no significant difference among the mean scores at 1- 2- and 3-pixel disparity (all $t(12)$ absolute values $<= 1.76$, all p values

> .1). This indicates that the above displays have the capability to show the difference in depth induced by these three levels of image disparity.

5.3.2 Response Time

Response time results are summarised in Table 5.4, where mean value and standard deviation are reported for each experimental condition. Means are also shown graphically in Figure 5.8.

Table 5.4: Response time results: mean, M, (ms) and standard deviation, SD.

Disparity	0		1		2		3	
	M	SD	M	SD	M	SD	M	SD
DTI	1858	704	1806	601	1008	863	769	216
SeeReal	1519	700	1666	711	1778	917	1005	734
Colorlink	1752	1213	1198	611	1041	869	801	375
Sharp	2281	1433	929	343	751	255	895	619
Iris3D	2284	1001	1119	408	935	381	789	192
Kodak	2276	1221	1527	1094	1096	575	1265	726
Shutter Glasses	2786	2222	1222	693	889	413	1018	637

Similarly to score data, response time data were also subjected to ANOVA, with *Disparity* and *Display* as within-subjects independent variables and *Response Time* as the dependent variable. Mauchly's test indicated that the assumption of sphericity had been violated for the main effect of disparity ($\chi^2(5) = 30.02, p < .001$) and the effect of display*disparity ($\chi^2(170)$ and p undefined). Therefore, degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity ($\varepsilon = .45$ for disparity and $\varepsilon = .22$ for display*disparity).

The ANOVA statistics revealed that there was a significant main effect of the amount of image disparity applied to the stimuli on participants' response time ($F(1.28, 15.32) = 51.22$, $p < .001$), while the main effect of the type of display on response time was non-significant ($F(6, 72) = .94$, $p = .469$). There was also a significant interaction effect

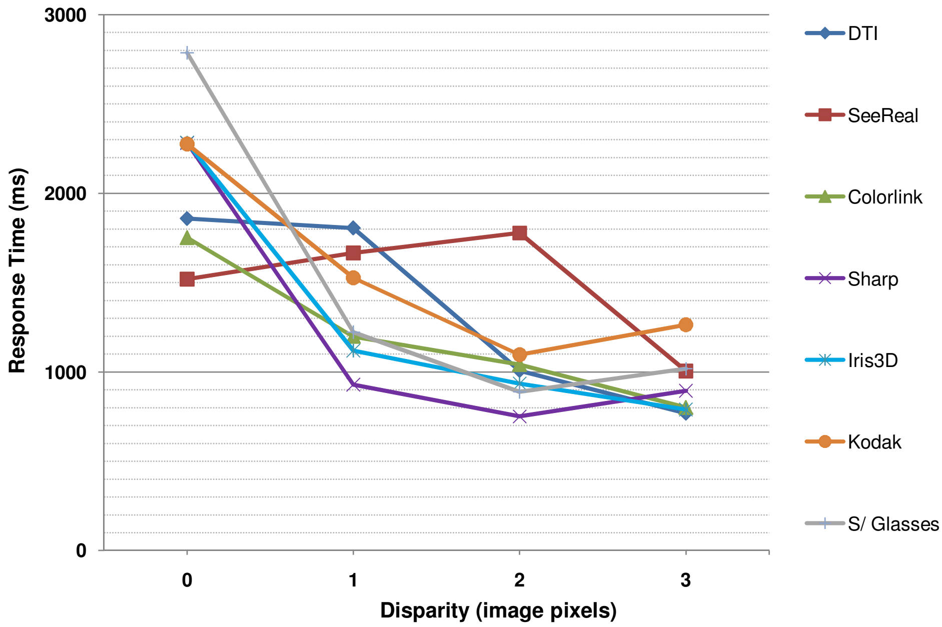


Figure 5.8: Response time means for all experimental conditions.

between the level of disparity used to treat the input image and the type of display used by participants to perform the task ($F(4.03, 48.38) = 2.73$, $p = .039$). This last point indicates that the seven displays tested during the experiment were affected differently by the amount of image disparity applied to the stimuli. In other words, this means that for some displays the level of disparity used had little or no effect on participants' response time, while for other displays it greatly affected how long it took participants to perform the task.

Given the non-significance of the display main effect, it was not possible to proceed with paired comparisons across experimental conditions with the same disparity level and different displays.

Performance for the DTI Display

The graph of Figure 5.9 shows mean response times and standard deviations for the DTI display.

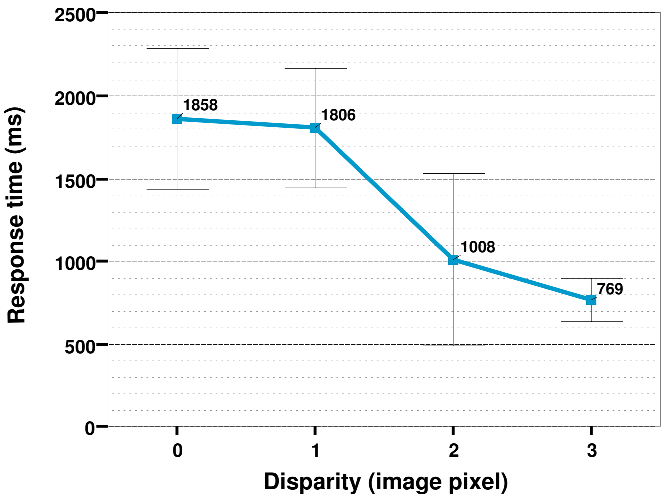


Figure 5.9: Mean response time and standard deviation for the DTI display.

Response time results for this display showed a trend that is not too dissimilar from the expected one, in that the mean response time was high at 0- and 1-pixel disparity with little numerical difference between these two conditions ($M(0 - pixel) = 1858$ ms, $M(1 - pixel) = 1806$ ms), and then dropped at 2-pixel disparity. However, there is an unexpected, evident difference between mean values at 2- and 3-pixel disparity respectively ($M(2 - pixel) = 1008$ ms, $M(3 - pixel) = 769$ ms).

To investigate if this apparent trend was statistically reliable, means at 0- and 1-pixel disparity were compared against means at 2- and 3-pixel disparity, for a total of four cross-condition paired comparisons. The t-test statistics revealed that at 0- and 1-pixel disparity participants were indeed significantly slower than they were at 2- and 3-pixel disparity (all $t(12)$ values $>= 2.46$, all p values $>= .029$). On the other hand, the mean difference between the 0- and 1-pixel disparity conditions was non-significant ($t(12) = .46$, $p = .652$), and so was the mean difference between the 2- and 3-pixel disparity conditions ($t(12) = .932$, $p = .370$).

These results suggest that participants had difficulty in performing the task at 1-pixel disparity, and that after that the level of ease increased with increased disparity. It is also worth noting though that despite the apparent anomalous trend of the mean values, the t-statistics indicate that response time data for the DTI display are according to predictions.

Performance for the SeeReal Display

Mean response times and standard deviations for the SeeReal display are shown in Figure 5.10.

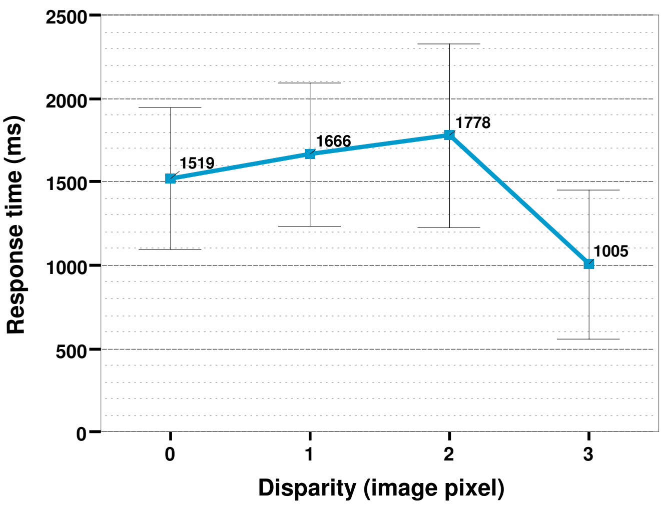


Figure 5.10: Mean response time and standard deviation for the SeeReal display.

Compared to the forecasted behavior, data for this display show an anomalous trend, with mean response times increasing slowly but steadily from 0-pixel disparity to 2-pixel disparity ($M(0-pixel) = 1519$ ms, $M(1-pixel) = 1666$ ms and $M(2-pixel) = 1778$ ms) and dropping only at 3-pixel disparity ($M(0-pixel) = 1005$ ms) rather than at 2-pixel disparity. Despite the fact that this last aspect of the response time data does not reflect the predicted trend, it is in agreement with the score data, which suggest that the SeeReal display failed to show an image disparity of 2 pixels.

The statistical analysis revealed that at 0-pixel disparity participants were marginally faster than at 1-pixel disparity ($t(12) = -1.82$, $p = .093$), while the mean numerical differences between the 0-pixel and the 2-pixel disparity conditions, and between the 1-pixel and the 2-pixel disparity conditions were both non-significant ($t(12) = -1.30$, $p = .217$; $t(12) = -.64$, $p = .535$ respectively). Furthermore, the t-tests confirmed that at 3-pixel disparity participants were significantly faster than at 2-pixel disparity ($t(12) = 2.25$, $p = .044$) and marginally faster than at 1-pixel disparity ($t(12) = 2.06$, $p = .062$).

Surprisingly, the mean difference between the 0-pixel and the 3-pixel disparity conditions was also reported to be non-significant. Following a detailed investigation this appeared to be due to noise in the data for the 3-pixel condition (see Appendix C for raw empirical data).

Performance for the Colorlink Display

Response time data for the Colorlink display are shown in Figure 5.11.

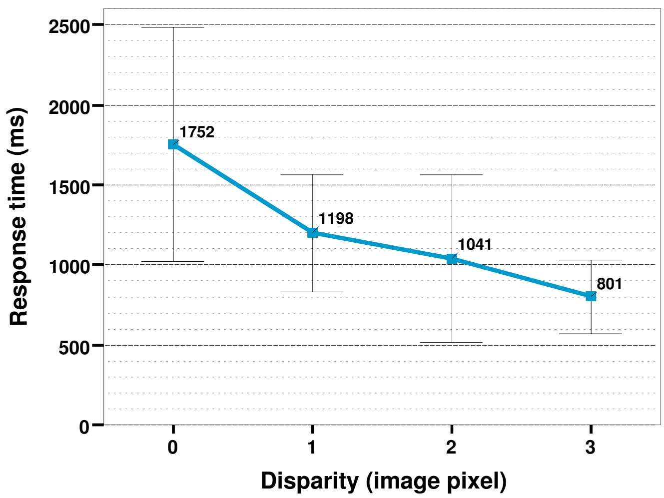


Figure 5.11: Mean response time and standard deviation for the Colorlink display.

According to the theoretical predictions, mean response times for this display should yield a high value at 0-pixel disparity, show a drop at 1-pixel disparity and then remain constant after that. As it can be seen from the graph, the data for the Colorlink display show the expected drop in mean value between the 0-pixel disparity ($M = 1752$ ms) and the 1-pixel disparity ($M = 1198$ ms) conditions, but after that they are characterized by an apparent, steady decrease in mean response time with increasing disparity ($M(2-pixel) = 1041$ ms, $M(3-pixel) = 801$ ms).

A pairwise comparison revealed that participants were indeed faster at 1-pixel disparity than at 0-pixel disparity ($t(12) = 2.52$, $p = .027$). On the other hand, the t-test statistics revealed also that the numerical mean differences between the 1-pixel and the 2-pixel disparity conditions, and between the 2-pixel and the 3-pixel disparity conditions were both non-significant ($t(12) = .73$, $p = .479$ and $t(12) = .97$, $p = .351$ respectively). However, participants were significantly faster when the stimuli were treated with 3-pixel disparity than when they were treated with 1-pixel disparity ($t(12) = 2.52$, $p = .027$).

Interestingly, the t-tests also indicated that at 0-pixel disparity participants were significantly slower than they were at 3-pixel disparity ($t(12) = 3.52$, $p = .004$), but only marginally slower than in the condition with 2-pixel disparity ($t(12) = 1.82$, $p = .094$). This last aspect of the results is quite surprising, but further investigation showed that it is due to the high variability that characterizes the data for this display (see Appendix C for an overview of the raw data).

Performance for the Sharp Display

The Sharp display yielded a data trend that is in disagreement with the theoretical predictions, in that it presents an apparent drop in mean response time at 1-pixel disparity instead of at 2-pixel disparity as was expected ($M(0-pixel) = 2281$ ms, $M(1-pixel) = 929$ ms, $M(2-pixel) = 751$ ms and $M(3-pixel) = 895$ ms). Similar to the SeeReal display though, this trend is in line with the score data and seems to reinforce the fact that the Sharp display does have the capability to reproduce a disparity of 1 image pixel.

The t-test statistics largely confirmed what was first observed at a glance in the data. In particular, at 1- 2- and 3-pixel disparity participants were indeed significantly faster than they were at 0-pixel disparity ($t(12) = 4.01$, $p = .002$; $t(12) = 4.41$, $p = .001$; $t(12) = 3.69$,

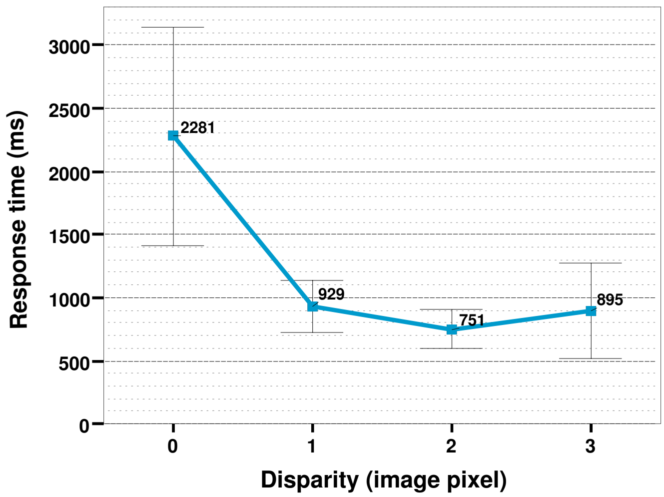


Figure 5.12: Mean response time and standard deviation for the Sharp display.

$p = .003$). Besides, participants' mean response time at 3-pixel disparity reported to be non-significantly different than means at 1- and 2-pixel disparity ($t(12) = 4.01$, $p = .002$; $t(12) = 4.41$, $p = .001$).

An interesting aspect of the data is that at 2-pixel disparity participants were significantly faster than at 1-pixel disparity ($t(12) = 2.41$, $p = .033$). This finding seems to suggest that it was easier to perform the task when the stimuli were treated with an image disparity of 2 pixels, however this effect could also be due to noise in the data.

A graphical representation of response time means and standard deviations for the Sharp display is given in Figure 5.12.

Performance for the Iris3D Display

Response time data for the Iris3D display are largely accordant with the expected trend and yielded the following mean values: $M(0-pixel) = 2284$ ms, $M(1-pixel) = 1119$ ms, $M(2-pixel) = 935$ ms and $M(3-pixel) = 793$ ms. As can be seen from Figure 5.13 the response time mean presented the expected drop in correspondence of the 1-pixel disparity condition, however after that it registered an unexpected, apparent steady decrease as the disparity increased.

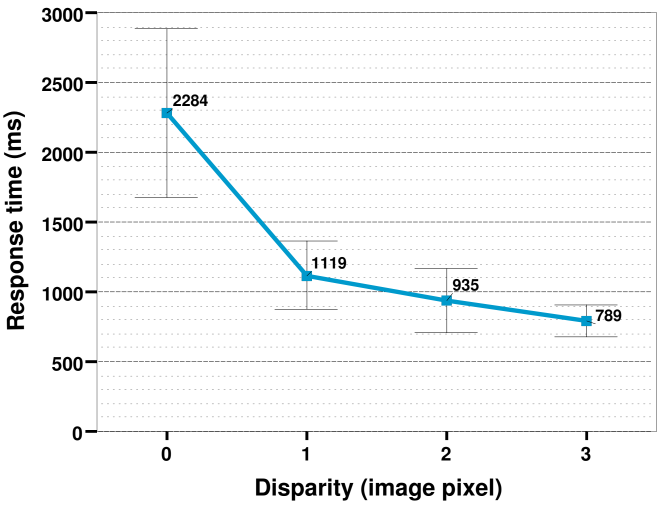


Figure 5.13: Mean response time and standard deviation for the Iris3D display.

The t-test statistics confirmed this trend and revealed that when using this display participants were significantly slower at 0-pixel disparity than they were at 1- 2- and 3-pixel disparity ($t(12) = 5.27$, $t(12) = 4.49$ and $t(12) = 6.07$ respectively; all p values $<= .001$). Furthermore, mean numerical differences between 1- and 2-pixel disparity, and between 2- and 3-pixel disparity were non-significant ($t(12) = 1.40$, $p = .189$; $t(12) = 1.38$, $p = .193$), while mean at 3-pixel disparity was significantly lower than mean at 1-pixel disparity ($t(12) = 4.27$, $p = .001$).

Performance for the Kodak Display

Response time means and standard deviations for the Kodak display are shown in Figure 5.14. As it can be observed from the graph, response time data for this display follow the expected trend for the 0- ($M = 2279$ ms) and 1-pixel ($M = 1527$ ms) disparity conditions, but after that they present some numerical differences ($M(2 - pixel) = 1096$ ms and $M(3 - pixel) = 1265$ ms) that, based on the theoretical predictions, were not expected.

The statistical analysis revealed that participants were significantly faster at 2- and

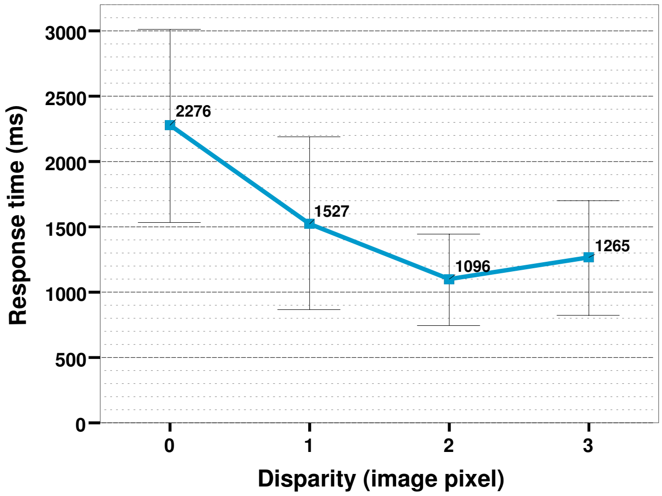


Figure 5.14: Mean response time and standard deviation for the Kodak display.

3-pixel disparity as compared to the 0-pixel disparity condition ($t(12) = 4.71$, $p = .001$ and $t(12) = 4.05$, $p = .002$ respectively) as expected, but the numerical difference between the 0-pixel and the 1-pixel disparity conditions was surprisingly non-significant ($t(12) = 1.76$, $p = .104$). Furthermore, mean differences between the 1-pixel disparity condition and the 2- and 3-pixel disparity conditions were also reported as non-significant ($t(12) = 1.55$, $p = .147$ and $t(12) = .89$, $p = .392$). On the other hand, the t-test statistics revealed that at 2-pixel disparity participants were significantly faster than at 3-pixel disparity ($t(12) = -2.51$, $p = .027$).

The statistical results for the Kodak display appear to be confusing at first. However, a more detailed analysis suggests that these apparently contradicting results are probably due to the rather high variability that generally characterizes the response time data for this display, and the noise present in the 1-pixel disparity condition (see Appendix C for more details).

Performance for the Shutter Glasses Display

The Shutter Glasses display yielded a response time data trend similar to the one of the Kodak display, in that the mean response time is high at 0-pixel disparity ($M = 2786$ ms), it then drops at 1-pixel disparity ($M = 1222$ ms), keeps on dropping until 2-pixel disparity ($M = 889$ ms) but then it shows an increase at 3-pixel disparity ($M = 1018$ ms). This can readily be observed in the mean and standard deviation graph of Figure 5.15.

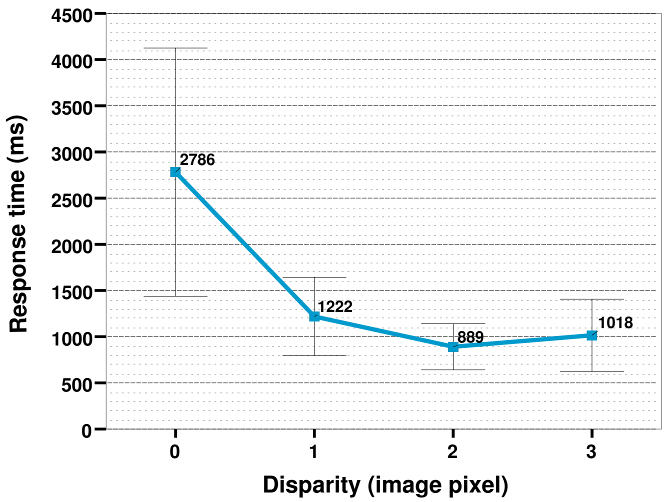


Figure 5.15: Mean response time and standard deviation for the Shutter Glasses display.

A number of paired t-test were performed in order to investigate the unexpected numerical differences among the non-null disparity conditions. These indicated that mean response time at 3-pixel disparity was non-significantly different than means at 1- and 2-pixel disparity respectively ($t(12) = .88$, $p = .398$ and $t(12) = -.731$, $p = .479$), as expected. However, the tests also revealed that at 2-pixel disparity participants were significantly faster than at 1-pixel disparity ($t(12) = 3.24$, $p = .007$).

Finally, the numerical differences between the 0-pixel disparity condition and the other three disparity conditions were all reported significant as expected (0-pixel versus 1-pixel: $t(12) = 2.79$, $p = .016$; 0-pixel versus 2-pixel: $t(12) = 3.12$, $p = .009$; 0-pixel versus 3-pixel: $t(12) = 2.72$, $p = .018$).

5.3.3 Subjective Results

Similar to the pilot experiment, participants were asked to express their personal opinion about the 3D displays used for the trials and subjective data were collected and analysed. In particular, at the end of each experimental session participants completed a detailed questionnaire concerning the display they had just used (level of ease seeing 3D, disturbing factors, level of discomfort, general comments about the display), while at completion of the whole experiment they were asked to fill in a general, more detailed questionnaire.

As part of the post-trial questionnaire, participants were asked to rank the seven displays from best (1) to worst (7) based on their personal experience during the experiment. Figure 5.16 shows a cumulative frequency histogram based on the participants' subjective ranking results, with average ranking score for each display shown across the top of the graph. On average, the Sharp display was ranked as best (score = 3.00) while the SeeReal was ranked as worst (score = 4.69).

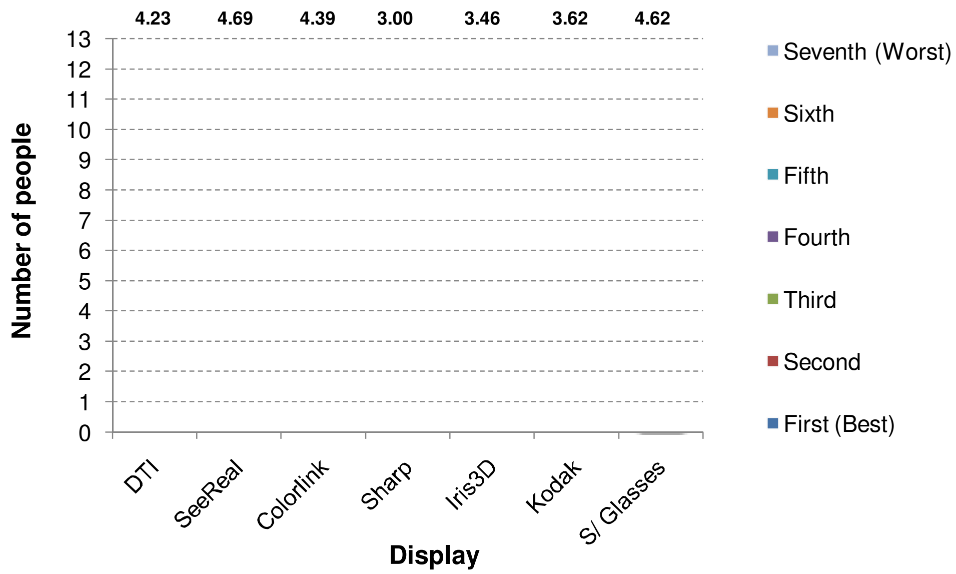


Figure 5.16: Subjective ranking of displays.

An interesting aspect of the ranking data is the divided opinion among participants

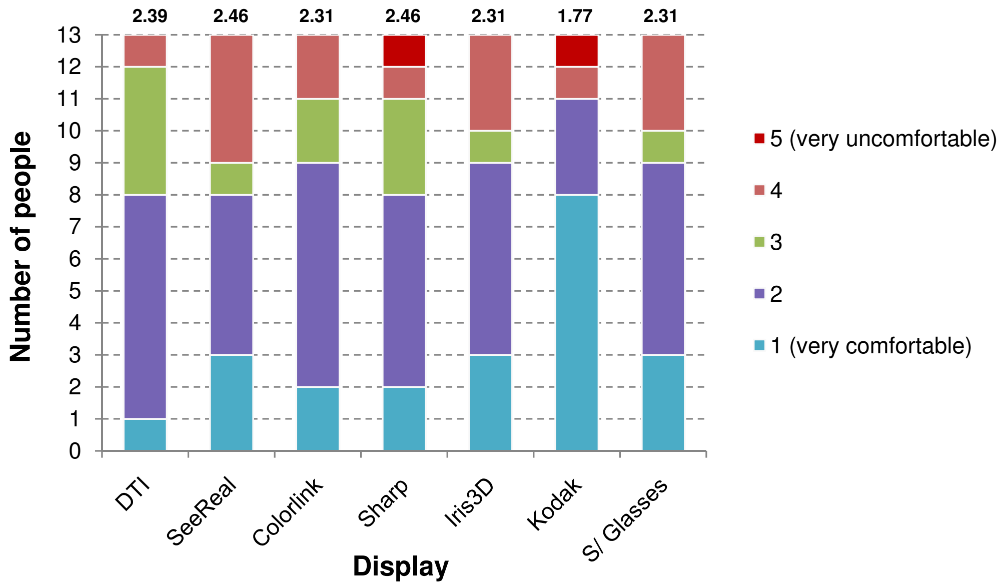


Figure 5.17: Subjective rating for level of comfort in seeing the 3D image on each display.

regarding some of the displays. For example the Shutter Glasses display, the Kodak display and the Iris3D display had several high rankings as well as several low rankings.

At the end of each display session, participants were asked to express how comfortable it was to see the 3D image on the specific 3D display and, if they experienced any discomfort, how strong it was on the specific display. In both cases answers were given via a 5-point Likert scale [Likert, 1932].

Figure 5.17 shows a cumulative comparison of results for the first question, with average level of comfort in seeing the 3D image on the specific display shown across the top of the graph. The scale used ranged from 1 = “very comfortable” to 5 = “very uncomfortable”. On average, participants found that the Kodak was the best display in terms of how comfortable it was to see the 3D stimuli (score = 1.77), while the Sharp and the SeeReal displays were the worst displays (both scores = 2.46).

The results regarding discomfort are shown in Figure 5.18 and 5.19 respectively.

Overall, these results suggest that most participants did not experience any discomfort with most displays. In other words, the percentage of people that experienced some

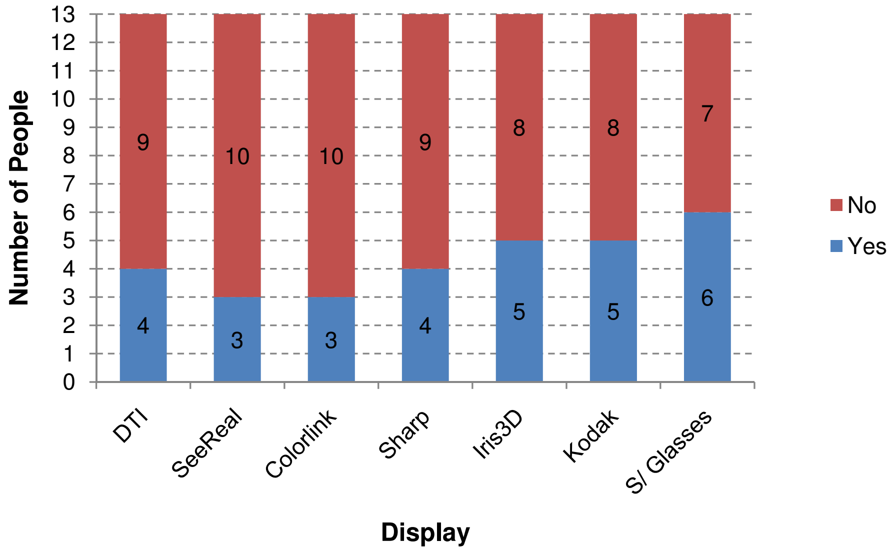


Figure 5.18: Subjective results for physical discomfort on each display.

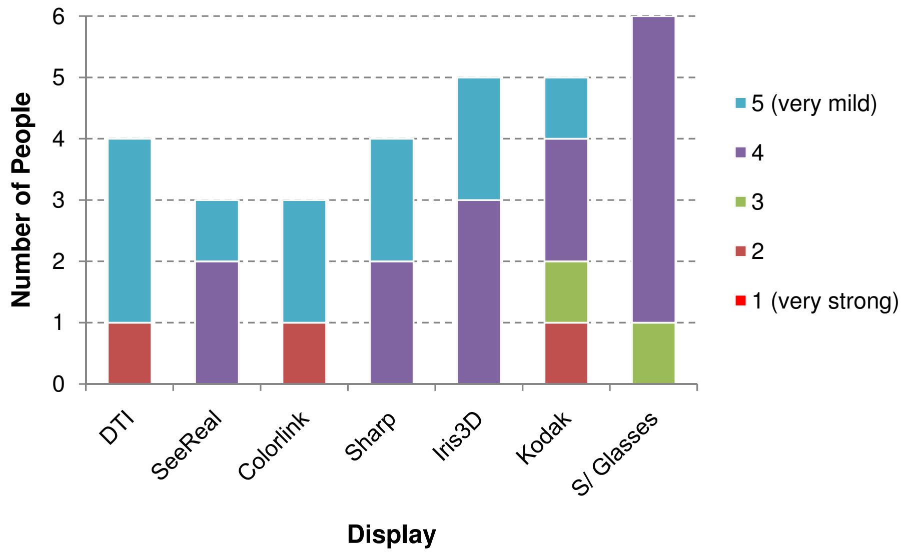


Figure 5.19: Subjective rating for level of discomfort on each display.

discomfort (e.g. headache, eye strain, etc.) at some point during the experiment was only 33%, against the 67% that did not experience any discomfort. Looking at the subjective data in more detail, it can easily be seen that when using the Shutter Glasses display six participants out of 13 experienced some form of discomfort, while when using the SeeReal and the Colorlink displays only three participants out of 13 perceived discomfort. On the other hand, in the cases where discomfort was perceived, the Kodak is the display with the strongest average level of discomfort (score = 3.6), while the Sharp is the one with the mildest (score = 4.5). Note that these scores are expressed on the same Likert scale used in the questionnaire, where 1 = "very strong discomfort" and 5 = "very mild discomfort".

5.4 Conclusions

In this chapter an improved methodology for assessing and comparing 3D displays has been presented. This includes a number of improvements as compared to the one used for the pilot experiment and it has successfully been used to perform a full-scale comparative experiment for the investigation of human binocular depth perception across a set of representative 3D displays [Froner et al., 2008]. During the main experiment seven different two-view desktop displays were tested across a 0- to 3-pixel range of image disparity, and participants' score and response time were recorded.

The results of the main experiment and the relative statistical analysis were presented in detail in this chapter. Empirical data were analysed using ANOVA and t-test statistics and reported accordingly, while subjective data were described using non-statistical analysis. These, together with the material considered in Chapter 3 and Chapter 4, form the basis for the discussion presented in the next chapter of this thesis.

Chapter 6

Discussion

6.1 Introduction

The previous two chapters of this thesis discussed a new methodology used to compare and assess different 3D displays and presented the empirical results of the 3D display comparative study. During the study seven different displays were tested across four different levels of image disparity for a total of 28 experimental conditions. Score and response time were recorded quantitatively as respective measures of task accuracy and task difficulty, while subjective data were collected via structured questionnaires. The statistical analysis of the data yielded a number of interesting points. The empirical results were often in accordance with the predicted outcome, however this was not always the case. The aim of this chapter is to present an interpretative discussion of the main findings emerging from the data analysis and examine in detail the reasons why in some instances the theoretical predictions were not supported by the results. The outcomes of the study are then used to clarify the effect of a number of technical factors on the perception of depth in stereoscopic 3D environments and to amend the 3D cursor design recommendations presented in Chapter 3.

This chapter is organised as follows. Section 6.2 discusses the main findings of the study from a purely 3D display technology point of view; score, response time and subjective results are interpreted independently in sections 6.2.1, 6.2.2 and 6.2.3 respectively. Section 6.3 presents a general discussion that summarises and ties together the main outcomes of the empirical study. Section 6.4 explains how the technical factors that characterise different 3D display systems can affect the reproduction of depth embedded in a stereo image and discusses the implications that these have on the design of a stereoscopic 3D cursor. Finally, section 6.5 concludes the chapter.

6.2 Results Interpretation

6.2.1 Score

Overall, the analysis of the score data showed that participants had high performance when a difference in depth between the two squares was perceivable. This provides evidence that the task was appropriate and clearly defined. It also suggests that the score data obtained during the study are reliable and provide a meaningful insight into the characteristics of different 3D displays when used in tasks requiring depth judgements.

The ANOVA statistics revealed that the three main effects (i.e. disparity, display and interaction between the two) were all extremely significant for both the pilot and the main experiments. There was a strong effect of disparity on participants scores, which indicates that the participants' ability to detect a depth difference between the two squares was highly dependent on the amount of image disparity applied to the stimuli. This effect was most clearly shown at image disparities of 3 pixels or higher, where a depth difference was always perceived, and at 0-pixel disparity, where no depth difference was perceived. The only exception is represented by the Iris3D display during the pilot experiment, as discussed in the next section of this chapter.

With regard to the display variable, the effect on score was also strong and the results support the prediction that the display chosen to carry out the task had an influence on task performance. An interesting finding is that while the tested displays often yielded the expected behavior, this was not always the case and displays whose specifications suggest they are the same may not in practice perform equivalently.

Additionally, the ANOVA test revealed a strong interaction between the two independent variables; that is, there was a strong combined effect of display and disparity on participants' score. Therefore it is not possible to fully consider the scores for perception of depth in this task without taking into account both the display used and the level of image disparity applied to the stimuli, for the impact of display on performance depends on the amount of image disparity. Practically this means that there are differences in performance across the disparity range between displays. Some but not all such differences were predicted from the display specifications and the following discussion explains why some of the predictions were not supported by the empirical results.

Performance at 0-pixel Image Disparity

The analysis of the main experiment score data for the 0-pixel disparity condition revealed that performance for all seven displays was not significantly different than chance and in accordance with the predictions. From a more pragmatic point of view, this means that all displays operated correctly for this condition and presented the scene accurately.

In regard to the pilot experiment, performance at chance was observed for all but one display. In particular, score performance for the Iris3D display was significantly different than chance and significantly different to any other tested display. This unexpected result clearly indicates that participants could perceive depth when no image disparity was present in the stimuli, and reasons why this would happen had to be diagnosed.

A detailed analysis of the Iris3D technical characteristics led to the identification of three factors that could have potentially affected the perception of depth during the experiment. The first is optical and is related to reflections produced by this display while the user is observing the 3D scene. A number of participants reported noticing secondary peripheral reflections of the actual stimuli, which could be distracting and could introduce a bias in their decision process. For example, it is possible that participants used the reflections as a cue to depth to discriminate between the two squares even when no actual difference in depth was shown. The second factor is electronic, as one channel from the driving PC was fed to the Iris3D display via a video splitter in order to simultaneously drive an external 2D monitor [McKay, 2005]. This could result in a delay to the signal to one eye therefore affecting the stereo perception of participants. The third possibility is a misalignment in the mechanical components that form the optical paths of the display itself. It turned out that the latter was the case and it was found that indeed there had been a small mechanical alignment error in the prototype display used in the pilot experiment. This was corrected prior to running the main experiment and as a result the Iris3D display then showed the predicted performance at chance.

Based on this discussion, a first important conclusion is that the evaluative methodology presented in this thesis is sensitive enough to detect this low level of alignment issue and provides an effective tool for testing the optical-mechanical alignment of 3D display systems.

Performance at 1-pixel Image Disparity

The results relative to the 1-pixel disparity condition are probably the ones that yielded the most interesting findings from a 3D display assessment point of view.

A first important point is that when using full horizontal view resolution displays, i.e. Colorlink, Iris3D, Kodak or Shutter Glasses, participants' score was generally high and significantly better than chance; this is strong evidence to confirm the hypothesis that Full-Resolution and Row-Interleaved 3D displays have the capability to reproduce perceivable depth induced by 1-pixel image disparity in a repeatable, reliable way. However, it is worth noting that there were some variations in performance among the above four displays. In particular, when using the Shutter Glasses and the Iris3D displays participants reached higher scores with lower variance than when using the Colorlink or the Kodak displays. These differences might be due to the variations in system design between the displays. For example, the Colorlink display implements half vertical screen resolution per view and displays stereo images in a row interleaved pattern, which results in a vertical misalignment between the left and right view. It is plausible to assume that these factors, lower vertical screen resolution and vertical misalignment, could affect the binocular vision matching process and could explain the variation in the results.

A detailed analysis of this last point led to a more precise conclusion. In literature there are a number of studies that investigated the effect of vertical misalignments on stereopsis and depth perception [Fukuda et al., 2009] [Kooi and Toet, 2004] [van Ee and Schor, 2000] [Prazdny, 1985] [Nielsen and Poggio, 1984] [Mitchell, 1970]. From these it has emerged that the human stereoscopic system is tolerant to some vertical misalignment in the two images seen by the eyes [Fukuda et al., 2009]. Nevertheless, when too prominent, vertical misalignments become disruptive to the detection of depth derived from horizontal disparity [Howard and Rogers, 2002b] and therefore can affect stereovision. In this regard results are very different among studies, with reported tolerance to vertical misalignments ranging from $3.4 - 3.5$ minutes of arc [Kooi and Toet, 2004] [Nielsen and Poggio, 1983] to 4 degrees [Mitchell, 1970] depending on the experimental conditions and the nature of the stimuli used during the experiments. Based on the calculations presented in Appendix D.2, the vertical misalignment between the two white square stimuli used for the 3D display comparative study when displayed on the Colorlink display is equivalent to 1.39 minutes of arc, which is well below the most conservative tolerance value of $3.4 - 3.5$ minutes of arc reported by Kooi and Toet [2004] and Nielsen and Poggio [1984]. This suggests that the vertical misalignment embedded in the Colorlink display most probably had no disruptive effect on task performance and that the variability registered in the data must have been due to the reduced vertical screen resolution and possibly crosstalk (see section 2.3.4).

Moving on to consider the Column-Interleaved and the Colour-Column-Interleaved displays, it was expected that for the 1-pixel image disparity condition no difference in depth between the two squares would be perceived. However, the score results showed that this was the case only for the SeeReal display, while with the DTI display participants could sometimes detect depth and with the Sharp display all participants achieved quasi-optimal performance. These variations in performance can be explained by the different drivers and interface electronics that characterise these three displays, as described below.

The SeeReal display uses the NVIDIA Vertical Interlace Monitor mode to generate the column-interleaved stereo image for display, which appears to eliminate alternate columns of pixels in the left and right channels that form the input stereo image. This means that any 1-pixel image disparity differences will be removed from the stereo pair before the electronic video signal is transmitted to the display. On the other hand, the DTI display operates using an analogue time-sequential video signal where full resolution left and right images are sent from the graphics card to the display; electronics in the display then generate the column-interleaved image for presentation. To achieve this the display must be sampling the incoming analogue signal exactly on alternate pixel columns. If these samples are not timed precisely then it is possible there will be sampling of information from the wrong pixel column, resulting in the possibility of information from 1-pixel disparities appearing in the final image. Since this is an analogue process it can generate a partial sample, which would explain why the participants only sometimes detected depth. Finally, the Sharp display uses a digital signal that is colour-column-interleaved, not column-interleaved. This means that within one view pixel there are components of both left and right input images, for example red and blue may be left and green right, or vice versa; this allows for a finer interleaving of the left and right images. One way to implement this in the driver is to mask the left and right images in alternate colour-columns, as shown in Figure 6.1. If this is the case, a partial retention of 1-pixel disparity

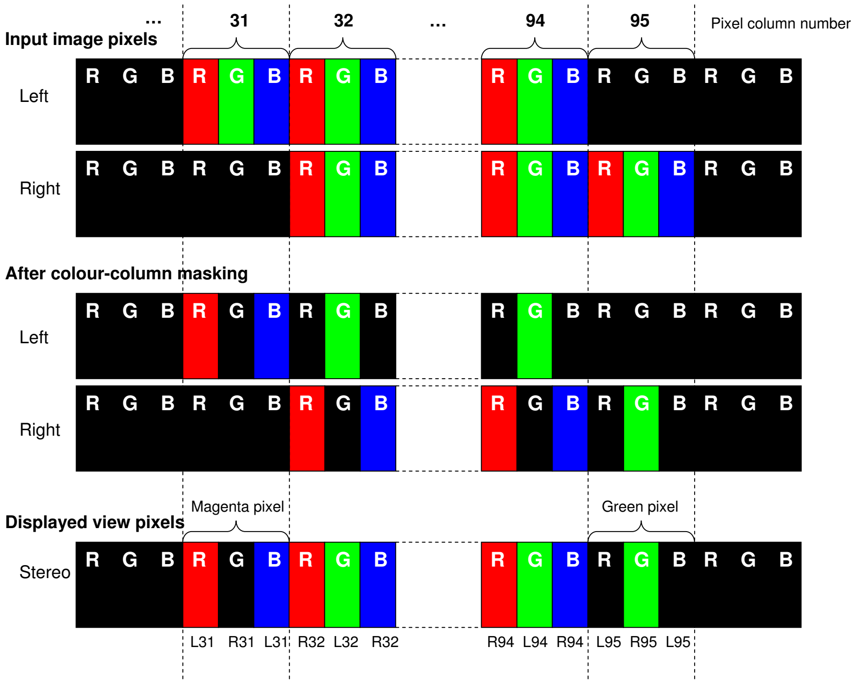


Figure 6.1: Colour-column interleaving pattern for a single row of pixels from a white square with 1-pixel image disparity. The interleaving pattern is shown in detail only for the left and right edges of the square respectively, as indicated by the pixel column counting.

information is possible (in either the red and blue, or the green channel). With the white test stimulus used in this study this should be enough to present detectable, but colour-rivalrous, disparity in the 1-pixel level condition. A detailed inspection of this matter suggested that this is the case; Figure 6.2 shows the *colour fringing artifact* as seen on the Sharp display when it is in 2D mode. The questionnaire results indicate that some participants observed this effect and reported noticing the colour fringing at the edge of the stimuli's white squares. The difficulties are that this behavior is not evident from the display specifications and in colour images it may not produce detectable disparity at 1-pixel for certain colours, for example green or magenta.

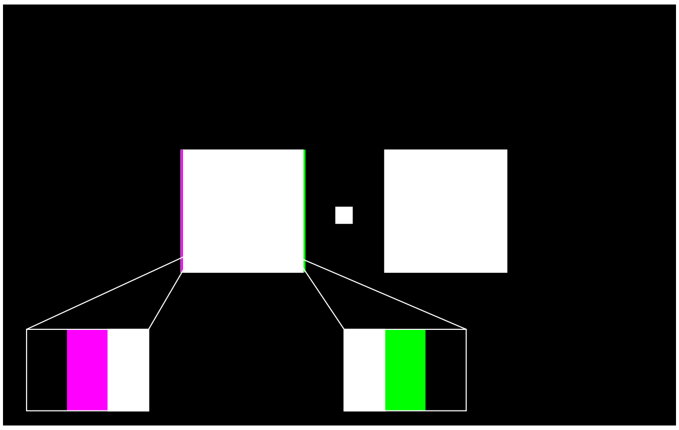


Figure 6.2: Colour fringing artifacts caused by the colour-column interleaving.

Performance at 2-pixel Image Disparity

In the condition of 2-pixel image disparity an increasing number of displays behaved as expected. According to the geometric predictions, all the tested displays should have the capability to reproduce this amount of image disparity as perceivable depth. The statistical analysis revealed that, despite some marginal numerical differences, this was the case for the DTI, Colorlink, Sharp, Iris3D, Kodak and Shutter Glasses displays. Score for these six displays was high and not substantially different than the expected optimum performance for both the pilot and the main experiment. However, it is worth noting the high level of variability that characterises the Colorlink score results in the pilot study ($SD = .28$ against a standard deviation of .08 or lower for the other displays) and the relatively high variability in the main experiment score results for the Kodak display ($SD = .14$ against a standard deviation of .05 or lower for the other displays).

On the other hand, the results for the SeeReal display were not consistent with the predictions. The statistical analysis clearly showed that with this display participants were unable to perceive depth for the 2-pixel image disparity condition and therefore performed at chance. After a detailed investigation it appeared that this effect was due to aliasing of the input image disparity resulting from the column-interleaving algorithm implementation

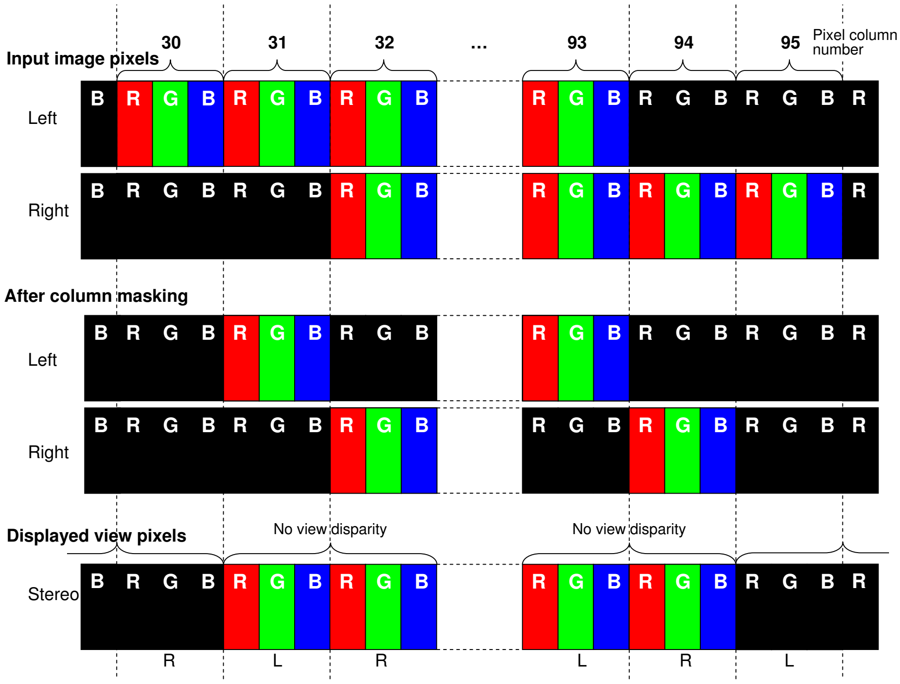


Figure 6.3: Column interleaving pattern and relative disparity masking effect for a single row of pixels from a white square with 2-pixel image disparity. The interleaving pattern is shown in detail only for the left and right edges of the square respectively, as indicated by the pixel column counting.

in the NVIDIA display driver. For the comparative study presented in this thesis, in adjusting the input image so as to obtain stimuli with 2-pixel image disparity, the test square was shifted by one image pixel in opposite directions in the left and right channels of the input stereo image respectively. The outcome was that both adjustments were masked by the interleaving process shown in Figure 6.3 and the result was no visible view disparity, and hence no perceivable depth difference between the two squares. This is because the column-interleaving algorithm can effectively delete the 2-pixel image disparity by removing one column of pixels from opposite sides of the left and right channel test square, resulting in a slightly narrower square in each view with no view disparity information. It is important to note in this regard that Figure 6.3 only shows the effect of the column interleaving pattern on the view disparity in the displayed interleaved image. In fact, on the SeeReal display the two stimuli's white squares did indeed lie on the same depth plane (i.e. the zero disparity plane, due to the lack of view disparity in the output image), but they appeared slightly in front of the physical screen plane. This is because on Column-Interleaved displays the zero disparity plane and the physical screen plane are not coincident, as explained in section 4.4.3.

Going back to the main discussion, an important observation is that the interaction between the input stereo image pair, the driver and the display varies depending on the driver and the choice of columns it shows to the user; this is probably why it only affected the SeeReal and not the DTI display. Detailed investigations on this matter showed that a different choice of disparity pattern would indeed affect the results obtained in this study and the DTI rather than the SeeReal would be unable to reproduce 2-pixel image disparity as view disparity. For example, if the adjustment in the input stereo image had been done by shifting the test square in one image channel by two image pixels (rather than by one image pixel in each input image channel), participants would have seen some view disparity on the SeeReal display, as illustrated in Figure 6.4. Furthermore, the same disparity aliasing effect obtained in this study for the SeeReal display could be generated for the DTI display simply by altering the starting image position of the test square from even to odd pixel columns. In both cases the source of this aliasing is the software drivers in combination with the electronic interface rather than the display optics.

One important conclusion that emerges from the above discussion is that, whatever the outcome, this type of disparity aliasing is clearly a problem for Column-Interleaved displays in applications requiring fine depth judgement. However, one possible solution to the issue is for the driver to average the two columns into one rather than simply deleting one of them.

On the other hand, the experiments showed that Colour-Column-Interleaved displays

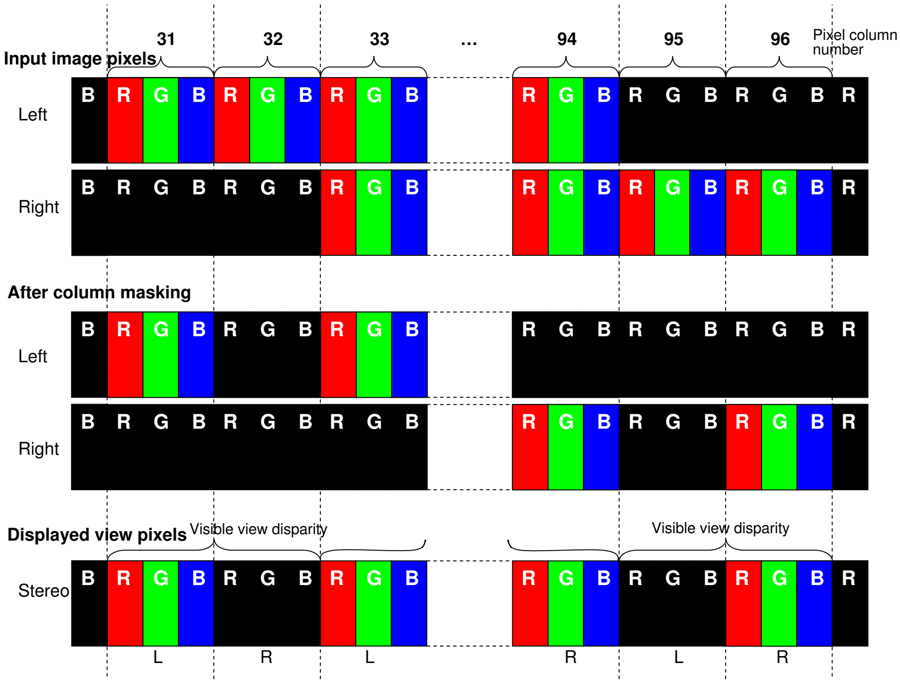


Figure 6.4: Column interleaving pattern with visible view disparity for a single row of pixels from a white square with 2-pixel image disparity. The interleaving pattern is shown in detail only for the left and right edges of the square respectively, as indicated by the pixel column counting.

are not affected by such disparity aliasing artifacts and have the capability to show all levels of view disparity tested in this study as perceivable depth. However they do suffer from colour fringing artifacts, which could affect the depth perception of objects of certain colours. For example, it is easy to anticipate that this would be the case for objects with strong green, or red and blue colour components.

Performance at 3-pixel Image Disparity and Above

In the conditions where the stimuli were treated with an image disparity of 3 pixels or higher, score performance was according to predictions for all displays. That is participants could reliably detect a depth difference between the two squares on all the displays.

Interestingly, in the main experiment's 3-pixel disparity condition the Kodak display reported again a notably larger variance than any other display. This is a curious fact that is worth trying to explain, given that the Kodak display, by design, provides two full resolution screens, very well aligned views and many viewers reported the display to be remarkably comfortable to use. The optical design of the Kodak has the distinguishing feature that the image as viewed is seen at optical infinity and also requires the viewer to verge towards parallel to match zero disparity homologous points in the stereo image. It is possible that for viewers with slight short-sightedness or mild strabismus these optical design factors are sufficiently different to their real world viewing experience to cause binocular matching difficulties. A detailed investigation of the main experiment results showed that one participant performed well on all displays except the Kodak, where they could not reliably detect depth even at 3-pixel disparity. Whether this might be due to focus, vergence or other issues would require further ophthalmic investigation.

In the pilot experiment, the score results for the Kodak display were not characterised by such high variability. However, among the three Full-Resolution displays tested in this study the Kodak was still the display with the highest standard deviation and lowest average score in all non-zero disparity conditions.

Another interesting point regarding the pilot experiment is the difference in the level of variability associated with the different displays, with interleaved displays (i.e. Row-Interleaved, Column-Interleaved and Colour-Column-Interleaved) registering on average a higher standard deviation than non-interleaved, Full-Resolution displays. This pattern is particularly evident in the higher image disparity conditions (i.e. 4 and 6 image pixels) and suggests that some participants might have found it more difficult to perform the task using the interleaved displays as compared to the non-interleaved displays. It is possible that the optical design and mechanics of the interleaved displays introduce a limit in the range of comfortably viewable depth that can be reliably "used" so as to ensure high levels of accuracy in tasks where depth judgements are critical. Nonetheless, the study leaves scope to further investigate the possible sources and effects of this variability.

6.2.2 Response Time

Overall, the response time data yielded the expected trends, with participants being on average significantly faster to perform the task when the depth difference between the two squares was detectable as compared to the conditions where no depth difference could be perceived. This reinforces the fact that the task was appropriate, clearly defined and understood by participants. However, the response time data were also noisier than the score data. Worthy of note in this regard is the fact that in the conditions where a depth difference could be perceived, the main experiment samples were characterised by lower standard deviation and lower mean values than the pilot experiment samples. Although this indicates a better control for biasing factors in the main experiment, the data suggest that there is scope for improvements in the technique used for time data collection; for example, a more sophisticated apparatus than a simple keyboard (e.g. a button box) could be used in order to minimise time overhead and increase accuracy in the measured response time.

The ANOVA statistics reflected this aspect and revealed that the three main effects (i.e. disparity, display and interaction between the two) were generally weaker than the ones observed for the score data. In particular, during the pilot experiment the amount of disparity applied to the stimuli had a strong effect on participants' response time, while the type of display used and the interaction between disparity and display both had a non-significant effect. Regarding the main experiment, the effect of disparity was again significant and so was the effect of the disparity-display interaction though to a lesser extent, whereas the effect of display was non-significant. These results suggest that the level of difficulty in performing the task was mainly dependent on the amount of image disparity used to treat the stimuli and, for the main experiment, on the combined effect disparity-display. On the other hand, the type of display used to perform the task seemed to have no significant influence on task difficulty.

Performance for Column- and Colour-Column-Interleaved Displays

The Column-Interleaved (DTI and SeeReal) and the Colour-Column-Interleaved (Sharp) displays yielded results that were in disagreement with the geometric predictions. This discrepancy is particularly evident at 1-pixel image disparity for the DTI and Sharp displays, and at 2-pixel image disparity for the SeeReal display. In fact, these results are perfectly in line with the score results and reinforce the finding that the DTI and Sharp have the capability to reproduce 1-pixel image disparity as perceivable depth and that, on the other hand, the SeeReal failed to show 2-pixel image disparity as view disparity.

The apparent increase in response time registered for the DTI and SeeReal displays at the higher disparity levels (i.e. 4- and 6- image pixels) was statistically not reliable, which suggests that, as expected, when using these displays the difficulty of the task did not increase with increasing image disparity.

Concerning the Sharp display, the results seem to suggest that participants found it easier to perform the task in the disparity range of 2 to 4 image pixels, while for lower image disparity levels (i.e. 1 image pixel) and higher disparity levels (i.e. 6 image pixels) they found it harder. Given this pattern, it appears that despite the fact that the Sharp display can support all the tested non-null disparity levels, the optimum depth range where people find it easiest to operate is restricted between 2 and 4 (possibly 5) pixels of image disparity. This is an interesting and intriguing finding. However, further, more accurate investigation is needed in order to prove its validity as this effect of image disparity on task difficulty could be due to the noise present in the Sharp response time data (see Appendix C for details).

Performance for Row-Interleaved and Full-Resolution Displays

Response time results for the Row-Interleaved and Full-Resolution displays were largely according to the geometric predictions, but revealed some subtle, unexpected differences among the non-null disparity conditions.

In particular, with the Colorlink and the Iris3D displays participants were reliably faster at 3-pixel image disparity than at 1-pixel image disparity, which is evidence that they may have found it harder to perform the task at low image disparity levels (i.e. 1 image pixel). On the other hand, high disparity levels (i.e. 4 and 6 image pixels) did not seem to have any detrimental effect on performance. This pattern suggests that with these displays task difficulty decreased slightly but steadily from 1- to 3-pixel image disparity, but after that it saturated.

In regard to the Kodak display, participants' response time performance appeared to be best at 2-pixel image disparity, and it became increasingly worse as image disparity increased. Furthermore, the data also seem to suggest that task difficulty was higher when the image disparity was only 1 pixel as compared to the 2-pixel condition, but the high level of noise in the Kodak data (especially for the 1-pixel disparity condition) does not allow a clear, reliable conclusion to be drawn about this point. A similar pattern was shown for the Shutter Glasses display, although for this display the increasing detrimental effect of high image disparity levels (i.e. 4 and 6 pixels) is less prominent. Note that response time data for the Shutter Glasses display were also affected by noise (see Appendix C for details).

6.2.3 Subjective Results

While subjective results are probably of limited help in understanding the variations in task performance, they also yielded a number of interesting points, with some displays dividing opinion among the participants and with substantial differences between the pilot and the main experiment.

A first interesting point is that the Shutter Glasses display received a surprisingly low average ranking score. In particular, some participants reported a low ranking for the Shutter Glasses even though their performance with this display was generally very high and in agreement with the predictions; this suggests that while users may not like the Shutter Glasses as a display device it can still provide good and reliable task performance as demonstrated by the score results. On the other hand, in both experiments the Kodak was voted as the best display by the highest number of people (also best on average in the main experiment and third best on average in the pilot experiment), while the empirical results suggest that this display did not always lead to optimal performance. The Kodak display is immediately followed by the Iris3D display (second best ranking score in both experiments) and the Sharp display, which was ranked as the best display in the main experiment (third worst in the pilot experiment) despite the colour fringing artifacts. Finally, it is probable that the ranking of the SeeReal display (worst display in both experiments) was unduly affected by the aliasing problems discussed earlier, as a number of participants noted they did not see any depth on this display during most of the trials.

Concerning the level of comfort in seeing the 3D image, participants had again divided opinions to some extent, but on average they reported the Kodak to be the most comfortable display while the SeeReal, together with the Sharp in the main experiment, was reported to be the least comfortable display. Overall, people seemed to find it more comfortable to see the 3D effect on full-resolution, non-interleaved displays than on their half-resolution, interleaved counterparts. This could be due to the fact that the mechanism used to implement the interleaving can bring side effects such as dimmer images and other artifacts that people noted during the experiments.

Finally, during both experiments a number of participants reported experiencing some form of physical discomfort at some point during the trials even though this was generally mild. Interestingly, this is the aspect of the subjective results where opinion was most divided and therefore it is not possible to identify a clear pattern in the data. The only conclusive point to be made is that the number of participants that experienced physical discomfort was notably lower in the main experiment than in the pilot experiment (33% against 48%). This could be due to two potential reasons. The first is directly related to the amount of disparity applied to the stimuli and the possible discomfort caused by the decoupling of the accommodation and vergence processes [Hoffman et al., 2008] when viewing a 3D image on a 3D display [Ware, 1995], as explained in section 2.3.4. During the pilot experiment the stimuli were treated with higher levels of image disparity which could have caused the participants to perceive a more pronounced sensation of discomfort. The second possibility is that during the main experiment factors that could cause physical discomfort (e.g. wrong viewing position) were better controlled for.

6.3 General Considerations

Overall, the analysis of the data collected during the 3D display comparative study suggests that the size of the sample and the number of repetitions used for the experiments were adequate in order to ensure reliable empirical results. Also, the fact that the score performance obtained in both experiments was generally high suggests that the task was appropriate and clearly defined. Furthermore, the results obtained in the main experiment are in agreement with the results collected during the pilot experiment; comparably, the score and the response time results are consistent with each other across experiments and support the same findings. These points all serve as evidence that the methodology used for the study is robust, reliable and effective in capturing important details regarding human depth perception on the different 3D display classes. Finally, the lower level of variance that characterises the main experiment data as compared to the pilot experiment is evidence that the methodology has been successfully improved.

The results that were of primary interest for the purpose of this thesis were the scores. The ANOVA revealed that the effects for this measure were all highly reliable. With respect to disparity, a distinct *threshold effect* has been identified in that the experiments clearly showed what is the minimum level of image disparity required on each display so depth can be perceived in the displayed image. The results also showed that the type of display used to perform the task had a direct effect on participants' score and that the amount of disparity used to treat the stimuli had a different effect depending on the type of display. In this regard, it emerged that Full-Resolution and Row-Interleaved displays can generally lead to good task performance and can reproduce image disparity as perceivable depth in a reliable way. However, 3D displays are not interchangeable; that is, for a given amount of image disparity different displays do not always produce the same depth percept, even if they belong to the same display class. This was particularly evident for Column-Interleaved and Colour-Column-Interleaved displays, which demonstrated a range of *disparity aliasing effects*. It is worth noting though that these aliasing effects appear to be dependent on the software drivers used by the graphics card and the display electronic interfaces, and not only on the displays' optical design.

Compared to score, response time produced effects that were statistically less reliable. In particular, the results for this measure suggest that the difficulty of depth-based tasks mainly depends on the amount of image disparity present in the input image and, to a lesser extent, on the disparity-display combined effect, while the type of display alone seems to be uninfluential. The numerical variations registered for the different level of image disparity suggest that for most 3D displays the optimum range of comfortably viewable depth is the one induced by an image disparity of 2 to 3 image pixels. At higher disparity levels, some displays did not show any change in performance while others led to an increase in response time. This behaviour may indicate that on these 3D displays, increasing the amount of image disparity beyond the range of 2-3 pixels can make depth-based tasks more difficult. A plausible explanation is that fusion and vergence-accommodation decoupling become gradually more challenging as the disparity between the left and right views increases [Ware, 1995]. Another factor that could have contributed towards this result is crosstalk (see section 2.3.4). Furthermore, the response time results also suggest that on some displays (e.g. Sharp) perceiving the depth induced by 1-pixel image disparity can at times be challenging, despite the fact that the display can reliably reproduce this disparity as perceivable depth. These results imply that even within the comfortable depth range of a 3D display, the human depth perception system seems to be sensitive to relatively small differences in image disparity, and in turn have implications for the design of a 3D cursor for stereoscopic 3D environments.

Moving on to the subjective results, participants' opinion on the different aspects of the 3D displays was at times very different, therefore identifying a common trend in the data was difficult. An interesting fact is that the subjective ranking of displays did not always reflect the empirical results, with displays receiving a relatively low ranking score when the associated task performance was high (Shutter Glasses) or, vice versa, being ranked highly when instead task performance was not as high as expected (Kodak) or artifacts were present in the output 3D image displayed to participants (Sharp). Performing a visual task on a 3D display is an experience inevitably characterised by a high perceptual component and as the results of this study suggest, the subjective opinion of the observers can be very different. However, subjective results can provide useful information about the technology that empirical data can not always reveal. This is especially true for applications where the general 3D experience of the user is equally important to task performance, like for example in 3D video games and other recreational applications.

A final interesting point that emerged from the study is that in both experiments some participants had lower task performance on some displays than predicted based on their stereo and vision acuity pre-trial screening test results. One of the causes could be the difficulty in fusing the two views that form the stereo image shown on the screen, which in turn could be due to a mild form of amblyopia or strabismus, or other focus-vergence issues. The identification of the exact reasons for this finding would require further ophthalmic investigation and is beyond the scope of this thesis. However, what these results do suggest from a practical point of view is that 3D display operators who are undertaking critical depth-judgement based tasks should be carefully screened for stereo vision.

6.4 Final Discussion and Recommendations

From the discussion above it has emerged that there are a number of technical factors that can affect the quality of a stereoscopic image delivered to the user via a 3D display. This is in line with what was anticipated in Chapter 3 based on the reviewed literature (section 3.3), however the study presented in this thesis has helped to understand how such factors affect the quality of the displayed 3D image and what are the direct implication in terms of designing a 3D cursor to support depth-based tasks in stereoscopic 3D environments.

The first and probably most obvious factor is the view resolution of the 3D display, and in particular its horizontal view resolution. In this regard the experiments showed that Column-Interleaved displays do not have the capability to reliably show 1-pixel image disparity as perceivable depth because of their reduced horizontal view resolution. However, the experiments also showed that Column-Interleaved displays may also fail to reproduce the depth associated with an image disparity of two pixels and that this is due to a combination of factors such as the specific implementation of the interleaving algorithm in the driver of the graphics card and the way the experimental stimuli had been produced; software drivers and input image generation both concurred in producing what were defined as *disparity aliasing artifacts*. Interestingly, one of the Column-Interleaved displays could occasionally show the depth induced by 1-pixel image disparity, but this unexpected behaviour was probably due again to disparity aliasing caused by sampling artifacts of the analogue input signal in the display's electronic interface.

On the other hand, the study suggests that Colour-Column-Interleaved displays do have the capability to reproduce all non-null image disparity levels, including 1-pixel image disparity, as perceivable depth. This result was rather surprising and it implies that this class of displays can overcome the limitations of having reduced horizontal image resolution by interleaving the input stereo image at a finer pitch, i.e. at sub-column colour component level. Besides, the experiments showed that Colour-Column-Interleaved displays are not affected by disparity aliasing artifacts. Nevertheless, due to their very interleaving mechanism they present *colour fringing artifacts* that for objects of specific colours can lead to the same issues as disparity aliasing artifacts.

Based on these main findings, there are a number of guidelines that should be taken into consideration when designing a universal stereoscopic 3D cursor and that complements what has already been presented in Chapter 3. These can be summarised as follows.

1. As anticipated in Chapter 4 (pg. 100), the comparative study presented in this thesis showed that the technical characteristics of a 3D display can affect the quality of the final stereoscopic 3D image and the visual properties of the objects contained in it. In particular, the experiments showed that the image interleaving pattern adopted by Column-Interleaved displays in combination with the way the input image stereo pair used during the trials had been produced resulted, in some cases, in narrower white squares in the output 3D image with no disparity information associated to them (see section 6.2.1. This result suggests that the display technical factors not only can affect the perceived depth of objects in the stereoscopic 3D scene shown to the viewer but also their size. Based on this discussion, it is recommended that a universal 3D cursor used to carry out depth-based tasks using a conventional 3D display has a minimum horizontal size of three image pixels; this requirement

should be sufficient to prevent the cursor from disappearing from the scene because of disparity aliasing artifacts and the limited horizontal resolution that characterizes Column-Interleaved displays.

2. In line with the previous point, it is also recommended that the cursor has a minimum vertical size of three image pixels, in order to avoid analogous potential problem on Row-Interleaved displays.

3. The results of the study reinforce and confirm what already emerged from the review of previous work on aliasing artifacts (see sections 3.3.2 and 2.4.2), in that an anti-aliasing algorithm should be adopted in order to avoid artifacts and inaccuracies in the reproduction of the cursor's shape and depth, and of the 3D scene in general.

4. The results also suggest that a universal 3D cursor should not be presented in pure magenta or green colour as on Colour-Column-Interleaved displays its depth will be masked out and its shape distorted due to colour fringing artifacts, as explained in section 6.2.1. Instead, it is recommended that the cursor is rendered using a colour that embeds all three RGB **primary colour** components. This requirement should suffice to ensure the retention of the cursor disparity, which is an essential prerequisite for tasks based on fine depth-judgements, even though on Colour-Column-Interleaved displays it will inevitably introduce a degree of *binocular colour rivalry* [Dutour, 1760] [O'Shea, 1999a] [Dutour, 1763] [O'Shea, 1999b]. During the trials, colour rivalry was present in the stimuli in the form of colour fringing at the vertical edges of the white boxes; despite the fact that such fringing artifacts were noticed by a number of participants, the results showed that they did not hinder task performance. Therefore it is believed that the design suggestion to avoid the use of pure primary colours is a valid and advantageous tradeoff for the type of tasks in question.

5. Finally, a general recommendation is that for tasks based on critical depth-judgements a Full-Resolution, non-interleaved display is to be preferred to any interleaved display, including Row-Interleaved displays. This is because, despite being characterised by full horizontal resolution and not being affected by disparity aliasing or colour fringing artifacts, Row-Interleaved displays present in fact an intrinsic vertical offset between the left and right view. For the conditions adopted in this study, the vertical misalignment presented in the images was not enough to affect the participants' stereovision and task performance, as explained in section D.2 of appendix D. However, under different viewing conditions (e.g. smaller viewing distance, bigger pixel pitch, different nature of the 3D scene) these vertical misalignments could hinder the stereo matching process of the viewer (see section 6.2.1) making the depth discrimination task more challenging, and therefore should be avoided. Based on this consideration, a final recommendation is that a universal 3D cursor to support depth-based tasks should present no interocular vertical misalignments between the left and right view.

Note that the suggested universal stereoscopic 3D cursor prototype could be optimized on a per-display basis by adapting the above design guidelines to the technical characteristics and limitations of the target 3D display. In particular, a stereoscopic 3D cursor for a Column-Interleaved or Row-Interleaved display should observe the minimum size recommendation of three horizontal and three vertical image pixels respectively; no colour limitation is required for these two types of 3D display. Conversely, pure magenta and green cursor shapes should be avoided for Colour-Column-Interleaved displays as specified above, with no limitation in minimum cursor size. Finally, Full-Resolution, non-interleaved displays do not require any of the minimum cursor size or colour design measures discussed in this section. However, the 3D cursor design guidelines based on the vision science literature review presented in Chapter 3 are still valid and should be taken into considerations independently from the type of target 3D display.

6.5 Conclusions

In this chapter, a detailed discussion of the empirical results obtained from the 3D display comparative study was presented. The interpretation of the data has allowed for a detailed understanding of the tradeoffs between different displays for tasks based on critical depth-judgements and of the implications that these have for the design of a 3D cursor for stereoscopic 3D environments.

Differences between the four classes of 3D display tested in the study have been clearly identified and the technical factors that directly affect the accuracy of the depth reproduction in the final 3D image thoroughly discussed. From the discussion it has emerged that Colour-Interleaved displays may present disparity aliasing effects, while Colour-Column Interleaved displays are affected by colour-fringing artifacts. Guidelines to overcome the side effects of such artifacts have been proposed in terms of the design of a universal stereoscopic 3D cursor.

Chapter 7

Conclusions and Future Work

7.1 Introduction

The aim of this thesis has been to investigate desktop stereoscopic 3D technologies to support visual tasks that require accurate depth judgements. The focus of the research was on task performance and human depth perception and the main interest was on stereoscopic 3D technologies that do not require the user to change their desktop working habits. For this purpose the focal point of this work was on desktop stereoscopic 3D environments, i.e. working environments where the desktop 2D display is substituted by a stereoscopic 3D counterpart and the classic 2D mouse cursor replaced by a stereoscopic 3D version. A key contribution towards addressing the research problem was to clarify how a stereoscopic 3D cursor suitable for supporting depth-based tasks in such environments should be designed.

These goals were pursued by undertaking a thorough analysis of the visual and technical factors that can affect depth perception in humans and by performing a quantitative, empirical assessment study of different classes of desktop 3D displays. The findings of this investigation provided the basis for the formulation of a number of guidelines that should be followed to assist operators with depth-based tasks in stereoscopic 3D environments and ultimately led to a novel design for a universal stereoscopic 3D cursor.

The research contributions and the novelty of this thesis are summarised in section 7.2. Section 7.3 outlines a number of directions for future work. Finally, section 7.4 concludes this chapter and the thesis.

7.2 Research Contributions

This thesis has presented a number of research contributions in the area of stereoscopic 3D technologies for visual tasks based on critical depth judgements, as summarised in the following paragraphs. The research achievements are discussed against the objectives outlined in section 1.3.

1. **Investigate influential factors to the perception of depth for depth-based tasks.**

   Chapter 3 presented a thorough and contextualised review of the visual and technical factors that can influence depth perception in natural and, in particular, synthetic 3D scenes presented via stereoscopic 3D displays. The discussion provides a novel understanding of the role that different visual cues play in the specific context of visual tasks in stereoscopic 3D environments that require accurate depth judgements; it is supported by the vast visual perception literature and it forms the basis for a number of stereoscopic 3D cursor design guidelines. In this regard, synthetic cuing as a means of supporting depth-based task performance was also discussed. Concerning the technical factors, the review highlighted a knowledge gap that the remaining work presented in this thesis attempted to fill.

2. **Develop a quantitative assessment framework for 3D display performance.**

   A new methodology was developed that uses human depth perception as a means of assessing the performance of conventional 3D displays, as discussed in Chapter

4. This includes a novel classification model for 3D display technologies based on their technical characteristics and optical design, and a detailed experimental design aimed at conducting quantitative empirical research on such displays. A number of technical factors that may affect the amount of disparity embedded in the final 3D image presented on the display, and therefore perceivable as depth by the observer, were identified and used to formulate performance hypotheses for four different classes of two-view 3D displays. The terminology used in this methodology, which represents the key contribution of the comparative study presented in this thesis, extends the one proposed by Holliman [2006].

3. **Perform an empirical comparative study on a set of representative 3D displays.**

   The methodology presented in Chapter 4 was successfully employed to compare the performance of seven two-view 3D displays across a range of image disparity levels [Froner et al., 2008] [Holliman et al., 2007]. This represents the first empirical study to investigate human depth perception on a set of representative electronic 3D displays. The methodology proved to be robust and sensitive enough to identify small differences between 3D display technologies and small alignment issues in the optics of the displays. The data collected during the experiments allowed for a detailed understanding of the tradeoffs between different display technologies for depth perception tasks, which could not be predicted from the manufacturers' specifications. Differences among four distinct display classes have been clearly identified and the implications in terms of depth perception and design of a stereoscopic 3D cursor have been thoroughly discussed in Chapters 5 and 6. Interestingly, the study has demonstrated that 3D displays are generally not interchangeable, even if they belong to the same class, and that interleaved displays may suffer from a number of *disparity aliasing* artifacts. Cursor design guidelines to overcome the side effects of such artifacts and other display-dependent technical factors have been proposed.

4. **Formulate a number of guidelines for the design of a stereoscopic 3D cursor.**

   The findings that emerged from the empirical study (Chapter 6) and the analysis of the effect of critical factors on depth perception (Chapter 3) have been used as a basis to propose a novel universal 3D cursor prototype suitable for supporting depth-based tasks in stereoscopic 3D environments. The contribution includes a number of both qualitative and quantitative guidelines that should be observed when designing such a cursor and that can be summarised as follows.

   Generally, the cursor should be represented by a simple 3D shape with a limited amount of detail (e.g. a box), in order to facilitate the fusion of the stereoscopic image and ensure that the visual artifacts introduced by potential aliasing errors are not exacerbated. Conversely, completely smooth 3D shapes (e.g. a sphere) and extremely detailed 3D shapes (e.g. a 3D star) should be avoided. The level of luminance contrast between the cursor shape and the scene background should be boosted in order to maximise stereoacuity, especially in the presence of high frequency content (5.0 c/deg and above), however it is suggested that this is no higher than 21 dB of the luminance contrast detection threshold of the cursor shape. Furthermore, no interocular difference in cursor vertical disparity, shape, luminance or colour should be shown as they would also have a detrimental effect on the user's depth perception and task performance. For similar reasons, the use of image blur as a cue to depth should be avoided and wireframe shapes should not be employed, especially in the presence of working scenes where depth cues are weak or absent. On the other hand, transparency could be applied to the faces of the 3D cursor shape and serve as an effective cue for guiding the user during target interaction.

   In order to avoid detrimental disparity aliasing effects introduced by some 3D display technologies, it is recommended that the cursor has a minimum horizontal and vertical size of three pixels and be rendered using a colour that embeds all three RGB primary colour components, avoiding pure primary colours. Where possible, the cursor's colour should also be complementary to the colour of the other objects in the 3D scene in order to provide a significant level of colour-contrast. This aspect guarantees that segregation is readily achieved and the cursor is easily identified, and is particularly important in working environments characterised by low spatial frequencies (2.0 c/deg and below).

It is furthermore recommended that an anti-aliasing algorithm be adopted in order to avoid artifacts and inaccuracies in the reproduction of the cursor's shape and depth. Lastly, aural and visual feedback should be used to provide the user with target spatial location information. This last requirement would also counterbalance the negative effects of potential chromostereoscopic phenomena on depth perception, which could hinder task performance.

The rationale behind these recommendations is fully explained in Chapters 3 and 6.

Note that the universal stereoscopic 3D cursor prototype suggested above could be optimized on a per-display basis by adapting the design guidelines to the technical characteristics and limitations of the target 3D display, as discussed in section 6.4.

7.3 Future Research Directions

The research presented in this thesis has provided meaningful insight into several aspects of stereoscopic 3D technologies in the context of depth-based visual tasks. Nevertheless, due to the complexity of the topic, this work could be expanded and progressed in a number of directions.

1. **Different Image Disparity, Same Perceived Depth?**

   The comparative study described in Chapters 4 and 5 was designed and successfully used to quantify human depth perception threshold levels on different 3D display classes. It is unsuitable, however, to investigate the differences in perceived depth induced by different levels of image disparity and answer questions such as: "On a specific 3D display, is the amount of depth induced by an image disparity of two and three pixels respectively perceived as the same by the observer"? To investigate this point a further study is required that employs an adapted version of the task used in this work in order to discern between different levels of perceived depth. This would provide important insight into the depth resolution of different 3D display classes, a crucial aspect if the display is to be used for applications that require accurate depth judgements.

2. **Aliasing Artifacts and Antialiasing Techniques**

   As discussed in Chapter 6, the empirical results showed that interleaved displays suffer from a number of disparity aliasing artifacts that can affect the disparity information, and therefore the depth, contained in the 3D image presented to the observer. The negative effect of aliasing was also discussed in theoretical terms in section 3.3.2. It would be interesting to repeat the same experiments using a scene depth-mapping method like that adopted by Froner and Holliman [2005] to generate the visual stimuli, control the image disparity in millimetres rather than pixels, and render the final 3D image using antialiasing. This should lead to an understanding of whether standard antialiasing techniques commonly available on commercial graphics cards are sufficient to counterbalance the negative effect of the disparity aliasing artifacts observed on Column-Interleaved displays such as the SeeReal and the DTI display families.

3. **Aliasing Artifacts and Multi-View Displays**

   The focus of this work was on two-view 3D displays and how their characteristics can affect the amount of depth embedded in the final 3D image presented on the screen. The outcomes of this research are valuable but can not be directly extended to another very common type of 3D displays, i.e. multi-view displays. The methodology presented in this thesis could be expanded to study human depth perception and aliasing artifacts on multi-view displays. Given the more complex nature of these displays, it is expected that the side effects of aliasing will be even more prominent, especially on displays using mixed designs as described in section 2.3.3.

4. **Response Time Accuracy**

   It would be worth repeating the 3D display comparative study using more accurate time measurement techniques. Despite being noisy, the collected response time data showed a number of interesting points that were useful in understanding the performance and the ease-of-use of different two-view 3D display classes, which the score data alone could not have highlighted. As discussed in Chapter 6, the response time results of this study seem to suggest that for most 3D displays the optimum range of comfortably viewable depth is the one induced by an image disparity of two to three image pixels and that for lower or higher image disparity levels the task difficulty on some 3D displays increases. It is believed that more accurate response time results will help to clarify these points and could provide additional valuable information.

5. **Empirical Evaluation of Cursor Prototype**

   The design of the cursor prototype proposed in this thesis is based on scientific evidence and on the results of the 3D display comparative study, however it has not been evaluated empirically. On the basis of the theoretical guidelines presented above it would be possible to design a number of experimental trials where the characteristics of the suggested cursor are tested against the performance of potential users in depth-based tasks. The literature review and the recommendations presented in Chapters 3 and 6 provide enough ground to formulate predictions and define the experimental design to adopt. Similar to the 3D display comparative study, the effects of the cursor design characteristics on task performance could be analysed using statistical models such as ANOVA and Student's t-Test. The stereoscopic 3D cursor could be implemented using an OpenGL-based 3D graphics toolkit, such as Open Inventor® or Java 3D. During the implementation care should be taken to assure that the cursor is always visible within the scene and that its depth is controlled accurately.

7.4 Conclusions

This research has focused on the investigation and design of 3D technologies to assist operators in performing visual tasks that require accurate depth judgements in desktop stereoscopic 3D environments. A contextualised review has been presented that analyses and discusses the effect of visual cues and display-dependent factors on human depth perception in such environments. An empirical study has been performed in order to assess different classes of desktop 3D displays and clarify how the display technical characteristics and the optical design affect the depth information contained in the final 3D image presented to the observer. The outcomes of this study and the review served as a basis to draw a set of guidelines for designing a universal stereoscopic 3D cursor to support depth-based tasks in desktop stereoscopic 3D environments. It is believed that the work presented in this thesis contributes valuable knowledge to the area of stereoscopic 3D technologies and their adoption for the visualisation, analysis and comprehension of three-dimensional data in a number of application domains.

Appendix A

Glossary

3D Display Electronic display device capable of conveying 3D images to the viewer.

Aliasing In signal processing aliasing refers to the distortions introduced by inadequate signal sampling, which can cause different signals to become indistinguishable when sampled or it can result in the signal reconstructed from the samples to be different from the original continuous signal.

Amblyopia Disorder of the visual system characterised by dimness and blurry vision in an eye that is otherwise healthy and that is caused by faulty transmission of signals between the eye and the brain. According to a recent study this condition affects 1 to 5% of the population. [Weber and Wood, 2005].

Anaglyph Type of **stereoscopic image** that provides a 3D effect when seen with glasses that have two chromatically opposite coloured lenses (typically red and cyan). The stereoscopic image consists of two superimposed colour images of the same subject but offset, printed or rendered together as shown in the example below (Figure A.1).

Figure A.1: Example of anaglyph image.
[From http://en.wikipedia.org/wiki/Anaglyph_image, last accessed: 27.12.2010]

ANOVA ANalysis Of VAriance or ANOVA is a statistical analysis technique used to analyse situations in which there are more than two experimental conditions. It is generally used in experiments with several independent variables (factorial ANOVA) or with a single independent variable with more than two levels (one-way ANOVA).

Atmospheric Perspective (or Aerial Perspective) Effect of the particles contained in the atmosphere on the appearance of an object, especially when viewed from a distance. As the distance between the observer and the object increases, colours become less saturated, details less clear and contrast decreases.

Attention The cognitive process of concentrating on one aspect of the surrounding environment while ignoring the rest. In the context of vision, attention refers to the ability of directing the visual processing resources towards a specific area of interest in the visual environment and is a fundamental process in many high level tasks (e.g. selection).

Augmented Reality Environment where the view of the real world has been enriched or *augmented* via the superimposition of virtual sensory inputs such as computer generated imagery or sound.

Beam Cursor Interaction techniques used to perform target selection or target acquisition in pen-based interaction environments [Yin and Ren, 2006].

Binocular Involving or relating to both eyes at the same time, e.g. binocular vision refers to vision that involves the use of both eyes simultaneously.

Binocular Disparity See **retinal disparity**.

Binocular Matching See **stereo matching**.

Binocular Rivalry Visual perception phenomenon that occurs when the two eyes are stimulated on the same retinal area by substantially different images and that results in either an alternation of perception, complete or partial, between the two images or, at times, in a constant dominance of one eye. This means that at a given time only one stimulus can be seen by the observer (dominant stimulus), while the other is suppressed and therefore can not be seen (suppressed stimulus). There are different types of binocular rivalry, depending on wether the two images differ only in their contours (binocular contour rivalry), colours (binocular colour rivalry), or other visual features. Figure A.2 is an example of a binocular rivalry stimulus.

Figure A.2: Example of binocular rivalry stimulus.
[From http://www.psy.vanderbilt.edu, last accessed: 27.12.2010]

Brightness The subjective perception of luminance. It is not a physical measure in the scientific meaning of the term but it is commonly used in visual perception to describe the perceived characteristics induced by the luminance of an object. The term brightness is also used to express the perceived intensity of colour.

Colour-Blindness Visual condition, usually genetic, that results in the inability to distinguish one or more chromatic colours that people with "normal vision" can.

Complementary Colour Two colours are said to be complementary if they are characterised by opposite **hue** in a specific colour model. The complementary of one of the three **primary colours** of the model is obtained by combining the other two primaries, e.g. in paints mixing yellow and red gives orange which is the complementary to blue.

Contrast Gain Control A tuning mechanism in the human visual system that allows the viewer to perceive contrast as approximately constant in visual environments characterised by changing contrast levels (i.e. from foggy low contrast conditions to clear high contrast conditions). It is an important process underlying the capability of the visual system to adapt to varying or extreme light conditions.

Contrast Sensitivity Function (CSF) Function that measures the sensitivity of the human spatial system. In experimental psychology studies it is often characterised by measuring the minimum amount of either luminance or colour contrast necessary to detect a stimulus (e.g. a sinusoidal grating) against the background at different spatial frequency levels.

Crosstalk In signal processing and electronics the term crosstalk refers to the interference of a signal (or circuit) with another signal (or circuit). In the context of stereovision, crosstalk indicates the interference of the information contained in the left view channel with the one contained in the right view channel or viceversa.

Cue A Stimulus or signal that evokes a certain perceptual response. Cues may interest different sensory systems, so for example there are verbal cues, physical cues or visual cues. In this thesis, if not otherwise specified the term cue refers to a visual cue, i.e. a signal that our visual system uses during the construction of a visual **percept**. Vergence and accommodation are examples of visual cues, and in particular visual cues to distance (or depth).

Decibel (dB) Logarithmic unit of measurement used to express the magnitude of a physical quantity relative to a reference quantity.

Defocus Blur Optical aberration that occurs when the eyes do not accommodate on the plane of focus. In real life focus and accommodation are normally coupled so the objects that are within the eye's plane of focus are sharp while objects that are out of this plane are characterised by defocus blur.

Depth Constancy Perceptual constancy that is related to the perception of depth.

Depth Cue Perceptual unit that can be used by the human sensory systems to infer the third dimension and allow us to perceive depth.

Diplopia (or Double Vision) Visual condition in which a single object is perceived as two images instead of a single, fused image.

Emmert's Law Optic law that states that the size of an afterimage projected onto a surface becomes smaller as the surface is brought nearer [Emmert, 1881]. This implies that objects that generate retinal images of the same size but appear to be located at different distances to the observer will look different in physical size.

Field Of View (FOV) Angular extent of what is seen, either through the naked eyes or through an optical instrument, at a certain moment in time. It is usually measured in degrees (deg or $^\circ$).

Fitts's Law Human-Computer interaction model that predicts that the time necessary to rapidly reach a target directly depends on the size and the distance of the target itself.

It is used to model pointing actions using a pointing device (e.g. a mouse) on computer displays and it was first proposed in 1954 by the American phycologist Paul Fitts [1954].

Flicker Fusion Threshold (or Flicker Fusion Ratio/Frequency) The flicker frequency at which a flickering visual stimulus is perceived by the observer as completely steady. In humans, this is approximately 50-60 Hz.

Fovea The fovea centralis, or simply fovea, is a small depression in the central part of the retina that is responsible for sharp vision and detailed colour discrimination. It constitutes less than 1% of the retinal surface (i.e. roughly 1.5 mm in diameter) but thanks to the high concentration of cone cells it represents the area of the retina with the best visual acuity and it can take up to 50% of the processing power of the human brain's **visual cortex**.

Fractal Image A fractal image is an image that shows self-similarity, i.e. the property of exhibiting features that appear the same or very similar at all magnification levels.

Full-Parallax 3D Display Type of 3D display that can reproduce variations in the images seen by the viewer with both horizontal and vertical head movements.

Fundus Image Image of the retina taken through the pupil using a special camera.

Fusion Limits Upper and lower levels of image disparity within which the left and right view can be fused and perceived in 3D.

Hue Perceptual attribute of colour that enables us to classify it and refer to it with a name, e.g. blue rather than green or orange.

Image Blur See **defocus blur**.

Independent Design When using an independent (or between-subjects) design, the independent variable is manipulated by using different participants. This means that different groups of subjects participate in different experimental conditions of the empirical study.

InterPupillary Distance (IPD) or Eye Separation The distance between the centre of the pupils of the eyes. IDP is a critical factor in the design and creation of stereoscopic viewing systems (including stereoscopic 3D displays) and stereoscopic content, as it affects the amount of depth perceived by the observer.

Isoluminance (or Equiluminance) Property of possessing the same luminance. In perceptual psychology, experiments are at times carried out using isoluminant stimuli, i.e. visual stimuli where no difference in luminance is present between background and foreground and that, for this reason, appear to be the same to our luminance processing pathway.

Latin Square A Latin square is a table formed by n rows and n columns and filled in with n different symbols so that each symbol occurs only once in any row and in any column.

Liquid Crystal Display (LCD) Type of digital display that uses a layer of liquid crystals, and their optical properties when stimulated by an electric field, in order to produce and electronically display information on the screen.

Likert Scale Subjective scoring technique typically used to measure preferences and likes. Widely employed in questionnaires and survey research, Likert scales consist of a rating scale that usually has five potential choices, even though these can go up to ten or more. By selecting one of the possible choices, questionnaire respondents express their level of agreement to a clear statement (e.g. Questionnaire item: "The display is easy to use." Answers: 1 = "Strongly agree", 2 = "Agree" 3 = "Neither agree nor disagree", 4 = "Disagree", 5 = "Strongly disagree"). In this way it is possible to assign a quantitative value to data that are intrinsically qualitative, so as to be able to subject them to statistical analysis. The Likert scale is named after its inventor, the psychologist Rensis Likert [1932].

Luminance (L) Physical measure used to quantify the level of brightness of an object. It describes the amount of light that is reflected by a specific area in a particular direction and according to the **SI** it is measured in units of luminous intensity, candelas, per squared metre (cd/m^2).

Luminance Contrast The relationship between the luminance of the object of interest and that of its background.

Monocular Involving or relating to one eye, e.g. monocular vision is the vision that involves the use of one eye only at a time.

Multistability The property of allowing two or more distinct perceptual interpretations. Examples of multistable stimuli are the Necker cube [Necker, 1832] and the Rubin vase [Rubin, 1921], shown in Figure A.3.

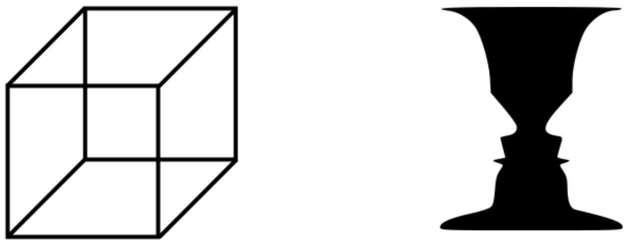


Figure A.3: Example of multistable stimuli: Necker cube and Rubin vase.
[From http://en.wikipedia.org/wiki/Multistability, last accessed: 15.03.2009]

Opacity The physical property of a material of not allowing light, or other sorts of magnetic radiation, to pass through. Opaque objects transmit a negligible amount of light and instead reflect, scatter or absorb it. Mirrors are an example of opaque objects.

Orthoscopic Test Test designed to ensure or produce correct vision, free from distortions. The term orthoscopic is formed by the prefix *ortho*, which means straight (or in this case correct), and *scopic* from the Greek word *skopien*, meaning to see or to view.

Percept A representation of what the sensory systems receive in input; a mental impression of what is perceived. In psychology and cognitive sciences a percept represents the basic unit in the formation of a concept.

Perceptual Constancy Ability to perceive objects as unchanged despite wide variations in the viewing conditions and in the retinal images. For example, when we walk towards another person we perceive their height as staying the same despite the fact that they appear larger and larger as we approach.

Pixel Abbreviation for **Pi**cture **el**ement, a pixel represents the smallest and most elementary unit of a raster image or a picture on a digital screen.

Pixel Pitch Distance between the centre of a **pixel** and the centre of the pixel adjacent to it. The smaller the pixel pitch, the higher the resolution, the sharper the image.

Planar 3D Display Type of 3D display that presents the output 3D image on a static, flat surface.

Primary Colour In a specific colour model, primary colours are basic pure colours that can be mixed to obtain all other colours. For example, in the RGB colour model the primary colours are represented by red, green and blue, while in paints the primary colors usually are red, blue and yellow (or magenta, cyan and yellow).

Random Dot Stereogram (RDS) Pair of images consisting of a collection of random dots that when fused produce a sensation of depth with objects appearing either in front or behind the actual image plane. This technique was first invented by Béla Julesz during the 1960s as part of his research work on human stereo vision [Julesz, 1971]. With RDS Julesz showed that binocular vision and stereopsis are sufficient for the perception of objects even when other depth cues are not present. Figure A.4 shows an example of a random dot autostereogram (i.e. **stereogram** where the depth information is combined into a single image) representing a shark.

Repeated-Measures Design In a repeated-measures (or within-subjects) design, the independent variable is manipulated by using the same participants. This means that the same group of subjects takes part in all experimental conditions of the empirical study.

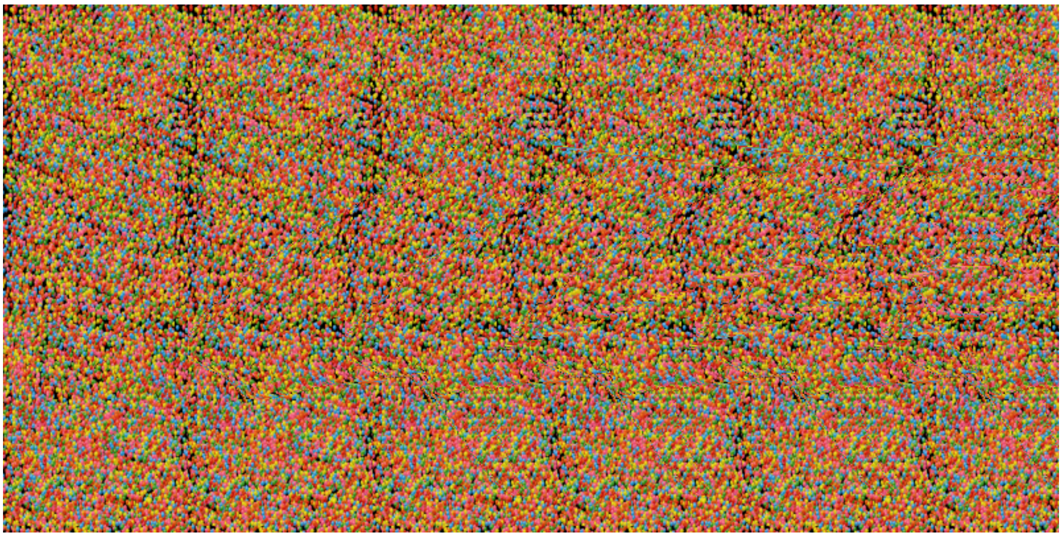


Figure A.4: Example of RDS.
[From http://en.wikipedia.org/wiki/Autostereogram, last accessed: 28.12.2010]

Retinal Disparity The lateral displacement between the left and right retinal image of an object. Retinal disparity is the stimulus for **stereopsis**.

Saturation Perceptual attribute that describes the degree to which a colour characterised by **hue** (i.e. a chromatic colour) differs from white.

SI The International System of Units.

Stereoacuity Ability to detect differences in distance and relative depth using **binocular disparities**. It can be defined as the smallest disparity in the stereoscopic images presented to the observer's eyes that can be reliably detected as depth and it is usually measured in seconds of arc. Patients can be screened for stereoacuity using a **stereo test**.

Stereoscope Optical binocular device used to display a stereoscopic image. It is generally used to view side-by-side stereo pairs (i.e. a type of **stereoscopic image**) and it contains a picture support and a pair of magnifying lenses arranged in a way that the left eye sees only the left image of the stereo pair and the right eye only the right image.

Stereoscopic Image (or 3D Image or Stereo Pair) An image that contains three-dimensional information and that provides the perception of depth to the observer. The illusion of depth is created from two flat 2D images by exploiting the innate ability of the human brain to detect depth via **stereopsis**.

Stereogram See **stereoscopic image**.

Stereo Matching Process through which a feature seen by one eye is "matched" with a similar feature in the other eye. When the matching occurs correctly, **binocular disparities** are computed, the two views are fused and a solid 3D image of the object containing that visual feature is perceived.

Stereopsis The ability to extract 3D information from two different 2D views of the same scene using **retinal disparity** information. The terms literally means "solid viewing", from the Greek *stereo*, meaning solid or three-dimensional, and *opsis*, meaning view or sight.

Stereo Test Visual test used in optometry and ophthalmology to assess the binocular stereo vision of a person. There is a wide range of techniques that can be used to screen stereo vision, from simple equipment to complex laboratory apparatus [von Noorden and Campos, 2001].

Student's *t*-Test Statistical method used to determine whether differences in the mean of two groups of data or between a mean and a specified value are significant. This test bases its logic on Student's t-distribution [Student, 1908] and is a type of analysis of variance. It was invented by William Sealy Gosset, a statistician that worked at Guinness and that published his work under the pseudonym "Student" so as to avoid being identified by his employer.

Titmus Test It is probably the most widely used **stereo test** to assess **stereoacuity**. The Titmus test consists of a polarisation-coded stereogram where the two views are polarised at 90 deg with respect to each other. When the patient is provided with properly oriented spectacles, the two views are seen separately in the two eyes and a 3D target can be perceived.

TNO Test Type of **stereo test** based on **anaglyph Random Dot Stereograms**. This test uses a pair of red-green spectacles and a booklet of test plates that consist of **stereograms** in which the two views have been superimposed and printed in complementary colours. When the test plates are viewed with the red-green glasses, a 3D image can be perceived.

Transient Stereopsis Impression of depth generated by the transient stereoscopic system, which is stimulated by briefly presented stimuli (e.g. for a few hundred milliseconds) characterised by large amount of disparity (up to 10 deg). This in contrast to the sustained stereoscopic system, which instead processes small disparities (within the singleness range) presented for a long duration.

Transparency The property of a material of being clear and transparent, allowing the passage of light and causing a clear see-through effect. Examples of transparent materials are clear glass and air. The opposite property to transparency is **opacity**, while translucency is the property of materials that allow some light to pass through but only diffusely and can not be seen through clearly. Mist and very thin cloth are examples of translucent materials.

Visual Angle The visual angle, or angular size of an object, is the angle subtended by the object at the eye. It is usually measured in degrees (deg or °).

Visual Cortex Part of the brain responsible for the processing of visual information. It is located at the back of the brain and it can be found in both hemispheres: the right hemisphere visual cortex receives information from the left visual field and the left hemisphere visual cortex from the right visual field.

Voxel By analogy with pixel, voxel is the abbreviation for a **vo**lume **el**ement. In 3D visualisation a stereoscopic voxel represents the smallest discernible element in a 3D space and can be used as a unit of measure for the resolution of a 3D display system.

Viewing Window (or Viewing Zone) Portion of space in front of a 3D display within which the left (or right) view of the stereoscopic image shown on the screen can be seen correctly by the observer's corresponding eye.

Appendix B

Materials

This appendix contains the questionnaires and the forms used during the main experiment.

3D DISPLAYS ASSESSMENT

Consent Form

The participant should complete the whole of this sheet himself/herself

	Please tick as necessary
Have you read the Participant Information Sheet?	YES [ ] NO [ ]
Have you had an opportunity to ask questions and to discuss the study?	YES [ ] NO [ ]
Have you received satisfactory answers to all of your questions?	YES [ ] NO [ ]
Have you received enough information about the study?	YES [ ] NO [ ]
Who have you spoken to? Dr/Mr/Mrs/Ms/Prof.	
Do you consent to participate in the study?	YES [ ] NO [ ]
Do you understand that you are free to withdraw from the study: • at any time and • without having to give a reason for withdrawing and • without affecting your position in the University?	YES [ ] NO [ ]

Signed .... **Date** ....

(NAME IN BLOCK LETTERS) ....

Signature of witness .... **Date** ....

(NAME IN BLOCK LETTERS) ....

3D DISPLAYS ASSESSMENT

Experiment Instructions for Participants

Introduction

Thank you for agreeing to participate to this experiment.

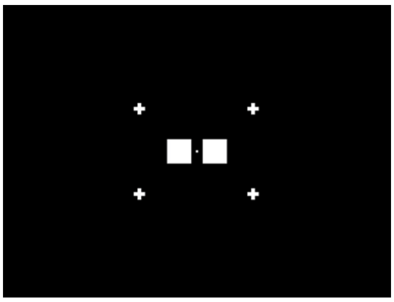
The aim of this test is to assess the quality of seven different autostereoscopic displays. In order to do so each of you will be asked to perform a task based on 3D-depth judgment as explained in the following instructions.

Please try to understand these instructions at the best of your ability.

Stimuli and Task

The image used for the test consists of two white squares on a black background. The squares are centred in the middle of the screen and are positioned horizontally one next to the other, as shown in the figure below. Between the two squares there is a white small square (fixation point) that marks the centre of the screen. You must maintain fixation on (i.e. stare at) this point throughout each trial as you make your depth judgment.

Around the central squares there are also four crosses, which will help you to maintain the right viewing position. When you are sitting in the correct viewing position, all four crosses should appear to be in front of the screen plane.



Your task is to identify which of the two squares is the closest to you. If the left square appears to be closer, press key “C” on the keyboard. If instead the right square appears to be closer, press key “M”.

Procedure and Schedule

You will be asked to perform this task a number of times on each one of the seven displays, in turn. Each display session will last about ten minutes, while the duration of the whole experiment will be approximately two hours. You can use the following table as a guideline.

Introduction	10 minutes
Session 1	10 minutes
Survey Session 1	10 minutes
Session 2	5 minutes
Survey Session 2	5 minutes
Session 3	5 minutes
Survey Session 3	5 minutes
Session 4	5 minutes
Survey Session 4	5 minutes
Break	20 minutes
Session 5	5 minutes
Survey Session 5	5 minutes
Session 6	5 minutes
Survey Session 6	5 minutes
Session 7	5 minutes
Survey Session 7	5 minutes
Final Survey	10 minutes
Debriefing Session	10 minutes

Important Notes

During the experiment it is important that you bear in mind the following points. They are crucial for the success of the experiment.

- Try to perform the task as quickly as possible **but try not to make mistakes**. If you do not know which square is the closest please just guess.
- Try to **keep the right viewing position**. If you are not sure, check the four crosses and remember that they should always appear to be in front of the screen plane.
- **Please keep both eyes open** at all times other than blinking. It is important that you do not shut one of your eyes while keeping the other open.
- Finally remember that while you perform the task, you must **maintain fixation on the fixation point** in the centre of the screen and you should **not tilt your head**.

3D DISPLAYS ASSESSMENT

Survey Session 7

Display Type: ………………………

Candidate ID: ………………………

Please tick as necessary

1) How comfortable was to see the 3D images with this display?

1. Very comfortable [ ]
2. Comfortable [ ]
3. Neither comfortable nor uncomfortable [ ]
4. Uncomfortable [ ]
5. Very uncomfortable [ ]

2) Was there any factor that was disturbing you during the experiment? YES [ ] NO [ ]

If YES, please explain.

……
……
……
……

3) Did you experience any discomfort (i.e headache, eye strain, etc.) during this session of the experiment? YES [ ] NO [ ]

If YES, how strong was the discomfort?

1. Very strong [ ]
2. Strong [ ]
3. Neither strong nor mild [ ]
4. Mild [ ]
5. Very mild [ ]

4) Any other comment concerning the display?

……
……
……
……

3D DISPLAYS ASSESSMENT

General Survey

All information and personal data provided will be processed anonymously and with high confidentiality.

Please tick as necessary

SESSION I – PERSONAL DATA

1) **Candidate ID:** ……………

2) **Candidate name:** ………………………………………….

3) **E-mail address:** …………………………………………….

4) **Age:** ………

5) **Gender:** F [ ] M [ ]

6) **Left Handed** [ ] **Right Handed** [ ]

7) **Do you use any corrective lenses?** YES [ ] NO [ ]
If you replied YES to this question, please also reply to question 8 and 9.

8) **If YES, you use them to correct:**

Nearsightedness	YES [ ] NO [ ]
Astigmatism	YES [ ] NO [ ]
Farsightedness (i.e. reading glasses)	YES [ ] NO [ ]
Other	………………….
I don't know	[ ]

9) **Did you use your corrective lenses during the experiment?** YES [ ] NO [ ]

SESSION II – FEEDBACK

1) Have you used 3D displays before this experiment? YES [ ] NO [ ]

2) How did you find the whole experience of using 3D displays?

1. Very good [ ]
2. Good [ ]
3. Neither good nor bad [ ]
4. Bad [ ]
5. Very bad [ ]

3) Below is the list of the displays you used during the experiment. Please sort them in order from 1 = the best to 7 = the worst.

DTI [ ]
SeeReal [ ]
Colorlink [ ]
Sharp [ ]
Iris3D [ ]
Kodak [ ]
Shutter Glasses [ ]

4) What was the feature/features that you liked BETTER on the display that you graded as the best?

………
………
………

5) What was the feature/features that was MOST UNPLEASENT on the display that you graded as the worst?

………
………
………

6) During or after the experiment did you experience any discomfort (i.e headache, eye strain, etc.)? YES [ ] NO [ ]

If YES, please explain.

………
………
………

7) Any other comments.

………
………
………
………

3D DISPLAYS ASSESSMENT

Debriefing From

Barbara Froner
RA and PhD Student in Computer Science
Supervisor: Dr Nick Holliman

Dept of Computer Science
Durham University
Science Laboratories
South Road
Durham
DH1 3LE
UK

During this study, all participants were asked to perform a 3D-depth based task repetitively on seven different stereoscopic displays. The purpose of this study was to assess the performance and quality of different 3D displays and investigate their capability to reproduce 3D space.
The advantage of 3D displays over 2D displays is that they can reproduce the third dimension like we see it in reality and give a more realistic representation of 3D objects. By sending slightly different images of the same object or scene to a viewer's left and right eye, stereoscopic depth may be perceived. Sometime though 3D depth can be reproduced only at the expense of lower 2D quality (i.e. lower image resolution). It is therefore important to determine to what extent a loss in 2D image quality is compensated by the added value of seeing stereoscopic depth, and how this vary across different stereoscopic systems. Our research focuses on this matter and particularly on how well a stereoscopic system is able to reproduce 3D depth, which gives a direct indication of the quality of the stereoscopic system itself.

If you would like more information on this issue, the following references may be of interest to you.

[1] N.S.Holliman, “3D Display Systems", draft report to be published, in "Handbook of Optoelectronics", IOP Press, ISBN 0 7503 0646 7, in Press.
http://www.dur.ac.uk/n.s.holliman/Presentations/3dv3-0.pdf

[2] D. F. McAllister (Ed.), “Stereo Computer Graphics and other True 3D Technologies”, Princeton U. Press, Princeton, NJ, Oct. 1993.

If you are interested in the results of this study, you may contact Miss Barbara Froner (e-mail: , mob:) at the completion of this study (April, 2009). Please note that only global results, not individual results, will be disclosed.

If you have any questions or concerns about this study, please contact Miss Barbara Froner (please see contact details above).

Thank you for your participation.

Appendix C

Empirical Data

This appendix contains the score and the response time raw data collected during the pilot and the main experiment and consists of four tables, one table per independent measure per experiment. In each table, each row represents a specific experimental condition, i.e. the combination of a specific 3D display with a specific level of image disparity, while each column represents the mean performance of a subject across the 28 experimental conditions tested during the experiments.

Score data are expressed with a real number between zero and one; a score of 1.000 indicates optimum performance (100%), while a score of 0.500 indicates performance at chance (50%). Response time data are expressed in milliseconds (ms).

In the statistical analysis of the pilot experiment data, samples relative to subjects S8 and S14 were excluded due to poor score performance (see Chapter 4). Likewise, the sample relative to subject S13 was excluded from the statistical analysis of the main experiment data (see Chapter 5).

Pilot Experiment - Score Data (prob)

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14
DTI_0	0.857	0.143	0.214	0.357	0.607	0.679	0.786	0.750	0.393	0.929	0.107	0.429	0.964	0.179
DTI_2	1.000	1.000	1.000	0.964	1.000	0.786	0.821	0.393	1.000	1.000	1.000	0.929	1.000	0.429
DTI_4	1.000	1.000	1.000	1.000	0.964	0.571	0.786	0.571	0.964	1.000	1.000	0.964	1.000	0.536
DTI_6	1.000	1.000	1.000	1.000	0.964	0.821	0.821	0.429	1.000	0.929	1.000	0.964	1.000	0.571
SeeReal_0	0.821	0.250	0.286	0.643	1.000	0.464	0.607	0.429	0.500	0.607	0.321	0.536	0.786	0.429
SeeReal_2	0.500	0.536	0.571	0.643	0.500	0.607	0.571	0.393	0.429	0.500	0.571	0.536	0.393	0.429
SeeReal_4	1.000	0.821	1.000	1.000	1.000	0.500	0.964	0.143	0.964	0.357	1.000	1.000	0.964	0.821
SeeReal_6	1.000	0.857	1.000	1.000	1.000	0.464	1.000	0.036	1.000	0.357	1.000	1.000	0.964	0.929
Colorlink_0	1.000	0.321	0.393	0.607	0.607	0.429	0.393	0.643	0.679	0.964	0.679	0.500	0.786	0.000
Colorlink_4	1.000	1.000	0.964	0.857	0.964	1.000	0.893	0.679	1.000	0.000	1.000	0.964	0.964	0.500
Colorlink_2	1.000	1.000	1.000	0.929	1.000	0.786	0.964	0.500	1.000	0.000	1.000	0.964	0.964	0.500
Colorlink_6	1.000	1.000	1.000	1.000	1.000	0.714	0.964	0.321	1.000	0.071	1.000	1.000	0.964	0.500
Sharp_0	0.893	0.679	0.571	0.607	0.464	0.643	0.321	0.786	0.643	0.821	0.643	0.571	0.714	0.179
Sharp_2	1.000	0.929	1.000	0.786	1.000	1.000	0.893	0.821	1.000	0.893	1.000	1.000	0.964	0.964
Sharp_4	0.964	0.964	1.000	1.000	1.000	0.857	0.821	0.321	1.000	0.857	1.000	1.000	0.964	0.964
Sharp_6	1.000	1.000	1.000	1.000	0.786	0.571	0.893	0.071	1.000	0.321	0.929	1.000	0.893	0.893
Iris3D_0	0.929	0.643	0.714	0.536	0.643	0.857	0.643	0.536	0.821	0.893	0.571	0.536	0.857	0.429
Iris3D_2	1.000	1.000	1.000	0.964	1.000	1.000	0.929	0.643	1.000	1.000	1.000	1.000	1.000	0.679
Iris3D_4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.786	1.000	1.000	1.000	1.000	1.000	0.464
Iris3D_6	1.000	1.000	1.000	1.000	1.000	0.964	0.964	0.643	1.000	1.000	1.000	1.000	1.000	0.607
Kodak_0	0.929	0.500	0.786	0.429	0.393	0.357	0.286	0.357	0.536	0.929	0.393	0.214	0.571	0.179
Kodak_2	1.000	0.964	1.000	0.750	1.000	1.000	0.893	0.607	1.000	0.964	1.000	1.000	1.000	0.643
Kodak_4	1.000	1.000	1.000	0.821	1.000	0.964	0.893	0.536	1.000	0.786	1.000	1.000	1.000	0.500
Kodak_6	1.000	1.000	1.000	0.786	1.000	1.000	0.857	0.571	1.000	0.786	1.000	1.000	1.000	0.536
ShutterGL_0	1.000	0.321	0.250	0.607	0.679	0.821	0.429	0.821	0.536	0.857	0.250	0.643	0.893	0.250
ShutterGL_2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.821	1.000	1.000	1.000	1.000	1.000	0.607
ShutterGL_4	0.964	1.000	1.000	1.000	1.000	0.893	1.000	0.857	1.000	0.964	1.000	1.000	1.000	0.500
ShutterGL_6	0.964	1.000	1.000	1.000	1.000	0.821	1.000	0.357	1.000	1.000	1.000	1.000	1.000	0.607

Pilot Experiment - Response Time Data (ms)

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14
DTI_0	3702	6001	1580	5387	1918	1160	2071	2954	2986	4728	2507	2099	2803	2400
DTI_2	864	2717	819	3856	863	1220	2212	2424	547	2207	982	2709	1154	2032
DTI_4	610	2647	845	2473	774	1769	1853	2208	507	3647	748	1743	1399	2489
DTI_6	605	2565	897	2118	1201	1738	1051	2961	677	5149	2852	2808	1773	2259
SeeReal_0	2225	3468	1720	3271	1448	1633	8090	1905	1316	3953	2175	2008	3037	5521
SeeReal_2	2204	3989	1670	3265	1177	2144	4344	1878	1135	4174	2528	1833	2732	5386
SeeReal_4	864	3037	663	1264	844	2158	882	2020	774	4628	1604	986	2100	5584
SeeReal_6	1257	3068	845	1039	1251	2446	1402	1891	879	4152	1460	1735	2268	4809
Colorlink_0	2123	3261	1856	6431	1535	4596	5486	1705	1790	3478	3039	3213	14060	5838
Colorlink_4	851	2140	1242	4445	690	2188	3284	1608	691	2506	698	2342	4374	6615
Colorlink_2	740	2449	1219	3093	681	4106	1788	1780	550	3145	780	1542	5243	6071
Colorlink_6	733	2015	1444	1715	1011	3638	2559	2004	480	4371	1293	1108	6274	6311
Sharp_0	2373	2971	2028	5017	2305	3827	2986	1518	1284	2302	2155	19090	3112	3186
Sharp_2	906	1738	1177	3790	1740	2028	2238	1698	464	1678	683	5816	1542	2545
Sharp_4	767	1572	1057	1855	1961	2780	2235	1606	500	2972	1024	4920	1692	2978
Sharp_6	754	1550	1078	1283	3486	3468	2617	1502	382	2455	2774	6420	2978	2962
Iris3D_0	1574	2973	2189	8336	2281	4029	3710	1898	1567	3845	3782	4861	3328	2931
Iris3D_2	934	1037	1112	4393	2360	1168	906	1931	910	1367	921	1469	1714	2754
Iris3D_4	745	1044	1321	4544	1301	1100	913	2092	521	1369	804	1371	1516	1914
Iris3D_6	804	1077	808	3630	1243	1224	1145	1786	621	1950	1210	1797	1845	2484
Kodak_0	1272	3696	3299	4234	2445	2466	4630	1801	2055	6787	4578	2652	3950	1946
Kodak_2	857	1358	813	3176	742	823	2947	1763	669	2510	1043	1429	1369	2669
Kodak_4	1132	1462	837	3170	1017	840	2922	1956	626	3858	1364	1266	1744	2263
Kodak_6	1073	1796	1155	3979	1417	881	2390	1888	865	3884	1069	1317	1869	2220
ShutterGL_0	1478	6671	2055	4531	2114	1486	2218	1685	1197	6898	3172	3241	27789	2627
ShutterGL_2	881	2474	638	2528	986	801	967	1633	682	1991	914	2105	800	2734
ShutterGL_4	830	2456	653	2905	1279	1230	849	1636	492	3733	1506	1255	1248	3269
ShutterGL_6	796	1764	556	3495	1088	1231	868	1835	750	3952	2222	2592	1296	3121

Main Experiment - Score Data (prob)

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	*S13*	S14
DTI_0	0.464	0.536	0.607	0.536	0.536	0.714	0.500	0.464	0.464	0.464	0.500	0.357	*0.357*	0.429
DTI_1	0.536	0.607	0.607	0.750	0.786	0.750	0.571	0.571	0.500	0.393	0.429	0.500	*0.571*	0.607
DTI_2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	*0.821*	1.000
DTI_3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	*0.964*	1.000
SeeReal_0	0.429	0.536	0.500	0.500	0.500	0.500	0.500	0.536	0.429	0.429	0.536	0.536	*0.500*	0.536
SeeReal_1	0.500	0.393	0.464	0.536	0.500	0.464	0.500	0.571	0.536	0.536	0.536	0.536	*0.500*	0.429
SeeReal_2	0.571	0.536	0.321	0.429	0.500	0.679	0.500	0.536	0.536	0.500	0.607	0.607	*0.464*	0.429
SeeReal_3	1.000	1.000	0.964	1.000	1.000	1.000	0.964	1.000	1.000	1.000	1.000	1.000	*0.857*	1.000
Colorlink_0	0.464	0.500	0.393	0.536	0.500	0.500	0.536	0.500	0.536	0.393	0.429	0.536	*0.464*	0.429
Colorlink_1	1.000	1.000	0.929	0.929	0.929	0.964	0.929	0.964	0.857	0.893	1.000	0.929	*0.786*	1.000
Colorlink_2	1.000	1.000	0.964	1.000	1.000	1.000	1.000	1.000	0.964	1.000	1.000	0.964	*0.893*	1.000
Colorlink_3	1.000	1.000	0.929	1.000	1.000	0.964	1.000	1.000	1.000	1.000	1.000	1.000	*0.893*	1.000
Sharp_0	0.500	0.500	0.607	0.500	0.500	0.607	0.500	0.464	0.571	0.607	0.571	0.500	*0.536*	0.429
Sharp_1	1.000	1.000	0.929	1.000	1.000	1.000	0.786	0.964	0.964	1.000	1.000	1.000	*0.821*	1.000
Sharp_2	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.964	1.000	1.000	1.000	*0.821*	1.000
Sharp_3	1.000	1.000	0.964	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	*0.964*	1.000
Iris3D_0	0.500	0.571	0.500	0.536	0.500	0.500	0.536	0.571	0.536	0.464	0.607	0.393	*0.607*	0.536
Iris3D_1	1.000	1.000	0.893	0.929	1.000	1.000	1.000	0.964	0.964	1.000	1.000	1.000	*1.000*	1.000
Iris3D_2	1.000	1.000	0.821	1.000	1.000	1.000	1.000	1.000	0.964	1.000	1.000	1.000	*1.000*	1.000
Iris3D_3	1.000	1.000	0.893	1.000	1.000	0.964	1.000	1.000	0.964	1.000	1.000	1.000	*1.000*	1.000
Kodak_0	0.500	0.679	0.464	0.536	0.500	0.679	0.464	0.536	0.321	0.321	0.607	0.571	*0.500*	0.500
Kodak_1	0.857	0.893	0.893	0.964	1.000	0.893	0.893	1.000	0.714	1.000	1.000	1.000	*0.964*	1.000
Kodak_2	1.000	0.893	0.929	1.000	1.000	0.893	1.000	1.000	0.714	1.000	1.000	1.000	*0.964*	1.000
Kodak_3	1.000	0.964	0.750	1.000	1.000	0.750	1.000	1.000	0.536	1.000	1.000	1.000	*1.000*	1.000
ShutterGL_0	0.393	0.536	0.429	0.536	0.500	0.679	0.393	0.429	0.536	0.464	0.643	0.464	*0.500*	0.464
ShutterGL_1	1.000	1.000	1.000	0.964	1.000	0.893	1.000	1.000	0.929	1.000	1.000	0.964	*1.000*	1.000
ShutterGL_2	1.000	1.000	1.000	1.000	1.000	0.893	1.000	1.000	1.000	1.000	1.000	1.000	*1.000*	1.000
ShutterGL_3	1.000	1.000	1.000	1.000	1.000	0.750	1.000	1.000	1.000	1.000	1.000	1.000	*1.000*	1.000

Main Experiment - Response Time Data (ms)

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	*S13*	S14
DTI_0	1876	2666	1701	1195	2492	1920	1112	2656	3151	1910	1094	1398	*1546*	989
DTI_1	1312	2255	1807	1798	2257	1838	1070	2838	2329	2490	929	1325	*1181*	1227
DTI_2	690	614	908	805	542	826	3821	1115	1042	844	708	588	*1021*	597
DTI_3	712	738	817	731	627	1350	558	983	865	776	727	547	*1211*	560
SeeReal_0	1151	1756	1474	726	889	1785	1133	2703	1835	1213	1045	990	*1281*	3048
SeeReal_1	1458	1770	1749	1313	741	1483	1184	2500	2379	1715	1073	974	*1346*	3318
SeeReal_2	1291	1672	1521	816	914	1514	1014	2733	2237	3777	1135	1379	*1400*	3107
SeeReal_3	594	757	721	492	3377	791	948	851	1074	1100	789	667	*985*	898
Colorlink_0	1747	1548	1424	1507	731	4642	842	4006	1102	1278	1006	845	*3776*	2092
Colorlink_1	595	1147	718	1266	1464	2931	982	1780	936	974	832	796	*2304*	1151
Colorlink_2	570	587	650	551	3638	1933	622	1027	1088	664	594	618	*1552*	995
Colorlink_3	1333	574	569	657	567	1828	599	896	743	564	616	813	*2342*	648
Sharp_0	1411	1085	1409	2116	4876	4540	801	2506	1493	1768	940	4492	*2077*	2218
Sharp_1	547	897	667	485	1595	1022	833	1061	1249	929	594	1448	*1330*	748
Sharp_2	505	461	717	438	830	1311	648	940	836	681	585	1108	*1083*	700
Sharp_3	495	479	596	437	817	1248	628	2758	828	1051	551	1134	*1167*	607
Iris3D_0	1780	1821	800	3748	3084	3252	937	3269	1185	1622	3059	3034	*1458*	2102
Iris3D_1	584	956	681	1876	1098	735	691	1502	951	1400	1683	1256	*838*	1137
Iris3D_2	566	535	736	825	793	618	614	1176	1897	968	954	1199	*792*	1276
Iris3D_3	603	525	597	977	803	692	660	991	606	986	944	1111	*818*	767
Kodak_0	735	2358	4773	2607	1795	3094	802	2792	1599	3856	1713	2722	*1990*	744
Kodak_1	974	1261	2958	882	963	2600	4354	1290	1109	1082	843	813	*1414*	721
Kodak_2	722	1596	2518	845	904	1949	675	1020	1023	927	787	706	*1073*	580
Kodak_3	580	2132	2824	1057	773	2208	752	1272	1683	1105	817	672	*1070*	565
ShutterGL_0	1488	8268	2935	2607	1795	3094	1087	2075	6766	2304	1288	1229	*1285*	1287
ShutterGL_1	2388	1548	1177	882	963	2600	542	743	1916	1192	760	677	*718*	493
ShutterGL_2	1395	682	978	845	904	1949	500	771	1144	686	707	548	*610*	449
ShutterGL_3	448	594	1047	1057	773	2208	580	611	1101	1006	771	2508	*555*	524

Appendix D

Geometric Perceived Depth on Planar 3D Displays

D.1 Pixel Size and Stereoacuity

A planar 3D display can only generate screen disparity in integer multiples of pixel pitch, therefore the smallest displayable screen disparity is the width of a single pixel of the underlying LCD panel. To ensure that the participants that took part to the experiments presented in this thesis could accurately detect depth during each trial, it was necessary to compare their measured stereoacuity against the stereoacuity needed to detect one pixel screen disparity on the display with the smallest pixel pitch.

All participants that took part in the 3D display comparative study had a threshold stereoacuity measured using the Titmus test of at least 40 seconds of arc. The display in the study with the smallest pixel pitch, viewed from the furthest distance, is the Iris3D display. The Iris3D is a Full-Resolution display, which makes it capable of reproducing an image disparity of one pixel as view disparity and hence as perceivable depth. This means that the experimental condition that required the best stereoacuity of all trials is the one that showed the stimulus treated with 1-pixel image disparity on the Iris3D display. Therefore the following calculations are based on the technical parameters of this display.

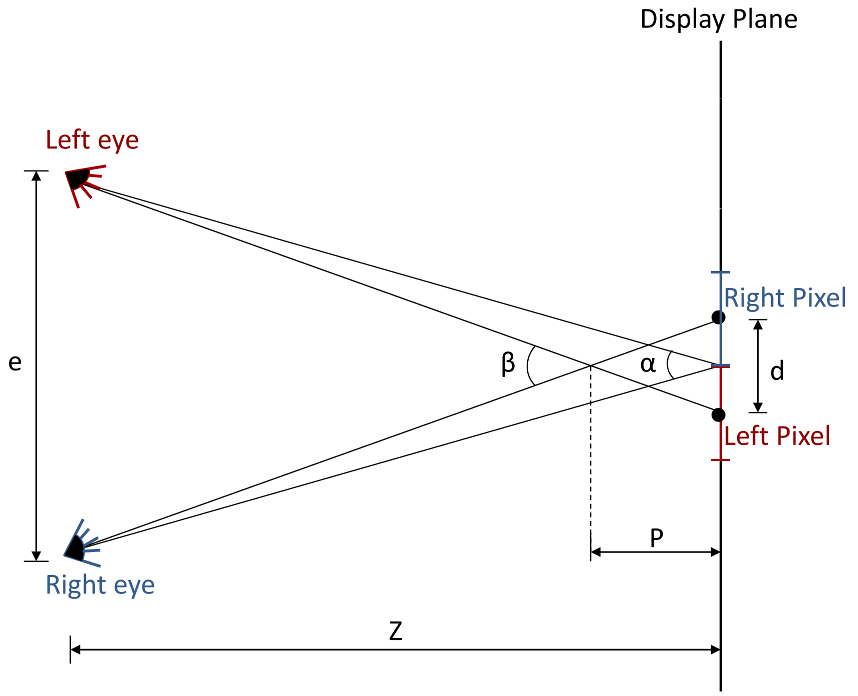


Figure D.1: Horizontal angular disparity due to one pixel screen disparity.

From the geometry shown in Figure D.1 it is possible to find the angular disparity equivalent to one pixel screen disparity, $\delta = |\alpha - \beta|$. The Iris3D display has a nominal viewing distance of $Z = 920$ mm and a pixel pitch of $d = 0.255$ mm. For this display the vergence angle α due to zero pixel screen disparity is given by:

$$\alpha = 2\arctan\left(\frac{\left(\frac{e}{2}\right)}{Z}\right) = 3.922^o \tag{D.1}$$

where e represents the observer's **InterPupillary Distance (IPD)** and is assumed to have a nominal value of 63 mm. While there is little agreement on what the value of e should be, a recent study by Dodgson [2004] suggests that the mean and median IPD for adult the human population lies around 63 mm; this study also shows how IPD varies greatly with age, gender and race.

To find the vergence angle β due to one pixel screen disparity it is necessary to calculate the perceived depth, P, due to one pixel crossed disparity using the equation for crossed disparity [Holliman, 2004]:

$$P = \frac{Z}{\left(\frac{e}{d}\right) + 1} = 3.709mm \tag{D.2}$$

By trigonometry the vergence angle β is then given by

$$\beta = 2 \arctan \left(\frac{\left(\frac{d}{2}\right)}{P} \right) = 3.938^o \tag{D.3}$$

The angular disparity due to one pixel screen disparity on the Iris3D display is therefore $\delta = 57.6$ seconds of arc. This is a larger disparity than the measured stereoacuity of all participants (i.e. 40 seconds of arc), which is a sufficient condition to guarantee that all participants in the trials were able to easily detect a one pixel screen disparity on all tested displays.

D.2 Vertical Misalignments and Depth Perception

The human visual system is tolerant to some vertical misalignment in the two images seen by the eyes [Fukuda et al., 2009]. However, when too prominent vertical misalignments can hinder the detection of depth derived from horizontal disparity [Howard and Rogers, 2002b] and therefore can affect stereovision. While there is little agreement on the exact threshold before vertical misalignments becomes disruptive, the most conservative limit suggested in literature is in the range of $3.4 - 3.5$ minutes of arc [Kooi and Toet, 2004] [Nielsen and Poggio, 1983].

In the 3D display comparative study, the only tested display that, given its optical design, introduced a vertical misalignment of one screen pixel between the left and right view was the Colorlink. In order to understand if this misalignment had a disruptive effect on participants' depth perception, it was necessary to compare the angular disparity induced by one screen pixel vertical disparity on the Colorlink display against the conservative threshold of $3.4-3.5$ minutes of arc assumed as the limit before stereopsis becomes affected.

The Colorlink display has a native resolution of 1280x1024 pixels and a physical screen size of 337.2x269.7 mm, which yields a pixel pitch of $d = 0.263$ mm. During the experiments this display was used with a nominal viewing distance of $Z = 650$ mm.

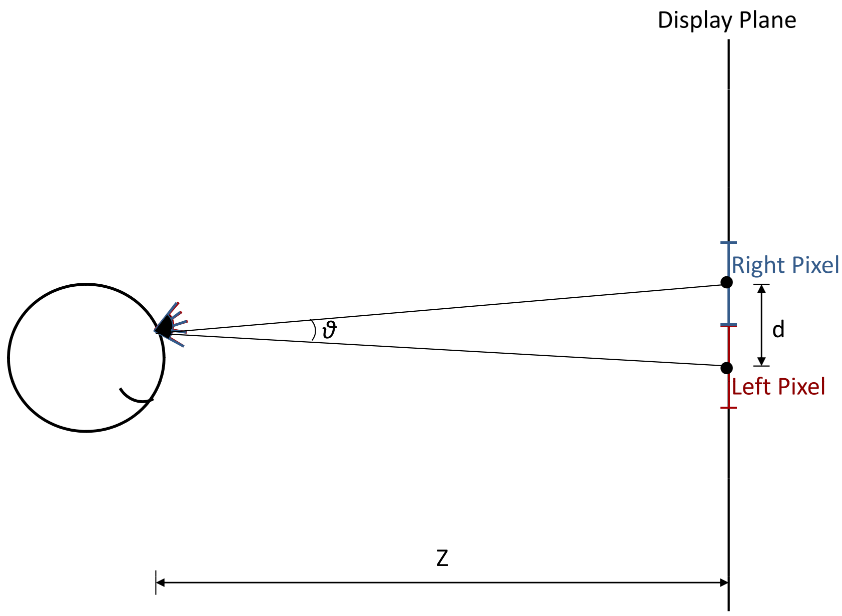


Figure D.2: Vertical angular disparity due to one pixel screen disparity.

Based on the geometry shown in Figure D.2, the angular vertical disparity due to one pixel screen disparity on the Colorlink display, θ, is given by:

$$\theta = 2\arctan\left(\frac{\left(\frac{d}{2}\right)}{Z}\right) = 1.391' \tag{D.4}$$

The vertical angular disparity due to one pixel screen disparity on the Colorlink display is therefore $\theta = 1.391$ minutes of arc, which is well below the threshold value of 3.4 minutes of arc. Based on this, it is safe to state that the vertical misalignment that inherently characterises the Colorlink display did not hinder participants' depth perception and had no detrimental effect on their task performance during the experimental trials presented in this thesis.

Note that given the same pixel pitch, $d = 0.263$ mm, and a viewing distance of $Z = 270$ mm the Colorlink display used in this study would yield an angular vertical disparity of $\theta = 3.349$ minutes of arc, which lies just below the critical threshold before vertical misalignments become disruptive to stereo vision. With a viewing distance of $Z = 250$ mm the angular vertical disparity on the Colorlink display would be $\theta = 3.6165$ minutes of arc; this is above the threshold value of 3.4 minutes of arc and suggests that in these viewing conditions the vertical misalignment embedded in the Colorlink display could potentially hinder the depth perception of users.

Bibliography

R.C. Allen and M.L. Rubin. Chromostereopsis. *Survey of Ophthalmology*, 26(1):22–27, 1981.

R.S. Allison, B.J. Gillam, and E. Vecellio. Binocular depth discrimination and estimation beyond interaction space. *Journal of Vision*, 9:1–14, January 2009.

Z.Y. Alpaslan and A.A. Sawchuk. Three-dimensional interaction with autostereoscopic displays. In *Stereoscopic Displays and Virtual Reality Systems XI, Proceedings of the SPIE*, volume 5291, pages 227–236, May 2004.

Z.Y. Alpaslan and A.A. Sawchuk. Three-dimensional interaction with autostereoscopic displays. United States Patent No. 0012675, January 2006.

Z.Y. Alpaslan, S. Yeh, A.A. Rizzo III, and A.A. Sawchuk. Effects of gender, application, experience, and constraints on interaction performance using autostereoscopic displays. In *Stereoscopic Displays and Applications XVII, Proceedings of SPIE*, volume 6055, page 6055A, 2006.

B.L. Anderson. The role of occlusion in the perception of depth, lightness, and opacity. *Psychological Review*, 110(4):785–801, 2003.

B.L. Anderson, M. Singh, and R.W. Fleming. The interpolation of object and surface structure. *Cognitive Psychology*, 44:148–190, 2002.

P. Anderson. Advanced display technologies. Technical report, JISC Technology and Standards Watch, Bristol, UK, August 2005. Available from World Wide Web: http://www.jisc.ac.uk. Last accessed: 25th July 2010.

D.A. Atchison, E.L. Markwell, S. Kasthurirangan, J.M. Pope, G. Smith, and P.G. Swann. Age-related changes in optical and biometric characteristics of emmetropic eyes. *Journal of Vision*, 8(4):1–20, April 2008.

H. Azari, I. Cheng, and A. Basu. Stereo 3D mouse cursor: A method for interaction with 3D objects in a stereoscopic virtual 3D space. *International Journal of Digital Multimedia Broadcasting*, 2010.

E. Bachelder and D. McRuer. Perception-based synthetic cueing for night-vision device rotorcraft hover operations. In *Proceedings of SPIE, Technologies, systems, and architectures for transnational defense*, volume 47453, pages 377–388. Society of Photo-Optical Instrumentation Engineers, 2002.

P.T. Barham and D.F. McAllister. A comparison of stereoscopic cursors for the interactive manipulation of B-splines. In *Stereoscopic Displays and Applications II, Proceedings of SPIE*, volume 1457, pages 18–26, August 1991.

C. Basdogan, C.-H. Ho, M.A. Srinivasan, and M. Slater. Visual discomfort and visual fatigue of stereoscopic displays: A review. *ACM Transactions on Computer-Human Interaction*, 7(4):443–460, December 2000.

J. Beck and R. Ivry. On the role of figural organization in perceptual transparency. *Perception & Psychophysics*, 44(6):585–594, 1988.

S.A. Benton. *Selected papers on three-dimensional displays*, volume 162 of Spie Press Monograph. SPIE Optical Engineering Press, 2001.

G. Berkeley. An essay towards a new theory of vision. In "A New Theory of Vision and other writings", Introduction by A.D. Lindsay. London: J.M. Dent & Sons Ltd. (1910), 1709.

B. Berlin and P. Kay. *Basic Color Terms: Their Universality and Evolution*. University of California Press, 1969.

R.N. Berry. Quantitative relations among vernier, real depth, and stereoscopic depth acuities. *Journal of Experimental Phychology*, 38(6):708–721, December 1948.

P.J. Bex, G.K. Edgar, and A.T. Smith. Temporal aliasing: Investigating multiple imaging. In *Proceedings of The Annual Meeting of the Applied Vision Association, "Simulated and Virtual Realities"*, 1993.

F. Biocca, A. Tang, C. Owen, and F. Xiao. Attention funnel: Omnidirectional 3D cursor for mobile augmented reality platforms. In *CHI 2006: Proceedings of the Conference on Human Factors in Computing Systems*, pages 1115–1122, 2006a.

F. Biocca, A. Tang, C. Owen, and F. Xiao. The omnidirectional attention funnel: A dynamic 3D cursor for mobile augmented reality systems. In *Proceedings of the 39th Hawaii International Conference on System Sciences*, volume 1, page 22c, 2006b.

R.S. Bolia, W.R. DAngelo, and R.L. McKinley. Aurally aided visual search in three-dimensional space. *Human Factors*, 41(4):664–669, 1999.

P. Bos and T. Haven. Field-sequential stereoscopic viewing systems using passive glasses. In *Proceedings of SID*, volume 39, pages 39–43, 1989.

H. Bossong, M. Swann, A. Glasser, and V.E. Das. Applicability of infrared photorefraction for measurement of accommodation in awake-behaving normal and strabismic monkeys. *Investigative Ophthalmology and Visual Science*, 50(2):966–73, 2009.

D.A. Bowman, E. Kruijff, J.J. LaViola, and I. Poupyrev. *3D User Interfaces: Theory and Practice*. Addison Wesley, 2005. ISBN 0-201-75867-9.

R.M. Boynton and C.X. Olson. Salience of chromatic basic color terms confirmed by three measures. *Vision Research*, 30(9):1311–1317, 1990.

M.F. Bradshaw and B.J. Rogers. The interaction of binocular disparity and motion parallax in the computation of depth. *Vision Research*, 36(21):3457–3468, November 1996.

D.H. Brainard, D.R. Williams, and H. Hofer. Trichromatic reconstruction from the interleaved cone mosaic: Bayesian model and the color appearance of small spots. *Journal of Vision*, 8(5):1–23, 2008.

V. Bruce, P.R Green, and M.A. Georgeson. *Visual perception: physiology, psychology & ecology*. Psychology Press, fourth edition, 2003.

N. Bruno and J.E. Cutting. Minimodularity and the perception of layout. *Journal of Experimental Psychology*, 117(2):161–170, June 1988.

D.R.W. Butts and D.F McAllister. Implementation of true 3-D cursors in computer graphics. In *Three-Dimensional Imaging and Remote Sensing Imaging, Proceedings of the SPIE*, volume 902, pages 74–84, 1988.

D.E. Carver and D.F. McAllister. Development of a stereoscopic three-dimensional drawing application. In *Stereoscopic Displays and Applications II, Proceedings of SPIE*, volume 1457, pages 54–65, August 1991.

O.M. Castle. *Synthetic image generation for a multiple-view autostereo display*. PhD thesis, Cambridge University (GB), 1995.

A. Chaparro, C.F. Stromeyer, E.P. Huang, R.E. Kronauer, and R.T. Eskew Jr. Colour is what the eye sees best. *Nature*, 361(28 January 1993):348–350, 1993.

H. Chun. *Auditory Sensory Feedback Tool to Supplement Visual Data Perception in Radiologic Imaging - A Demonstration Using MR Mammography.* PhD thesis, Graduate School of The Ohio State University, 2006.

W.S. Chun, J. Napoli, O.S. Cossairt, R.K. Dorval, D.M. Hall, T.J. Purtell II, J.F. Schooler, Y. Banker, and G.E. Favalora. Spatial 3-D infrastructure: Display-independent software framework, high-speed rendering electronics, and several new displays. In *Stereoscopic Displays and Virtual Reality Systems XII, Proceedings of SPIE*, volume 5664, pages 302–312, 2005.

J.M. Cobb. Autostereoscopic desktop display: an evolution of technology. In A.J. Woods, M.T. Bolas, J.O. Merritt, and I.E. McDowall, editors, *Stereoscopic Displays and Virtual Reality Systems XII, Proceedings of SPIE*, volume 5664, pages 139–149, 2005.

L.K. Cormack, S.B. Stevenson, and C.M. Schor. Interocular correlation, luminance contrast and cyclopean processing. *Vision Research*, 31(12):2195–2207, 1991.

L.K. Cormack, S.B. Stevenson, and D.D. Landers. Interactions of spatial frequency and unequal monocular contrasts in stereopsis. *Perception*, 26(9):1121–1136, 1997.

F.C. Crow. The aliasing problem in computer-generated shaded images. *Communications of the ACM*, 20:799–805, 1977.

C. Cruz-Neira, D.J. Sandin, and T.A. DeFanti. Surround-screen projection-based virtual reality, the design and implementation of the cave. In *Proceedings of SIGGRAPH*, volume 27, pages 135–142. ACM Press, 1993.

C.A. Curcio, K.A. Allen, K.R. Sloan, C.L. Lerea, J.B. Hurley, I.B. Klock, and A.H. Milam. Distribution and morphology of human cone photoreceptors stained with anti-blue opsin. *Journal of Comparative Neurology*, 312(4):610–624, 1991.

M.H.E. de Lussanet and M. Lappe. Bistability of point-light biological motion and necker cubes under stereoscopic and perspective depth cues. Perception 37 ECVP Abstract Supplement, page 70,, 2008.

K.K. De Valois. *The visual neurosciences*, volume 2, chapter 59 - The role of colour in spatial vision, pages 924–935. MIT Press, November 2003.

C.M.M. de Weert and K.J. Sadza. *Colour Vision: Physiology and Psychophysics*, chapter New data concerning the contribution of colour differences to stereopsis, pages 553–562. Academic Press, 1983.

H.E.M. den Ouden, R. van Ee, and E.H.F. de Haan. Colour helps to solve the binocular matching problem. *Journal of Physiology*, 567(2):665–671, 2005.

M. Dengler and W. Nitschke. Color stereopsis: a model for depth reversals based on border contrast. *Perception & Psychophysics*, 53(2):150–156, 1993.

N.A. Dodgson. Variation and extrema of human interpupillary distance. In *Stereoscopic Displays and Virtual Reality Systems XV, Proceedings of SPIE*, volume 5291, pages 36–46. SPIE, 2004.

D. Drascic and P. Milgram. Positioning accuracy of a virtual stereographic pointer in a real stereoscopic video world. In *Stereoscopic Displays and Applications II, Proceedings of SPIE*, volume 1457, pages 302–313, August 1991.

É.-F. Dutour. Discussion d'une question d'optique [discussion on a question of optics]. *l'Acadmie des Sciences. Mmoires de Mathmatique et de physique prsents par Divers Savants*, 3:514–530, 1760.

É.-F. Dutour. Addition au mmoire intitul, discussion d'une question d'optique, imprim dans le troisime volume des mmoires des savan[t]s trangers, pages 514 suivantes. [addition to the memoir entitled, discussion on a question of optics printed in the third volume of memoirs of foreign scientists, pages 514 and following]. *l'Acadmie des Sciences. Mmoires de Mathmatique et de physique prsents par Divers Savants*, 4:499–511, 1763.

M. D'Zmura. Color in visual search. *Vision Research*, 31(6):951–966, 1991.

J. B. Eichenlaub. Developments in autostereoscopic technology at dimension technologies inc. In *Stereoscopic Displays and Applications IV, Proceedings of SPIE*, volume 1915, pages 177–186, 1993.

J.B. Eichenlaub. Further advances in autostereoscopic technology at dimension technologies inc. In *Stereoscopic Displays and Applications III, Proceedings of SPIE*, volume 1669, pages 163–175, 1992.

J.B. Eichenlaub. Lightweight, compact 2D/3D autostereoscopic LCD backlight for games, monitor, and notebook applications. In *Stereoscopic Displays and Virtual Reality Systems IV, Proceedings of SPIE*, volume 3012, 1997.

J.B. Eichenlaub and R.W. Gruhlke. Reduced thickness backlighter for autostereoscopic display and display using the backlighter. United States Patent No. 5897184, April 1999.

E. Emmert. Grössenverhältnisse der Nachbilder. *Klinische Monatsblätter für Augenheilkunde*, 19:443–450, 1881.

R. Engle. Beowulf 3D: a case study. In A.J. Woods, N.S. Holliman, and J.O. Merritt, editors, *Stereoscopic displays and applications XIX, Proceedings of SPIE*, volume 6803, pages 23–30, 2008.

A.X. Falcão and J.K. Udupa. A 3D generalization of user-steered live-wire segmentation. *Medical Image Analysis*, 4(4):389–402, 2000.

S.M. Faris. Novel 3D stereoscopic imaging technology. In *Stereoscopic Displays and Virtual Reality Systems, Proceedings of SPIE*, volume 2177, 1994.

S.M. Faris. Multi-mode stereoscopic imaging system. United States Patent No. 6195205, February 2001.

G.E. Favalora. Volumetric three-dimensional display architecture. United States Patent No. 6487020, November 2002.

G.E. Favalora, J. Napoli, D.M. Hall, R.K. Dorval, M.G. Giovinco, M.J. Richmond, and W.S. Chun. 100 million-voxel volumetric display. In *Cockpit Displays IX: Displays for Defense Applications, Proceedings of SPIE*, volume 4712, pages 300–312, 2002.

A. Field. *Discovering Statistics by Using SPSS - Second Edition*. SAGE Publications, 2005.

P.M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47(6):381–391, June 1954. Reprinted in Journal of Experimental Psychology: General, 121(3):262–269, 1992.

R. Fleming and B.L. Anderson. *The visual neurosciences*, volume 2, chapter 86 - The perceptual organization of depth, pages 1284–1299. MIT Press, November 2003.

B. Froner and N. Holliman. Implementing an improved stereoscopic camera model. In *Procedings of Eurographics UK Chapter, Theory and Practice of Computer Graphics 2005*, Canterbury, June 2005. Eurographics Association.

B. Froner, N.S. Holliman, and S.P. Liversedge. A comparative study of fine depth perception on two-view 3D displays. *Displays*, 29(5):440–450, December 2008.

K. Fukuda, L.M. Wilcox, R.S. Allison, and I.P. Howard. A reevaluation of the tolerance to vertical misalignment in stereopsis. *Journal of Vision*, 9(2):1–8, February 2009.

J.W. von Goethe. *Zur Farbenlehre*. Tübingen, 1810.

M.M. Goodsitt, H.P. Chan, and L. Hadjiiski. Stereomammography: Evaluation of depth perception using a virtual 3D cursor. *Medical Physics*, 27(6):1305–1310, October 2000. ISSN 0094-2405.

R.T. Goodwin and P.E. Romano. Stereoacuity degradation by experimental and real monocular and binocular amblyopia. *Investigative Ophthalmology and Visual Science*, 26:917–923, 1985.

D.R. Griffin. *Echoes of bats and men.* Doubleday Publishing Group - Anchor Books, Garden City, N.Y., 1959.

P. Grossman. Depth from focus. *Pattern Recognition Letters*, 5(1):63–69, 1987.

T. Grossman and R. Balakrishnan. Pointing at trivariate targets in 3D environments. In *CHI 2004: Proceedings of the Conference on Human Factors in Computing Systems*, pages 447–454, 2004.

T. Grossman and R. Balakrishnan. An evaluation of depth perception on volumetric displays. In *AVI 2006: Proceedings of the working conference on Advanced Visual Interfaces*, pages 193–200, 2006a.

T. Grossman and R. Balakrishnan. The design and evaluation of selection techniques for 3D volumetric displays. In *Proceedings of the 19th annual ACM symposium on User interface software and technology*, pages 3–12, Montreux, Switzerland, 2006b. ACM.

C.R.C. Guibal and B. Dresp. Interaction of color and geometric cues in depth perception: When does red mean near? *Psychological Research*, 69(1-2):30–40, 2004.

M.S. Habib, J. Lowell, D. Vaideanu, N.S. Holliman, A. Hunter, and D.H.W. Steel. Assessment of qualitative stereo viewing and quantitative mapping of optic disc using polarized goggles verses autostereoscopic screen. In *Proceedings of the Annual Conference of the American Academy of Ophthalmology*, October 2005.

M.S. Habib, J.A. Lowell, N.S. Holliman, A. Hunter, D. Vaideanu, A. Hildreth, and D.H.W. Steel. Assessment of stereoscopic optic disc images using an autostereoscopic screen - experimental study. *BMC Ophthalmology*, 8(13), 2008.

D.L. Halpern and R.R. Blake. How contrast affects stereoacuity. *Perception*, 17(4):483–495, 1988.

F.J. Harris. *Multirate Signal Processing for Communication Systems.* Prentice Hall, 2004.

J. Harrold, A.M.S. Jacobs, G.J. Woodgate, and D. Ezra. 3D display systems hardware research at sharp laboratories of europe: an update. *Sharp Technical Journal*, (8), 1999.

J. Harrold, A. Jacobs, G.J Woodgate, and D. Ezra. Performance of a convertible 2D and 3D parallax barrier autostereoscopic display. In J. Morreale, editor, *20th International Display Research Conference, Proceedings of the Society for Information Display*, pages 280–283. Society of Information Display, September 2000.

L.O. Harvey and H.W. Leibowitz. Effects of exposure duration, cue reduction, and temporary monocularity on size matching at short distances. *Journal of the Optical Society of America*, 57(2):249–253, 1967.

R.S. Harwerth, M.C. Moeller, and J.M. Wensveen. Effects of cue context on the perception of depth from combined disparity and perspective cues. *Optometry & Vision Science*, 75(6):433–444, 1998.

D. Hassaine, N.S. Holliman, and S.P. Liversedge. Investigating the performance of path searching tasks in depth on multiview displays. *ACM Transactions on Applied Perception*, 8(1), October 2010.

G.M. Heider. New studies in transparency, form and color. *Psychologische Forschung*, 17 (1):13–55, 1933.

R.T. Held, E.A. Cooper, J.F. O'Brien, and M.S. Banks. Using blur to affect perceived distance and size. *ACM Transactions on Graphics*, 29(2):19:1–16, March 2010.

R.F. Hess, C.H. Liu, and Y.-Z. Wang. Luminance spatial scale and local stereo-sensitivity. *Vision Research*, 42:331–342, 2002.

R.F. Hess, C.H. Liu, and Y.-Z. Wang. Differential binocular input and local stereopsis. *Vision Research*, 43:2303–2313, 2003.

Y. Higashida, Y. Hirata, R. Saito, S. Doudanuki, H. Bussaka, and M. Takahash. Depth determination on stereoscopic digital subtraction angiograms. *Radiology*, 168(2):560–562, August 1988. ISSN 0033-8419.

L.F. Hodges. Basic principles of stereographic software development. In *Stereoscopic Displays and Applications II, Proceedings of SPIE*, volume 1457, pages 9–17, August 1991.

L.F. Hodges and E.T. Davis. Geometric considerations for stereoscopic virtual environments. *Presence*, 2(1):34–43, 1993.

D.M. Hoffman, A.R. Girshick, K. Akeley, and M.S. Banks. Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision*, 8(3): 1–30, 2008.

N.S. Holliman. Mapping perceived depth to regions of interest in stereoscopic images. In *Stereoscopic Displays and Virtual Reality Systems XI, Proceedings of SPIE*, volume 5291, pages 117–128, 2004.

N.S. Holliman. *Handbook of Optoelectronics*, volume II, chapter Three-Dimensional Display Systems. Taylor & Francis, May 2006. ISBN 0 7503 0646 7.

N.S. Holliman, B. Froner, and S.P. Liversedge. An application driven comparison of depth perception on desktop 3D displays. In *Stereoscopic Displays and Virtual Reality Systems XIV, Proceedings of SPIE*, volume 6490A, January 2007.

M. Hou. User experience with alignment of real and virtual objects in a stereoscopic augmented reality interface. In *CASCON '01: Proceedings of the 2001 conference of the Centre for Advanced Studies on Collaborative research*, page 6. IBM Press, 2001.

H.J. Howard. A test for the judgement of distance. *Transactions of the American Ophthalmological Society*, 17:195–235, 1919.

I.P. Howard. *Seeing in Depth*, volume 1 - Basic Mechanisms, chapter Vergence eye movement. I. Porteous, 2002a.

I.P. Howard. *Seeing in Depth*, volume 1 - Basic Mechanisms, chapter Binocular fusion and rivalry, pages 284–297. I. Porteous, 2002b.

I.P. Howard and B.J. Rogers. *Seeing in Depth*, volume 2 - Depth Perception, chapter Tokens for Stereopsis, pages 139–141. I. Porteous, 2002a.

I.P. Howard and B.J. Rogers. *Seeing in Depth*, volume 2 - Depth Perception, chapter Stereoscopic acuity. I. Porteous, 2002b.

J. Hsu, Z. Pizlo, D.M. Chelberg, C.F. Babbs, and E.J. Delp. Issues in the design of studies to test the effectiveness of stereo imaging. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 26(6):810–819, November 1996.

K. Isakovic, T. Dudziak, and K. Kchy. X-rooms: a PC-based immersive visualization environment. In *3D technologies for the World Wide Web, Proceeding of the seventh international conference on 3D Web technology*, pages 173–177. ACM SIGGRAPH, 2002. ISBN:1-58113-468-1.

W.H. Ittelson. Size as a cue to distance: Static localization. *American Journal of Psychology*, 64(1):54–67, January 1951.

K.-M. Jää-Aro and L. Kjelldahl. Effects of image resolution on depth perception in stereo and non-stereo images. In *Stereoscopic Displays and Virtual Reality Systems IV, Proceedings of SPIE*, volume 3012, pages 319–326, May 1997.

A. Jacobs, J. Mather, R. Winlow, D. Montgomery, G. Jones, M. Willis, M. Tillin, L. Hill, M. Khazova, H. Stevenson, and G. Bourhill. 2D/3D switchable displays. *Sharp Technical Journal*, (4), April 2003.

K.A. Jameson, S.M. Highnote, and L.M. Wasserman. Richer color experience in observers with multiple photopigment opsin genes. *Psychonomic Bulletin & Review*, 8(2):244–261, 2001.

J.-S. Jang, Y.-S. Oh, and B. Javidi. Spatiotemporally multiplexed integral imaging projector for large-scale high-resolution three-dimensional display. *Optics Express*, 12(4): 1077–1086, March 2004.

Z.X. Jin, Y.J. Zhang, X. Wang, and T. Plocher. Evaluating the usability of an autostereoscopic display. In *Proceedings of the 12th international conference on Human-Computer Interaction: interaction platforms and techniques*, pages 605–614, 2007.

E.B. Johnston, B.G. Cumming, and A.J. Parker. Integration of depth modules: stereopsis and texture. *Vision Research*, 33(5–6):813–826, March-April 1993.

E.B. Johnston, B.G. Cumming, and M.S. Landy. Integration of stereopsis and motion shape cues. *Vision Research*, 34(17):2259–2275, September 1994.

G. Jones, D. Lee, N. Holliman, and D. Ezra. Controlling perceived depth in stereoscopic images. In *Stereoscopic Displays and Virtual Reality Systems VIII, Proceedings of SPIE*, volume 4297, pages 42–53, 2001.

D.B. Judd and K.L. Kelly. Method of designating colors. *Journal of Research of the National Bureau of Standards*, 23:355–366, 1939.

B. Julesz. *Foundations of Cyclopean Perception*. University of Chicago Press, 1971. ISBN 0-226-41527-9.

G. Kanizsa. *Organization in Vision: Essays on Gestalt Perception*. Praeger Publishers, 1979. ISBN 0275903737.

L. Kaufman. *Sight and Mind: An Introduction to Visual Perception*. Oxford University Press, 1974.

L. Kaufman and J.H. Kaufman. Explaining the moon illusion. In *Proceedings of the National Academy of Sciences*, volume 97, pages 500–504, 2000.

R. Kayyali and S. Shirmohammadi. A benchmarked automated progress measurement system for haptic motor rehabilitation. *International Journal of Advanced Media and Communication*, 3(1-2):179–196, 2009.

F.A.A. Kingdom and D.R. Simmons. Stereoacuity and colour contrast. *Vision Research*, 36(9):1311–1319, 1996.

F.A.A. Kingdom, D.R. Simmons, and S. Rainville. On the apparent collapse of stereopsis in random-dot-stereograms at isoluminance. *Vision Research*, 39(12):2127–2141, 1999.

F.A.A. Kingdom, H.C.O. Li, and E.J. MacAulay. The role of chromatic contrast and luminance polarity in stereoscopic segmentation. *Vision Research*, 41(3):375–383, 2001.

B.N. Kishto. The colour stereoscopic effect. *Vision Research*, 5(6-7):313–329, 1965.

K. Knoblauch and S.K. Shevell. *The Visual Neurosciences*, volume 2, chapter 57 - Color Appearance, pages 892–907. MIT Press, November 2003.

K. Koffka. *Principles of Gestalt Psychology.* Harcourt, Brace & Company, New York, 1935.

L.H. Koh and W.N. Charman. Size constancy and angular size matching in size perception of near objects. *Optometry and Vision Science*, 76(4):241–246, 1999.

F. Kooi and A. Toet. Visual comfort of binocular and 3D displays. *Displays*, 25(2-3): 99–108, August 2004.

J. Krauskopf and J.D. Forte. Influence of chromaticity on vernier and stereo acuity. *Journal of Vision*, 2(9):645–652, 2002.

M.T.M. Lambooij, W.A. IJsselsteijn, M. Fortuin, and I. Heynderickx. Visual discomfort and visual fatigue of stereoscopic displays: A review. *Journal of Imaging Science and Technology*, 53(3):030201–(14), May/June 2009.

G.E. Legge and Y. Gu. Stereopsis and contrast. *Vision Research*, 29(8):989–1004, 1989.

Y. Li, A. Leithb, and J. Frank. Tinkerbell a tool for interactive segmentation of 3D data. *Journal of Structural Biology*, 120(3):266–275, 1997.

R. Likert. A technique for the measurement of attitudes. *Archive of Psychology*, 22(140): 1–55, 1932.

L. Lipton. *Stereo Computer Graphics and Other True 3D Technologies*, chapter 2 - Composition for Electrostereoscopic Displays, pages 11–25. Princeton University Press, second edition, May 1993.

L. Lipton. Factors affecting ghosting in a time-multiplexed plano-stereoscopic CRT display system. In *Three-dimensional visualization and display technologies, Proceedings of SPIE*, volume 761, pages 75–78, 1987.

L. Lipton and M. Ackerman. Liquid crystal shutter system for stereoscopic and other applications. United States Patent No. 4967268, October 1990.

M.S. Livingstone and D.H. Hubel. Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience*, 7(11):3416–3468, 1987.

C. Lu and D.H. Fender. The interaction of color and luminance in stereoscopic vision. *Investigative Ophthalmology*, 11(2):482–490, 1972.

H.A. Mallot. Spatial scale in stereo and shape from shading: Image input, mechanisms, and tasks. *Perception*, 26(9):1137–1146, 1997.

I. Mariko and N. Sachio. Evidence for the formal equivalence between emmert's law and size-distance invariance hypothesis. *Vision*, 17(4):237–241, 2005.

J.J. Marotta and M.A. Goodale. The role of familiar size in the control of grasping. *Journal of Cognitive Neuroscience*, 13(1):8–17, 2001.

D. Marr. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.* W.H. Freeman & Company, 1982. ISBN 0716712849, 0716715678. Foreword by S. Ullman, Afterword by T. Poggio.

J.A. Marshall, C.A. Burbeck, D. Ariely, J.P. Rolland, and K.E. Martin. Occlusion edge blur: a cue to relative visual depth. *Journal of the Optical Society of America A*, 13(4): 681–688, 1996.

G. Mather. Image blur as a pictorial depth cue. In *Proceedings of the Royal Society of London: Biological Sciences*, volume 263 of *B*, pages 169–172, 1996.

G. Mather and D.R.R. Smith. Depth cue integration: stereopsis and image blur. *Vision Research*, 40(25):3501–3506, 2000.

G. Mather and D.R.R. Smith. Blur discrimination and its relation to blur-mediated depth perception. *Perception*, 31(10):1211–1219, 2002.

G. Mather and D.R.R. Smith. Combining depth cues: effects upon accuracy and speed of performance in a depth-ordering task. *Vision Research*, 44(6):557–562, 2004.

J.W. Mauchly. Significance test for sphericity on a normal n-variate distribution. *Annals of Mathematical Statistics*, 11(2):204–209, June 1940.

J.G. May and D.R. Badcock. *Handbook of Virtual Environments: Design, Implementation, and Applications*, chapter Vision and Virtual Environments, pages 29–63. Lawrence Erlbaum Associates, 2002.

D.F. McAllister. *Stereo Computer Graphics and Other True 3D Technologies*, chapter 1 - Introduction, pages 1–10. Princeton University Press, second edition, May 1993.

D.F. McAllister. *Wiley Encyclopaedia on Imaging Science and Technology*, volume II, chapter Display Technology: Stereo & 3D Display Technologies, pages 1327–1344. Wiley Interscience, 2005.

K.J.W. McCaffrey, R.R. Jones, R. Holdsworth, R. Wilson, P. Clegg, J. Imber, N.S. Holliman, and I. Trinks. Unlocking the spatial dimension: digital technologies and the future of geoscience fieldwork. *Journal of the Geological Society*, 162(6):927–938, 2005.

J.E. McClain, A.J. Cacioppo, J.M. Reising, and R.J. Koubek. The interaction of chromostereopsis and stereopsis in stereoscopic CRT displays. *Applied Ergonomics*, 21(2): 101–106, 1990.

D.W. McCready. The moon illusion explained. Web site last visited on 30/03/2009, 2004.

S. McKay. Autostereoscopic 3D displays: bringing imax to the desktop? *First Break*, 23 (3):59–64, March 2005.

S. McKay, S. Mason, L.S. Mair, P. Waddell, and S.M. Fraser. Stereoscopic display using a 1.2-m diameter stretchable membrane mirror. In J.O. Merritt, M.T. Bolas, and S.S. Fisher, editors, *Stereoscopic Displays and Virtual Reality Systems VI, Proceedings of SPIE*, volume 3639, pages 122–131, 1999.

S. McKay, S. Mason, G. Mair, and C. Harrison. Stereoscopic display. United States Patent No. 0097319, May 2007.

U. Meier, F.J. Garca, N.C. Parr, C. Monserrat, J.A. Gil, V. Grau, M.C. Juan, and M. Alcaiz. 3D surgery trainer with force feedback in minimally invasive surgery. In *International Congress Series, Computer Assisted Radiology and Surgery*, volume 1230, pages 32–37, 2001.

C. Mello-Thoms, S.M. Dunn, C.F. Nodine, and H.L. Kundel. The perception of breast cancers: What differentiates missed from reported cancers in mammography? *Academic Radiology*, 9(9):1004–1012, September 2002.

C. Mello-Thoms, S.M. Dunn, C.F Nodine, and H.L. Kundel. The perception of breast cancers: A spatial frequency analysis of what differentiates missed from reported cancers. *IEEE Transactions on Medical Imaging*, 22(10):1297–1306, October 2003.

F. Metelli. An algebraic development of the theory of perceptual transparency. *Psychological Review*, 13(1):59–66, 1970.

F. Metelli. The perception of transparency. *Scientific American*, 230(4):90–98, 1974a.

F. Metelli. Achromatic color conditions in the perception of transparency. Perception: Essays in Honor of J.J. Gibson, Cornell University Press, 1974b.

A.A. Michelson. *Studies in Optics.* University of Chicago Press, 1927.

D.E. Mitchell. Properties of stimuli eliciting vergence eye movements and stereopsis. *Vision Research*, 10(2):145–162, February 1970.

C.N. Moller and A.R.L. Travis. Correcting interperspective aliasing in autostereoscopic displays. *IEEE Transactions on Visualization and Computer Graphics*, 11(2):228–236, 2005.

D.J. Montgomery, G.J. Woodgate, and D. Ezra. Parallax barrier for an autostereoscopic display. United Kingdom Patent No. 2352573, January 2001a.

D.J. Montgomery, G.J. Woodgate, A. Jacobs, J. Harrold, and D. Ezra. Analysis of performance of a flat panel display system convertible between 2D and autostereoscopic 3D modes. In *Stereoscopic Displays and Virtual Reality Systems VIII, Proceedings of SPIE*, volume 4297, pages 148–159, 2001b.

J.R. Moore, N.A. Dodgson, A.R.L. Travis, and S.R. Lang. Time-multiplexed color autostereoscopic display. In *SPIE Symposium on Stereoscopic Displays and Applications VII, Proceedings of SPIE*, volume 2653, pages 10–19, 1996.

L.A. Necker. Observations on some remarkable phenomena seen in Switzerland: and on an optical phenomenon which occurs when viewing a figure of a crystal or geometrical solid. *London and Edinburgh Philosophical Magazine and Journal of Science*, 3:329–337, 1832.

S.F.W. Neggers, T.R. Langerak, D.J.L.G. Schutter, R.C.W. Mandl, N.F. Ramsey, P.J.J. Lemmens, and A. Postma. A stereotactic method for image-guided transcranial magnetic stimulation validated with fMRI and motor-evoked potentials. *NeuroImage*, 21(4): 1805–1817, 2004.

K.R.K. Nielsen and T. Poggio. Vertical image registration in stereopsis. Internal Report, Massachusetts Institute of Technology, 1983.

K.R.K. Nielsen and T. Poggio. Vertical image registration in stereopsis. *Vision Research*, 24(10):1133–1140, April 1984.

H.C. Nothdurft. The role of features preattentive vision: Comparison of orientation, motion and color cues. *Vision Research*, 33(14):1937–1958, 1993.

G.E. Favalora N.S. Holliman, N.A. Dodgson and L. Pockett. Three-dimensional displays: A review and applications analysis. *IEEE Transactions on Broadcasting*, 57:362–371, 2011.

H. Nyquist. Certain topics in telegraph transmission theory. *Transactions of the American Institute of Electrical Engineers*, 47:617–644, 1928. Reprint as classic paper in Proceedings of IEEE, Volume 90, No. 2, February 2002.

R.P. O'Sea, S.G. Blackburn, and H. Ono. Contrast as a depth cue. *Vision Research*, 34 (12):1595–1604, 1994.

R.P. O'Shea. Translation of dutour (1760) [on-line]. Available: http://psy.otago.ac.nz/r_oshea/WebPubs/dutour60.html, 1999a. Web site last visited on 31/05/2010.

R.P. O'Shea. Translation of dutour (1763) [on-line]. Available: http://psy.otago.ac.nz/r_oshea/WebPubs/dutour63.html, 1999b. Web site last visited on 31/05/2010.

R. Over. Size and distance estimates of a single stimulus under different viewing conditions. *American Journal of Psychology*, 73(3):452–457, 1963.

S.E. Palmer. *Vision Science: Photons to Phenomenology.* MIT Press, 1999. ISBN 0262161834.

B. Parkin, G. Shuttleworth, M. Costen, and C. Davison. A comparison of stereoscopic and monoscopic evaluation of optic disc topography using a digital optic disc stereo camera. *The British journal of ophthalmology*, 85(11):1347–1351, November 2001.

S. Pastoor. Human factors of 3D imaging: Results of recent research at Heinrich-Hertz-Institut Berlin. In *Proceedings of the 2nd International Display Workshop*, pages 69–72, 1995.

A.P. Pentland. A new sense for depth of field. *in IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(4):523–531, July 1987.

J.D. Pfautz. Depth perception in computer graphics. Technical Report based on a PhD thesis submitted May 2000 by J.D. Pfautz to the University of Cambridge, 2002.

G.F. Poggio and T. Poggio. The analysis of stereopsis. *Annual Review of Neuroscience*, 7:379–412, March 1984.

J. Pokorny and V.C. Smith. *The visual neurosciences*, volume 2, chapter 58 - Chromatic discrimination, pages 908–923. MIT Press, November 2003.

K. Prazdny. Properties of stimuli eliciting vergence eye movements and stereopsis. *Perception and Psychophysc*, 37(1):81–83, January 1985.

D. Raghavarao and L.V. Padgett. *Block Design: Analysis, Combinatorics And Applications*. World Scientific Publishing, 2005.

S. Reichelt, R. Haussler, G. Futterer, and N. Leister. Depth cues in human visual perception and their realization in 3D displays. In B. Javidi and J.Y. Son, editors, *Three-Dimensional Imaging, Visualization, and Display 2010, Proceedings of SPIE*, volume 7690, 2010. doi: 10.1117/12.850094.

W.F. Reinhart. *Effects of Depth Cues on Depth Judgements Using a Field-Sequential Stereoscopic CRT Display*. PhD thesis, Department of Industrial Engineering and Operations Research, Virginia Polytechnic Institute & State University, 1990.

W.F. Reinhart. Depth cueing for visual search and cursor positioning. In *Stereoscopic Displays and Applications II, Proceedings of the SPIE*, volume 1457, pages 221–232, August 1991.

W.F. Reinhart, R.J. Beaton, and H.L. Snyder. Comparison of depth cues for relative depth judgements. In *Stereoscopic Displays and Applications, Proceedings of the SPIE*, volume 1256, pages 12–21, September 1990.

W. Richards. Stereopsis and stereoblindness. *Experimental Brain Research*, 10:380–388, 1970.

A.M. Rohaly and H.R. Wilson. The effects of contrast on perceived depth and depth discrimination. *Vision Research*, 39(1):9–18, 1999.

P. Rosen, Z. Pizlo, C. Hoffman, and V. Popescu. Perception of 3D spatial relations for 3D displays. In *Stereoscopic Displays and Virtual Reality Systems XI, Proceedings of SPIE*, volume 5291, pages 9–16, 2004.

H.E. Ross and C. Plug. *The mystery of the moon illusion: Exploring size perception.* Oxford University Press, 2002.

S. Roth. Ray casting for modeling solids. *Computer Graphics and Image Processing*, 18 (2):109–144, 1982.

E Rubin. *Visuell wahrgenommene Figuren - Studien in psychologischer Analyse*. Gyldendal, first edition, 1921.

J. Sands, S.W. Lawson, and D. Benyon. Do we need stereoscopic displays for 3D augmented reality target selection tasks? In *Proceedings of the Eighth International Conference on Information Visualisation (IV'04)*, pages 633–638. IEEE Computer Society, 2004.

L.V. Scharff and W.S. Geisler. Stereopsis at isoluminance in the absence of chromatic aberrations. *Journal of the Optical Society of America A*, 9(6):868–876, 1992.

C.M. Schor and T. Heckmann. Interocular differences in contrast and spatial frequency: effects on stereopsis and fusion. *Vision Research*, 29(7):837–47, 1989.

C.M. Schor and I. Wood. Disparity range for local stereopsis as a function of luminance spatial frequency. *Vision Research*, 23(12):1649–1654, 1983.

C.M. Schor, T. Heckmann, and C.W. Tyler. Binocular fusion limits are independent of contrast, luminance gradient and component phases. *Vision Research*, 29(7):821–835, 1989.

C.M. Schor, M. Edwards, and D.R. Pope. Spatial-frequency and contrast tuning of the transient-stereopsis system. *Vision Research*, 38:3057–3068, 1998.

A. Schwerdtner and H. Heidrich. Optical system for the two- and three-dimensional representation of information. United States Patent No. 5774262, June 1998a. Filed in Germany in 1993.

A. Schwerdtner and H. Heidrich. Dresden 3D display (D4D). In *Stereoscopic Displays and Virtual Reality Systems V, Proceedings of SPIE*, volume 3295, pages 203–210, 1998b.

R. Sekuler and R. Blake. *Perception*. McGraw Hill, 1994.

Sensics Inc. A unique research platform, 2008. Available from the World Wide Web: http://www.sensic.com, last accessed: 10th October 2010.

P.J.H. Seuntiëns, L.M.J. Meesters, and W.A. IJsselsteijn. Perceptual attributes of crosstalk in 3D images. *Displays*, 26(4-5):177–183, October 2005.

C.E. Shannon. Communication in the presence of noise. In *Proceedings of the Institute of Radio Engineers*, volume 37, pages 10–21, 1949. Reprint as classic paper in Proceedings of IEEE, Volume 86, No. 2, February 1998.

P. Shirley and S. Marschner. *Fundamentals of Computer Graphics*. A.K. Peters, third edition, 2005. ISBN 1-56881-269-8.

H. Siegl and A. Pinz. A mobile AR kit as a human computer interface for cognitive vision. In *Proceedings of WIAMIS'04: the 5th International Workshop on Image Analysis for Multimedia Interactive Services*, Lisbon, 2004.

H. Siegl, M. Hanheide, S. Wrede, and A. Pinz. An augmented reality human-computer interface for object localization in a cognitive vision system. *Image and Vision Computing*, 25(12), 2007.

D.R. Simmons and F.A.A. Kingdom. On the independence of chromatic and achromatic stereopsis mechanisms. *Vision Research*, 37(10):1271–1280, 1997.

D.R. Simmons and F.A.A. Kingdom. Interactions between chromatic- and luminance-contrast-sensitive stereopsis mechanisms. *Vision Research*, 42(12):1535–1545, 2002.

M. Singh and B.L. Anderson. Toward a perceptual theory of transparency. *Psychological Review*, 109(3):492–519, 2002.

W.J. Sinsteden. Über ein neues pseudokoptisches bewegungs phanomen. *Poggendorffs Annalen der Physik*, 187(10):336–339, 1860.

M. Slater. An algorithm to support 3D interaction on relatively low performance graphics systems. *Computers & Graphics*, 16(3):311–315, 1992. ISSN 0097-8493.

H.S. Smallman and S.P. McKee. A contrast ratio constraint on stereo matching. In *Proceedings of the Royal Society of London: Biological Sciences*, volume 260, pages 265–271, 1995.

T. Stein and S. Coquillart. The metric cursor. In *Proceedings of the 8th Pacific Conference on Computer Graphics and Applications*, pages 633–638. IEEE Computer Society, 2000.

K.A. Stevens and A. Brookes. Integrating stereopsis with monocular interpretations of planar surfaces. *Vision Research*, 28:371376, 1988.

G.W. Stuart, M. Edwards, and M.L. Cook. Colour inputs to random-dot stereopsis. *Perception*, 21(6):717–29, 1992.

Student. The probable error of a mean. *Biometrika*, 6(1):1–25, 1908.

A. Sullivan. 3-deep: New displays render images you can almost reach out and touch. *IEEE Spectrum*, 42:30–35, April 2005.

F.C. Sun, L. Stark, A. Nguyen, J. Wong, V. Lakshminarayanan, and E. Mueller. Changes in accommodation with age: static and dynamic. *American Journal of Optometry and Physiological Optics*, 65(6):492498, 1988.

G. Sun and N.S. Holliman. Evaluating methods for controlling depth perception in stereoscopic cinematography. In *Stereoscopic Displays and Virtual Reality Systems XX, Proceedings of SPIE*, volume 7237, 2009.

K. Sung, A. Pearce, and C. Wang. Spatial-temporal antialiasing. *IEEE Transactions on Visualization and Computer Graphics*, 8(2):144–153, 2002.

I.E. Sutherland. Sketchpad: A man-machine graphical communication system. In *Annual ACM IEEE Design Automation Conference, Proceedings of the SHARE design automation workshop*, pages 6329–6346, 1963a.

I.E. Sutherland. *Sketchpad: A Man-Machine Graphical Communication System*. PhD thesis, Lincoln Laboratory, Massachusetts Institute of Technology, 1963b.

K. Takaki, D. Arita, S. Yonemoto, and R.-I. Taniguchi. 3D direct manipulation interface by human body posture and gaze. In *Proceedings of International Conference on Human-Computer Interaction*, 2005.

R.S. Tannen, W.T. Nelson, R.S. Bolia, J.S. Warm, and W.N. Dember. Evaluating adaptive multisensory displays for target localization in a flight task. *International Journal of Aviation Psychology*, 14(3):297–312, 2004.

A.M. Truchard, I. Ohzawa, and R.D. Freeman. Contrast gain control in the visual cortex: Monocular versus binocular mechanisms. *Journal of Neuroscience*, 20(8):3017–3032, 2000.

R. Turner, E. Gobbetti, and I. Soboroff. Head-tracked stereo viewing with two-handed 3D interaction for animated character construction. In *in Computer Graphics Forum, Proceedings of EUROGRAPHICS '96*, volume 15, pages 197–206, 1996.

E.L. Ty. Investigation of surface characteristic effects on real-virtual object alignment in stereoscopic augmented reality. In *45th Human Factors and Ergonomic Society Annual Meeting, Proceedings of HFES*, 2001.

R. van Ee and C.M. Schor. Unconstrained stereoscopic matching of lines. *Vision Research*, 40(2):151–162, January 2000.

C.W. van Overveld. Application of a perspective cursor as a 3D locator device. *Computer-Aided Design*, 21(10):619–628, December 1989.

L. Vanacken, T. Grossman, and K. Coninx. Multimodal selection techniques for dense and occluded 3D virtual environments. *International Journal of Human-Computer Studies*, 67(3):237–255, 2009. Special Issue on Current Trends in 3D User Interface Research.

D. Venolia. Facile 3D direct manipulation. In *Proceedings of CHI '93: the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, pages 31–36. ACM, 1993.

G.K. von Noorden and E.C. Campos. *Binocular Vision and Ocular Motility: Theory and Management of Strabismus - Sixth Edition.* Mosby Inc., 2001.

J.J. Vos. The color stereoscopic effect. *Vision Research*, 6(1-2):105–107, 1966.

C. Ware. *Information Visualization: Perception for Design.* Morgan Kaufman Publishers, second edition, May 2004. ISBN 1 55860 819 2.

C. Ware. Dynamic stereo displays. In *CHI 95: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 310–316. ACM Press/Addison-Wesley Publishing Co., 1995.

C. Ware. *Visual Thinking for Design.* Morgan Kaufmann Series in Interactive Technologies. Morgan Kaufman Publishers, April 2008.

C. Ware and K. Lowther. Selection using a one-eyed cursor in a fish tank VR environment. *ACM Transactions on Computer-Human Interaction*, 4(4):309–322, 1997.

J.L. Weber and J. Wood. Amblyopia: Prevalence, natural history, functional effects and treatment. *Clinical and Experimental Optometry*, 88(6):365–375, 2005.

M.A. Webster, V.E. Raker, and G. Malkoc. Visual search and natural color distributions. In B.E. Rogowitz and T.N. Pappas, editors, *Human Vision and Electronic Imaging III, Proceedings SPIE*, volume 3299, pages 498–509, 1998.

S. Westland, O. Da Pos, and C. Ripamonti. Conditions for perceptual transparency. In B.E. Rogowitz and T.N. Pappas, editors, *Proceedings of SPIE, Human Vision and Electronic Imaging VII*, volume 4662, pages 315–323, 2002.

C. Wheatstone. Contribution to the physiology of vision. - Part the first. On some remarkable, and hitherto unobserved, phenomena of the binocular vision. *Philosophical Transactions of the Royal Society of London*, 128:371–394, 1838.

G.J. Woodgate, J. Harrold, A.M.S. Jacobs, R.R. Moseley, and D. Ezra. Flat panel autostereoscopic displays: characterisation and enhancement. In *Stereoscopic Displays and Virtual Reality Systems VII, Proceedings of SPIE*, volume 3957, 2000a.

G.J. Woodgate, R.R. Moseley, D. Ezra, and N.S. Holliman. Autosterescopic display. United States Patent No. 6055013, April 2000b.

A. Woods, T. Docherty, and R. Koch. Image distortions in stereoscopic video systems. In *Stereoscopic Displays and Applications IV, Proceedings of SPIE*, volume 1915, pages 36–48, 1993.

J. Wright and D.F. McAllister. Interactive design of rational bezier tensor product surfaces in stereo. In *Stereoscopic Displays and Applications IV, Proceedings of SPIE*, volume 1915, pages 82–94, September 1993.

S.P. Wu and C. Ho. Head-mounted diplay and optical engine thereof. United States Patent No. 6847489, Jauary 2005.

Y.Y. Yeh. *Stereo Computer Graphics and Other True 3D Technologies*, chapter 4 - Visual and Perceptual Issues in Stereoscopic Color Displays, pages 50–69. Princeton University Press, second edition, May 1993.

Y.Y. Yeh and L.D. Silverstein. Limits of fusion and depth judgements in stereoscopic color displays. *Human Factors*, 32(1):45–60, February 1990.

J. Yin and X. Ren. The beam cursor: A pen-based technique for enhancing target acquisition. In *In Proceding of HCI 2006*, pages 119–134. Springer, 2006.

T. Yoshimura, Y. Nakamura, and M. Sugiura. 3D direct manipulation interface: Development of the Zashiki-Warashi system. *Computers & Graphics*, 18(2):201–207, 1994. ISSN 0097-8493.

M.J. Young, M.S. Landy, and L.T. Maloney. A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision Research*, 33(18):2685–2696, May 1993.

S. Zhai, W. Buxton, and P. Milgram. The "silk cursor": investigating transparency for 3D target acquisition. In *Proceedings of ACM CHI'94 Conference on Human Factors in Computing Systems*, volume 1, pages 459–464, 1994. ISBN:0-89791-650-6.

S. Zhai, W. Buxton, and P. Milgram. The partial-occlusion effect: utilizing semitransparency in 3D human-computer interaction. *ACM Transactions on Computer-Human Interaction*, 3(3):254–284, September 1996. ISSN:1073-0516.