Bangor University

DOCTOR OF PHILOSOPHY

Vergence-accommodation conflicts and visual performance in stereoscopic 3-D imagery

Ryan, Louise

Award date:
2015

Awarding institution:
Bangor University

Vergence-accommodation conflicts and visual performance in stereoscopic 3-D imagery.

Louise Ryan, BSc, MRes

Submitted for the degree of Doctor of Philosophy

2015

School of Psychology

Bangor University

Acknowledgements

Creating a PhD thesis is not an individual experience; rather it involves several persons, whom I would like to thank sincerely.

I will be forever grateful for the opportunity that was afforded to me by Dr. Simon Watt in allowing me to undertake this PhD. Without his help, guidance, and insatiable patience, this PhD thesis would not have reached fruition. I feel lucky to have been so expertly guided by Dr. Watt, with whom I have shared the past few years of discovery.

I would like to thank my second-supervisor Dr. James Intriligator, for being along with me on this interesting and challenging journey. There are few who are more willing to invest so much time and effort into their students. I would also like to thank my committee chair, Dr. Emily Cross, who was a great source of support and guidance since the beginning.

I would also like to thank Dr. Sarah Shea for her input and advice on some of the clinical measures used in this thesis.

My appreciation also extends to my laboratory colleagues. A special thanks goes to Dr. Kevin MacKenzie, who was invaluable at the beginning stages of my PhD career and without whom, I would not be where I am today. A huge thanks goes to Ruth Dickson, her encouragement, moral and academic support as well as her willingness to help made her invaluable throughout this journey.

I am extremely grateful to the School of Psychology at Bangor University. This exceptional department have always been a great source of support and has provided me with a wealth of opportunity. I would also like to thank my participants, who sat through hours of psychophysical tasks, mostly without complaint.

I would like to take this opportunity to thank my friends and family. I would especially like to mention my parents, Colette and Jimmy, and my sisters, Berni and Lorraine. I am extremely lucky to have such an amazing family, who have always been there to advise, listen and support me. I am deeply indebted to you all.

To my friends, who have listened and helped, thank you. Thank you to everyone who has read my thesis and offered guidance and encouragement. Finally, to my partner, Peter Tomsett, who has tirelessly encouraged me, and helped me in every way possible.

ABSTRACT

Stereoscopic three-dimensional (stereo 3-D) devices continue to be popular in a multitude of applications. If safe and effective 3-D content is to be developed, a better understanding of how the visual system responds to stereoscopic media needs to be established. Technological advances have improved stereo media, but problems remain. One of the more fundamental problems with conventional stereoscopic 3-D displays is termed the 'vergence-accommodation conflict'. This refers to the inappropriate stimulus to the eye's focusing response (accommodation) at the screen surface while the image points are presented stereoscopically nearer and farther than the screen. This causes a mismatch or conflict between the stimuli to accommodation and vergence. The vergence and accommodation response systems are neurally coupled (Fincham & Walton, 1957; Martens and Ogle, 1959), which means that changes in vergence drives changes in accommodation and vice versa. Therefore a clear, stereoscopic percept requires decoupling these normally coupled responses. This PhD thesis investigates several aspects of how vergence-accommodation conflict affects perception of stereo 3-D. We first explored the effectiveness, and practicality, of a proposed solution to vergence-accommodation conflicts known as multiple-focal-planes displays (Chapter 3). Unfortunately, these displays are more suited to individual viewing, such as gaming or laparoscopic surgery in medicine, however, using such displays is infeasible for situations such as cinema and television. We therefore also investigated several aspects of tolerance of perception of stereo 3-D to vergence-accommodation conflicts in conventional displays. We measured how an individual's stereo acuity varies as a function of vergence-accommodation conflict, and whether viewing such stimuli causes an after-effect on subsequent 'real-world' viewing (without conflict; Chapter 4). We also examined several factors that might predict individual differences in perception of stereo 3-D with vergence-accommodation conflicts, most notably age-related changes in the ability to accommodate (presbyopia; Chapter 5). Our findings suggest that multiple-focal-plane displays can be used to create stereo images that are comparable to real-world images while matching the vergence- and accommodation-specified distances. For the instances where these displays cannot be used, vergence-accommodation conflict causes adverse effects on stereo performance. In chapters 4 and 5, we found that there are large individual differences in stereo performance when viewing 3-D stereo content. However, generally, as the vergence-accommodation conflict increases, so too do the negative effects on stereo performance. In chapter 5, we found that the ophthalmological factors of the audience, including degree of presbyopia and phoria can act as an indicator for individual tolerance to viewing stereo 3-D images on conventional displays. These findings should contribute both to the development and specification of future stereo 3-D displays, and to developing content for conventional displays that better match the properties of the user's visual system.

CHAPTER 1

1 GENERAL INTRODUCTION

1.1 Seeing in depth

When we view our environment, our perceptual experience is of a three dimensional (3-D) world. Depth perception refers to our ability to perceive the third dimension: the distance to an object, its depth relief, and the separation in depth of different objects. The fundamental difficulty with this is that the 3-D environment is projected onto 2-D surfaces (the retinas of the two eyes). Our visual system then needs to reconstruct the depth and distance from various ‘depth cues’, and assumptions about scene structure. In order to accomplish this, our visual system relies on a mixture of retinal cues (information within the retinal images) and extra-retinal signals (oculomotor information about eye position, for example). Binocular cues are those requiring both eyes, and signals available from one eye are referred to as monocular cues. The principal depth cues are briefly summarised below.

Binocular disparity

One of the most compelling experiences of depth comes from binocular disparity: the perception of depth from two slightly different projections of the world onto the retinas of the two eyes. Because our eyes are separated laterally, each eye receives a slightly different perspective of the same scene. The displacement of scene points on the retina of the two eyes with respect to each other is referred to as binocular disparity. Binocular disparity provides information about the relative depth of images away from the object of interest. When we fixate on an object of interest in the world, our eyes rotate in an equal and opposite direction so that the perceived images lie on corresponding points on the retina. The fixated object lies on what is called the horopter. The horopter is the surface in space wherein all images will stimulate corresponding points on the retinas of the two eyes (i.e. zero retinal disparity). Points that lie in front of or behind the horopter have non-zero binocular disparity. Figure 1.1 demonstrates binocular disparity with respect to a fixation point in the world (red dot). When an object is closer than the fixated image (blue dot), this results in crossed disparity. An image that is farther than fixation results in uncrossed disparities. The magnitude of the disparity increases with increasing depth between fixation and the object. Therefore, binocular disparities can specify both the sign and magnitude of depth separation. Binocular disparity alone is a relative cue to depth, as it compares the relative distance between points on the retinas. To recover absolute depth and distance, binocular disparity must be scaled by an absolute estimate of the distance to fixation (Garding, Porrill, Mayhew, & Frisby, 1995). The eyes' vergence posture (see below) is thought to be a principal source of this estimate (Howard, 2002; Howard & Rogers, 2002), but other distance signals can also contribute under some circumstances, including familiar size and perspective (O'Leary & Wallach, 1980) and accommodation (Watt et al., 2005).

Binocular cues can be controlled in displays by separating the images received by each eye. The principal techniques to achieve this are discussed in Section 1.2.

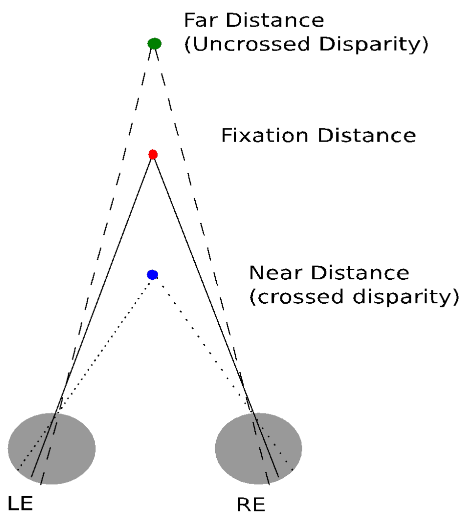


Figure 1.1 – Illustrating binocular disparity from a scene. The points at near and far distances stimulate disparate points on the retina of the left (LE) and right (RE) eyes.

Monocular retinal cues: pictorial cues and motion parallax

There are many sources of depth information in the static monocular retinal image, which generally provide a relative perception of depth and distance. It is clear that they are effective, as they are the only cues to depth and distance in paintings and photographs (2-D displays), and they are often referred to as pictorial cues. The following are some brief examples. Full discussions can be found in various textbooks (e.g. Howard, 2002; Palmer, 1999).

One of the more powerful cues is occlusion, where closer objects obscure objects that are farther away. This provides a rank ordering of proximity of objects in a scene that is a strong cue to distance and shape, but no information about the distance between objects.

Relative size refers to the comparison of the retinal sizes of objects of similar size in the world; if two objects of a similar size are viewed in the world, and one subtends a larger visual angle on the retina than the other, the object with the larger angle appears closer. This is closely related to another monocular cue, familiar size, where prior knowledge of the size of the object being viewed allows us to interpret the retinal size to recover distance. Shadows and shading can also indicate depth structure, and are especially important to the perception of the shape of objects. The projection of the texture of objects onto the retina also provides a general class of cues to 3-D surface structure, which includes linear perspective (where parallel lines converge with distance). Other pictorial cues include aerial perspective ('bluing' of the image with increasing distance due to atmospheric effects), and angle of elevation. All of these cues can of course be controlled in static 2-D images on electronic displays.

Another important static monocular depth cue in the context of this thesis is the retinal blur gradient. In natural viewing, objects nearer and farther than the distance at which the eye's optics are focused are differentially blurred in the retinal image, with greater relative distances causing increased blur. Camera optics behave similarly, and this cue is exploited in photography (where depth-of-focus is manipulated) and in painting. Recently, it has been shown that blur makes a larger contribution to depth perception than has previously been thought (Held, Cooper, & Banks, 2012; Vishwanath & Blaser, 2010). As described in Section 1.2.2, retinal blur is typically incorrect in 3-D (and 2-D) displays. Exploring ways to present correct blur cues is important and is likely to be an active area of future research (Banks, Read, Allison & Watt, 2012). This thesis does not examine this aspect of 3-D displays. It does, however, concentrate on the effects of an incorrect stimulus to accommodation.

Another powerful monocular source of depth information comes from retinal motion. In particular, motion parallax refers to the relative motion of points in the retinal image with observer motion. Geometrically, this is equivalent to the information provided by binocular vision (multiple views separated in time, as opposed to space). Akin to binocular vision, provided the magnitude of observer motion and of any eye movement is known, motion parallax can also be used to recover metric 3-D scene structure (Palmer, 1999). Motion parallax can also result in a compelling sense of depth, similar to that from binocular vision (Rogers & Graham, 1982). Note, however, that presenting motion parallax cues on a display requires real-time knowledge of the observer's viewpoint (e.g. from motion capture), and appropriate updating of the image on the display. This requires special hardware to implement.

It is important to note that the pictorial cues described above generally require some assumptions about the structure of the world (e.g. that lines in the world are straight, in the case of linear perspective) to be useful. In contrast, motion parallax and binocular vision can be used to 'measure' scene structure with few assumptions.

Oculomotor Cues

The primary oculomotor depth cues are accommodation and vergence. Accommodation is the process through which the ciliary muscles attached to the eye control the optical focus of the lens by changing its shape. In order to provide absolute depth information, the object of interest must be optimally focused on the retina (Palmer, 1999). Therefore, the visual system must be able to detect when an object is or is not in focus. The best indication of proper focus is the perception of sharp images as opposed to blurry ones. Retinal blur in the image is a cue to depth (Mon-Williams & Tresilan, 2000; Wallach & Norris, 1963). The accommodation response is therefore triggered by optical blur on the retina and aims to maximise retinal-image contrast (Alpern, 1958; Fender, 1964; Heath, 1956; Kotulak & Schor, 1986a; Manny & Banks, 1984; Owens, 1980; Raymond, Lindblad, & Leibowitz, 1984; Switkes, Bradley, & Schor, 1990; Toates, 1972). To achieve this, the ciliary muscles change the shape of the lens until the object of interest is viewed as clear and in focus, as opposed to blurry. Nearby objects are brought into focus by thickening the lens, increasing its curvature and therefore its focusing power (Bruce, Green & Georgeson, 2003). For far objects, the lens is made increasingly thin, until the object is perceived as clear and sharp.

Vergence is the equal and opposite rotational movements of the two eyes to fixate an object with both eyes. Inward movements (convergence) are made to fixate nearer objects and outward movements (divergence) are made to fixate farther objects. The principal stimulus to vergence is binocular disparity; the eyes move so that the object of interest lies on corresponding points on the two eyes retinas. The eyes' vergence angle is defined as the angle between the two lines of sight, and decreases non-linearly with increasing distance between the observer and the fixation point (Palmer, 1999). Presuming the brain knows the distance between the eyes, the sensed eye position can be used to recover the distance to fixated objects.

The combined information from the vergence and accommodation responses provide the visual system with absolute cues to distance and are therefore, essential for the ability to perceive depth (Mansson, 1998).

Vergence and accommodation distances can be expressed in terms of different metrics, two of which are meter angles or dioptres (Howard and Rogers, 2002). Meter angles and dioptres are units of stimulus distance, and are mathematically equivalent. For clarity, the unit 'dioptre' (abbreviated as D) is used throughout this thesis. Dioptres (and metre angles) are defined as the reciprocal of distance (from the interocular midpoint to the object of interest) in metres, dioptre = 1/distance or distance = 1/dioptre. Another possible metric that is used in optometry and ophthalmology is the prism dioptre (p) defined as p = IOD/distance in metres, where IOD is the interocular distance of the viewer. It is sensible to define accommodation and vergence in dioptric units for two reasons. Firstly, there is an inverse, non-linear relationship between the distance to the object and the accommodation and vergence responses. For near viewing, small changes in object distance results in large vergence angles and focusing responses. Whereas, for far viewing, small changes in object distance results in relatively small changes in vergence angle and focusing response. These response patterns are proportional in units of dioptres (Watt & MacKenzie, 2013). For example, if a near object changed distance by 30cm (from 30cm to 60cm) this relates to a dioptric change of 2.33D, whereas if a far object changed by 30cm (3m to 3.3m), this results in a dioptric change of 0.03D. Secondly, the dioptre is a direct measure of image distance unlike the prism dioptre, which depends on the interocular distance of the viewer and is therefore not a direct measure of image distance (Shibata et al., 2011).

Is binocular vision 3-D vision?

Perhaps because binocular disparities and vergence together, without any monocular cues, can be used to recover metric 3-D structure of a scene viewed by a stationary observer (Garding et al., 1995), binocular vision results in a very compelling sense of 'solid depth'. This is the case even when viewing random-dot stereograms, which contain no discernible structure in the monocular images (see Chapter 2: General Methods). This perceptual experience, referred to as stereopsis, is qualitatively distinct from that perceived in normal '2-D' images such as static photographs (Vishwanath & Hibbard, 2013). Perhaps for this reason, in common speech, as well as in the visual media industries, the terms '(stereoscopic) binocular vision' and '3-D' are often used synonymously. This is misleading because, as shown above, the brain is sensitive to many signals to 3-D structure. Indeed, an emerging idea from the last 15 years or so is that the brain combines or integrates information from *all* available sources to estimate depth as precisely as possible (Knill & Saunders, 2003; Hillis et al., 2004). In keeping with recent practice in the literature, we therefore use the term 'stereo 3-D' to refer to the techniques used to present different images to the two eyes (to manipulate binocular disparities and the stimulus to vergence), and the perception of depth that results.

1.2 Stereo Displays and Technique

Current stereo 3-D displays are used in a multitude of applications and have become more commonplace in applications like medicine (including laparoscopic surgery and imaging), scientific visualization, education, video conferencing and more. Arguably, the most common application of stereo 3-D displays is their use in entertainment (TV, cinema, and gaming).

Since their release in 2011, 3-D televisions have suffered a decline in popularity; issues with technology, lack of content and adverse visual experiences have been cited as the reasons for this decline (Easdown, 2013). This has not been experienced by the other entertainment platforms such as gaming and cinema. Over the past decade, there has been resurgence in the amount of stereo 3-D movies being created. According to the *Theatrical Marketing Statistics* report (2014), the number of 3-D movies viewed globally from 2010-2014 has, on average, remained consistently high. Consistent with this, stereo 3-D gaming is gaining in popularity with the re-introduction of the stereo 3-D, virtual-reality (VR) gaming headsets into the gaming market. Evidence of their popularity comes from the VR headset, Oculus Rift, which started as a campaign on the 'crowd funding ' website, *Kickstarter*, the aim being to raise $250,000 in order to build and develop the VR headset. The campaign closed with $2.4m raised, 974% more than its intended target; such was the popularity of the idea from gamers and developers (Webber, 2014). This suggests that, generally, interest in stereo 3-D has remained popular since its reintroduction into the entertainment market. This is despite reports of eyestrain, headaches, dizziness and impaired motor coordination when using conventional stereo 3-D displays.

Therefore, given the range of applications, the use with mass audiences (education, television, cinema, and gaming) and their popularity, it is likely that a wide range of individuals will be exposed to this technology. Consequently, it is important that researchers establish the adverse effects caused by conventional stereo 3-D displays and how to address them. There are a number of possible causes of these adverse effects (Lambooij, IJsselsteijn, Fortuin, & Heynderickx, 2009; Banks, et al., 2012; Watt & MacKenzie, 2013). Several of these causes are technical, some of which are explored briefly below. In this thesis, the more fundamental issues with conventional stereo 3-D displays are explored and researched. We start by discussing the different techniques used in conventional stereo displays.

1.2.1 Conventional Stereo Display Techniques

The aim of the 3-D display is to enhance visual experience, by creating an immersive sense of depth in the portrayed scene. Conventional stereo 3-D displays attempt to replicate the vergence and binocular disparity cues that provide a vivid sense of depth and texture by projecting two different images on the same screen in such a way that one image is viewed by the right eye and the other by the left. In principle, these images (which might be captured by multiple cameras, or computer generated) can reproduce the geometrical patterns found when viewing any real world scene.

The principle of stereoscopic presentation can be illustrated by considering a single image point in a presented scene. Figure 1.2 shows how conventional stereo 3-D displays present binocular cues to depth. In the 3-D industry, the horizontal separation between image points is referred to as screen parallax, which results in disparities at the retinas consistent with binocular disparity. When there is no screen parallax, the image is simulated at the distance of the screen. Otherwise, the separation of the images on the screen means each image point lies on different locations on the retinae of the two eyes. The idea is that the vergence system responds to this screen parallax and converges or diverges to an image either in front of, or behind the screen, as shown in Figure 1.2.

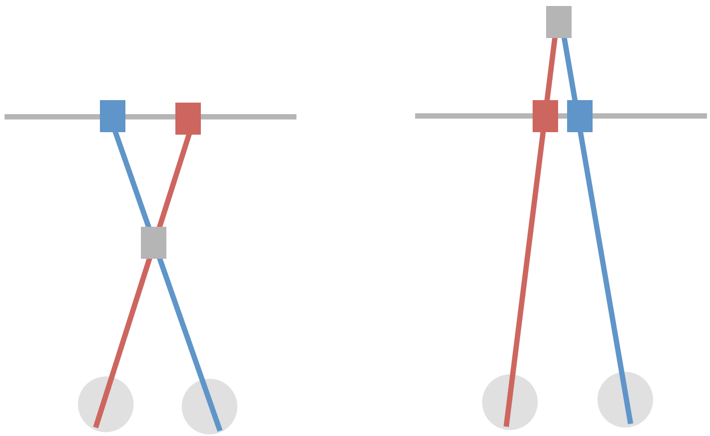


Figure 1.2. How conventional stereo 3-D displays present binocular depth cues. Both graphs show two of the same image being projected onto a single screen (grey solid line). The right and left eyes view the red and blue squares respectively. The location of the images on the screen and their corresponding points on the retina of the two eyes dictate the perceived distance to the resultant single image.

The pattern of binocular disparities at the eyes is not only a product of screen parallax but also of the techniques used to create them. There is a variety of techniques used to separate the two eye's images so that the left eye sees only its image, and vice versa:

Stereoscopes: Stereoscopes were the first method of separating the left and right eyes images. Wheatstone initially proposed mirror stereoscopes in 1838 and 1852. The mirror stereoscope allowed the targets for the left and right eyes to be presented by reflection through mirrors, each placed at a 45-degree angle to the line of sight of the viewer. The mirror for the left eye reflected the target for the left eye and vice versa for the right eye. Brewster (1848) later proposed a lens stereoscope. This stereoscope involved placing two images side-by-side in a box. They were viewed through prisms, which altered the horizontal separation and magnified the images (Howard, 2012). There are common advantages to both stereoscopes: they allow complete separation of the two eye's images, and each eye receives images with the full range of luminance, contrast, colour, temporal resolution, and spatial resolution that the display is capable of. The essence of these stereoscopes is that it allows control of the input to the two eyes separately. Therefore, researchers can isolate binocular variables and study their effects (Howard & Rogers, 2002). These displays are therefore very useful to researchers and continue to be used today. As they are a fixed-viewpoint display, allowing only one viewer to observe the images at a time, they are not practical as a display for multiple audiences and therefore for television, cinema or other applications where multiple viewpoints are required. The concept of these displays, however, have been adapted and used for some single-viewer, head-mounted display applications.

The principle techniques currently used can be categorised according to whether they use active or passive glasses. Active glasses include the use of shutter glasses; these systems are defined as active as they require a power source to present the corresponding image to each eye. This technique presents the left and right eyes images on alternate frames, the shutters or liquid-crystal lenses are synchronised with the projector such that each eye views the corresponding image only. There are a number of passive techniques, which include:

Anaglyph filters: The anaglyph filter system was the first multi-viewer stereo display technique. Here, the eyes images are presented using different portions of the visible wavelength spectrum. Left- and right-eye images are different colours (typically red and green, or red and cyan) and the viewer wears respective colour filter glasses designed to allow each eye to see its intended image while filtering out the other (Pastoor & Wopking, 1997). This has been a popular technique because it is cheap, and can be used with conventional 2-D ‘displays’ such as print media, and normal television. The main inherent disadvantages of this technique, however, are that it results in retinal rivalry (see below) from the different coloured images in the two eyes, and images cannot be presented with the full colour gamut (Sexton & Surman, 1999).

Polarisation: This technique works by projecting two images that are superimposed onto the same screen through different polarising filters (circular or linear). The viewer wears glasses, which also contain polarising filters, each filter passes only light that is similarly polarised and blocks the light polarised differently so that each eye views a different image. Circular polarization has an advantage over linear polarization such that the viewer does not have to have their head upright and aligned with the screen to view the images correctly. This is because with linear polarization, tilting the head affects the amount of light that passes through the filters and causes the non-corresponding eye to see the other eyes images more easily. A disadvantage of this system is cost, despite using relatively cheap filter glasses, silver or aluminised screens are required to reflect the polarized light. Also, in principle, two projectors are required for this technique, which would add to the costs of displaying stereo 3-D images using this technique.

Line-By-Line polarisation: This technique also uses differently polarised left and right eye images, but in a different way to that described above. On certain computer monitors and television sets, the left and right eye's images are presented on alternate rows of pixels (i.e. spatially interlaced). A modulator in front of the display acts as a polarising filter, such that all the 'odd' rows (corresponding to one eye's image) are polarised one way, and the 'even' rows another way. As above, passive polarized glasses then 'decode' the corresponding polarized light so each eye sees the intended image.

Interference filter technology (or spectral comb filtering): This system is used by *Dolby 3D Digital Cinema*, and involves using specific red, green and blue (RGB) wavelengths for one eye's image and a different RGB wavelength for the other eye's image. The viewer wears glasses that are similar to the polarization filters, blocking the wavelengths of the other eyes image so that each eye views a corresponding image. The advantage over the polarising technique is that silver screens are not required, however, the filter glasses are more costly.

Glass-free: There are also a number of glass-free displays. In these displays, two images are imposed on the same screen in narrow, alternating strips. To view a stereoscopic 3-D image, the viewer has to be positioned so that each eye views one of the images presented on the screen. Either parallax barriers or lenticular lenses are used to create this effect.

As previously mentioned, shutter glass systems use a field-sequential pattern, in which the different images to the two eyes are presented in a sequence of frames or fields. Whereas the polarization or wavelength systems present left and right eyes images simultaneously. In principle, many cinemas (with the exception of IMAX) use a combination of both techniques. They use the polarization or wavelength technologies and present the images in a field-sequential pattern, placing an active, switchable filter over a single projector. This system has its advantages, including cost reduction through using one projector and not two. It also eliminates the need to cross-calibrate luminance, colour and the spatial properties of the two projectors (Watt & MacKenzie, 2013).

1.2.2 Problems With Conventional Stereo Displays

Anecdotal evidence, media articles and manufacturers guidelines all report an array of adverse effects from viewing conventional stereo 3-D displays (Read et al., 2015). These adverse effects include symptoms such as eyestrain, headache, dizziness and more (Read & Bohr, 2014; Banks, et al., 2012; Shibata et al., 2011; Ukai & Howarth, 2008). Research has confirmed the existence of adverse effects when viewing conventional stereo media (Yang et al. 2012; Yang & Sheedy, 2011; Shibata et al., 2011; Lambooij et al., 2009; Hoffman et al., 2008). There are a number of possible causes of these adverse effects, all of which stem from an unnatural stimulus to the visual system in a number of ways (Lambooij, IJsselsteijn, Fortuin, & Heynderickx, 2009; Banks, et al., 2012). Some of these unnatural stimuli are technical, which, despite being difficult to address, are possible to eliminate. Other causes are more fundamental, and harder to solve as they are a construct of how conventional stereo displays work. They are almost impossible to eradicate using conventional display methods.

Technical issues with conventional stereo 3-D displays

Some of the technical issues associated with adverse effects when viewing conventional stereo 3-D displays are briefly discussed:

Cross Talk: Cross talk (sometimes referred to as 'ghosting') happens when the image meant for one eye is visible to some extent by the other. This phenomenon produces double contours (Kooi & Toet, 2004; Pastoor, 1995) of the same image features. As well as degrading stereoscopic depth perception, double contours can result in eyestrain and headaches for the viewer (Kooi & Toet, 2004). Cross talk is worst with anaglyph stereo 3-D, but can occur with all techniques apart from those using separate displays for each eye (mirror stereoscopes, head-mounted displays).

Low image luminance and contrast: These stereo 3-D displays present images with a reduced luminance and contrast relative to conventional '2-D' displays. There are two main reasons for this. Firstly, unlike conventional '2-D' media, often two cameras are used to record stereo 3-D content. Beam-splitters are regularly used, which allow the separation between the cameras (the inter-axial distance) to be kept small but also result in a reduction of light entering the cameras. Secondly, regardless of the technique used, luminance is lost at the projector and/or the glasses. For filter technology (polarisation and wavelength) the filters over the projector and at the glasses leads to a reduction in luminance at the eye. For shutter filter techniques, the duty cycle of each frame is a maximum of 50%, resulting in low levels of luminance and contrast.

Unnatural Timing: For displays that present content in a time-sequential manner, there is a possibility that unnatural timing between the left and right eyes may produce flickering, depth artefacts or the perception of motion blur or judder (Read & Bohr, 2014; Hoffman, Karasev, & Banks, 2011; Banks et al., 2013).

Vertical Misalignment: If the images are either recorded, or projected, with vertical misalignment between the two eye's images, small vertical vergence eye movements (looking at different elevations with the two eyes) are required to fuse the stereoscopic images, which can cause visual fatigue or eyestrain (Banks et al, 2013). If the vertical offset is too large, the images will appear diplopic (double vision), and stereoscopic depth perception will not occur.

Conflict between vestibular and visual cues to motion: When viewing stereo 3-D displays, observers can report feeling motion sick, presumably because the viewer receives powerful and convincing visual cues to motion, that conflict with vestibular signals that indicate the true state that the viewer is stationary.

The use of glasses: A recent study by Read and Bohr (2014) showed that around 8% of their participants reported adverse effects which were not due to viewing stereo 3-D displays but instead due to wearing glasses. A possible explanation for their findings are ill-fitting glasses, this can lead to issues such as cross talk as well as headaches and general discomfort. This is especially true with use of shutter glasses due to their relatively heavy and bulky nature. Another possible explanation is the reduction in luminance mentioned previously. However, it is important to note that negative preconceptions about stereo 3-D may have also lead observers to report discomfort (the "nocebo effect"; Read and Bohr, 2014).

These are just a few of the technical issues associated with current stereo 3-D displays. For a more comprehensive discussion about issues relating to stereo 3-D displays refer to: Watt and MacKenzie (2013) and Patterson (2015).

Fundamental issues: Incorrect focus cues in 3-D displays

One of the most fundamental problems with conventional stereo 3-D displays, and the issue being addressed in this thesis, is the unnatural stimulus to accommodation. As mentioned earlier, the aim of the stereo 3-D display is to create an enhanced visual experience, by creating an immersive sense of depth in the perceived scene. However, consider viewing a scene either naturally (a real scene viewed binocularly) and on a conventional stereo 3-D display. The scene on the display may have all the same geometric patterns (binocular disparity, occlusion, shading, etc.) as a natural scene. But, because the image is presented on a single (planar) surface, the light from the display is fixed, meaning there are no variations in focal distance. This provides inappropriate depth cues in two ways.

First, retinal blur doesn't change as it normally would with changes in depth in the portrayed scene (Howarth, P., 2011). Instead, blur in the retinal image is consistent with the distance to the display and not with the depth in the depicted scene. Therefore, given that blur in the retinal image is a cue to depth in the scene (Mon-Williams & Tresilan, 2000; Wallach & Norris, 1963), this depth cue is wrong, signalling a flat surface.

The second problem with incorrect focus cues in stereo 3-D displays is the unnatural conflict between the stimulus to vergence and accommodation. While conventional stereo 3-D displays allow the stimulus to vergence to vary according to the depth in the depicted scene, the cue to accommodation remains at the location of the screen. This is contrary to what happens in the natural world where these two systems co-vary (see below). This results in a mismatch between the cue to vergence and accommodation, which is popularly referred to as the 'vergence-accommodation conflict' (herein conflict).

In this thesis, the adverse effects of conflict are investigated for two reasons. Firstly, they are present in every conventional stereo 3-D display. Secondly, they are very difficult to eliminate; proposed solutions are promising, but are not suitable for multiple viewing applications (Banks et al., 2013).

1.2.3 Vergence-Accommodation conflict

In conventional stereo 3-D displays, in order to converge and accommodate accurately to image points at distances other than the screen, the viewer must accommodate at the distance of the screen while converging at another (Watt & MacKenzie, 2013). This can be a problem as the accommodation and vergence systems are neurally coupled, acting synergistically to make responses more efficient in the real world (Martens & Ogle, 1959; Mays & Gamlin, 2000; Yang & Sheedy, 2011; Heron, Charman & Schor, 2001; Fincham & Walton, 1957; Schor, 1992). As such, when one system responds to a stimulus, the other system also changes to a corresponding degree that is never fully accurate (Tait, 1951). These neural links allow for the binocular images to be sharply focused and lie on the fovea of both eyes over a range of object distances (Heron et al., 2001). Decoupling these responses can be effortful and not always possible.

However, both systems can tolerate a certain amount of retinal defocus and disparity without adjusting their responses - the viewer is still able to perceive a sharp and single image. Although it is best to have an object positioned optically conjugate with the retina, an image that is slightly defocused will be perceived as clear as long as it is positioned within the depth-of-focus of the eye (Wang & Ciuffreda, 2006), which has been estimated to be ~0.25 to 0.3 D. Objects in the environment do not have to be located on the horopter to be fused perceptually into a single image. The area around the horopter within which disparate images are fused is called Panum's fusional area (approximately 0.25-0.5 degrees). Therefore, both systems can respond to different distances and still perceive a clear and single percept of a scene. However, while these systems can tolerate some difference in response, their neural coupling means that large enough conflict between the responses can lead to the risk of a blurry retinal image, diplopic images or both (Banks et al., 2013; Shibata, Kim, Hoffman, & Banks, 2011).

Much of the research conducted into our ability to decouple the vergence and accommodation response comes from ophthalmology and optometry (Shibata et al., 2011; Kim, Kane & Banks, 2013). Although applicable as it relates to conflicting cues to distance between the two responses, these measures relate to fixed changes using prisms and lenses over a long period of time. Stereo 3-D displays, however, are constantly varying the stimulus to vergence over short periods of time. Nevertheless, the work already conducted can provide some insight into our tolerance to decoupling these responses. Fry (1937) examined the 'zone of clear single binocular vision' (ZCSBV), which is the set of vergence and focal stimuli that can be perceived clearly and as a single image. For a particular focal distance, if you increase the stimulus to vergence, an observer would increase their vergence response without adjusting their focal response, up to a point, after which they would adjust their accommodation response and report blur in the image (Fry 1937; see Grosvenor, 2007). Fry's estimated ZCSBV provides the range of conflicts we can perceive as clear and single, but ignores whether the action of decoupling at those distances is comfortable. Percival (1892), however, proposed the range of vergence and accommodation responses that resulted in comfortable viewing for optical corrections. Figure 1.3 illustrates the measured ZCSBV and Percival's zone of comfort. This shows that although there is a range of possible vergence and accommodation tolerance limits, comfortable viewing lies within that range. Thus indicating that although it is possible converge at one distance and accommodate to another, the effort involved might not be comfortable. However, as stated these measures are for optical correction, therefore, it might be easier for the observer to adapt to the fixed conflict induced by optical correction as opposed to the varying conflict induced by stereo 3-D displays (Shibata et al., 2011).

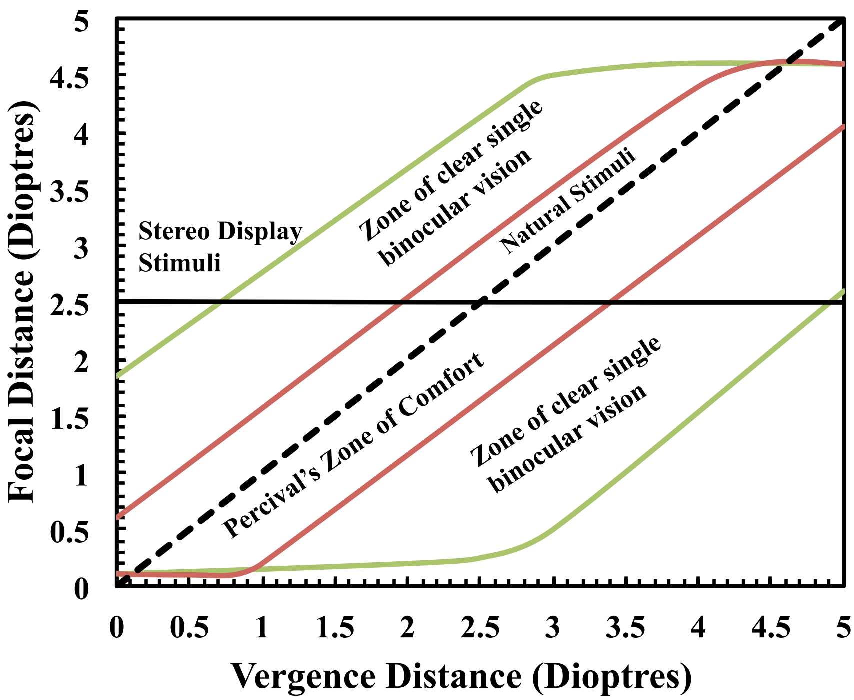


Figure 1.3 – ZCSBV and Percival's zone of comfort. The zones are outlined in green (ZCSBV) and red (Percival's zone of comfort). The dashed diagonal line is the natural stimulus. The horizontal line is the cue to vergence and accommodation with stereo 3-D displays. The x and y axes are the stimulus to vergence and accommodation respectively in diopters. This figure has been adapted from Banks, Akeley, Hoffman and Girshick (2008).

1.3 Current Research

The aim of this thesis is to address the issues associated with conflicts such as visual discomfort and fatigue (Shibata et al., 2011; Lambooij et al., 2009; Hoffman et al., 2008; Ukai and Howarth, 2008; Ukai, 2007; Hakkinen, Polonen, Takatalo & Nyman, 2006; Emoto, Nida and Okana, 2005; Menozzi, 2000; Wann & Mon-Williams, 2002; Peli, E. 1998; Howarth & Costello, 1997) along with distortions in stereo performance (Watt, Akeley, Girshick & Banks, 2005; Hoffman et al., 2008). There are two approaches to doing this. The first is to attempt to eliminate the conflict, by investigating possible display solutions to the conflict that match the stimulus to vergence and accommodation distances. Where using these displays is not feasible, the second option is to better understand the causes and effects of conflicts, so that content can be optimised to provide an enhanced stereo 3-D experience, while minimising adverse visual effects.

The following sections involve discussions about the possible display solutions to the conflict, the display used in this thesis (section 1.4), the techniques used relating to this display and the previous research examining the efficacy of using a fixed-viewpoint volumetric display. This is followed by a discussion of the previous research examining the effects of conflict on visual discomfort and the perception of depth from stereo images (section 1.5).

1.4 Volumetric Displays

Many displays exist that offer the possibility to present stereoscopic images with correct or near correct focus cues. In particular, there is a class of displays called ‘volumetric displays’, which are defined as “a transparent physical volume (image space) in which static and animated image components may be placed” (Blundell, 2012). Such images inherently produce a varying stimulus to accommodation because image points are displayed at different locations in a physical volume. Such displays can, if appropriately configured, present 3D effects with matched cues to vergence and accommodation, eliminating the conflict.

There are two types of volumetric display: autostereoscopic and fixed-viewpoint, both of which are discussed. First, the use of autostereoscopic volumetric displays is explored. Unfortunately, they have a number of fundamental disadvantages, making them unsuitable as a general solution (see below). The second part of this section explores the efficacy of fixed-viewpoint volumetric displays. The conceptual background to this is also discussed.

It should be noted that there are some software techniques which aim to reduce the effects of conflict by creating an accurate depth-of-field in the images according to where the viewer is fixated in the scene (Duchowski, et al., 2014; Vinnikov & Allison, 2014; Maiello, Chessa, Solari & Bex, 2014). These techniques work by using eye-tracking cameras, which inform where the viewer is fixated, the software then blurs the images according to the distance of the image points from the fixation point in the scene. The advantage of this technique is a possible reduction in discomfort when viewing stereo 3-D content (Duchowski et al., 2014) and improved stereoscopic vision (Maiello et al., 2014). However, despite reducing discomfort, these are not solutions to the conflict for a number of reasons. Fundamentally, they do not address the inaccurate focus cues; the stimulus to accommodation is still at the surface of the screen while the stimulus to vergence is stimulated to distances away from the screen. Also, research has found a general dislike for the resultant images created by this technique, most likely due to the temporal lag in real-time data transmission between the gaze-contingent and software components (Vinnikov and Allison, 2014; Wann, Rushton & Mon-Williams, 1995). They are therefore not considered here as a possible solution to the conflict.

Autostereoscopic Volumetric Displays

Multiple types of autostereoscopic displays have been produced such as; Holographic, swept-volume, and static-volume displays. In all cases, the concept is the same, image points are presented within a 3D volume, which provide matched cues to vergence and accommodation for the range of distances within the display (Akeley et al., 2004).

Holographic displays operate by displaying a recorded or computer generated inference pattern of light, reproducing the properties of light waves in terms of their luminance, wavelength and phase differences with great accuracy, allowing them to generate a realistic 3-D scene (Pastoor and Wopking, 1997). When suitably lit, the inference pattern of light created by the hologram either diffracts or phase modulates the light into a reproduction of the original light field with the cues to depth and distance maintained, these change realistically with any change in the position of the viewer. Static holographic displays, where the images are not updated in real-time, are currently used in medical imaging, military imaging, industrial design and marketing (Tay, 2007).

Swept-volume displays match the cues to vergence and accommodation by using a rapidly moving display element, such as a screen that moves back and forth in distance, or a projection screen that spins on a central axis. The display of information on the screen is synchronised with the position of the display element such that an image is built up in a time-multiplexed fashion (Favalora et al., 2002; Schowengerdt & Seibel, 2006). Because each image point is positioned at an appropriate location in physical space, the location of the stimulus to accommodation is inherently correct.

Static-volume displays use multiple display planes to 'fill up' a volume. For example, the DepthCube (LightSpace Technologies; Sullivan, 2004) uses liquid-crystal scattering shutters, only one of which is active at any given time, acting as a scattering rear-projection screen. The 3D image is continuously projected to each screen in turn, like a 'sweep' through the stack of screens, again creating a 3D image in appropriate locations in physical space.

The principal advantages of autostereoscopic displays are that multiple viewers can view them from multiple angles, without the need for stereo glasses, and that conflicts are reduced significantly. Unfortunately, however, there are numerous noteworthy disadvantages. Practical problems such as the cost and complexity of building highly specific display hardware (and software) can potentially be overcome. There are, however, more challenging problems with these approaches that limit their potential as general display solutions. Perhaps most obviously, the range of accommodation (and vergence) distances that can be presented is limited to the physical size of the display. It is implausible that large scenes could be displayed in this way (with the recent exception of holographic displays; Tay, 2007). Also, these displays must have equal spatial resolution in all dimensions (x, y and z), so that image quality and the resolution of 3-D information (including the stimulus to accommodation) is the same wherever the display is viewed from. This number of voxels (a unit of graphic information that defines a point in 3-D space) for which image information must be computed is therefore enormous. More fundamentally, in the static volume and fixed-volume displays, because voxels emit the same light in all directions, these displays are incapable of correctly representing view-dependent lighting, such as occlusions and reflections (Akeley et al., 2004), which are a fundamental part of compelling and accurate 3-D graphics. For holographic displays, several issues including limited capabilities of the recording materials and the computation required to update the images have hindered their use.

Fixed-viewpoint volumetric displays

A related but more promising approach is to use volumetric displays in which the viewing position is fixed with respect to the display. This precludes potentially desirable features such as multi-viewer and multi-viewpoint capability. This approach does, however, have two fundamental advantages. First, because the position of the viewer is known, view-dependent lighting effects can be correctly computed (and presented) for each eye's single line of sight. More importantly for our purposes, fixing the viewpoint eliminates the need to have equal spatial resolution in all dimensions. The resolution of the visual system to focal depth is orders of magnitude poorer than its spatial resolution (Wandell, 1995; Campbell, 1957; Charman & Whitefoot, 1977). It is theoretically possible therefore to achieve the necessary spatial resolution (to present "geometrical" depth cues accurately, including binocular disparity) and focal-depth resolution with a display consisting of a relatively small number of sparsely spaced "focal planes" (Rolland et al., 1999; Akeley et al., 2004). That is, voxels are very non-uniformly distributed, dramatically reducing the computational overhead, and complexity of the display over the autostereoscopic approaches described above. Note that it is also possible, with appropriate optics, to present distances up to optical infinity. The various implementations and the implications of this approach, as well as some important emergent questions, are explored below.

There are several versions of fixed-viewpoint volumetric displays. All approaches have the same conceptual idea, to present stereo 3-D content with matched cues to vergence and accommodation distance, but vary in their implementation. One of the more straightforward approaches, and the one used here, is to use multiple 'screens' (referred to as a multiple-focal-plane display; Akeley, et al., 2004). This method is briefly described (for a full description see the General Methods chapter), followed by a description of the other proposed methods:

<u>Multiple-focal-plane displays</u>: Multiple-focal-planes displays (multi-plane displays) are a fixed-viewpoint display in which the eye sees the sum of images presented at various focal distances through the use of multiple screens (or focal-planes). These focal-planes are comprised of a series of beam-splitters and mirrors located at different focal distances and

viewed simultaneously. This effectively creates a stack of transparent screens, at different focal distances. In this display, there are two of the same arrangement, each viewed by a separate eye ensuring stereoscopic viewing. An example of the arrangement for one eye in our display is shown in Figure 1.4.

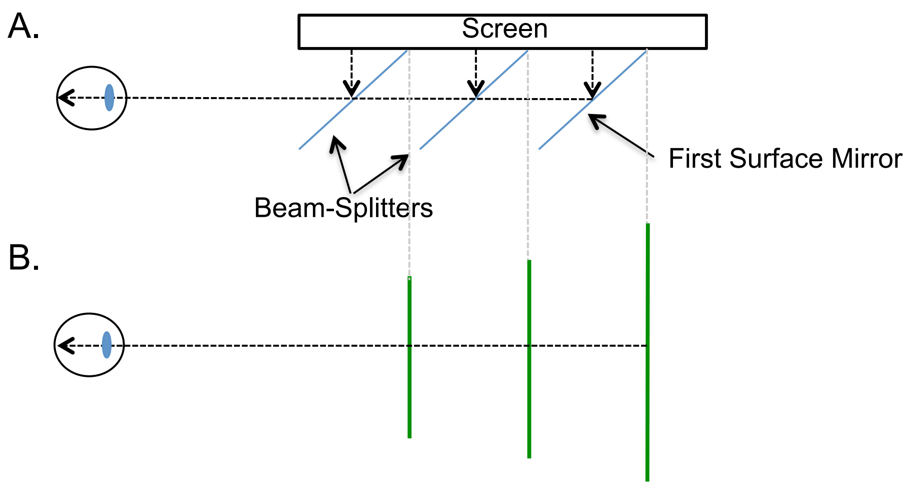


Figure 1.4 - Multiple-focal-planes display. This is a side view of one section in our display. (A) Shows the monitor and the beam-splitters/mirror arrangement. The black dashed lines demonstrate the optical path for each focal plane. (B) Is an illustration of the resultant transparent focal planes; the grey dashed lines demonstrate the distance for each focal-plane in the display.

<u>Telecentric Optical System</u>: In this display, images are viewed through polarizing filters, which separate the left and right eyes images, positioned in front of a telecentric optical system. The stimulus to accommodation is altered by controlling the position of the LCD screen, through mechanical movement, with respect to a lens system in a front-back direction. The position of the screen is varied as per the desired vergence distance in the scene. (Shibata et al., 2005).

<u>Moveable relay lens system</u>: The viewer's vergence change to images in the presented scene is detected through a gaze-tracker. The position of relay lenses within the display is then adjusted so that the screen surface appears at the same distance as the convergent point. (Shiwa, Omura & Kishino, 1996).

<u>Deformable Mirror Display</u>: In this display, the scene is presented in the form of a modulated laser beam that is reflected from a deformable membrane and raster scanned (the beam sweeps horizontally left-to-right at a steady rate, creating a full image). The scan is converged at the entrance to the pupil of the viewer's eye, creating a Maxwellian view of the displayed

image. As the beam is scanned, the deformable membrane mirror dynamically changes the angle of the beam, to present images at different focal distances that align with the vergence distances in the image. (Schowengerdt & Seibel, 2006).

Adaptive lens approach: There are multiple versions of this display method, reported by Suyama, Date, and Takada (2000), Liu and Hua (2009), and Love et al., (2009). Each of these display types use a stationary, switchable lens, placed in front of the eye, which is synchronised to the graphics of the display (Love et al., 2009). This switchable lens creates a series of focal distances in a time-multiplexed manner; presenting an image at one focal distance, and then presenting the same image at the next and so on. Thus, these displays create a temporary multiplexed image with matched cues to vergence and focal distance. Both displays by Suyama et al., (2000) and Liu and Hua (2009) use variable focal length lenses, however, these displays are limited by the response time of the lenses used. Therefore, I will briefly describe the display by Love et al. (2009). The display uses two CRT monitors, arranged in a haploscope configuration (one monitor for each eye); the switchable lenses are placed in front of each eye. The lenses are comprised of calcite material with a birefringent refracting element. Birefringent lenses have two indices of refraction, one for light polarised along the crystalline axis and the other for light polarised along the orthogonal axis, depending on the polarisation of the incident light (Love et al., 2009). Thus, a single lens has two focal lengths, which is selected with a polarisation modulator. In their display, they use two lenses in front of each eye, effectively creating four transparent screens at various focal distances. For a full description, refer to Love et al., (2009) and Shibata et al., (2011).

In theory, these displays are capable of presenting correct or near correct focus and vergence cues. However, some require mechanical movement, and all vary in physical size. It is currently not possible to reduce some of these displays into a convenient and portable device, with the exception of the adaptive lens approach. Unfortunately then, these displays may not be suitable as a replacement for conventional stereo displays in their current form. However, some can be adapted to be used in fixed-viewpoint applications, such as gaming and medicine. Importantly, as the vergence and accommodation distances can be manipulated independently while holding all other stimuli properties constant, they are very effective in visual research and understanding the effects of conflict on stereo 3-D viewing (Akeley et al., 2004; Hoffman et al., 2008; Shibata et al., 2011; Kim, et al., 2014).

1.4.1 Considerations with Volumetric Displays

As with all fixed-viewpoint volumetric displays, the multi-plane display allows us to independently manipulate the spatial and focal resolution; the display monitor (Figure 1.4) is a conventional display with an appropriate spatial resolution, which is projected onto the focal planes (herein planes), the spacing between the planes determines the focal-depth resolution. However, a key design parameter of these displays is the spacing between (and therefore the number of) planes. As previously discussed, the resolution of the visual system to focal depth is orders of magnitude poorer than its spatial resolution (Wandell, 1995; Campbell, 1957; Charman & Whitefoot, 1977). If we want to present correct and continuous focal cues, the planes only need to be positioned within the depth of focus of the eye (~0.25 to 0.3 D), which is relatively large (Rolland, Kryeger and Goon, 1999).

Rolland et al. calculated that for a display with a range of 50 cm to optical infinity (around 10 m), 14 planes would be required to have planes within the depth-of-focus of the eye at all distances within the display. However, as beam-splitters (used in multi-plane displays) allow for planes that are both reflective and can allow images projected behind them to be viewed overall luminance is lost with only a portion reflected from the plane. Therefore, it can be deduced that the more planes used in the display, the lower the overall luminance (MacKenzie et al., 2010; 2012). Luminance is also lost in variable optic displays with increased planes. Each plane is displayed sequentially for a portion of the time (Love et al., 2009). Therefore, doubling the planes used effectively halves the duty cycle of individual planes, reducing their perceived luminance. The refresh rate needs to be high enough to present images at each plane without flicker. Finally, the more planes in the display, the larger the computational requirements needed to render images on the display.

Therefore, it is optimal to have as few planes as possible in these displays. However, decreasing the number of planes in the display necessarily increases the spacing between them. On initial analysis, this approach therefore appears unsuited to presenting a continuous range of accommodation distances (required to match changes in vergence distance) because it can only provide a very coarse approximation to the continuous variations in focal distance for natural scenes.

1.4.2 Depth Filtering

Given that we cannot position planes at every focal distance in a given scene, it is likely that most points in a scene will not correspond with the location of the planes. Consequently, we need a solution in which images can be presented between the planes. One solution is to assign image intensities to each focal plane. The simplest rule would be to assign all the luminance of an image point to the closest focal plane; this is referred to as 'box-filtering' (Akeley at al., 2004). However, for continuous surfaces in depth, like a slanted image (Figure 1.5, left panel) if this assignment rule is applied, half the luminance will be presented on the near plane and the other on the far plane. For cases where there is a separation of more than the depth-of-focus of the eye between the planes, the accommodation system cannot focus on both planes at once. It has been shown that when viewing images using this rule, there are visible artefacts at the transition between where the image is presented on one plane and where is it presented on the other (Akeley et al., 2004).

To address this problem, Akeley et al. propose a similar but inherently different technique, a depth-weighted blending technique referred to as depth filtering (Figure 1.5, right panel). With this rule, for each image-point, luminance is assigned according to the relative distance (in dioptres) of the image point from each of the nearest planes, computed along the line of sight (see Figure 1.5; Akeley et al., 2004). When the image intensities are summed and combined by the visual system, there are no visual discontinuities using this technique (MacKenzie et al., 2010).

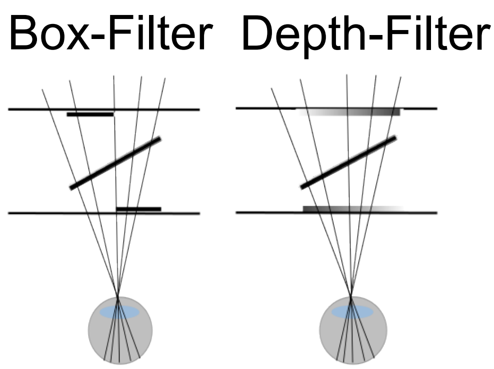


Figure 1.5 – Rules for assigning image intensity in multi-plane displays. Both images show an illustration of an object slanting away from the viewer, positioned between two planes. Luminance intensity for each point of the image is distributed across the planes according to the depth-filtering rule being applied. The bars in front of the planes represent the intensity of luminance projected onto the planes. Darker colours indicate more luminance. The box-filter (left image) does not vary the intensity of the luminance at either plane. The tent filter (right image) varies luminance intensity as a function of image-point distance from the plane.

Do depth-filtered images present a continuous stimulus to accommodation?

As well as eliminating visual artefacts, it has been suggested that the depth-filtering approach might be able to produce continuous variations in the stimulus to accommodation from relatively few, discrete focal planes (Hoffman et al., 2008; MacKenzie, Hoffman & Watt, 2010). The accommodation system aims to optimise retinal image contrast (Alpern, 1958; Fender, 1964; Heath, 1956; Kotulak & Schor, 1986a; Manny & Banks, 1984; Owens, 1980; Raymond, Lindblad, & Leibowitz, 1984; Switkes, Bradley, & Schor, 1990; Toates, 1972). Therefore, whether depth-filtered images act as a stimulus to accommodation can be determined by examining the retinal-image contrast that results from viewing depth-filtered images.

MacKenzie et al., (2010) used modelling of retinal-image contrast to compute the contrast ratio (between the retinal-image contrast and the incident contrast) for various focal-plane separations and spatial frequencies. They computed the contrast ratio based on the aberrations, pupil size and accommodative state of a real-eye's optics (MacKenzie et al., 2010). Figure 1.6 shows the results of their computations for the situation of viewing a 'real-world' stimulus (100% of the luminance is on the plane; dashed lines) and a depth-filtered simulation (solid lines) at the same focal distance. The contrast ratio was computed for a range of focal-plane separations (0.6D and 1.1D; top and bottom panels respectively) and spatial frequencies (2 – 16 cycles per degree; cpd) as a function of where the eye is focused. Examining the graphs, they show that peak retinal-image contrast for 'real-world' images occurred at the location of the image (denoted by '0' on the x-axis). For plane separations of 0.6 D (top panel) peak retinal-image contrast for depth-filtered images occurred at the distance of the simulated image for spatial frequencies of up to 8 cpd. This suggests that depth-filtered images can provide an accurate stimulus to accommodation. However, when you examine larger plane separations (1.1D, bottom panel) there is no peak retinal-image contrast at the location of the simulated image for spatial frequencies of 8cpd, rather peak contrast occurs at one or other focal plane.

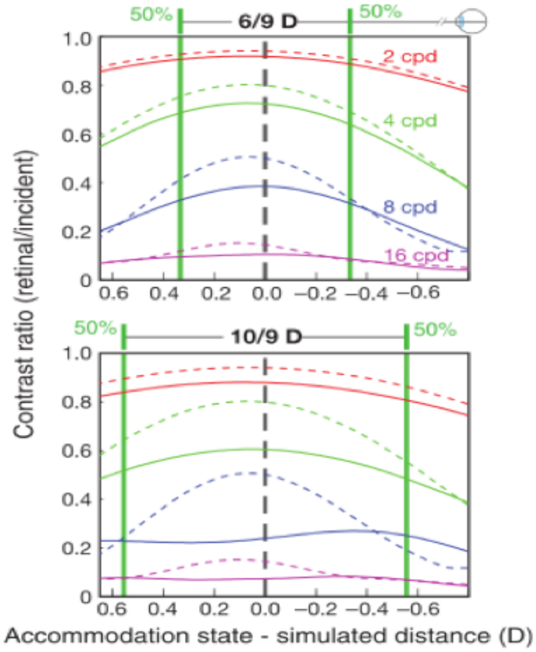


Figure 1.6 – Retinal-image contrast functions for real targets and depth-filtered images with differing levels of defocus. The x-axis is the accommodative state relative to the simulated focal distance in diopters and the y-axis is the retinal-image contrast ratio. The dashed lines represent real-world viewing and the solid lines represent depth-filtered viewing. The vertical green lines are the near and far planes, the grey dashed line is the stimulus location. The different colours show the results for the various spatial frequencies. The titles over the graphs show the focal-plane separation. This graph is adapted from MacKenzie et al (2010).

MacKenzie et al. also measured the accuracy and reliability of individuals' monocular accommodation response to depth-filtered images. They measured this as a function of plane separation to investigate how the stimulus to accommodation in depth-filtered images matched the stimulus to vergence with increasing distances between the light sources. Their findings suggest that depth-filtered images result in continuous, accurate variations in accommodation response for focal-plane separations up to and including 1.1 D. However, this is not a true representation of the viewing conditions of these displays, which are normally viewed binocularly. It is important to examine this as the vergence and accommodation systems are neurally coupled, which means that a change in vergence will necessitate a change in accommodation (Tait, 1951). It is possible then that coupled with an accurate vergence response, the accommodation response would be as or even more tolerant to large plane separations.

MacKenzie, Dickson and Watt (2012) measured the accuracy and reliability of individuals' binocular accommodation response to depth-filtered images. Surprisingly, they found that the depth-filtered simulations resulted in similar steady-state vergence and accommodation responses as 'real-world' targets for plane separations of up-to and including

0.9 D. Hence, the accommodation response was accurate for a slightly smaller, yet still fairly large plane separation.

Taken together, these findings suggest that a display with a range of 33 cm to around optical infinity (0.0 D) would require only four focal planes, with separations of up to 0.9 D. With these plane separations, the depth-filtered simulations can act as an accurate stimulus to accommodation, such that the focal response is driven to the distances of the simulated images. However, Figure 1.6 shows that the contrast ratios for the depth-filtered images were consistently lower than the ‘real-world’ equivalents. The reduced retinal-image contrast might suggest an attenuation of fine detail in the resultant image. This may affect the appearance of the resultant images regardless of plane separation magnitude and accommodation response. Therefore, although the research demonstrates that focal-plane separations of up to 0.9 D results in accurate accommodation response, a viewer’s ability to identify fine detail in depth-filtered images may differ from ‘real-world’ equivalents. In this thesis, the extent of this possible difference and whether this has an effect on the perception of detail and depth from depth-filtered images is examined.

1.5 Adverse Effects of Viewing Vergence-Accommodation Conflict

Even if fixed-viewpoint volumetric displays could produce images that continuously vary the stimulus to vergence and accommodation at all distances within the display, similar to natural viewing, they may not be suitable as a replacement for conventional stereo displays in their current form. As already stated, our next option is to better understand the effects of viewing conflict as well as the underlying causes of these effects.

Researchers and manufacturers have long speculated about conflicts causing adverse viewing effects. A lot of the early work investigating the effects of viewing stereo 3-D was conducted using head-mounted displays (HMDs: Mon-Williams, Wann, & Rushton, 1993; Costello, 1997; Mon-Williams & Pascall, 1995). In their study, Mon-Williams, Wann, and Rushton (1993) reported physiological symptoms in a number of subjects following immersion in a HMD. They found that viewing a stereo 3-D environment caused binocular stress and, amongst other possible causes, isolated the demand placed on the vergence and accommodation systems as the greatest source of this stress. Since then, studies have been conducted investigating the potential adverse effects of conflict on visual discomfort and the perception of depth from stereo images. The majority of this research has examined the effects of conflict on visual discomfort and fatigue (Shibata et al., 2011; Lambooij et al., 2009; Hoffman et al., 2008; Ukai and Howarth, 2008; Ukai, 2007; Hakkinen, Polonen, Takatalo & Nyman, 2006; Emoto, Nida and Okana, 2005; Menozzi, 2000; Wann & Mon-Williams, 2002; Peli, E. 1998; Howarth & Costello, 1997). Less research has been directed into the other aspect potentially affected by conflicts – the effect on the perception of depth and fine detail from stereo images (stereo performance). If comprehensive guidelines are to be produced, both aspects need to be comprehensively understood (Banks, et al., 2012; Watt & MacKenzie, 2013). In this section, the research into the effects of conflict on discomfort and fatigue are discussed along with what is known about effects on stereo performance, which is the main focus of this thesis.

Effects of conflict on visual discomfort and fatigue:

The initial evidence in support of a link between conflict and visual discomfort failed to prove that conflicts *per se* caused these adverse effects for two reasons. Firstly, the studies compared the effects of viewing stereoscopic and non-stereoscopic images using conventional stereo 3-D displays and '2-D' displays (Emoto et al., 2005; Hakkinen et al., 2006; Jin, Zhang, Wang, & Plocher, 2007; Yamazaki, Kamijo, & Fukuzumi, 1990; Yano, Ide, Mitsuhashi & Thwaites, 2002). For example, Emoto et al. (2005) measured the effects of v-a conflict using prisms on a conventional 2D display. However, prisms change the whole visual field, and not just the binocular disparity between certain images, as is changed with conventional S3D displays (Lambooij et al., 2009). Therefore, it was not possible to isolate the effect of conflict from the other technical issues that cause adverse effects (mentioned previously; cross talk, vertical misalignment, etc). Secondly, and most importantly, comparing the effect of viewing 3-D displays and '2-D' displays means that there was no control condition in which the observers viewed the same stereo imagery under the same conditions with and without a conflict. These kinds of experiments were common amongst experiments on conflict using HMDs (such as Peli, E., 1998; Wann et al., 1993; Costello, 1997; Mon-Williams & Pascall, 1995). This means that other factors such as changes in vergence demand and not the conflict between the vergence and accommodation responses may have contributed to the findings.

Research by Hoffman et al. (2008) provided the first demonstration that conflict *per se* can cause fatigue and discomfort by isolating the effect of the incorrect stimulus to accommodation. To do this, they used a multiple-focal-planes display, which allowed the vergence and accommodation distances to be manipulated independently while holding all other stimuli properties constant (Akeley, Watt, Girshick & Banks, 2004; see above and the General Methods section). By so doing, they could directly compare the effects of viewing stimuli in which accommodation and vergence distance varied together (simulating natural viewing) with effects of viewing the same stimuli, but with a fixed stimulus to accommodation (simulating conventional stereo 3-D viewing). Their results show that conflict between the vergence and accommodation responses leads to an increase in visual discomfort and fatigue relative to viewing stereo images with no conflict.

Shibata et al., (2011), using an adaptive lens display, expanded on this research by examining the effect of varying accommodation distance and sign of the conflict (stimulus to vergence was either nearer or farther than the screen) on visual discomfort and fatigue. By doing so, they were able to examine the range of conflict that results in comfortable viewing as a function of accommodation distance – termed the *zone of comfort* (ZoC; Shibata et al., 2011).

Effects of conflict on Stereo performance:

The research into the effects of viewing conflict on visual discomfort and fatigue is a valuable start in order to understand how conflict causes adverse effects. However, it falls short if comprehensive guidelines for the creation of content that is both comfortable and easy to view (for most individuals) are to be created. This is because less is understood about how stereo performance is affected by conflict. The research that has been conducted has shown a degraded perception of depth, leading to increased fusion time (Watt, et al., 2005; Hoffman et al., 2008), and a reduced ability to discriminate fine detail in stereoscopic depth (stereo resolution; Hoffman et al., 2008). Therefore, this research only demonstrates there is a problem. However, it does not provide a comprehensive description of the relationship between conflict and stereo performance.

Conclusions

The work described above has shown that conflict negatively affects measured levels of visual fatigue and discomfort (Hoffman et al., 2008; Shibata et al., 2011). It also leads to reductions in depth perception (Hoffman et al., 2008; Watt et al., 2005; Wann et al., 1995). Although research has been conducted into perceptual issues with conflict, these experiments just prove there is a problem. Little is known about the relationship of stereo performance with changes in conflict magnitude and focal distance. Even less is known about how individual differences (such as age and degree of presbyopia) affect a person's' stereoacuity with conflict. The estimated 'zone of comfort' by Shibata et al. (2011), goes someway to outlining the limits of conflict tolerance. However, if comprehensive guidelines are to be produced, more specific knowledge regarding stereo performance both during and after viewing needs to be established, as well as the factors that affect it.

CHAPTER 2

2 GENERAL METHODS

In this work we planned to measure a possible solution to the conflict as well as measure the effect of conflict on stereo performance under various viewing conditions. This chapter describes the general aspects of the methods used to achieve these aims. Specifically, Section 2.1 discusses the method used in most of the empirical studies to present stereoscopic stimuli, and manipulate the stimulus to accommodation. This section also describes some general principles of its use (such as calibration). Section 2.2 discusses the stimuli and techniques used to characterise stereo performance.

2.1 Multi-Plane Display

We used a multi-planes display to manipulate the stimulus to accommodation in stereoscopic images. The principle of the display was based on that of Akeley et al. (2004), in their novel display, static mirrors and beam-splitters were used to create multiple focal planes. However, this display is different to the display used here in a number of ways (for a full description see Akeley et al., 2004).

A schematic of our multi-plane display is shown in Figure 2.1.1. Stereoscopic presentation is achieved by each eye viewing its own multi-plane display. Within each, images were displayed on 30” Samsung 305T TFT monitors (2560 x 1600 pixel resolution, 0.25mm dot pitch) projecting onto two beam-splitters and a first-surface mirror. These give rise to focal planes at three different distances. The mirrors and beam-splitters are mounted on optical rails, allowing the position of, and separation between, focal planes to be adjusted continuously. This allows for independent control over the vergence and focal distances. Using this display, therefore, we can measure the effects of viewing a conventional stereo display against a control where the same vergence distances are matched in focal distance.

The display is configured in a haploscope arrangement, with the left eye viewing its display via a first-surface mirror near the eye, and the right eye viewing its display directly (Figure 2.1.1 B).

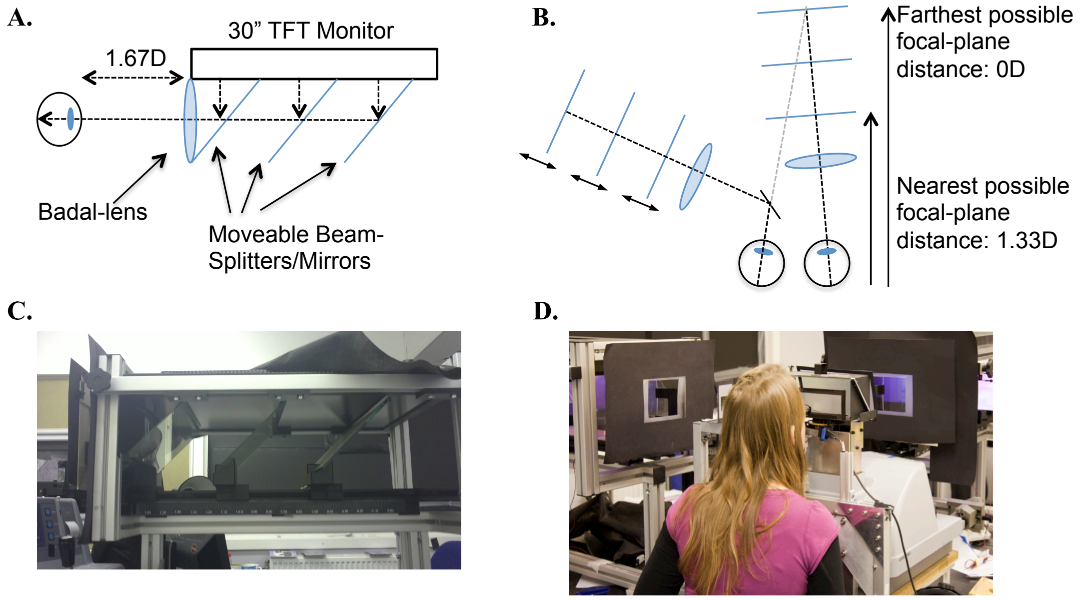


Figure 2.1.1. The multi-plane display used in this thesis. (A) A side view of one eye’s display, showing the arrangement of beam-splitters, first-surface mirror and Badal lens. (B) Plan view of the arrangement of the left- and right-eye’s displays. The left eye viewed its display via a first-surface mirror; the right eye viewed the display directly. (C) The side view of the display showing the screen projecting onto the beam-splitters and first-surface mirror positioned on top of an optical rail. (D) An observer on the display.

Each multi-plane display is mounted on ‘ball wheel’ casters which, when loosened, allow the displays for the two eyes to move horizontally. Thus, the haploscope separation was adjusted for the inter-ocular distance (IOD) of the observer. This means that the vergence distances generated could be matched to the IOD of the observer, creating correct geometric locations of the left and right eyes images on the focal planes. The casters also allowed each eye’s display to rotate around the eye’s center of rotation. This was particularly important when measuring stereo performance with vergence-accommodation conflict as it allowed us to optimise the screen position for a range of vergence specified distances.

Each eye’s display was configured as a Badal optical system. That is, the display was viewed through a positive spherical lens, placed at a distance from the eye equal to its focal length (here, 60 cm or 1.67 D). This configuration has a number of significant benefits for our purposes over a system with no such lens. First, a given image size on the monitor has a constant angular size at the retina, regardless of the focal distance it is presented at (the principle of the Badal Optometer; MacKenzie, et al., 2010). This means all calculations for determining geometrical image properties can be made with respect to a single ‘virtual screen surface’ at the distance of the lens. Moreover, the spatial resolution of the display (the angular size of pixels) remains the same at any focal distance. In this case, the Nyquist limit was around 21 cpd. Second, in a Badal optical system, objects positioned at the lens' focal length beyond the lens are at optical infinity (0 D). Thus, the Badal lens allows us to present a large range of focal distances (up to 0 D) in a physically small display. The nearest focal distance we could present was 1.33 D (physical constraints evident in Figure 2.1.1. (A) mean the near focal plane could not be coincident with the lens, at 1.67 D). Third, in a Badal optical system variations in dioptric distance are linear in physical distance (i.e. metres and centimetres). This means that a constant physical distance between focal planes corresponded to a constant focal-plane separation in dioptres, regardless of the overall focal distance to the planes. This greatly simplified setting up the display, particularly in experiments in which inter-plane separation was varied.

The size of the Badal lenses was the theoretical limiting factor for the field-of-view of the display. The lenses were 15 x 12cm high, providing a maximum possible field-of-view of 14.3 x 11.4 degrees. In our experiments, however, it was critically important to prevent any visible parts of the apparatus (lens mounts, edges of lenses etc.) from acting as an additional, uncontrolled stimulus to accommodation. This was possible because the TFT displays used in the display emit significant amounts of background light even when no image is displayed, which could directly illuminate (or silhouette) parts of the apparatus. Conventional apertures could not be used to prevent this because the sharp edges of the apertures could act as a stimulus to accommodation. We therefore constructed a series of translucent baffles, made from lighting gel (a diffusing sheet material used on stage lights) placed at three different distances from the observer (with apertures matched for angular size). These baffles eliminated any visible sharp edges that could drive accommodation, and resulted in a gradual luminance 'ramp' between the display surface and the surround. The larger apertures can be seen in front of the displays in Figure 1.2.1 (D).

Positioning the observer

Knowing the position of the observers with respect to the device is important for calculating and presenting stimuli at known locations in the left and right eyes display. For Experiments 1-7 this was controlled using a bite-bar mount. For these experiments, it was important that the position of the observer was exact relative to the display as depth-filtered stimuli were used, which require precise alignment across different focal planes.

The bite bars were fit once and were individually constructed. A sighting device was used to adjust the mounts to place the participant in a known location, relative to the display

(Hillis and Banks, 2001). The sighting device accounted for the observers IOD (inter-ocular distance) and allowed us to calibrate the bite bars such that the nodal points for each eye were centered around an origin point on a line relative to the bite bar mount. The constructed bite bar, combined with the IOD adjustment of the multi-plane display, positioned the centre of rotation of the observers' eyes in a known distance along the surface normal from the centre of each eye's display.

For Experiments 8 and 9 knowing the exact location of the observer was less important. In these experiments, depth-filtered stimuli were not used. Also, we measured using a larger number of participants; as such constructing individual bite bars was impractical. Therefore, chin and forehead rests were used to fix the participants relative to the display in these experiments.

The observers were kept stable using a chin and forehead rest (using soft-plastic goggles). Images of this arrangement are shown in Figure 2.1.2. The height of the chin rest and goggles could be adjusted. Due to the combination of the chin and forehead rest, any movements made were minimal and could be felt by the observers. They were instructed to move as little as possible while completing the tasks.

Figure 2.1.2. – The chin and forehead rest. The image on the left shows the goggles straight on. A black cardboard septum ensured that the right eye could not view the images for the left eye and vice versa. The image on the right shows the side of the glasses, illustrating the chin rest below.

Software

The multi-plane display is controlled using custom-written Matlab software (The Mathworks Inc., Natick, Massachusetts), and stimulus presentation used the Psychophysics Toolbox library, version 3.0.8. (Brainard, 1997; Pelli, 1997).

2.1.1 Luminance Calibration

It was essential for our experiments that we had accurate control over the luminance at the eye from each focal plane in the display, for two reasons. First, depth-filtered stimuli (Experiments 1-5) require accurate specification of the relative luminance of images on two (or more) focal planes. Second, we wanted to match the luminance of stimuli at different distances in *all* experiments, so that the quality of the stimulus to accommodation was not confounded with other manipulations. This was relatively non-trivial in our display due to the presence of the various intervening mirrors and beam-splitters between the eye and the display surface, which not only reduce luminance, but are also wavelength non-neutral.

To calibrate the output of each focal plane, as seen by the eye, we measured an overall 'gamma function' by determining (via trial-and-error) the red, green and blue (RGB) triads required to produce a white square (CIE chromaticity coordinates = 0.333, 0.333, 0.333), at luminance intervals of approximately 0.5-1 cd/m^2. We then fit polynomial equations to each colour function allowing us to interpolate between measured luminance values to produce the target luminance required in a given instance. Luminance measurements were made using a Minolta CS-100 Chroma Meter, located at the eye's viewing position. Separate calibrations were performed for each focal plane (near, mid, and far) for each eye's display. We also measured this at the focal-plane positions used in a particular experiment. The relatively small stimulus sizes used (typically ~4 degrees visual angle, or 4.2 cm on the screen) meant that a single calibration point on each plane was sufficient.

The highest luminance used in our experiments was determined by the maximum luminance of the dimmest focal plane (the farthest one). The lowest luminance (i.e. the black level) used was determined by the luminance when no images were displayed. The actual luminance values used (and resultant contrasts) are reported in the individual experiments because they fluctuated depending on the position of the focal planes in the display, and over the years of data collection.

2.1.2 Alignment

It was important that images at different focal planes were aligned with one another for depth-filtered images. Moreover, the location of images needed to be accurate to present geometric information correctly (i.e. the disparity stimulus to vergence).

As described above, this was partially achieved by control of the observers' viewing position. To compensate for any additional inaccuracies, however, a monocular alignment task was carried out at the beginning of each experimental session to determine the centre of each focal plane with respect to the viewing position. This was carried out separately for each eye's display. A sighting device was placed in front of the Badal lens, consisting of two thin wires that intersected at the centre of the display's field-of-view. A white dot (diameter 10 arc min) was projected onto the near plane. The task was to centre the dot at the intersection of the wires to find the centre of the nearest focal plane. Observers then completed a vernier alignment task twice, to align first the mid focal plane, and then the far focal plane, to the near plane. This task consisted of aligning horizontal and then vertical gratings shown on each plane (Akeley et al., 2004). Each task was completed three times for each focal plane, and alignment was considered acceptable if the standard deviation of the three centring values was less than 0.75 pixels (when using the bite-bar) and 1.25 pixels when using the chin rest.

2.2 Measuring Stereoscopic Performance

Throughout this work we directly measured the effects of viewing stimuli under various viewing conditions, including viewing stimuli with and without conflict as well as depth-filtered images, on stereoscopic performance (stereo performance). Stereo performance here relates to stereoscopic depth perception, including the detection of fine details in the stimuli, the time taken to fuse the stimuli and being able to perceive depth in the image. This section discusses the stimuli and techniques used to characterise stereo performance.

2.2.1 Stimuli – Random Dot Stereograms

Because we wished to determine the effects of various display parameters on stereoscopic performance, specifically, we needed to use a stimulus that isolated stereoscopic depth perception. This can be achieved by using random-dot stereograms (first proposed by Julesz in 1960). In these stimuli, each eye's image is composed of randomly placed dots, with no apparent (monocular) depth structure. Instead, depth is specified only by the disparities introduced between the positions of individual dots (Poggio, Motter, Squatrito, Trotter, 1985; Palmer, 1999).

The sensitivity of the visual system to stereoscopic depth is typically assessed using random-dot stereograms depicting a sinusoidal corrugation in depth (Bradshaw & Rogers, 1999; Banks, Gepshtein & Landy, 2004 – see end of doc). By using a sine-wave depth profile, such stimuli allow measurements of the minimum detectable disparity-specified depth in the stimulus (i.e. peak-to-trough disparity), at a given corrugation frequency. Alternatively, one can measure the maximum resolution in disparity-specified depth that can be seen (i.e. the highest corrugation frequency that can be seen, expressed in cycles per degree of visual angle), at a given peak-to-trough disparity. The former is referred to as stereoacuity, and the latter as stereoresolution. Such stimuli have also been used to measure other aspects of stereo performance that are relevant here, such as the time required for stereoscopic fusion to occur (Hoffman et al., 2008). We measured stereo performance in terms of all three of these parameters (stereoacuity, stereoresolution, and time-to-fuse) in different experiments.

It should be borne in mind that these measures are not in fact independent, and variations in one parameter can affect measures in another, for fundamental reasons. For instance, the maximum detectable corrugation frequencies, and minimum peak-to-trough disparities, are interrelated. If the peak-to-trough disparity in the stimulus is below the threshold of the viewer, the corrugation cannot be seen, regardless of the corrugation frequency. Therefore, when manipulating one measure, one must ensure the other is above threshold such that increasing its value would not affect a viewer's ability to see depth in the stimulus.

Also, increasing peak-to-trough disparity (and corrugation frequency) increases the disparity gradient of portions of the surface (the 'flanks' of the corrugation have a steeper gradient in depth with respect to the eyes), which ultimately results in degraded depth perception due to the so-called disparity gradient limit (Burt & Julesz, 1980). Thus, there is also an *upper* limit on resolvable depth in a random-dot corrugation, as well as a lower one, and the amount of disparity in the stimulus must be in a 'sweet spot' to measure the effects of varying corrugation frequency. Similarly, there is a low (as well as high) cut-off in the visible corrugation frequency visible in random-dot corrugations. This reflects both the sensitivity of the system and the fact that, given a certain stimulus size, increasingly small proportions of a cycle are presented as corrugation frequency is reduced (Bradshaw & Rogers, 1999; Banks, Gepshtein & Landy, 2004). So, again, corrugation frequency must be set in a certain range for effects of peak-to-trough disparity to be measured. It follows logically that both parameters must be set at suitably intermediate values when measuring stereoscopic fusion time.

Precise details of stimulus parameters are specified in the descriptions of the individual experiments, as appropriate. Unless stated otherwise, a circular aperture clipped all stimuli, with a diameter of 4 degrees visual angle, dot density was 45 dots/deg^2, dots were 2 pixels in width (which equates to 0.05 visual angle), and the luminance of the dots was 25 cd/m^2 (white) on a 'black' background at 0.21 cd/m^2.

2.3 Measuring thresholds for stereo performance

Stereo performance was determined in each case using a classical *forced-choice* psychophysical method, where the observer had to respond to the stimuli according to a criterion, the results of which inform whether or not he or she can see the stimulus. In the majority of our experiments, the task was to report the orientation of the disparity-defined corrugation, which could be +/-15 degrees from horizontal (see Figure 2.1.3). The parameter of interest (e.g. peak-to-trough disparity) is varied across trials, and the relationship can be established between the proportion of correct responses and the value of the parameter (see below). Often, the task is to compare across two or more stimuli, known as an M- alternative forced-choice (AFC) task, where 'M' is the number of stimuli the participant has to compare. In our experiments, instead of comparing across multiple images, we present the observers with a single stimulus. They then have to discriminate whether the stimulus is orientated to the left or right of horizontal. Therefore, our measure is technically a 1-AFC task.

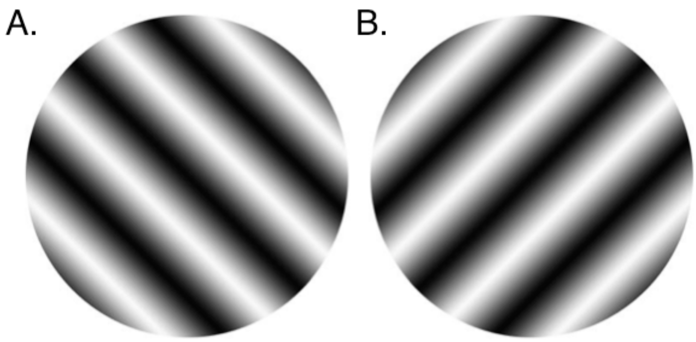


Figure 2.1.3 –Example of the stimulus orientations. The stimuli were composed of random dot and not lines (as shown here for clarity). The perceived corrugations were orientated either -45 degrees (A) or +45 degrees (B) from horizontal.

A threshold in psychophysics can be described as the stimulus strength that relates to a determined value of sensitivity to the stimulus (normally defined in percent correct). The stimulus strength that relates to ~75% correct is a common definition of threshold (Klein, 2001). We mainly use this as the threshold value in our experiments. The 75% correct point in a 1AFC task is the mid-point in accuracy between chance performance (50% correct) and perfect performance.

Generally, we estimated threshold values by fitting a psychometric function to an individual's data collected at various stimulus levels. There are several methods for controlling the stimulus levels presented on each trial. We used an adaptive staircase procedure, that varied the stimulus level according to the individual's response. This is more efficient than the traditional Method of Constant Stimuli, which presents a fixed number of trials at pre-determined stimulus levels, some of which might be uninformative (Howard, 2002; Rose, 1995).

Staircase procedures adjust the stimulus level on subsequent trials according to the observer's response. For example, if the observer answers correctly on a trial, the difficulty can be increased until the observer answers incorrectly. The difficulty can then reverse again, getting easier, until the observer is accurate in their response and so on. In this way, performance can be 'titrated' around a particular level. It is termed a 'reversal' every time the staircase changes direction (changing from the stimulus level getting higher on subsequent trials to getting lower, or vice versa). Different staircase reversal rules can be used, causing the staircase to converge on different performance levels. We typically used a 1-up, 2-down, reversal rule, and/or a 1-up, 3-down rule. In the former case, one incorrect response causes the next trial to be easier, but two correct responses in a row are required to make the stimulus level go 'down' (i.e. make the task harder). In the latter case, three correct responses in a row are required to make the stimulus level get harder (so the staircase would be expected to converge at a higher percent correct). Sometimes staircases are used to directly determine a threshold, based on statistical assumptions about the performance level at which they converge (Wetherill & Levitt, 1965). Here, we generally used staircases only to sample points at different places on the steepest part of the slope of the psychometric function, where they are most informative (based on the procedure used by Hillis et al., 2004), and calculated thresholds from the fitted function (see below).

Unless otherwise indicated, staircases were terminated after 12 reversals had occurred. We also varied the staircase step size – the size of the jump in stimulus level that occurred when there was a change. If the step size is too big, the measurements may not be close enough for the threshold to accurately determined. If step sizes that are too small, however, it can take many trials for the staircase to converge at 'threshold' for that stimulus (Gescheider, 1997). In our experiments, we initially used large step sizes, so the staircase would quickly adjust to the appropriate point early on. Then, after four reversals of the staircase, the step size was halved.

2.4 Analysis

Stereoacuity, stereoresolution, or fusion-time thresholds were estimated as follows. In each experiment, and for each condition, an individual's data were fitted with a cumulative Gaussian function, using a maximum-likelihood criterion (Wichmann and Hill, 2001; Kingdom, 2009). Usually, multiple staircases were measured simultaneously in one block of trials, the data from these staircases were combined and fitted with the cumulative Gaussian function. An example stereoresolution function is shown in Figure 2.1.4. This figure shows the proportion of correct trials (correctly identifying the orientation of the corrugation), as a function of corrugation frequency. The data points show the observer's responses (the size of the data points is proportional to the number of trials at each stimulus level). The plot shows how, using this staircase procedure, the stimulus levels relating to 100% accuracy were not tested a lot. However, the levels on the steep part of the function (which help to determine its shape) were heavily sampled. The 75% correct point is read from the psychometric fit, which we define as the threshold value for stereoresolution in this case.

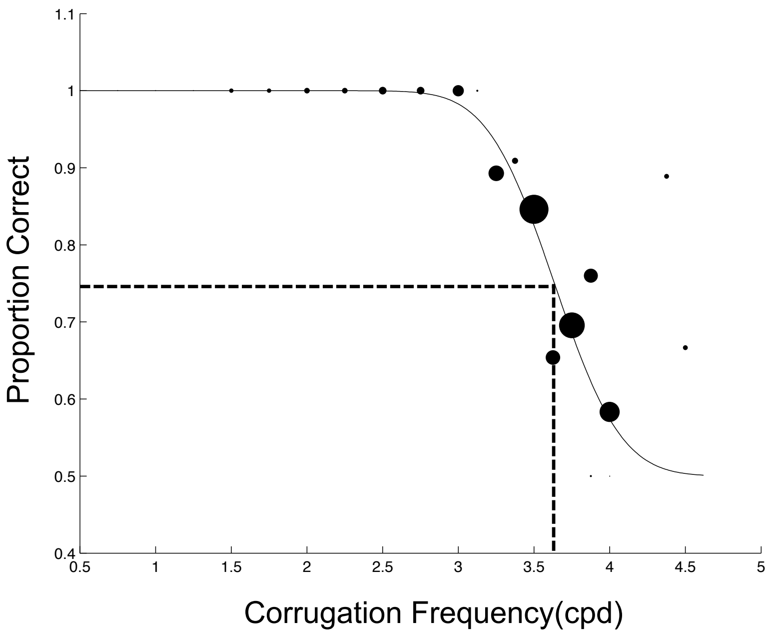


Figure 2.1.4. – Example psychometric plot. The percentage of correct responses (x-axis) is plotted against the stimulus level (y-axis). The size of the dots is proportional to the number of times that stimulus level was viewed. The solid line is a non-linear regression line through the data. The dashed line shows the threshold at 75% correct.

2.4.1 Error Bars

Our psychometric approach allowed us to measure individual performance precisely and accurately. However, it is very time consuming when using large sample sizes as each individual is tested for many hours each. This is unfortunate, because we are also interested in group-level effects, not only effects on individuals. Nonetheless, for the individual data, we calculated the standard error of the fit to the psychometric function (see above). For group data, we calculated either the standard error of the mean of the sample, which is the standard deviation of the sample means, or confidence intervals.

2.5 Stimuli – Artefact

For Experiments 1-4 in this thesis, the data from the farthest distances measured had to be discarded due to a stimulus artefact. This artefact provided the participants with monocular cues to the orientation of the stimuli when viewed at the far distances in the display (0.25D and 0.1 D). This was due to a 'banding' of dots along the flank of the gratings, as seen in Figure 2.1.5 (top and middle panels), this shows an example of the stimulus at the far distance (0.1 D).

This 'banding' happens as a result of the increased physical steepness of the flanks of the corrugation at the much large than usual viewing distances used in our study. This occurs because retinal size and binocular disparity scale differently with distance. If an object's distance doubles, its physical size in x and y must be doubled to project the same size in the retinal image. Disparity varies with a $1/d^2$ relationship with distance, however. Thus, if the distance to a stimulus is increased by a factor of two, the physical depth in the stimulus needs to be multiplied by a factor of 4 for the peak-to-trough disparity to remain the same. Or, if we were to move a stimulus from 1m to 10m in the display, its physical width would need to grow by a factor of 10, whereas its physical depth would need to increase by 100 times, to preserve the same retinal size and peak-to-trough disparity. For this reason, even with threshold level peak-to-trough disparities, the physical depth of the stimulus becomes very large at large distances, resulting in the banding illustrated in Figure 2.1.5 (C and D).

Previous research using this stimuli and manipulations did not use such large distances (Hoffman et al., 2008), those who have tested at such large distances in the past have used simpler stimuli such as single light-emitting-diodes (LED) lights (Allison, Gillam, & Vecellio, 2009). We were therefore previously unaware of this issue and its solution before testing Experiments 1-4.

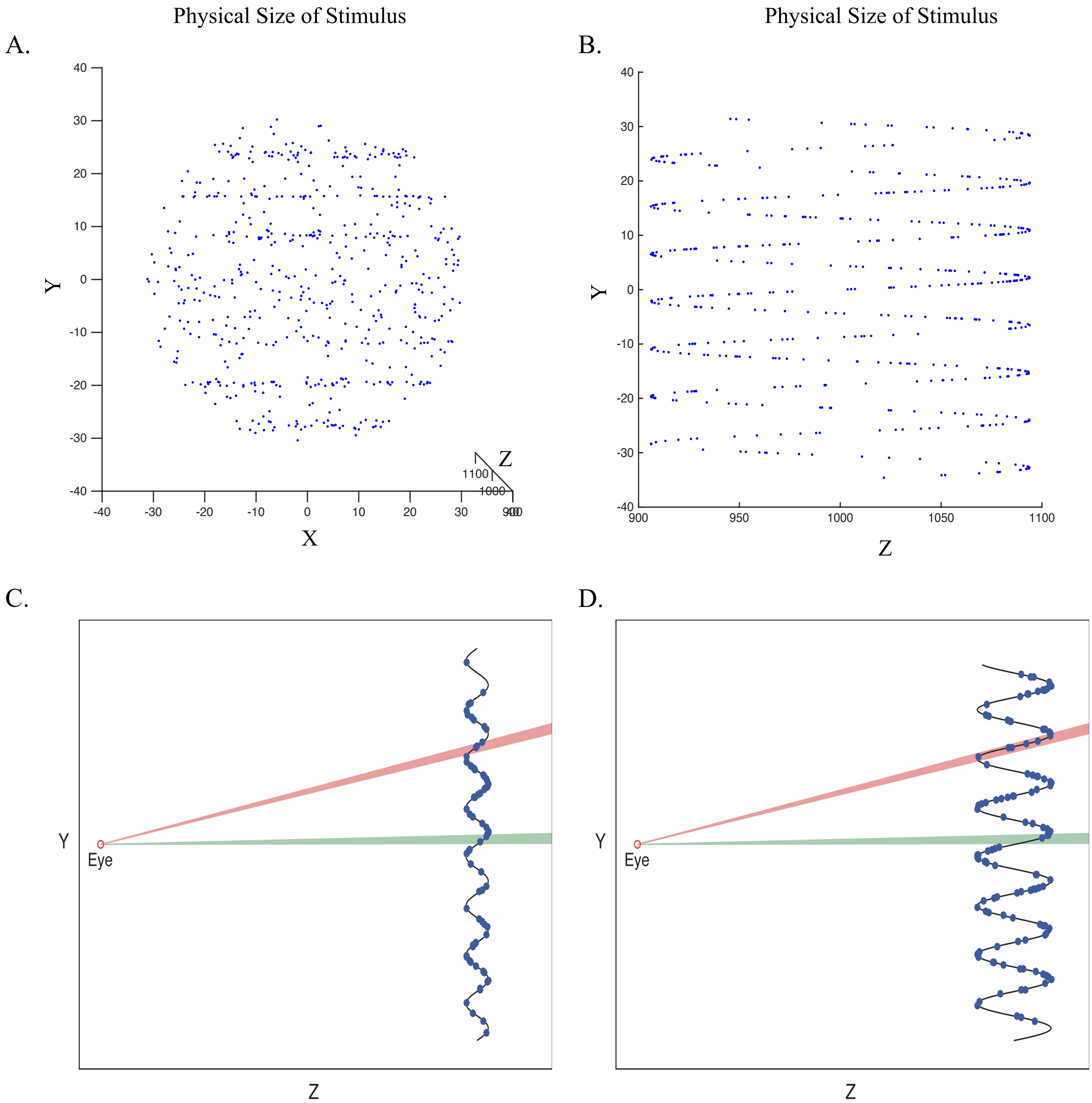


Figure 2.1.5 – Illustrations of the stimuli in Experiments 1-4 at the far distance (0.1 D). Graphs A and B show a random dot stereogram, which is projected with respect to the cyclopean eye. Graph A shows the resultant stimulus viewed as in the display and the right graph shows the stimulus from the side (in 'z' space). The image has an angular size of four degrees, a peak-to-trough disparity of 240 arcseconds and a corrugation frequency of 2 cycles per degree (cpd). The bottom panel is an example illustration of the stimulus with different peak-to-trough disparities but the same corrugation frequencies. The blue dots are the random dots in the stimulus; the lines are included for clarity and outline the corrugations in the stimulus. The green and red lines are the lines of sight, from the cyclopean eye. The stimulus on the left (C) has more depth than the stimulus on the right (D).

In later studies we solved this problem by moving the dots not purely in z, but along the line of sight to the cyclopean eye. This results in a corrugation whose depth is not sinusoidal in Cartesian space, but is sinusoidal in terms of the angular elevation of the dots with respect to the eye. That is, the corrugation appears to 'fan' out with respect to the cyclopean eye. The resultant monocular image is shown in Figure 2.1.6.

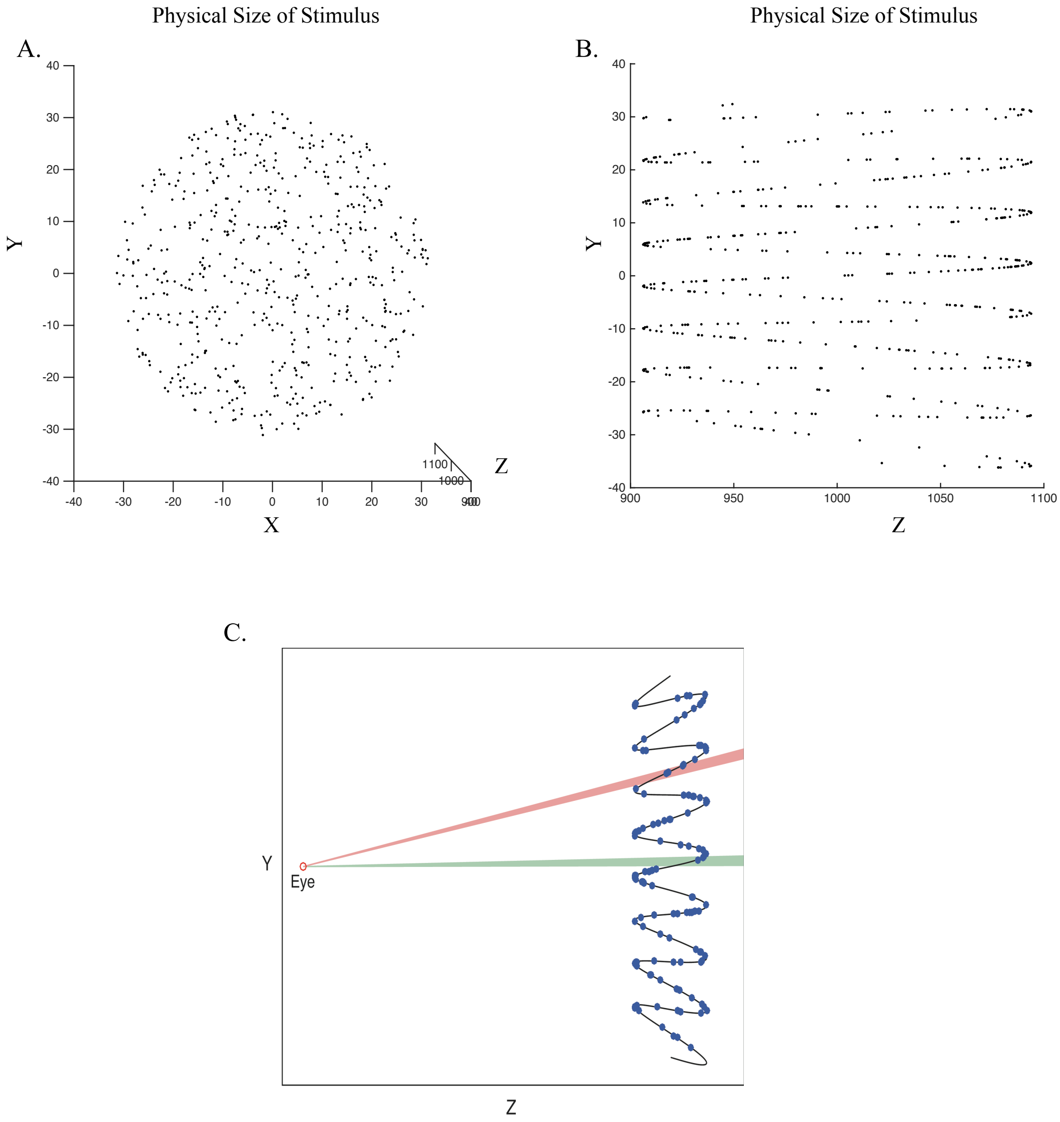


Figure 2.1.6 – Illustrations of the corrected stimuli at the far distance (0.1 D). Graphs A and B show the same as in Figure 2.1.5, with the dots being projected in terms of azimuth and elevation. The image has an angular size of four degrees, a peak-to-trough disparity of 240 arcseconds and a corrugation frequency of 2 cycles per degree (cpd). The illustration (C) is an example illustration of the stimulus with a large peak-to-trough disparity and corrugation frequency. The blue dots are the random dots in the stimulus; the lines are included for clarity and outline the corrugations in the stimulus. The green and red lines are the lines of sight, from the cyclopean eye. The stimulus 'fans' out with respect to the cyclopean eye as the stimuli are calculated in terms of angular elevation of the dots with respect to the eye.

2.6 Experiment constants

Ethics

In all studies reported, the observers were volunteers who were reimbursed for their time. They were aware they could leave the experiment at any time. All experiments followed the tenets of the Declaration of Helsinki and were approved by Bangor University's Ethics Committee. Informed consent was obtained after the nature of the study was explained to observers.

Pre-Testing

In all the experiments reported, we screened observers to ensure they were not stereoblind. For experiments 1-7, this involved the observers viewing a random dot stereogram in the multi-plane display, at all focal distances used in the experiment. The stimulus had a corrugation frequency of 1.5 cycles per degree (cpd) and a peak-to-trough disparity of 50 arc seconds. We used these values as they are minimum detection threshold for normal observers. The observers completed five trials in which they responded by reporting the orientation of the random-dot stereogram to the experimenter. There was no limit on the stimulus presentation time. For Experiments 8 and 9, we used the 'circles' part of the Randot stereo test (Stereo Optical Company, Inc.). Observers with stereoacuity >= 60 arc secs did not proceed to the main experiment. We changed the pre-screening procedure for these experiments, as we wanted to have a clearer understanding of the observers' stereoacuity.

Part of the pre-testing proceedure for all experiments involved measuring the observers' inter-ocular distance (IOD). This was measured using an "Essilor Varilux" ruler.

CHAPTER 3

3 ARE MULTI-FOCAL-PLANE DISPLAYS A PRACTICAL SOLUTION TO THE VERGENCE-ACCOMMODATION CONFLICT?

3.1 Introduction

As previously discussed, one of the more fundamental issues with current stereoscopic 3-D displays is the conflict between the stimuli to vergence and accommodation. This arises because images are presented on a single surface, independent of the depth structure of the portrayed scene. Because vergence and accommodation systems are synergistically coupled (Martens & Ogle, 1959; Mays & Gamlin, 2000; Yang & Sheedy, 2011; Heron et al., 2001; Fincham & Walton, 1957; Schor, 1992), decoupling them sufficiently can be difficult, or even impossible. The effort involved in fusing images with vergence-accommodation conflict has been shown to result in reduced stereoacuity (Hoffman, et al., 2008; Watt, et al., 2005), an increased time to fuse (Akeley et al., 2004; Hoffman et al., 2008; Watt et al., 2005), visual discomfort and fatigue (Hoffman et al., 2008; Shibata et al., 2011), and distortions in perceived depth (Hoffman et al., 2008; Watt et al., 2005).

3.1.1 Accommodation to depth-filtered images

A proposed solution to the vergence-accommodation conflict is a volumetric multiple-focal-plane (multi-plane) display. As described in Chapter 1 (section 1.4) this approach is capable of matching the stimulus to vergence and to accommodation. Presenting the eye with the sum of images projected on a number of discrete focal planes, in conjunction with the depth-filtering technique—in which the luminance ratio of image points is varied across the focal planes, to simulate focal distances between them—can accurately stimulate the accommodation response to the location of the simulated image (MacKenzie et al., 2010; 2012). Importantly, accommodation (and vergence) responses were accurate to spatially broadband depth-filtered images for focal-plane separations up to 0.9 D (MacKenzie et al., 2010; MacKenzie et al., 2012). This is encouraging for multi-plane displays because it suggests that a display with just four focal planes, with 0.9 D spacings, could eliminate vergence-accommodation conflicts across a useful range of presentation distances spanning 2.7 – 0 D (37 cm to optical infinity).

3.1.2 Stereo performance with depth-filtered images

The analysis of MacKenzie et al.'s (2010) demonstrate, however, that accommodation to depth-filtered images is accurate because accommodation is primarily driven by mid-range spatial frequencies (4-8 CPD; see also Mathews & Kruger, 1994; Owens, 1980; Tucker, Charman, & Ward, 1986; Ward, 1987), which are preserved in depth-filtered images up to quite large plane separations. As noted in Chapter 1 (Figure 1.6), however, spatial frequencies that are higher, but well within the visible range, are attenuated even at small focal-plane separations. Essentially, this is because when the eye is 'accurately' focused at a simulated distance between two focal planes, it is defocused with respect to both planes, and a relatively small amount of defocus significantly attenuates higher spatial frequencies. Depth-filtered images therefore cannot present the same level of fine detail (high spatial frequencies) as images on a single focal plane. Thus, we might expect that depth-filtered images will result in reduced precision of stereoscopic vision (Legge & Gu, 1989; Simons 1984; Smallman & McKee, 1995) (as well as a reduction in perceived image quality, due to blur) as a function of image-plane separation. Stereo performance, rather than vergence and accommodation responses, may therefore be the limiting factor in determining the maximum usable focal-plane separation for stereoscopic depth-filtered images.

Here we explore the relationship between focal-plane separation and stereo performance, to examine whether reasonable performance can be achieved in a practical multi-plane display. To do this we establish how stereo performance is affected by focal-plane separation against a number of criteria: stereoresolution, time-to-fuse (the time required for stereoscopic fusion to occur), and stereoacuity. We also measured at what focal-plane separation noticeable decrements in image clarity occurred. Critically, in each case we compared performance for depth-filtered stimuli to a real-world equivalent (single-plane stimulus, at the same distance) holding all other image properties constant.

3.1.3 Previous Research

A limited amount of previous research has examined stereo performance with depth-filtered images. Time-to-fuse and stereoresolution have been found to be comparable for depth-filtered images (focal-plane separation = 0.67 D) and images presented on a single focal plane (equivalent here to a real-world stimulus, because all image points can be equally focused, and there is no vergence-accommodation conflict) (Akeley et al., 2004; Hoffman et al., 2008). However, these studies used a display with fixed focal-plane positions, and therefore a single inter-plane spacing. This meant they could not measure the relationship between focal-plane separation and stereo performance. It also meant that they could not match the distance to the depth-filtered and 'real-world' stimuli (i.e. the real-world stimulus was not the perfect control condition), or measure stereo performance with depth-filtered stimuli at different distances. If performance on depth-filtered images is affected by location, their comparison of plane-separations with stimuli at a fixed distance may not be representative of all uses of the display.

3.2 Experiment 1: Stereoresolution with Depth-Filtered Images.

Experiment 1 examines stereoresolution with depth-filtered images, as a function of focal-plane separation. As discussed in General Methods, here, stereoresolution is defined as the finest spatial scale of modulations in depth from disparity that can be perceived. The stereoresolution threshold was operationalized as the highest corrugation frequency of a sinusoidal corrugation in depth for which orientation could be identified at 75% correct (Hoffman et al., 2008).

Particular care was taken in these experiments to control for the stimulus to vergence on each trial, with respect to fixation. We used fixation positions away from the location of the upcoming stimulus so the eyes could not be correctly positioned/focused beforehand (which is unrepresentative of real-world applications of stereo displays). Thus, our task required participants to make vergence and accommodation responses to the stimulus, prior to making a depth judgement. Obviously, it was necessary to match the magnitude of the required responses across different conditions when measuring time-to-fuse, to avoid a confound. This was also necessary when measuring stereoresolution and stereoacuity, however, as these are measured with a fixed stimulus presentation time and so delayed eye movements (or indeed accommodation) could also result in reduced performance on these measures. Unfortunately, for some sessions in the experiment, matching the magnitude of responses was not possible. This resulted in some blocks of trials having one focal-plane positioned farther from fixation than the other (Blocks 1-5, Figure 3.2.1). However, we were able to match the distances of the focal-planes in the display, across all blocks of trials (see Figure 3.2.1). We were therefore able to examine the effect of focal-plane distance from fixation on all measures of stereo performance (i.e. TTF, stereo resolution and stereoacuity) to examine how this might affect the results, if at all.

We also considered the sign of the required vergence (and accommodation responses). The control processes for convergent and divergent eye movements are thought to differ, leading to differences in overall response time (Semmlow, Hung, & Ciuffreda, 1986; Semmlow & Wetzel, 1976; Hung, Zhu, & Ciuffreda, 1996). Specifically, divergent responses normally have slower velocities and overall responses to the stimuli (Semmlow et al., 1986). This has been associated with differences in muscular control of the two systems. For divergent movements, the lateral recti control the response of the eyes' in an outward direction (Mon-Williams & Tresilian, 1998). However, they do this passively (i.e. they receive no command); these muscles simply relax. As the image becomes increasingly fixated, the medial recti muscles apply tension to slow the response until it stops at an appropriate distance for the image being viewed. When a convergent eye movement is made, the medial rectus is controlled and made to shorten, drawing the eyes towards each other. At the same time, the lateral rectus muscles are lengthening. Therefore, the medial recti have an active control command causing the muscles to contract, creating opposing tension in the lateral recti, this tension increases with increasing convergence (Mon-Williams and Tresilian, 1998). To address the difference in fusion times based on the movement made to create fusion, we controlled the number of trials requiring divergent and convergent responses in each experiment. As detailed below, this required the use of 'external' light-emitting diodes (LEDs) as fixation targets.

Importantly, we initially analyse the data using the combined divergent and convergent trials, to get an overall impression of stereo performance. We then analysed the data separately to examine any effects of divergent or convergent responses.

3.2.1 Method

Observers

There were five observers in this experiment: AA (21 years old), AB (20 years old), AC (35 years old), LR (author, 22 years old), and AD (20 years old). At the time of testing, none of the observers required optical correction.

Experimental Sessions

The observers were tested during seven separate sessions, each measuring a separate block of trials (herein referred to as 'blocks'). These blocks are shown in Figure 3.2.1. We examined 1) whether the distance of the depth-filtered image from the observer affected stereo performance and 2) the effect of inter-plane spacing on stereo performance. To examine the effect of distance we varied simulated distance (depth-filtered images) while holding image-plane separation constant at 0.6 D (blocks 1, 2, 3, and 4). To test the other subset, we varied focal-plane separation while holding simulated distance constant (5, 6, and 7). The largest image-plane separation (1.2 D) spanned the range of the display with the focal-planes positioned at 1.3 D and 0.1 D, and so the stimulus distance chosen for the simulated mid-point was 0.7 D. We used the dioptric mid-point because MacKenzie et al.'s (2010) analysis of retinal-image formation showed that retinal-image contrast is maximally attenuated for these stimuli, relative to real-world stimuli. This should, therefore, be a 'worst-case scenario' for stereo performance from depth-filtered images. For the real-world comparison (equivalent single-plane stimulus), we had two sessions (3 and 4) where focal planes were located at 0.7 D.

In addition to the experimental trials, observers completed 'filler' trials (100% image intensity on the nearest or farthest focal plane) in each block, so that observers could not anticipate the oculomotor responses that would be required on upcoming trials. These trials also served as controls for the depth-filtered stimuli presented at the same distance (see Figure 3.2.1). Therefore, each block included trials in which depth-filtered stimuli were presented, interleaved with stimuli presented entirely on one or other image plane. The order of blocks was randomised across observers.

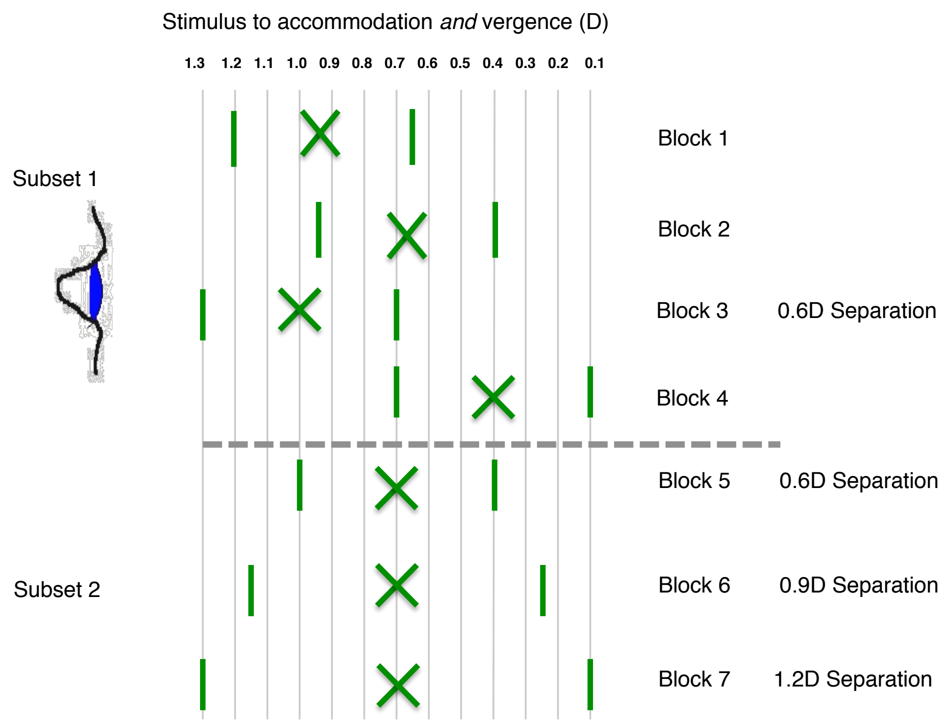


Figure 3.2.1 – The blocks in Experiment 1. The figures at the top denote the distance of the stimulus (i.e. stimulus to accommodation *and* vergence) from the eyes (shown on left of the diagram). The vertical green lines represent the positions of focal planes in each block. The green crosses denote a depth-filtered image location. The focal-plane separations in each case are shown in the far-right column. Blocks 1-5 had the same focal-plane separation (0.6 D) but at different distances. Depth-filtered stimuli in blocks 5, 6 and 7 were at the same simulated distance, but with varying focal-plane separations.

Stimuli

This experiment used the stimuli (random-dot stereograms, depicting sinusoidal corrugations in depth) and task (report corrugation orientation) described in General Methods (Chapter 2). The peak-to-trough disparity was 1.8 arcmin at all viewing distances and corrugation frequency was varied by the experiment procedure. All stimuli were presented binocularly and the room lights were turned off, so that only the stimulus was visible.

During pilot testing, a monocular artefact was identified for the farthest focal-plane positions (0.25D and 0.1 D), whereby as the corrugation frequency increased, a 'line' of dots could be seen, which indicated the orientation of the stimulus (see General Methods). Because these were filler trials, included to match oculomotor demand, we did not need to measure stereoresolution at these distances. We therefore simply fixed the corrugation frequency of these stimuli at 2.5cpd, and the peak-to-trough disparity at 1 arcmin, so the artefact was never visible. The stimuli at these distances did not vary and the number of trials was capped at 40 so that the numbers of trials in a block were, on average, the same compared with the other blocks. We tested for the artefact at all distances by testing discrimination performance under monocular viewing with both naïve and experienced observers. Performance was at chance for this measure, indicating that no artefact was present in the stimuli at any distance.

Fixation

As noted earlier, we matched the magnitude of the required vergence (and accommodation) response with stimulus onset, across blocks. We also balanced the direction of the required response (divergent and convergent trials were randomly ordered). Because the focal planes were in different places in different blocks, they could not be used to present the matched fixation targets. We therefore used physical near and far fixation targets, positioned at 1.3 D and 0.1 D. Each fixation target was comprised of two LEDs, one for each eye. The lateral positions of the LEDs were carefully adjusted for each observer (on continuously adjustable rods) such that the vergence-specified distances were accurate. The fixation targets appeared 3 deg to the right of the stimulus position, so as not to occlude the stimulus. In order to ensure the LED fixation light were positioned at the correct location in the display for the observer a centering technique was administered. This was a monocular task, in which a cross was projected first onto the nearest focal-plane, the centre of which was positioned 3 deg to the right of the centre of the stimulus. The experimenter adjusted the LED light according to the observers instruction so that the LED light was at the centre of the projected cross for the observer. This was repeated for the fixation LED lights in the left and right eyes display, for the near and far distance.

Procedure

In all the experiments in this chapter, observers were positioned using the bite-bar setup, and completed the display calibration/alignment procedure, as described in the General Methods. Observers in this experiment initiated each trial with a button press. The fixation marker was illuminated for a random time interval (uniform distribution) between 600 and 1400 msec, to prevent anticipatory eye movements. Following the fixation, the stimulus appeared for 2 sec. On each trial, the orientation of the stereoscopic depth corrugation was either +/-15 deg from horizontal, selected at random. Observers indicated the orientation of the corrugations using a button press. They received no feedback.

For each trial type, the corrugation frequency of the depth corrugation was varied using a 2-up, 1-down staircase. In this experiment, 'up' refers to increasing the task difficulty by increasing the corrugation frequency. For this to happen, the observer had to be correct on two consecutive trials for that staircase. The corrugation frequency used in the first trial was chosen at random between 0.5 and 3.5 cycles per degree (cpd). The step-size was 0.25 cpd until after the fourth staircase reversal, when it was halved to 0.125 cpd. Each staircase quit after 12 reversals.

Each block of trials consisted of multiple interleaved staircases. There were separate staircases for each image distance (two for each focal-plane location and one depth-filtered image), and fixation position (requiring a divergent or convergent response), making six staircases in all. Each block was repeated four times in a single session.

Analysis

To measure the effect of distance and of inter-plane spacing on stereo performance, we fit psychometric functions using the data from the two staircases for the divergent and convergent responses at each distance. We then analysed the data separately, using the data from the individual staircases, to examine any effects of divergent or convergent responses.

3.2.2 Results

Stereoresolution thresholds were computed for each observer, for each stimulus type, according to the methods described in Chapter 2.

We wished to directly compare performance for depth-filtered stimuli and equivalent real-world (i.e. single-plane) stimuli. Our experiment design meant that we had real-world stimuli at the same location in space as depth-filtered stimuli. For some instances, this resulted in us measuring performance for the relevant real-world stimuli at 0.7, 1.15 and 1.3 D in two separate blocks (see Figure 3.2.1). In principle, because they are two measurements of the same thing, it should be possible to simply average across these repetitions, to generate a single real-world stereoresolution threshold for a given distance. To ensure this was meaningful, however, we first checked that stereoresolution thresholds for the different blocks were indeed similar. The thresholds for each distance were calculated using the combined divergent and convergent staircase data. Figure 3.2.2 shows the two threshold measurements for each of the three distances, averaged across all observers.

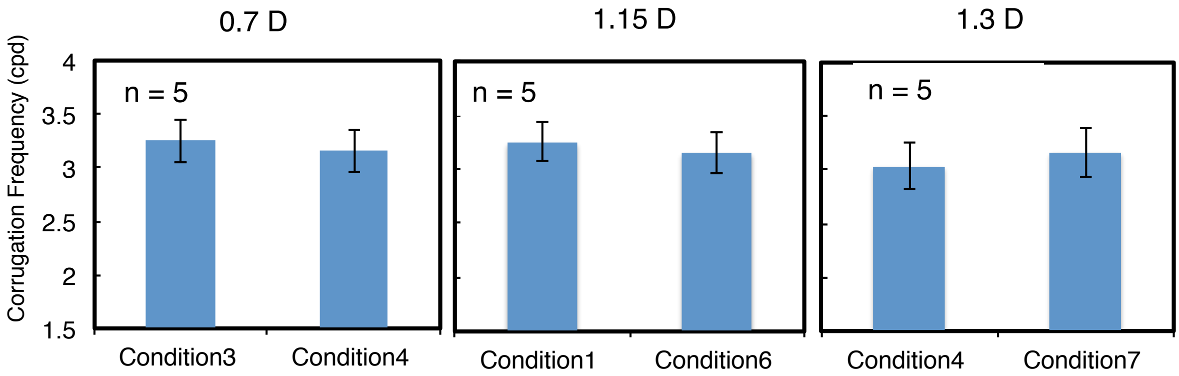


Figure 3.2.2. – Stereoresolution thresholds for repeated measurements of real-world stimuli**.** The data are the average of all observers' data in each block (n = 5). The title of each panel denotes the different stimulus distances. The bars show different blocks (labelled on the x-axis) in which a given distance was presented. The y-axis is the stereoresolution threshold, in terms of corrugation frequency. The error bars represent 95% confidence intervals.

Paired sample t-tests were conducted to compare the thresholds of each condition. There were no significant differences between the conditions for all distances (see Table 3.2.1). We therefore computed a single threshold for the real-world stimuli in each case by averaging across the two measurements.

Condition	Mean	SD	Significance
3	3.21	0.40	t(4), -0.308, p = .77
4	3.17	0.38	
1	3.2	0.40	t(4), -0.883, p = .43
6	3.13	0.38	
4	3.07	0.36	t(4), -1.174, p = .31
7	3.18	0.38	

Table 3.2.1. – Table showing t-test results for each condition**.**

Effect of Focal-plane Separation

Figure 3.2.3 plots the stereoresolution thresholds as a function of focal-plane separation. Again, the data for each focal-plane separation is the data from both the convergent and divergent eye-movement staircases, combined to fit the psychometric function for either the real-world images or the simulated, depth-filtered images at the same location in space. The top-left panel shows the average across all observers, and the remaining panels show each individual's data. A higher threshold indicates better performance (finer modulations in depth could be reliably discriminated). The results show that stereoresolution was highest for the real-world stimuli (0 D focal-plane separation) with a consistent reduction in stereoresolution thereafter, with increasing focal-plane separation. All the observers show the same general trend, albeit to different degrees.

A Repeated Measures ANOVA showed a significant main effect of focal-plane separation size on stereo resolution thresholds, $F(3,12) = 15.63$, $p = <.01$. Bonferroni-Corrected pair-wise comparisons across all conditions showed that stereo resolution thresholds were significantly reduced for the 1.2 D separation condition ($p = .04$) relative to the real-world condition. Whereas there was no significant differences between the real-world condition and the 0.6 D ($p = .55$) and 0.9 D ($p = .20$) conditions. This suggests that depth-filtered images with a focal-plane separation of up to and including 0.9 D result in stereo resolution thresholds that are comparable to real-world equivalent images. It is important to note, however, that although they are statistically comparable, we can see that there is a constant reduction in threshold with focal-plane separation as well as large individual differences. Therefore, the focal-plane separation used to create depth-filtered images should be adjusted according to the task.

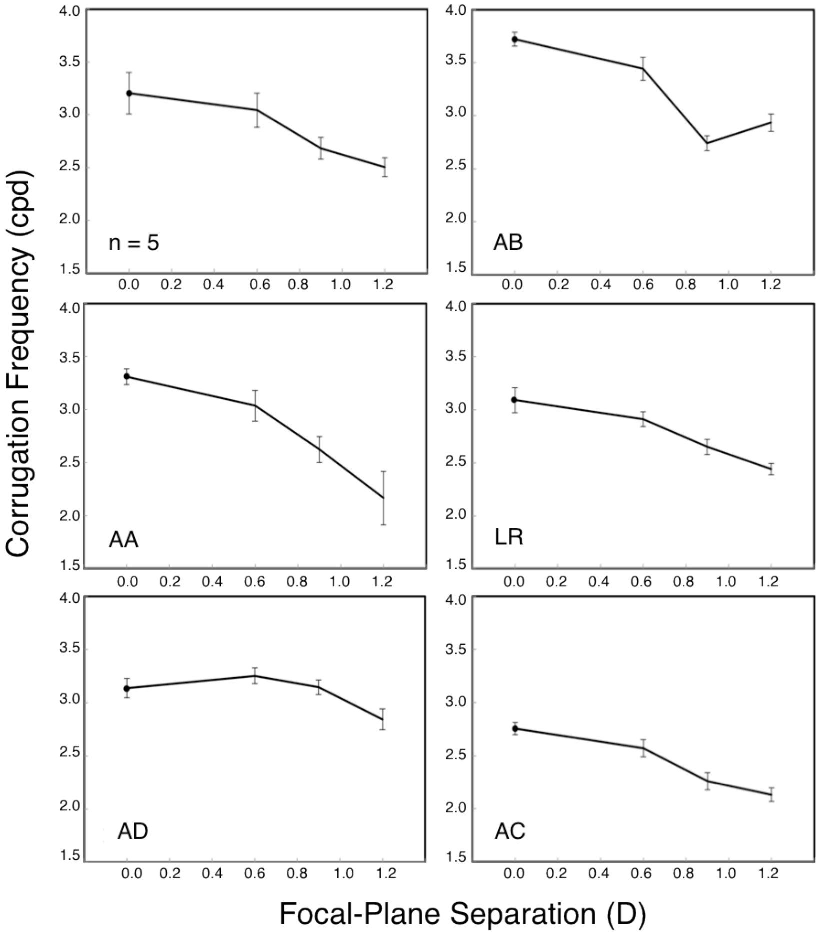


Figure 3.2.3 – Stereoresolution thresholds as a function of focal-plane separation. The top left panel shows the mean thresholds, across all observers, and the other five panels show individual thresholds. The x-axis shows the focal-plane separation in dioptres. Zero on the x-axis and the solid data point denotes single-plane (i.e. real-world equivalent) stimuli. The y-axis is the stereoresolution threshold in terms of corrugation frequency at 75% correct. Error bars for the average graph represent 95% Confidence Intervals. The error bars for the individual graphs denote standard error of the threshold estimate (i.e. standard error of the psychometric function fit).

We also examined stereoresolution as a function of focal-plane separation separately for convergent and divergent response. The results are shown in Figure 3.2.4(a), averaged across all observers. In order to establish whether there was any statistical differences between the thresholds for the divergent and convergent responses, we calculated the mean difference between the thresholds at each focal-plane separation. The results are shown in Figure 3.2.4 (b). The error bars are 95% confidence intervals. As all of the error bars intersect ‘zero’ on the x-axis, we conclude that direction of the required response did not appear to affect stereoresolution measurements.

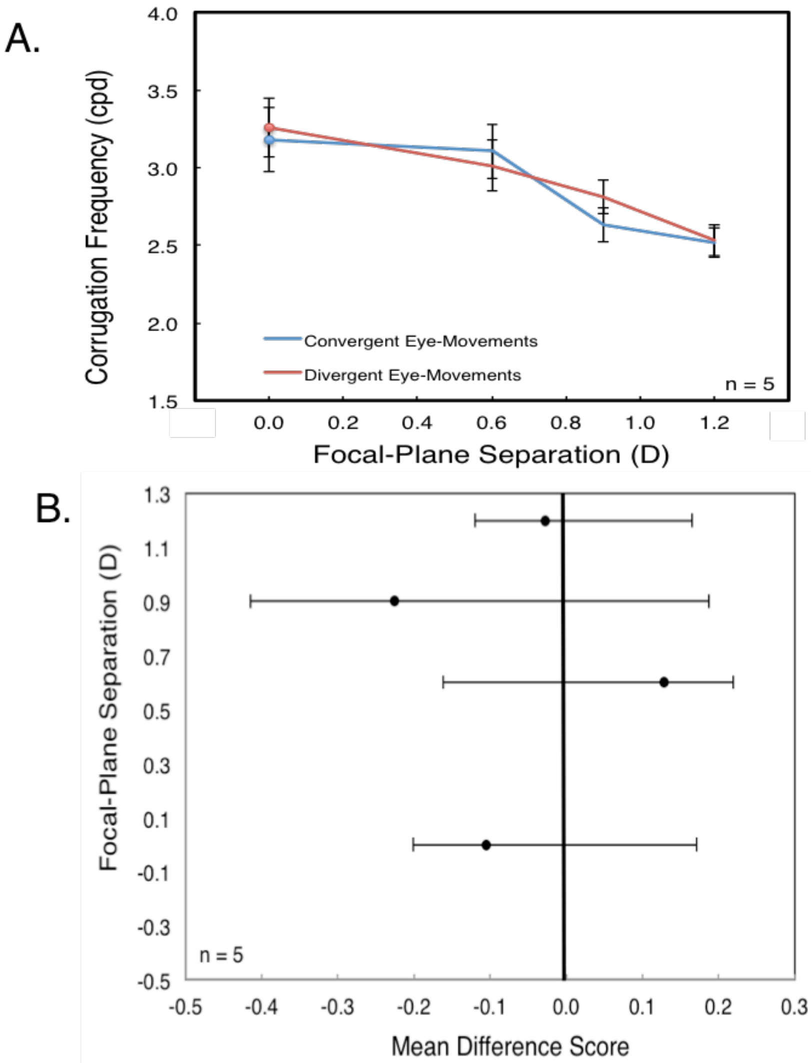


Figure 3.2.4 – (A) The effect of depth filtering on stereoresolution with direction of response (inward vs. outward). Divergent (red lines) and convergent (blue lines) trials with focal-plane separation. The x-axis shows the focal-plane separation in dioptres (D) and the y-axis shows the stereoresolution thresholds. Zero on the x-axis and the solid data point denotes single-plane (i.e. real-world equivalent) stimuli. (B) The mean difference scores between the divergent and convergent response, across the focal-plane separations. The x-axis shows the mean difference score, where zero indicates the same threshold for both response types. The y-axis shows focal-plane separation in dioptres (D). In both graphs, error bars represent 95% Confidence Intervals.

Effects of viewing distance

Experiment 1 also examined whether the distance to depth-filtered stimuli (0.6 D focal-plane separation) affected stereo resolution. As a control, we examined stereoresolution for real-world stimuli at the same distances (and two nearer distances). Figure 3.2.5 shows the average of the observers' stereoresolution thresholds for 'real-world' (green) and depth-filtered (purple) stimuli at the same distances. The results show that stereoresolution was largely unaffected by distance, either for depth-filtered, ($F(4,16) = 0.86$, $p = .51$), or real-world ($F(6,24) = 0.43$, $p = .85$) stimuli. It is also evident that, as seen earlier, there were no significant differences between performance in the depth-filtered and equivalent real-world blocks ($F(4,16) = 1.37$, $p = .30$; with 0.6 D image-plane separation). This suggests our previous results are relevant at other distances within the display.

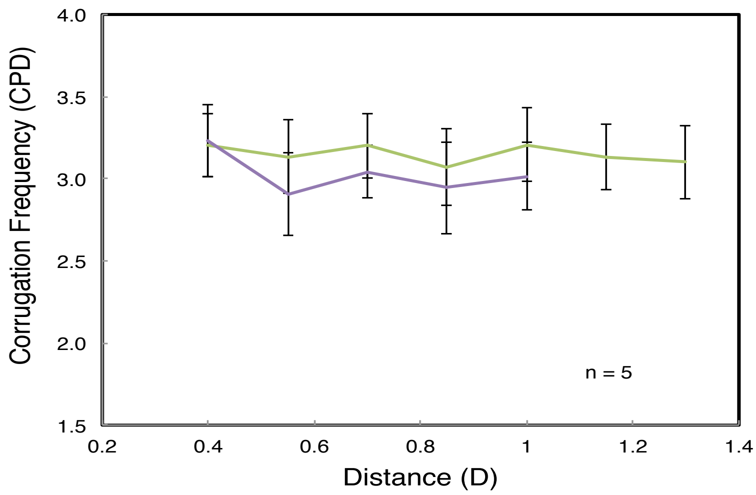


Figure 3.2.5. Effect of viewing distance on stereoresolution. 'Real-world' (green lines) and Depth-filtered (purple lines) trials with distance of the image from the viewer. The x-axis shows the distance of the images from the viewer in dioptres (D) and the y-axis shows the stereoresolution thresholds in cycles per degree. Error bars represent 95% Confidence Intervals.

We also compared the effects of stimulus distance from fixation on stereo performance for convergent and divergent response directions. This was to test whether the response type has an effect on performance throughout the task. Figure 3.2.6 shows the results for 'real-world' (top left graph) and for depth-filtered stimuli (bottom left graph). Due to the artefact at the farthest distances, there are no data for the 'real-world' data at the nearest divergent distances and the farthest convergent distances. However, the stereoresolution thresholds remain relatively similar for both response types and across distances from fixation. Similar to Figure 3.2.4, to establish whether there were any statistical differences between the thresholds for the divergent and convergent responses, we calculated the mean difference between the thresholds at each focal-plane separation. This was done for the real-world (top right) and depth-filtered (bottom right) stimuli. For the real-world stimuli, the mean difference was calculated for the distances where both divergent and convergent thresholds could be obtained. The error bars are 95% confidence intervals. As all of the error bars intersect 'zero' on the x-axis, we conclude that direction of the required response did not appear to significantly affect stereoresolution measurements. For the real-world stimuli, the error bars only just intersect the line at the 0.75 D distance from fixation. A paired-samples t-test showed that the thresholds for the convergent and divergent responses were non-significant, t(4) = -0.51, $p = 0.64$.

Therefore, we conclude, there was no effect of response direction or stimulus distance on stereoresolution thresholds for both types of images. This is a relatively unsurprising result as the stimulus presentation time during this experiment (2 seconds) is adequate fusion time for both the convergent and divergent response systems regardless of distance. Thus, the fast fusional response should already have been made irrespective of whether the response was convergent or divergent (Schor, 1979).

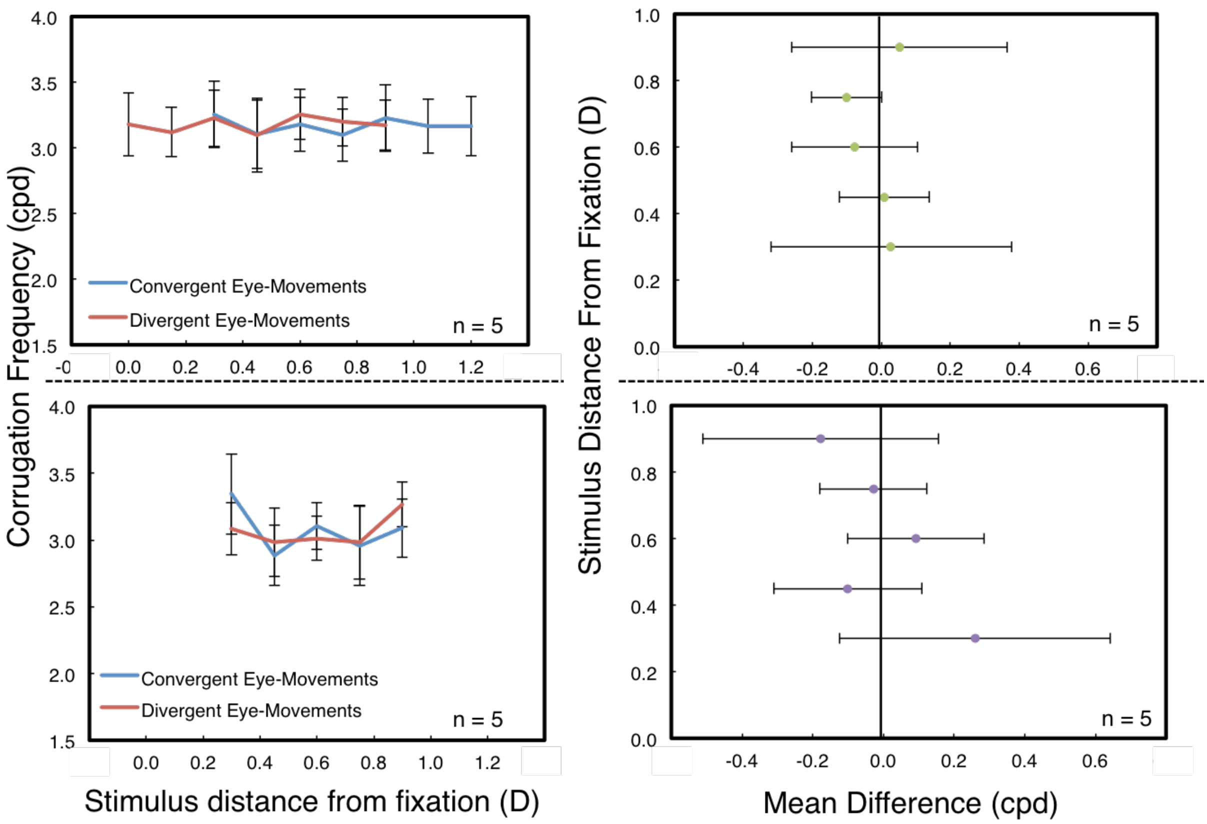


Figure 3.2.6 – The left graphs show the effect of stimulus distance from fixation on stereo resolution for different response directions. The top left graph shows the data from 'real-world' images, the bottom left graph shows the depth-filtered data. The blue lines denote convergent responses and the red lines denote divergent responses. The x-axis shows the distance of the stimuli from the viewer in dioptres (D) and the y-axis shows the stereoresolution thresholds in cycles per degree. The graphs on the right show the mean difference score between the divergent and convergent response, across the focal-plane separations, for the corresponding graph on the left. In all graphs the error bars represent 95% Confidence Intervals.

Summary of results:

Overall, it appears that the increasing reduction in retinal-image contrast caused by increasing focal-plane separation systematically decreases the finest modulations in stereoscopic depth that can be perceived. Our results are also consistent with those of Hoffman et al. (2008), who showed that a 0.67 D focal-plane separation results in comparable thresholds to a single-plane image (albeit at a different viewing distance). Our experiment sampled focal-plane separations fairly coarsely, and so it is difficult to know precisely at what separation a significant reduction in stereoresolution occurs, compared to real-world equivalent images. We revisit this question in Experiment 4.

We also found that for focal-plane separations of 0.6 D stereo resolution was relatively unaffected by the distance of the depth-filtered stimuli to the observer. This suggests that our main findings are relevant for other distances, not just 0.7 D. Finally, we found that stimulus distance (real-world and depth-filtered) had no effect on response type with convergent and divergent response types having very similar thresholds.

3.3 Experiment 2: Time-to-Fuse with depth-filtered images.

Here we examined the time required to fuse stereoscopic depth-filtered images as a function of focal-plane separation, and viewing distance. The blocks were the same as those used in Experiment 1 (see Figure 3.2.1). We again included both convergent and divergent response directions, as this was especially important here given the previously observed differences in the overall response times of convergence and divergence eye movements (Semmlow et al., 1986), which may affect our overall measure fusion time.

3.3.1 Method

Observers

There were four observers in this experiment, three of which took part in the previous experiment (LR, AB, and AD) along with AE (21 years old). At the time of this experiment, none of the observers required optical correction.

Procedure

The stimuli, procedure and task were the same as Experiment 1, with the following exceptions. To measure time-to-fuse we fixed the stimulus corrugation frequency and peak-to-trough disparity, and varied the stimulus duration. The corrugation frequency was set at 1.35 cpd and the peak-to-trough disparity at 1.8 arcmin. These values were chosen to ensure the stimulus was above threshold with sufficient viewing time (i.e. the limiting factor was the time taken to fuse the stimulus). The initial stimulus presentation time was chosen at random between 20 and 120 msec (uniform distribution). The stimulus presentation time varied according to a 2-up, 1-down adaptive staircase. In this experiment, ‘up’ refers to making the task more difficult by decreasing the response time. For this to happen, the observer had to be correct on two consecutive trials for that staircase. There was a step size of 30 ms that was halved after four reversals. All other details, including the fixation target, remained the same.

Analysis

As with Experiment 1, we fit psychometric functions using the combined data from the two staircases for the divergent and convergent responses as well as analysing the data separately, to examine any effects of divergent or convergent responses.

3.3.2 Results

We again computed a single threshold for the duplicated ‘real-world’ (single-plane) stimuli at 0.7 D, 1.15 D and 1.3 D by averaging across thresholds measured in multiple blocks (Figure 3.2.1.). Figure 3.3.1 compares time-to-fuse for these duplicate blocks.

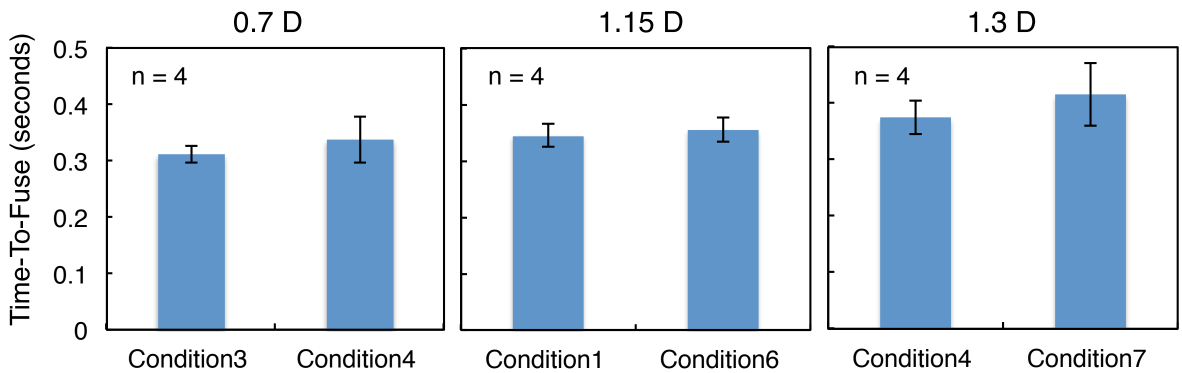


Figure 3.3.1. – TTF thresholds for repeated measurements of real-world stimuli**.** The data are the average of all observers’ data in each block (n = 4). The title of each panel denotes the different stimulus distances. The bars show different blocks (labelled on the x-axis) in which a given distance was presented. The y-axis is the TTF threshold, in terms of seconds. The error bars represent 95% confidence intervals.

As with the previous experiment, a paired samples t-test showed that there were no significant differences across conditions at each distance. The results are shown in Table 3.3.1. It can be seen that, as with stereoresolution, time-to-fuse did not differ substantially across duplicates, and so computing a single average was meaningful.

Condition	Mean	SE	Significance
3	0.311	0.098	t(3), -1.019, p = .38
4	0.347	0.026	
1	0.345	0.014	t(3), -0.90, p = .44
6	0.359	0.013	
4	0.374	0.019	t(3), -1.969, p = .14
7	0.416	0.036	

Table 3.3.1. – Table showing t-test results for each condition**.**

Effect of focal-plane separation

We analysed the data for Experiment 2 in a similar manner to Experiment 1. Figure 3.3.2 plots the TTF thresholds as a function of focal-plane separation. The top-left panel shows the average across all individuals, and the remaining panel shows each individual's data. A lower threshold indicates better performance, as less time is required to discriminate depth in the stimulus. A repeated-measures ANOVA showed a significant difference of focal-plane separation size on time-to-fuse, F(3,9) = 7.252, p = .01. However, Bonferroni-Corrected pair-wise comparisons showed non-significant differences across all conditions. This could be a result of reduced statistical power as a result of our small sample size. We would therefore, recommend further testing with larger samples to establish significance.

Examining the graphs, the results show no real change across focal-plabe separations in the time taken to fuse the images. The data for the different individuals are very similar, except for 'AD' whose time-to-fuse thresholds were appreciably longer at the largest focal-plane separation.

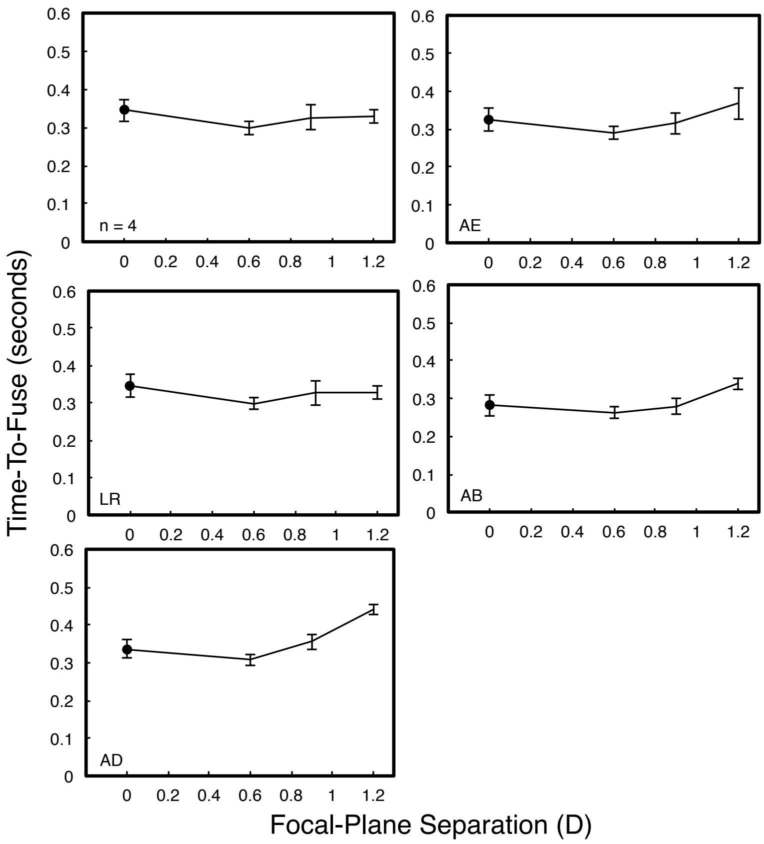


Figure 3.3.2 – TTF thresholds as a function of focal-plane separation. The top left panel shows mean thresholds, across all observers and the other four panels show individual thresholds. The observers' initials are displayed on the bottom left of each graph. The x-axis shows the focal-plane separation in dioptres (D). Zero on the x-axis and the solid data point denotes single-plane (i.e. real-world equivalent) stimuli. The y-axis is the stereoresolution threshold in terms of corrugation frequency. Error bars for the average graph represent 95% Confidence Intervals. The error bars for the individual graphs denote standard error of the threshold estimate (i.e. standard error of the psychometric function fit).

We also examined the effect of focal-plane separation on time-to-fuse thresholds separately for convergent and divergent response directions. The results are shown in Figure 3.3.3, plotted in the same format as the previous figure. Here, we plotted the individual data as well as the average data (Figure 3.3.3, top left panel). We did this here because of the increase in time-to-fuse at the largest plane separation from observer 'AD'. The increased fusion time from observer 'AD' may have been driven by one or other response type (divergent/convergent). However, it can be seen that there was no effect of response direction on time-to-fuse (see below). This is also true for observer 'AD', who experienced elevated fusion times for both divergent (red) and convergent (blue) responses at the largest plane separation. As in Experiment 1, we examined whether there was a difference in convergent and divergent responses by plotting the mean differences in TTF for each focal-plane separation. This is plotted beside the average data (Figure 3.3.3, top right panel). The error bars are 95% confidence intervals. As all of the error bars intersect 'zero' on the x-axis, we conclude that direction of the required response did not appear to affect stereoresolution measurements.

These results are encouraging for multi-plane displays because they suggest that stereoscopic fusion time is not affected by the reduced retinal-image contrast caused by depth filtering. We know from Experiment 1 that depth filtering results in reduced stereoresolution thresholds with focal-plane separations larger than 0.6 D. In this experiment, the depth-filtered images created with focal-plane separations of 0.9D resulted in similar fusion times to 'real-world' stimuli. This suggests that TTF is not the limiting factor for determining focal-plane separation.

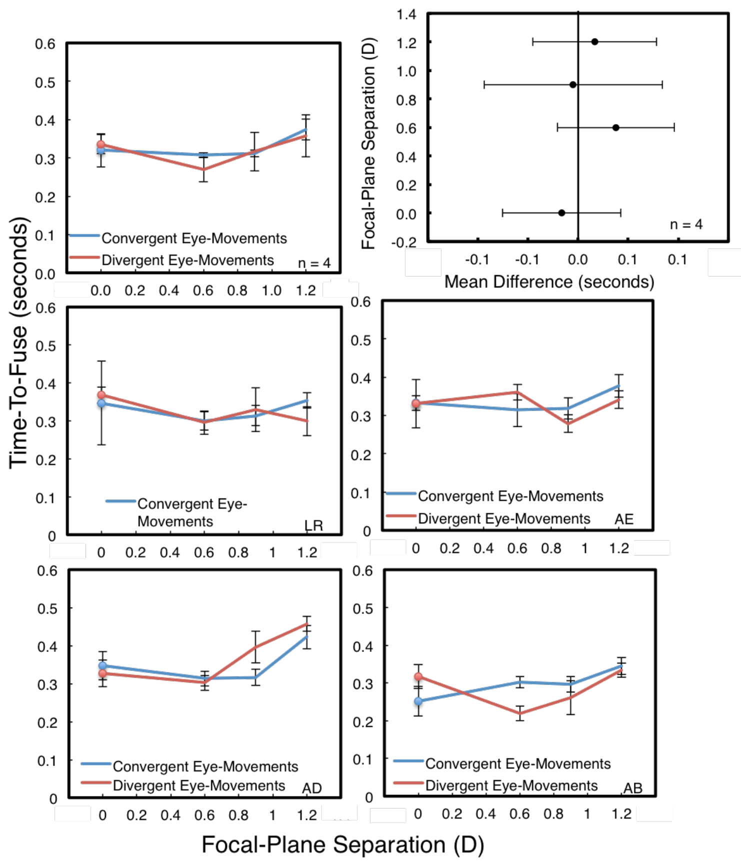


Figure 3.3.3 – The effect of depth filtering on TTF with direction of response (inward vs. outward). The top left panel shows mean thresholds, across all observers, while the top right graph shows the mean difference scores between the divergent and convergent response, across the focal-plane separations. The other four panels show individual thresholds. The observers' initials are displayed on the bottom right of each graph. Mean and individual graphs show the divergent (red lines) and convergent (blue lines) trials with focal-plane separation. The x-axis shows the focal-plane separation in dioptres and the y-axis shows the TTF thresholds in seconds. Zero on the x-axis and the solid data point denotes single plane (i.e. real-world equivalent) stimuli. For the mean difference graph, the x-axis shows the mean difference score, where zero indicates the same threshold for both response types. The y-axis shows focal-plane separation in dioptres (D). In both graphs, error bars represent 95% Confidence Intervals. While the error bars for the individual graphs denote standard error of the threshold estimate (i.e. standard error of the psychometric function fit).

In some ways, these results are not surprising. As the spatial frequency of the images was above threshold (1.35 cpd), only the retinal-image contrast attenuation as a result of increased focal-plane separation could have affected the fusion time when viewing depth-filtered stimuli. This decrease in retinal-image contrast with increasing focal-plane separation can also be thought of as increasing retinal blur (attenuation of high spatial frequencies). Limits of stereoscopic fusion (Panum's fusion area) have in fact been shown to *increase* with retinal blur resulting in a larger fusional area for lower spatial frequencies (Tyler, 1975; Schor & Tyler, 1981; Schor, Wood, & Ogawa, 1984). One would therefore not necessarily expect worsened stereoscopic fusion performance with depth-filtered images. Perhaps a more surprising result, however, is that we did not see slower times in the divergent response direction, this is explored in light of all the data below.

Effects of viewing distance

We also examined the effect of stimulus distance on fusion time, for real-world and depth-filtered images (0.6 D focal-plane separation). The results are plotted in Figure 3.3.4. It appears there is a very slight trend for longer fusion times with nearer viewing distances. This effect is not very large, however, with the difference in fusion time between the farthest (0.4 D) and the closest (1.3 D) distance being less than 0.1 seconds. Again, there were also no appreciable differences here between time-to-fuse thresholds for the real-world ($F(6,18) = 0.57$, $p = .75$) and depth-filtered blocks ($F(4,12) = 1.04$, $p = .41$), with a 0.6 D focal-plane separation. It is also evident that, as seen earlier, there were no significant differences between performance in the depth-filtered and equivalent real-world blocks ($F(4,12) = 0.66$, $p = .63$; with 0.6 D image-plane separation). These results suggest that the time required for stereoscopic fusion to occur is largely unaffected by viewing distance.

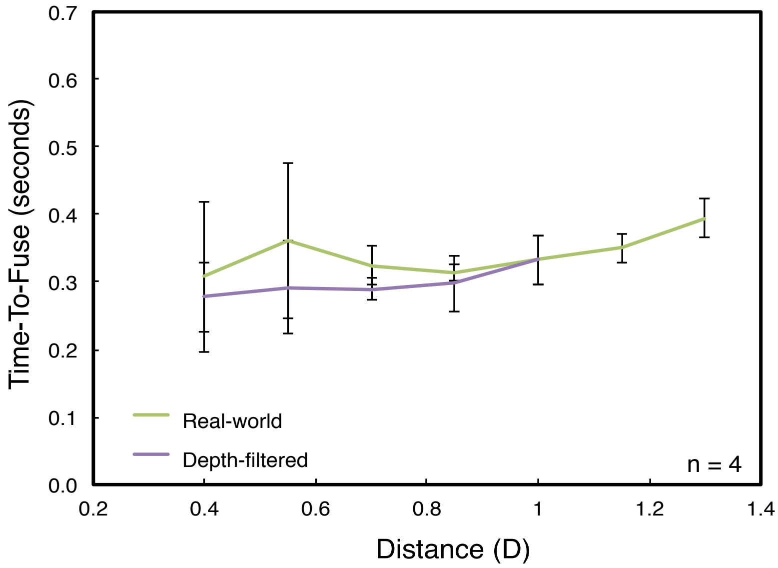


Figure 3.3.4 – Effect of viewing distance on TTF. The green lines represent the 'real-world' image locations, the purple represent the depth-filtered image locations. The x-axis shows the distance of the images from the viewer in dioptres (D) and the y-axis shows the time-to-fuse thresholds in cycles per degree. Error bars represent 95% Confidence Intervals.

Finally, we compared the effects of stimulus distance from fixation on fusion time for convergent and divergent response directions. Again, this tests whether the response type has an effect on performance throughout the task. We would expect response type to have an effect in this experiment, given the differences in neural control between the convergent and divergent responses, leading to differences in overall response time (Semmlow, Hung, & Ciuffredo, 1986; Semmlow & Wetzel, 1976; Hung, Zhu, & Ciuffreda, 1996). As with Experiment 1, we measured this for both response types and for the 'real-world' (top panel Figure 3.3.5) and the depth-filtered stimuli (bottom panel). We plotted the mean differences between the response types in order to establish any significant differences between them. This was done for the real-world (top left graph) and depth-filtered (bottom left graph) stimuli, for the distances at which both response types were measured. For the real-world data, there are no significant differences between response types, as all of the error bars intersect 'zero' on the x-axis. However, for the depth-filtered stimuli, not all error bars overlap. For the 0.9 D distance from fixation, there is a significant difference between the convergent ($M = 0.36$, $SD = 0.49$) and divergent ($M = 0.26$, $SD = 0.071$) responses, $t(3) = 3.591$, $p = .037$. This was an unexpected result given the previously reported research stating that the divergent response tends to be slower than the convergent (Semmlow et al., 1986). The result is also surprising as this was not the case for the real-world data at the same distances. It should be noted, however, that the difference between the two response types is less than 0.2 seconds. This, taken with the previous results, might indicate that although significantly different at far distances, the difference in response types may not have affected performance throughout the main measure.

Unfortunately, the artefact at the far distances limits the 'real-world' data (top graph). Therefore, it is unknown whether, similar to the convergent (blue) responses, the fusion time for the divergent response (red) increases with increased stimulus distance from fixation.

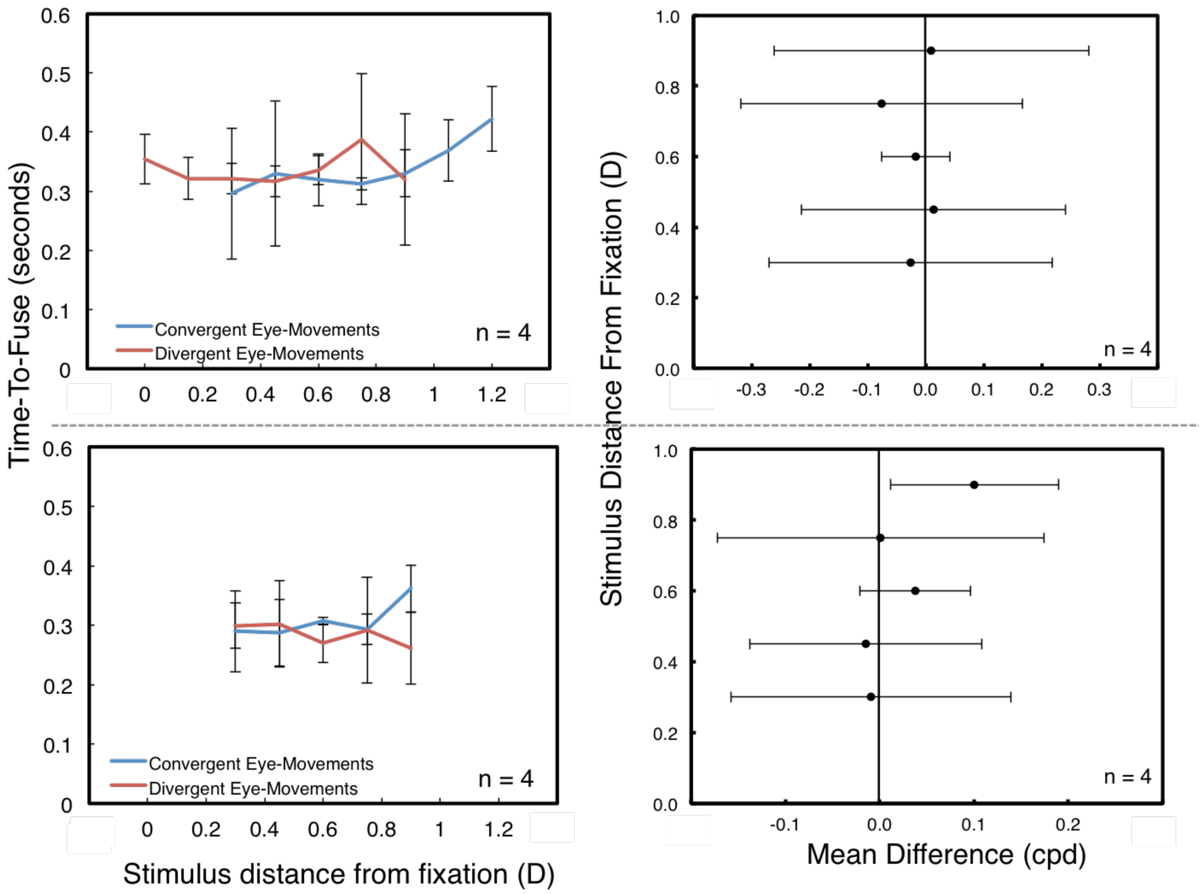


Figure 3.3.5 – The left graphs show the effect of stimulus distance from fixation on stereo resolution for different response directions. The top left graph shows the data from 'real-world' images, the bottom left graph shows the depth-filtered data. The blue lines denote convergent responses and the red lines denote divergent responses. The x-axis shows the distance of the stimuli from the viewer in dioptres (D) and the y-axis shows the stereoresolution thresholds in cycles per degree. The graphs on the right show the mean difference score between the divergent and convergent response, across the focal-plane separations, for the corresponding graph on the left. In all graphs the error bars represent 95% Confidence Intervals.

Summary of results:

Overall, there is little evidence to suggest that when focal-plane separations increase, the time to fuse depth-filtered images also increases. This was also true for the convergent and divergent responses with focal-plane separation. We also examined the effect of stimulus distance on time-to-fuse. We found no significant effect of stimulus distance on the depth-filtered or real-world stimuli. Again, this indicates that our results are relevant for all distances within the display. Finally, we looked at the effect of response type (divergent/convergent) on time-to-fuse as a function of stimulus distance from fixation. We found no significant differences for the real-world stimuli. However, there was a significant difference at a distance of 0.9 D for the depth-filtered stimuli, with the convergent response being significantly slower than the divergent. This is surprising given the previous reported results but also the previous research, which states that divergent responses are slower than convergent (Semmlow et al., 1986).

3.4 Experiment 3: Stereoacuity with Depth-Filtered Images.

In this experiment, we again investigated the effect of focal-plane separation for depth-filtered images, and viewing distance, on stereo performance. Here, however, we assessed performance in terms of stereoacuity, which is the minimum detectable depth from binocular disparity. Our primary aim is to evaluate the efficacy of depth-filtered images. However, we again measured effects of distance, in part because there is an existing literature on this question, suggesting stereoacuity may change with distance (see Discussion).

Binocular disparity is often thought of as only being useful at near distances (e.g. Palmer, 1999), because of the previously mentioned rapid fall-off in the magnitude of disparities created by a given interval in depth (see General Method). Large depths, at large viewing distances, can however, result in binocular disparities that are, in principle, detectable. Researchers have pondered, therefore, whether disparity sensitivity at far distances is limited by low-level disparity-based mechanisms (i.e. the magnitude of disparities per se), or other factors including the ability to represent depths perceived as being at far distances. In the former case, one would predict identical stereoacuity thresholds *in units of binocular disparity* (Bradshaw & Glennerster, 2006). In the latter case, predictions are harder to specify, but one might expect differences in stereoacuity at different distances. For instance, if disparity sensitivity were limited by both disparity and the perceived depth information in the stimulus, then minimum detection thresholds plotted in terms of disparity would be smaller at farther viewing distances and larger at nearer viewing distances (Bradshaw & Glennerster, 2006).

We therefore intend to measure the amount of peak-to-trough disparity required to recognize depth in the stimulus and how that changes with distance. Given we measure the ‘real-world’ response at a range of distances when measuring the effect of depth-filtered images with focal-plane separation, we thought this would be an interesting question to answer, although not directly related to our main research.

Thus, in this experiment, we measure the effect of depth filtering on stereoacuity as a function of focal-plane separation. We also measure the effect of distance on the binocular disparity required to recognize depth in 'real-world' images.

3.4.1 Method

Observers

There were three observers in this experiment: LR (author, 23 years old), AF (19 years old), and AG (22 years old). At the time of the experiment, none of the observers required optical correction.

Stimuli

In this experiment, we wanted to measure stereo performance at the far distances, 0.25 D and 0.1 D. Given the artefact in the random-dot stereograms at this distance (which we later solved; see General Methods chapter), we measured stereoacuity here using two fronto-parallel, rectangular planes, defined by random dots. The task was to report whether the smaller (target; 3 deg high and 8 deg wide) rectangle was in front of, or behind the larger (reference; 8 deg high and 8 deg wide) rectangle. A cartoon of the stimulus is shown in Figure 3.4.1. Panel A shows the target rectangle in front of the reference rectangle while panel B shows the target rectangle behind the reference. A virtual mask was applied to the stimulus, creating a stimulus with a width of 4 deg. This prevented visible differences in the outline contours when the vergence distance to the smaller rectangle changed, which could have provided a monocular cue to the depth sign. Monocular artefacts in the stimulus were again tested for using naïve and experienced observers, and monocular performance was always at chance.

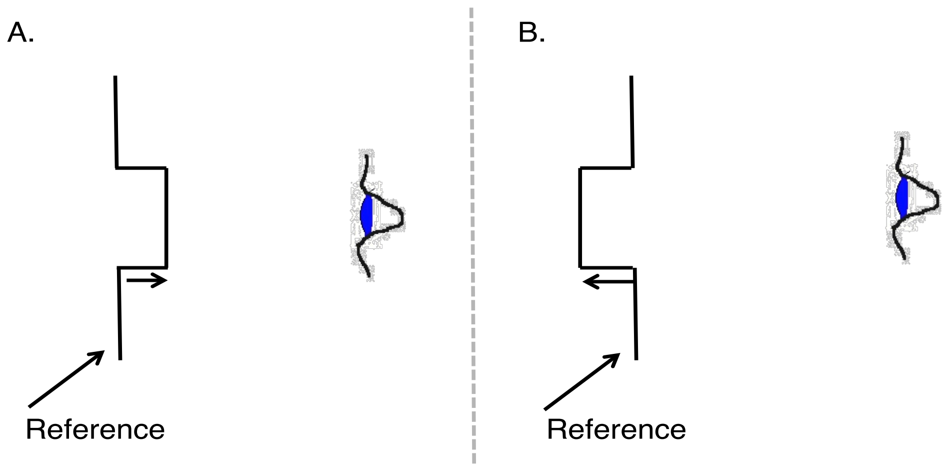


Figure 3.4.1. –Example of the stimulus used in Experiment 3. A. The smaller rectangle appears in front of the reference rectangle. B. The smaller rectangle appears behind the reference rectangle.

Sessions

The image presentation sessions used in this experiment are shown in Figure 3.4.2. These are a subset of those used in Experiments 1 and 2. We measured stereoacuity for depth-filtered images at the same focal-plane separations as previously examined. However, given our findings in the previous experiments relating to the effect of stimulus distance on stereo performance, we did not examine in detail the effect of viewing depth-filtered stimuli at different distances within the display. Instead we included three blocks (1, 2 and 3; Figure 3.4.2) in which the location of the depth-filtered stimulus varied with a focal-plane separation of 0.6 D. Real-world stimuli at 0.7 D were measured in blocks 1 and 2, serving as a control for the depth-filtered images.

As before, where the focal-plane spacing was non-zero, we used depth filtering to simulate a focal distance at the dioptric midpoint between the two focal planes (that is, a 50:50 ratio of image intensity on each focal plane). Again, we used the dioptric mid-point because this should be a 'worst-case scenario' for stereo performance from depth-filtered images. Finally, in every condition, we presented 'real-world' images on the near and far focal-planes, acting as 'filler' trials as well as controls for depth-filtered stimuli at the same location.

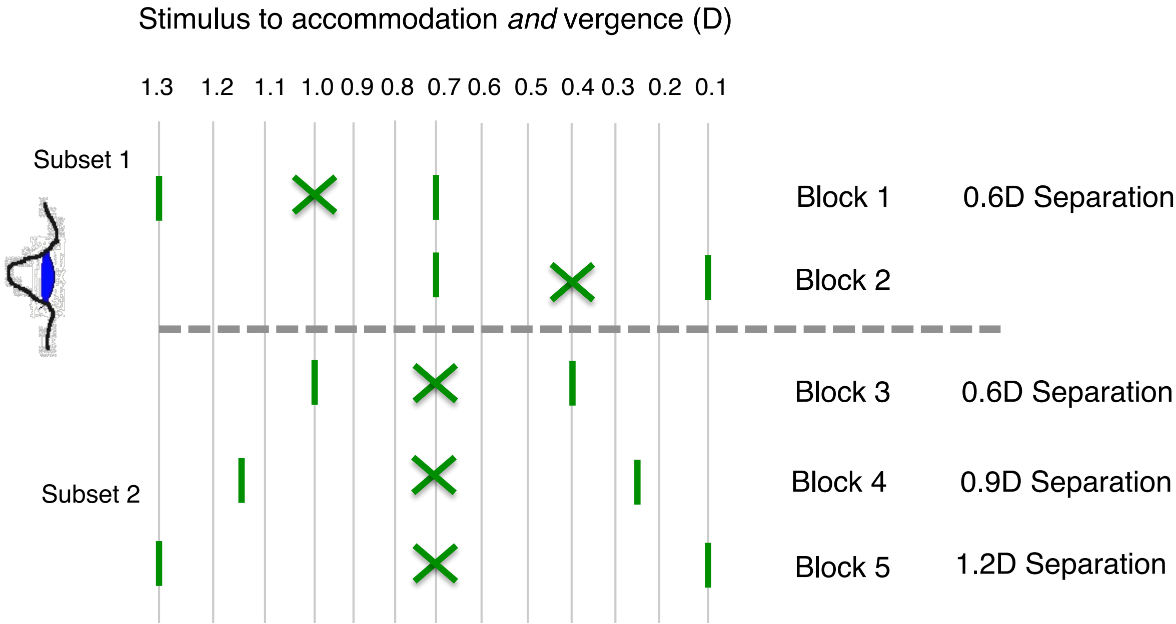


Figure 3.4.2 – The blocks in Experiment 3 The figures at the top denote the distance of the stimulus (i.e. stimulus to accommodation *and* vergence) from the eyes (shown on left of the diagram). The vertical green lines represent the positions of focal planes in each block. The green crosses denote a depth-filtered image location. The focal-plane separations in each case are shown in the far-right column. Blocks 1 and 2 had the same focal-plane separation (0.6 D) but at different distances. Depth-filtered stimuli in blocks 3, 4 and 5 were at the same simulated distance, but with varying focal-plane separations.

Procedure

The procedure was similar to Experiments 1 and 2 except here we fixed stimulus duration, and varied the disparity step in the stimulus to measure the minimum detectable level. Each block of trials consisted of multiple interleaved staircases. There were separate adaptive staircases for each image distance (two for each focal-plane location and one depth-filtered image) and for fixation position (requiring a divergent or convergent response), making six staircases in all. For each stimulus type, the level of depth in the stimulus was controlled using an adaptive staircase procedure, using a 1-up, 2-down reversal rule. In this experiment, 'up' refers to increasing the disparity and therefore decreasing the difficulty, this occurred after one incorrect response. The initial step-size was 5 arcsec, which was halved after four reversals. Threshold was defined as the stereoacuity (in seconds of arc) at which the perceived direction (in front of or behind the larger rectangle) was correctly identified on 75% of trials. Each block was repeated four times in a single session.

3.4.2 Results and Discussion

As with the previous Experiments, we computed a single threshold for the duplicated 'real-world' (single-plane) stimuli at 0.7 D, 1.3 D and 0.1 D by averaging across thresholds measured in multiple blocks (Figure 3.4.2.). Figure 3.4.3 shows differences between all duplicates.

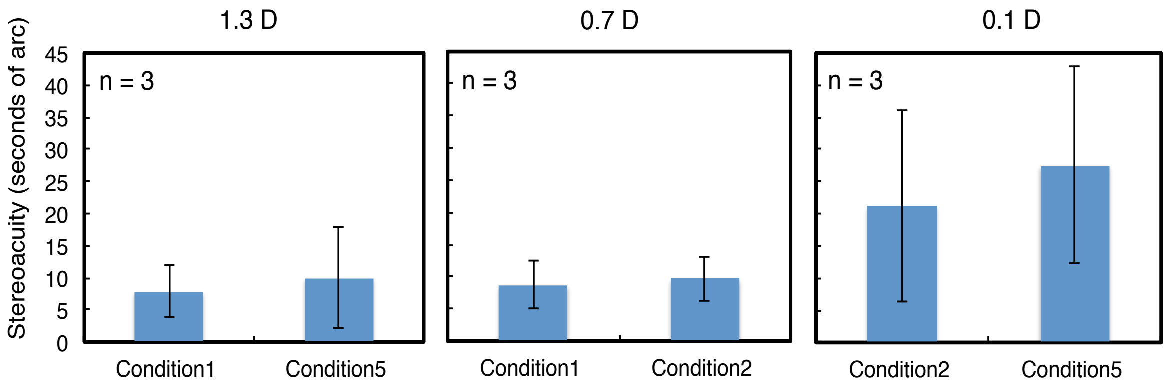


Figure 3.4.3 – Stereoacuity thresholds for repeated measures of real-world stimuli. The data are the average of all observers' data in each block (n = 3). The title of each panel denotes the different stimulus distances. The bars show different blocks (labelled on the x-axis) in which a given distance was presented. The y-axis is the stereoacuity threshold, in terms of seconds of arc. The error bars represent 95% confidence intervals.

Paired sample t-tests were conducted to compare the thresholds of each condition. There were no significant differences between the conditions for all distances (see Table 3.4.1). We therefore computed a single threshold for the real-world stimuli in each case by averaging across the two measurements.

Condition	Mean	SD	Significance
1/1.3 D	7.87	3.44	t(3), -0.801, p = .51
5/1.3 D	9.99	1.92	
1/0.7 D	8.70	5.52	t(3), 1.15, p = .37
2/0.7 D	9.70	4.43	
2/0.1 D	21.23	12.16	t(3), -1.44, p = .28
5/0.1 D	27.53	10.82	

Table 3.4.1. – Table showing t-test results for each condition.

Effect of focal-plane separation

Figure 3.4.4 plots the average (top left graph) and individual stereoacuity thresholds for depth-filtered and real-world stimuli as a function of focal-plane separation. Lower thresholds indicate better performance, as it means the observers required less disparity in order to perceive depth in the stimuli. The data show that for focal-plane separations up to 0.9 D, there was very little change in the average stereoacuity thresholds. However, thresholds are slightly elevated for the 1.2 D focal-plane separation, which appears to be driven by two of the three observers. The difference between the 'real-world' and 1.2 D separation for these two observers equates to a maximum threshold difference of around 13 seconds of arc. A repeated measures ANOVA showed no significance across focal-plane separation, $F(3, 6) = 3.57$, $p = .09$. This indicates that stereoacuity is not the limiting factor for depth-filtered images. However, depending on the application, stereoacuity may still have to be considered when using this depth-filtering technique at larger (greater than 0.9 D) focal-plane separations.

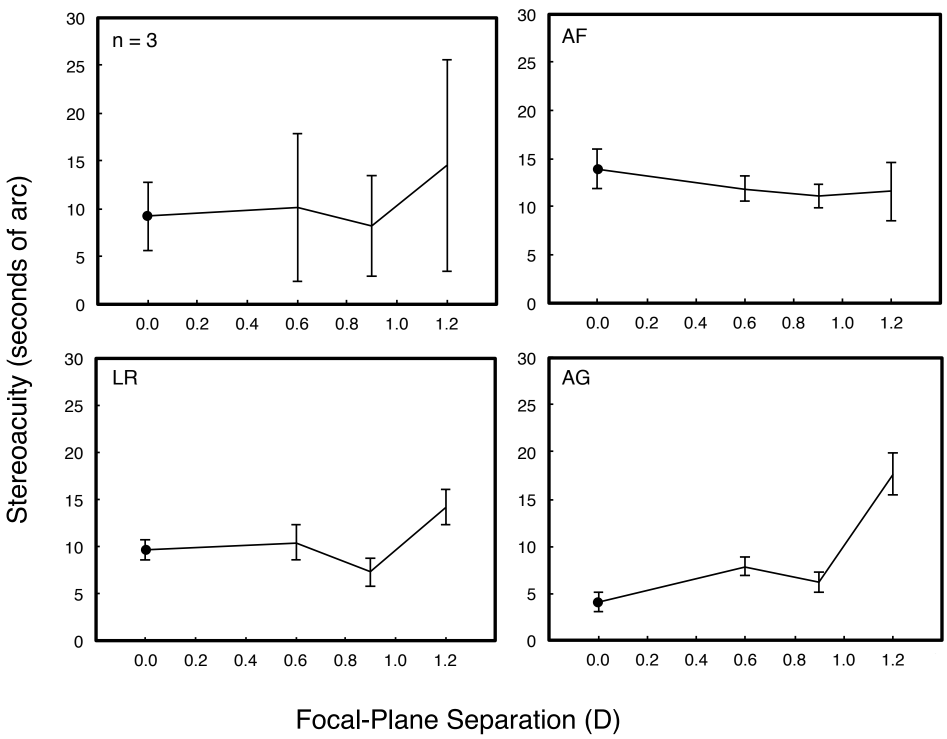


Figure 3.4.4 – Stereoacuity thresholds as a function of focal-plane separation. The top left panel shows the average stereoacuity thresholds and the other three show individual data. The observers initials are displayed on the top left of each graph The x-axis shows the focal-plane separation in dioptres (D). Zero on the x-axis and the solid data point denotes single-plane (i.e. real-world equivalent) stimuli. The y-axis is the stereoresolution threshold in terms of corrugation frequency. Error bars for the average graph represent 95% Confidence Intervals. The error bars for the individual graphs denote standard error of the threshold estimate (i.e. standard error of the psychometric function fit).

We then examined the effect of response direction (convergent vs. divergent) on stereoacuity at each focal-plane separation. As in previous experiments, we examined this for the average of the observers' data. In order to establish whether there was any statistical differences between the thresholds for the divergent and convergent responses, we calculated the mean difference between the thresholds at each focal-plane separation. These results plotted on the right in Figure 3.2.4 (b), suggest there was no difference between the stereoacuity thresholds for convergent vs. divergent responses.

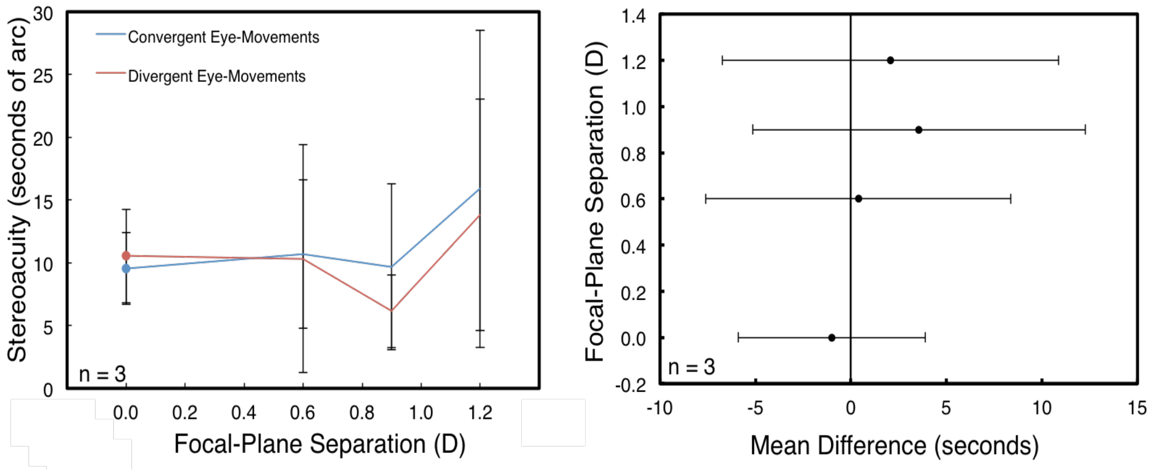


Figure 3.4.5 – Left Graph: Effect of focal-plane separation on stereoacuity with direction of response (inward vs. outward). The graph shows the convergent (blue) and divergent (red) trials with focal-plane separation. The x-axis shows the focal-plane separation in dioptres and the y-axis shows the stereoacuity thresholds in seconds of arc. Zero on the x-axis and the solid data point denotes single plane (i.e. real-world equivalent) stimuli. Right Graph: The mean difference scores between the divergent and convergent response, across the focal-plane separations. The x-axis shows the mean difference score, where zero indicates the same threshold for both response types. The y-axis shows focal-plane separation in dioptres (D). In both graphs, error bars represent 95% Confidence Intervals.

Effect of distance on stereoacuity

We also examined the effect of distance on stereoacuity thresholds, for the real-world stimuli (solid green lines, Figure 3.4.2). Figure 3.4.6 illustrates the average of the observers' stereoacuity thresholds (top left), as well as the individual data, as a function of distance. There is a general trend in the average data of increasing stereoacuity thresholds with increasing distance from the observer, which is reflected in the individual graphs to varying degrees.

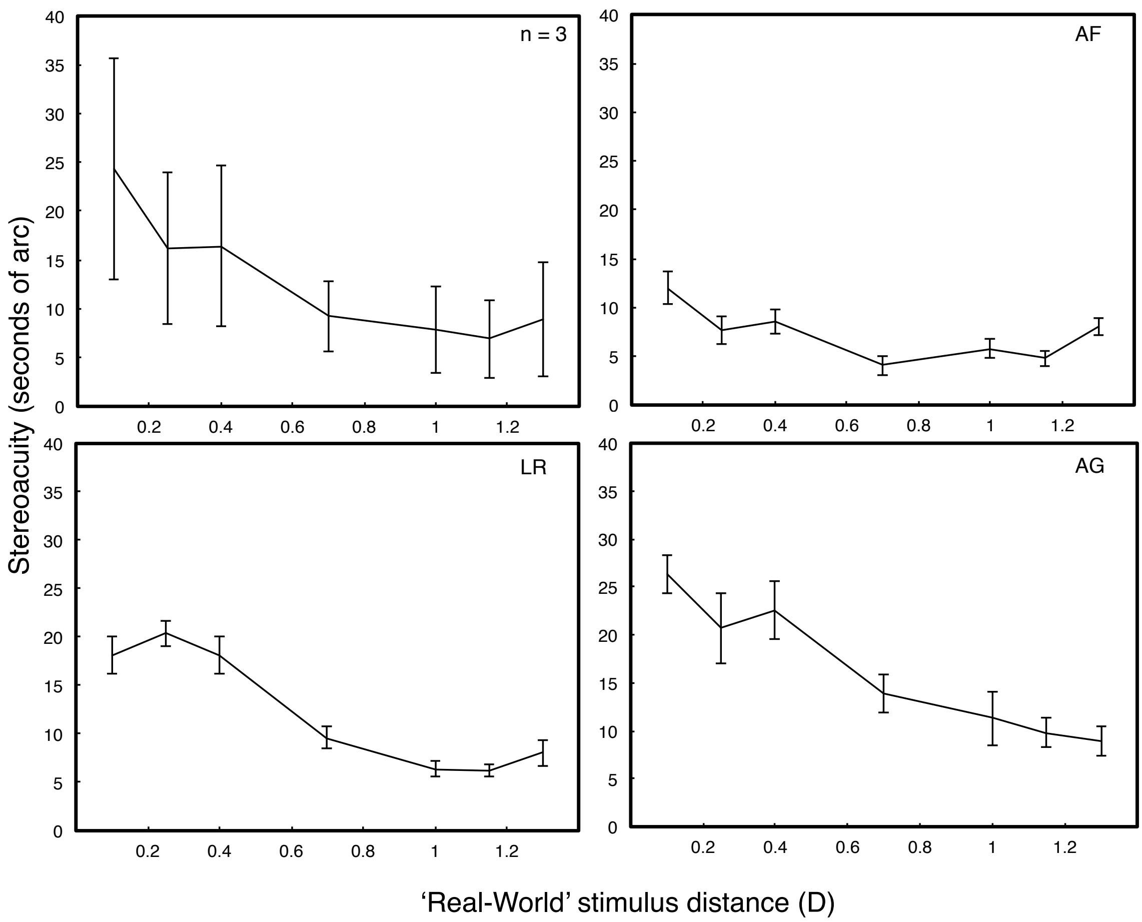


Figure 3.4.6 – Effect of distance on stereoacuity. The top left panel shows mean thresholds, across all observers and the other four panels show individual thresholds. The observers' initials are displayed on the top right of each graph. The x-axis shows the image location in dioptres and the y-axis shows the stereoacuity threshold in seconds of arc. Average error bars denote 95% confidence intervals. The error bars for the individual graphs denote standard error of the threshold estimate (i.e. standard error of the psychometric function fit).

For the average data, there is a significant effect of distance on stereoacuity threshold, $F(6, 12) = 8.33$, $p = <.01$. We conducted bonferoni corrected paired sample t-tests to make post hoc comparisons of the data. Table 3.4.2 show the significant comparisons from this test. These data show that significant differences occur between the near (within 0.7 D) and furthest (0.1 D) distances. When we consider the average data, this is an unsurprising result. This suggests that the amount of disparity in an image required to perceive depth increases at far distances. We consider the implications of this further, in the Discussion section.

Comparison	Mean	SD	Significance
1.3 D – 0.1 D	1.3 D: 8.93 0.1 D: 24.38	1.3 D: 1.58 0.1 D: 10.87	p <.01
1.15 D – 0.1 D	1.15 D: 9.20 0.1 D: 24.38	1.15 D: 2.58 0.1 D: 10.87	p <.01
1.0 D – 0.1 D	1.0 D: 7.82 0.1 D: 24.38	1.0 D: 3.03 0.1 D: 10.87	p <.01
0.7 D – 0.1 D	0.7 D: 9.20 0.1 D: 24.38	0.7 D: 4.95 0.1 D: 10.87	p <.01

Table 3.4.2. – Table showing significant post hoc t-test comparisons.

Summary of Results:

We examined the effect of focal-plane separation for depth-filtered images, and viewing distance, on stereo performance, specifically stereoacuity. We found no significant effect of focal-plane separation on stereoacuity, with only a slight elevation in thresholds at the 1.2 D separation. We also found no effect of convergent or divergent responses on stereoacuity thresholds with focal-plane separation.

Finally, we examined the effect of distance on stereoacuity thresholds. For the real-world stimuli. There was a significant effect of distance on stereoacuity thresholds, with increasing disparity required to recognise depth in the image with increasing distance.

3.5 Experiment 4: Stereo Spatial Resolution with Depth-Filtered Images: A Closer Investigation.

Across Experiments 1-3, the clearest effects of increasing image-plane separation were on stereoresolution. Effects on time-to-fuse were negligible, and effects on stereoacuity were rather small. It therefore seems that this aspect—the ability to resolve fine modulation in stereo depth—could be considered a limiting factor of stereo depth perception when specifying multi-plane displays. As discussed earlier, Experiment 1 sampled focal-plane separation rather coarsely, and so while we identified that it can cause impairments to performance, we could not precisely map out the function relating these factors, and so establish the maximum focal-plane separation that results in stereo performance that could be considered comparable to when viewing real-world stimuli. In this experiment, we therefore again investigated stereoresolution with depth-filtered stimuli as a function of focal-plane separation, but using much smaller increments in focal-plane separation compared to Experiment 1.

3.5.1 Method

Observers

There were four observers in this study, two of whom took part in Experiment 1 (LR (author, 22 years old; AD, 20 years old) as well as AF (18 years old) and AH (21 years old). At the time of this experiment, none of the observers required optical correction.

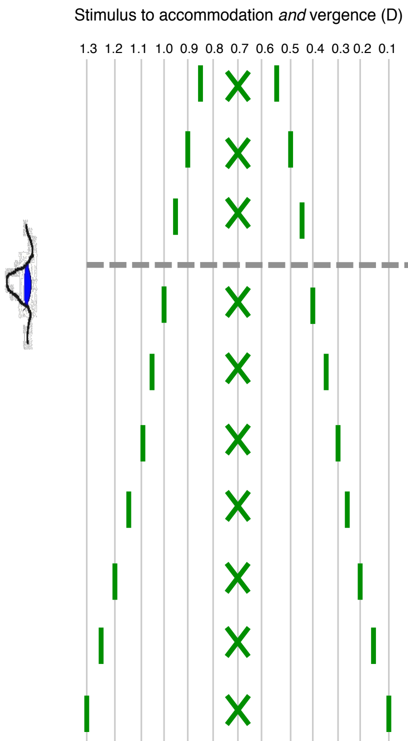
Stimuli

The stimuli and task were similar to Experiment 1. Based on the findings of Experiment 1, we knew that the viewing distance of the depth-filtered image did not affect the stereoresolution thresholds. As such, we presented all the stimuli ('real-world' and depth-filtered) at 0.7 D (the dioptric mid-point for the largest focal-plane separation in our display). We also used a Maltese cross as a fixation target in this experiment, presented at the same distance (0.7 D). By so doing, we hoped to eliminate eye movements and accommodation responses, thereby isolating the effects on stereoresolution of retinal-image contrast attenuation due to depth filtering.

Experimental Blocks

For blocks 1-3, (Figure 3.5.1) the separation between the focal-planes was physically too small to fit a third focal-plane. Therefore, to create the fixation in these blocks we used a depth-filtering technique (which we know from previous work supports accurate accommodation and vergence responses up to larger focal-plane separations; MacKenzie et al., 2010, 2012).

The image presentation blocks in this experiment are illustrated in Figure 3.5.1. We measured the effect of focal-plane separation with incremental steps of 0.1 D.



Blocks	Separation size	No. Focal-Planes
1	0.3 D	2
2	0.4 D	2
3	0.5 D	2
4	0.6 D	3
5	0.7 D	3
6	0.8 D	3
7	0.9 D	3
8	1.0 D	3
9	1.1 D	3
10	1.2 D	3

Figure 3.5.1 – The blocks in Experiment 4. The figures at the top denote the distance of the stimulus (i.e. stimulus to accommodation *and* vergence) from the eyes (shown on left of the diagram). The vertical green lines represent the focal planes in each block. The green crosses denote the simulated depth-filtered image location. The values to the right of the diagram state the block, focal-plane separation and the number of planes used in each case. Blocks 1-3 (above the dashed line) use 2 focal-planes while blocks 4-10 (below the dashed lines) use three focal-planes, on of which is at the 0.7 D location (at the crosses).

Procedure

The procedure was the same as Experiment 1 with some exceptions. In blocks 4-10 we used all three focal planes allowing us to interleave depth-filtered stimuli simulating 0.7 D (presented using the near and far planes) and real-world stimuli presented on the middle plane, positioned at 0.7 D. For blocks 1-3, only depth-filtered stimuli were presented. For all stimulus types, corrugation frequency was again varied using interleaved, 2-up, 1-down staircases. In this experiment, ‘up’ refers to increasing the task difficulty by increasing the corrugation frequency. For this to happen, the observer had to be correct on two consecutive trials for that staircase. The corrugation frequency used in the first trial was chosen at random between 0.5 and 3.5 cycles per degree (cpd). The step-size was 0.25 cpd until after the fourth staircase reversal, when it was halved to 0.125 cpd. Each staircase quit after 12 reversals and was repeated four times.

3.5.2 Results and discussion

The real-world (i.e. single plane) performance for images at 0.7 D was measured multiple times (blocks 4-10). To generate a 'real-world' stereoresolution threshold for the results, we examined the average of the observers' thresholds across blocks (see Figure 3.5.2). The results of a repeated measures ANOVA show that there was no significant difference in stereoresolution threshold across duplicate blocks, $F(6,18) = 1.14$, $p = .38$. We therefore averaged across the blocks to get a value for the 'real-world' stimuli.

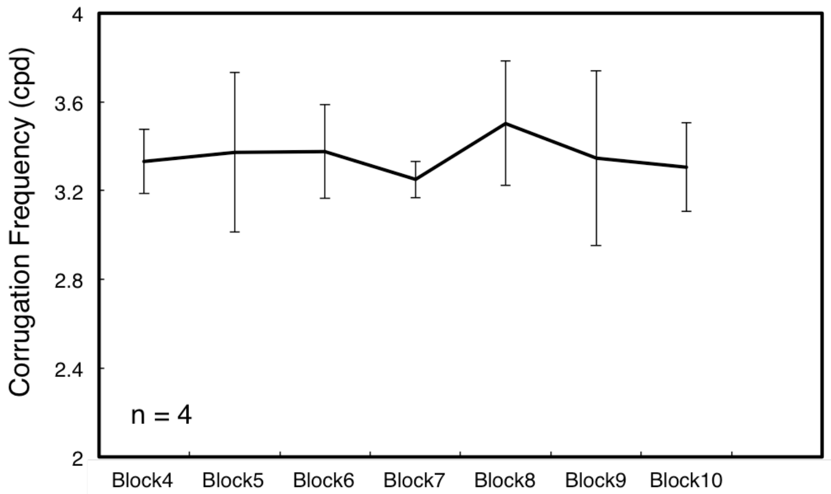


Figure 3.5.2 – Stereoresolution thresholds for repeated measures of 'real-world' stimuli. The data are the average of all observers' data in each block (n = 4). The x-axis shows the block and the y-axis shows the corrugation frequency thresholds for each block in cycles per degree. The error bars denote 95% confidence intervals.

Effect of Focal-plane Separation

Figure 3.5.3 plots stereoresolution thresholds as a function of focal-plane separation. Again, a higher threshold denotes better stereo performance as it indicates that finer modulations in depth could be reliably seen. This figure shows a similar overall decrement in performance to that seen in Experiment 1. Stereoresolution thresholds were highest for the real-world block (0 D focal-plane separation) and decreased thereafter. The red zone in the figure shows the 95% confidence intervals around the real-world threshold, for comparison to other blocks. It can be seen that, at least for these observers, the stereoresolution thresholds remained at a similar level between 0.6 and 0.9 D focal-plane separations and only fell off significantly for larger focal-plane separations. Of course, these data must be treated with caution, given the small number of observers. However, a repeated measures ANOVA was conducted to measure significant differences across focal-plane separation. The results suggest a significant effect of focal-plane separation, $F(10,30) = 3.326$, $p < .01$. To understand what focal-plane separations were these significant, a bonferoni corrected paired sample t-test was conducted to make post hoc comparisons of the data. The results show that there were significant differences between the real-world (M = 3.35, SD = 0.23) and 1.0 D (M = 2.63, SD = 0.27, $p = .03$) and 1.2 D (M = 2.698, SD = 0.36, $p = .05$) separations. This suggests that stereoresolution may not fall off appreciably until focal-plane separations larger than 0.9 D, and that at 0.6 D (which would still permit a practical display) it remains equivalent to real-world performance.

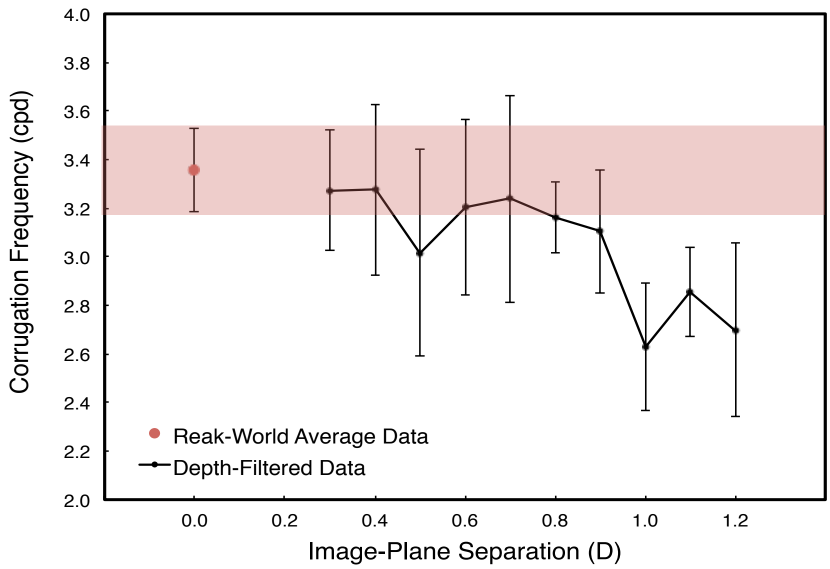


Figure 3.5.3 – Stereoresolution threshold as a function of focal-plane separation in Experiment 4. The blue line shows performance on the depth-filtered images while the single red data point represents the averaged data for the 'real-world images, across the multiple blocks. The data are the average of the observers' threshold for each block. The x-axis shows the focal-plane separation in dioptres and the y-axis shows the corrugation frequency thresholds for each block in cycles per degree. The error bars denote 95% confidence intervals.

Summary of results:

In this experiment, we re-examined stereo resolution as a function of focal-plane separation with smaller increments than what was examined in Experiment 1. Our results show that there are significant differences in threshold when viewing the real-world images compared to depth-filtered images with focal-plane separations larger than 0.9 D. This is good news for our display and the depth-filtering technique, meaning that performance is only affected relative to real-world equivalents for quite large separations.

3.6 Experiment 5: How does Depth-Filtering Affect Appearance?

So far the efficacy of depth filtering has been explored in terms of its effect on vergence and accommodation responses, and stereo performance (stereo resolution, time-to-fuse, and stereoacuity). Here, we conclude our investigations into depth filtering by examining the appearance of the resultant images in terms of the level of blur perceived. The aim is to measure at what focal-plane separation depth-filtering results in a detectable difference compared to 'real-world' images.

There has been no research conducted measuring the effect of the depth-filtering technique and focal-plane separation on the perception of blur in the resultant images. We know from MacKenzie et al.'s (2010) analysis of retinal-image formation with depth-filtered images that retinal contrast is increasingly attenuated when viewing depth-filtered images, particularly at high spatial frequencies. That is, fine/sharp detail is lost. Given this analysis, we expect blur in depth-filtered images may be evident for focal-plane separations smaller than 0.6 D, before the effects on stereo performance appear.

Just because the depth-filtered image may appear detectably blurry relative to the real-world images, it does not mean the resulting blur is necessarily troublesome or objectionable. Indeed, for many stereo 3-D applications there may need to be a trade-off between eliminating the vergence-accommodation conflict, while presenting acceptable image quality, and the required fine detail to complete the task being carried out. To explore this, we measured both objective and subjective aspects of the perceived appearance of depth-filtered images, as a function of focal-plane separation. We measured at what point depth-filtered images could be discriminated from real-world (single-plane) images, and asked subjective questions about viewing comfort, satisfaction, and perceptual quality of simple images. Thus, we examine at what focal-plane separation the depth-filtered images are noticeably different to real-world images, and at what separation this difference becomes objectionable.

3.6.1 Method

Observers

There were seven observers in this experiment: AI (20 years old), AJ (20 years old), AK (26 years old), AL (19 years old), AD (22 years old), AM (24 years old), and AN (19 years old). All observers were naïve to the hypothesis and wore any optical corrections required.

Stimuli

We wished to measure the appearance of blur in depth-filtered images compared with real-world images. Therefore, we needed to match all other aspects of the different image types, so that depth filtering was the only difference. This is highly non-trivial for complex, full-colour images, and even greyscale images (such as photographs of natural scenes). Recall that our display has a relatively truncated luminance range (due to the combination of the shortcomings of TFT monitors, and the presence of mirrors/beamsplitters). Moreover, the fact that the mirrors and beamsplitters are different, and non-neutral with respect to wavelength transmittance/reflectance. Through extensive testing, we discovered that these factors, combined with the bit depth available (8 bits per colour channel), meant that we could not perceptually match the colour properties of complex depth-filtered and real-world images. Specifically, even for greyscale images, there were typically visible changes in colour of the stimulus that allowed the depth-filtered stimulus to be detectable, beyond changes in blur/image quality. These problems are not fundamental to depth-filtered images, but result from limitations of our particular display. Indeed, they are potentially solvable in our equipment using techniques from computer graphics such as colour dithering (Heckbert, 1982). The time and difficulty involved in implementing these techniques rendered them impractical for our purposes, however, and so we chose instead to use only high-contrast black-on-white and white-on-black images, for which accurate calibration was possible.

We used ten different (planar) images, two each from five different image 'classes': an arrangement of same-sized letters, an arrangement of differently sized letters, similar to a conventional eye chart, repeating patterns, shapes, and text quotations. Examples of each type of image are shown in Figure 3.6.1. The images were viewed monocularly, using the left eye's display. We measured monocular viewing so that we could isolate retinal-image contrast, eliminating the effect of other varying factors such as accurate vergence and individual stereoacuity thresholds etc.

Typically for these types of images, anti-aliasing techniques are used (for example applying Gaussian blur) to remove any pixellation of the stimulus edges. This effectively adds a blur 'pedestal' to the stimulus, and so might make it harder to detect blur due to depth filtering. We therefore did not apply anti-aliasing to our stimuli, and so the images appeared pixelated. These images should therefore provide a highly sensitive test of the visibility of blur in depth-filtered images.

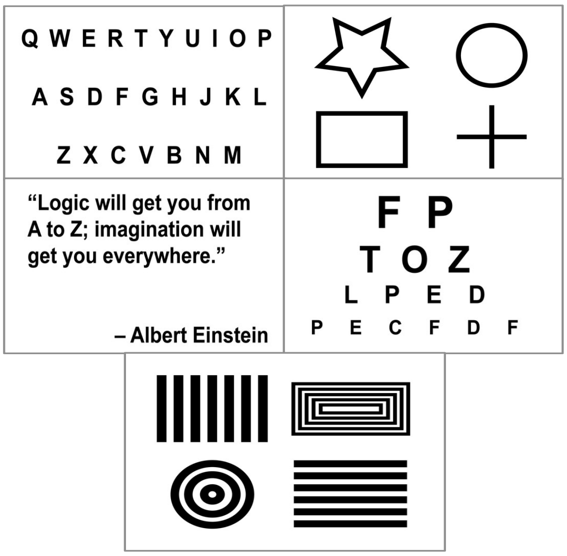


Figure 3.6.1. Five of the images used in the experiment. Two versions of each image type were used with the other five being the inverse of these shown (white letters on black background).

We calibrated the luminance and colour balance of the display using the normal calibration procedure (see General Methods), such that the luminance and colour of both black and white regions of the depth-filtered and real-world (single-plane) images could not be distinguished. The average white luminance value was 24.06 cd/m^2 for the real-world images and 24.08 cd/m^2 for the depth-filtered images. The 'black' luminance was 4.08 cd/m^2 for the real-world images and 4.06 cd/m^2 for the depth-filtered images. The background luminance was higher than in other experiments (normally 0.21 cd/m^2) because the luminance at a black region in our display was significantly affected by the surrounding pixels. Typically we used sparse random-dot stereograms, where the effects of this are negligible. However, when using large regions of white, the black regions became lighter. This is due to a combination of the properties of TFT displays (light and electrical leakage) and of our display (where 'ambient light' form the display can reflect back from its surface).

Blocks

The blocks were the same as those used in Experiment, 4 (Figure 3.5.1; though note, again, that stimuli were viewed monocularly in the current Experiment). However, unlike in Experiment 4, we needed to present 'real-world' stimuli for all blocks in this experiment. For blocks 4-10, the real world stimuli were presented onto the focal-plane, located at 0.7 D. For blocks 1-3, the real-world stimuli were presented onto the nearest focal-plane for that block.

The fixation target (a Maltese Cross) had the same constraints as Experiment 4: the Maltese cross was presented using the depth-filtering technique. For conditions 4-10, the fixation was presented onto the focal-plane at 0.7 D.

Procedure

This experiment used a three-interval forced-choice (3-IFC), or oddity task. On each trial the fixation target appeared for 1.5 seconds, followed by three of the same images, presented sequentially. These three intervals consisted of two real-world images, one of which was presented on the near and the other on the far focal-planes, and one depth-filtered equivalent. The observers indicated via a button-press which one of three was the odd-one-out. The position of the depth-filtered image in the sequence (first, second, or third) was controlled so that, for each image, it appeared three times in each position (in a randomised order). Therefore, each of the ten images was presented to the participant nine times. This meant there were a total of 90 trials per block.

Observers were given no specific instructions as to what to look for as different between the images. With this method, the participant can use any difference they observe to complete the task. Observers were positioned using the bite-bar setup, and completed the display calibration/alignment procedure, as described in the General Methods. Their right eye was covered with an eye-patch. The order of the blocks (Figure 3.6.1) was chosen at random for each observer so none had the same testing order as another.

After each block the observers took a five-minute break. We then asked them to complete a similar, shorter test to ensure that the luminance and/or colour of the depth-filtered images was not a cue to the odd-one-out. This task was the same as the main experiment, with the only difference being that the images were all white squares. These white squares were comprised of real-world images where all the luminance was on the near or far focal plane and depth-filtered white squares where the luminance was evenly distributed across focal planes. Because these images contained no discernible structure, if our calibration procedures were correct, observers should have performed at chance at the oddity task. Alternatively, if the luminance or colour of the depth-filtered stimuli were perceivably different to that of the real-world images, they would perform above chance. Six of the seven observers took part in this control block ('SJ' did not complete this section). After a 10-minute break, the observers were then given the questionnaire designed to subjectively measure the perception of image quality.

Questionnaire

Once the luminance check was complete, the observers viewed the depth-filtered simulation of each image once. They were then asked to respond to the questionnaire. The real-world stimuli were also measured in this way but only once. This was measured at the beginning of a randomly chosen block, after the centring for that block was complete and before the main discrimination task was conducted.

We administered a line-scale questionnaire. We chose the line-scale so that observers could not remember their response relative to the previous block. Observers marked on a 15 cm line their responses to a series of questions designed to measure various aspects of satisfaction, a copy of which is included:

Question 1. How difficult would you find it to view/read these images for long periods of time (for example at work or in lectures)?

Very easy — Very difficult

Question 2. How satisfied with your display device (computer, kindle, smartphone etc.) would you be if it reproduced images at this quality?

Very unsatisfied — Very satisfied

Question 3. How easy was it to see fine details in the images?

Very difficult — Very easy

Question 4. How uncomfortable was your eye when looking at these images?

Very comfortable — Very uncomfortable

Question 5. How comfortable did your head and neck feel while viewing these images?

Very comfortable — Very uncomfortable

Question 5 was intended as a control question, to assess whether observers were responding to the specific symptoms being measured by each question, or simply answering each one according to a general sense of how they were feeling, or to detectable differences in the stimuli (e.g. more blur) (see Hoffman et al., 2008). If they were responding to the level of specific symptoms present we would expect no effect of focal-plane separation on ratings of head and neck comfort, because this should be equal in all blocks. We also included reversed questions (far right response could be good or bad) to check for observers simply marking all responses in the same manner, without attending to the specific question being asked.

3.6.2 Results

Performance for the main experiment (and the luminance/colour control) was assessed by computing percent correct scores, where choosing the interval showing the one depth-filtered image was recorded as ‘correct’. For the questionnaire, the response positions along the 15 cm line were measured. The reversed questions were adjusted so that for all questions, a higher score indicates a positive rating of the images (i.e. better/preferred). To create confidence-interval error bars for 1) the discrimination task, 2) the luminance/colour test and 3) questionnaire results, we ran a bootstrapping function on all observers’ data for each block, using SPSS “Version 20.0”.

To ensure the results of the main manipulation are not due to inaccurate luminance matching across blocks, we first examined the results of the luminance/colour control experiment (all-white images). Figure 3.6.2 plots percent correct at identifying the depth-filtered stimulus as a function of image-plane separation. The grey dashed line shows chance performance (33%). A repeated measures ANOVA confirmed that observers performed close to chance, independent of focal-plane separation, $F(9,18) = 0.90$, $p = .55$. This indicates that we satisfactorily removed luminance or colour artefacts from our stimuli.

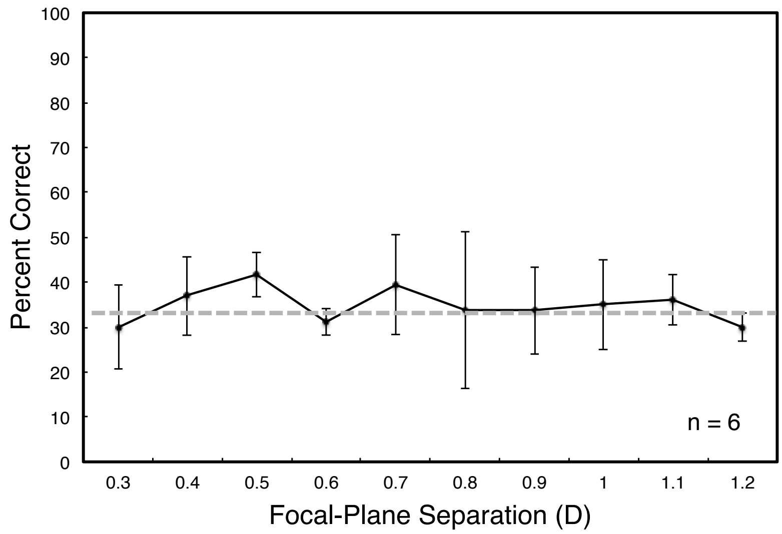


Figure 3.6.2. Performance in the luminance/colour control experiment. The graph shows percent correct performance at the oddity task as a function of focal-plane separation. The dashed grey line shows chance performance (33% correct). The x-axis shows the focal-plane separation in dioptres (D). The y-axis shows the percent correct achieved. Error bars are 95% confidence intervals.

Effect of focal-plane separation on detectability of depth-filtered images

We examined if depth filtering produces an image that is noticeably different to real-world equivalent images and whether this difference scales with focal-plane separation. Figure 3.6.3 shows the percent correct on the oddity task as a function of focal-plane separation. If there were no discernable differences between the depth-filtered and real-world images, we would expect performance to be at 33% (the guessing rate of the task) in all blocks. The results show, however, that as the focal-plane separation increased, so did the ability to discriminate between the real-world images and a depth-filtered equivalent. Indeed, even at the smallest focal-plane separation tested (0.3 D) the depth-filtered images were accurately recognised as different on almost 60% of the trials, which is higher than chance. These results are surprising. We would expect no differences to be detected for focal-plane separations of up to 0.5 D. This is because focal-plane separations smaller than this are reportedly within the depth-of-focus of the eye (Campbell, 1957).

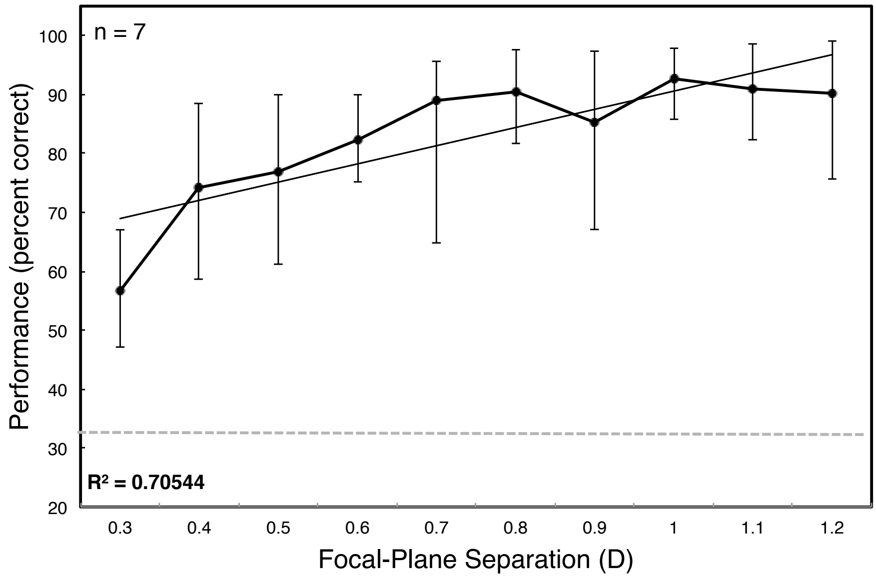


Figure 3.6.3 – Percent Correct scores as a function of focal-plane separation in Experiment 5. The data are the average of the observers' performance for each block. The x-axis shows the focal-plane locations in dioptres. The y-axis shows the performance in percent correct. Error bars are 95% confidence intervals.

Since the results of our control experiment rule out luminance/colour calibration as a possible cause of this effect (Figure 3.6.2) it seems most likely that the attenuation of retinal-image contrast with defocus, even for small image-plane separations, is the cause of above-chance detectability of depth filtered images. Mackenzie et al. (2010) postulated that for focal-plane separations of 0.6 D and above, reduced contrast as a result of one or more focal planes being defocused could result in depth-filtered images looking different to real-world equivalents. Our results suggest that this may happen for smaller focal-plane separations, despite the smallest separation being within the depth of focus of the eye. This has implications for using the technique in applications such as medical imaging, where clear and sharp edges may be of paramount importance.

There are caveats to this interpretation. First, it is possible that a contributory factor to identifying the depth-filtered images was the change in focal distance between the depth-filtered image and the 'real-world' images in the first three experimental blocks. Given that the observers were instructed to respond to any differences they perceived, the change in accommodation could therefore have indicated the difference between the images. We think this is unlikely, however, because the change was very small (1.5, 2.0 and 2.5 D; blocks 1-3 respectively), and experiments on perceived depth from accommodation have shown that sensitivity to depth/distance from this signal is very coarse (Fisher & Ciuffreda, 1988; Mon-Williams & Tresilian, 2000). A second possible factor is an imperfect spatial alignment of the two focal planes with respect to the observer. This could result from incorrect static positioning (creating ‘ghosted’ edges to stimuli features) caused either by completing the spatial calibration procedure inaccurately, or by small differences in the positioning of the mouth on the bite bar across different trials inaccurate spatial calibration. Alternatively, a small amount of movement is possible while biting on the bite bar, which would cause a form of motion parallax cue as to the identity of the depth-filtered stimulus. If this is the cause (and we do not know) it could nonetheless be problematic for depth filtering, because it would demonstrate how difficult it is to achieve perfect alignment of the focal planes even in our context, let alone in a real-world application.

Second, as noted earlier, our stimuli were likely to be a particularly tough test for depth filtering. ‘Tolerance’ to image-plane separation might be higher for more naturalistic images, that contain blurry elements.

Finally, as discussed previously, the point at which depth-filtered stimuli can be detected, and when they cause problems for viewers, may not be the same. We turn to this next.

Questionnaire results

Figure 3.6.4 shows the mean responses to each question in our questionnaire as a function of image-plane separation. The red band shows the 95% confidence intervals for the ratings of the real-world images. It can be seen that there was a systematic reduction in how positive the ratings were with focal-plane separation for Questions 1-3, which asked about readability, fine detail, and satisfaction with the device, respectively. Question 4 had a gradual and slight reduction that was comparable to the real-world ratings for focal-plane separations up to 1.2 D. However, there was or no appreciable change in rating of head and back comfort (Question 5) with focal-plane separation, suggesting that observers were responding to the specific symptoms queried in each case.

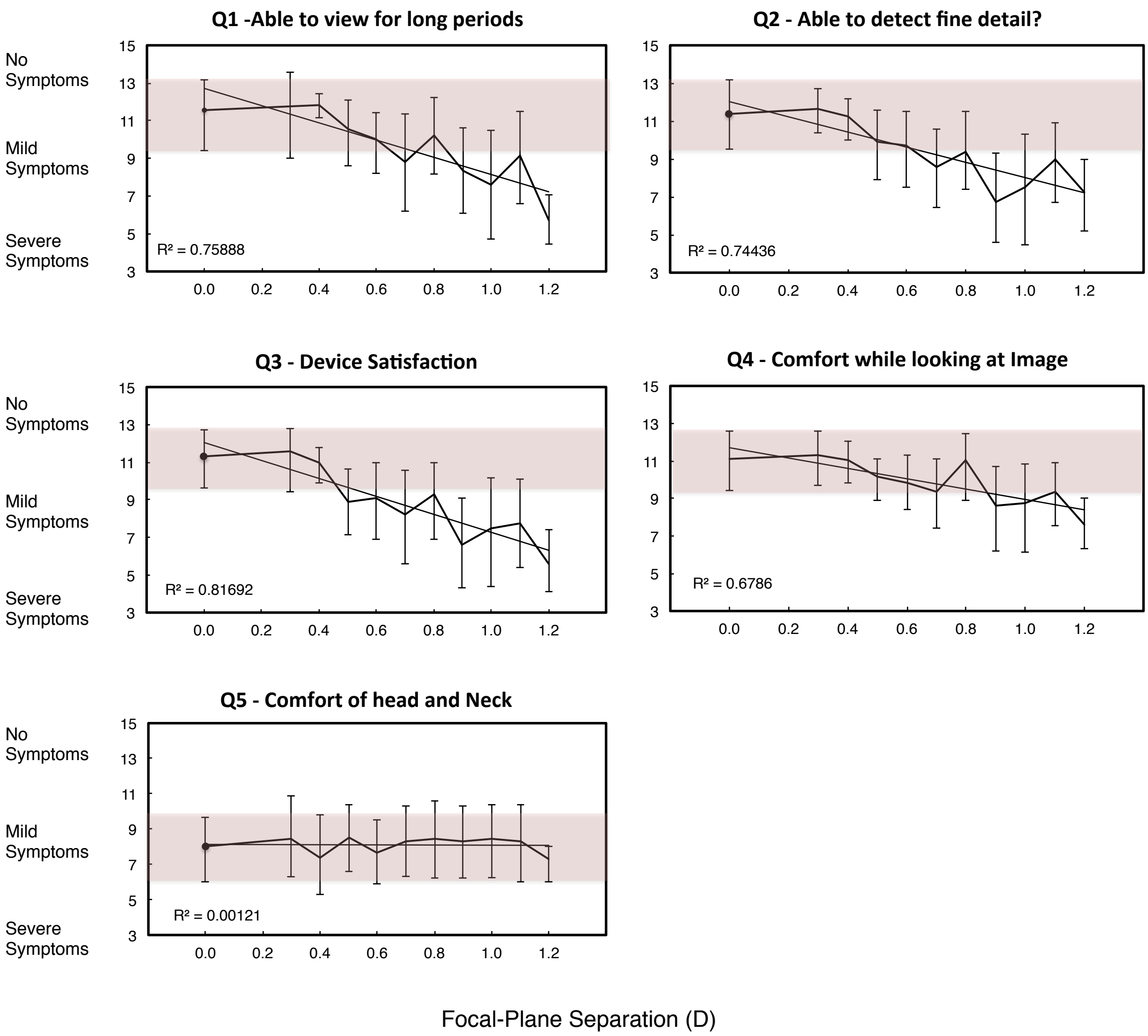


Figure 3.6.4 – Results of the questionnaire data. The red band on each graph is within the range of error bars for the real-world responses. A linear regression line was fit to the data, showing the severity of decline in perception of quality from blur. The x-axis shows focal-plane separation sizes in dioptres (D). Zero on the x-axis and the solid data point denotes single plane (i.e. real-world equivalent) stimuli. The ordinate shows the score in cm, the higher the score, the more positive the response. Error bars are 95% confidence intervals.

When compared with the RW ratings in each case, we found that there was a significant difference in ratings with focal-plane separations for Questions 1 ($F(10,60) = 4.10$, $p < .01$), Questions 2 ($F(10,60) = 3.13$, $p = .003$), Question 3 ($F(10,60) = 4.16$, $p = .001$) and Question 4 ($F(10,60) = 3.15$, $p = .003$). However, there was no significant difference for Question 5, $F(10,60) = 0.54$, $p = .85$. For Questions 1-4, we plotted the mean differences between the real-world ratings and focal-plane separation in order to establish any significant differences between them. The results are shown in Figure 3.6.5. The question number is at the top right of each graph. In each case there are significant differences between the real-wolrd and depth-filtered stimuli as the error bars do not overlap for all focal-plane separations. For Question 1 (top right graph) there are significant differences between the real-world and depth-filtered images with focal-plane separations of 1.2 D (p <.01). For

Question 2 (top left graph) there are significant differences between the real-world and depth-filtered images with focal-plane separations of 0.9 D (p = .04). For Question 3 (bottom right graph) there are significant differences between the real-world and depth-filtered images with focal-plane separations of 0.9 D (p = .04) and 1.2 D (p <0.1). Finally, for Question 4 (bottom left graph) there are significant differences between the real-world and depth-filtered images with focal-plane separations of 1.2 D (p = .03).

These results suggest there is a gradual decline in ratings for criteria such as reading, detecting fine details, device satisfaction, and comfort when viewing depth-filtered images with increasing focal-plane separation. Statistically, the severity of decline, and the focal-plane separation that results in substantially different scores than 'real-world' images, differs across the three measures. For "device satisfaction" (Q3) this decline occurs at a focal-plane separation of 0.9 D, whereas for "able to view for long periods", "able to detect fine detail", and comfort (Qs, 1, 2, & 4), this appears to occur at focal-plane separations of 1.2 D.

It is important to note that although there are significant differences between the real-world data and depth-filtered equivalents, similar to Experiment 1, the focal-plane separation used should depend on the task at hand.

Summary of Results:

We examined if depth filtering produces images that are noticeably different to real-world equivalent images and whether this difference scales with focal-plane separation. In the oddity task, we found that all focal-plane separations produce images that are noticeably different to real-world equivalents. This is a surprising result for separations smaller than 0.6 D. For the questionnaire data, we found that the questions relating to reading, detecting fine details, device satisfaction, and comfort are all increasingly affected with increases in focal-plane separation. Specifically for "device satisfaction" (Q3) this decline occurs at a focal-plane separation of 0.9 D, whereas for "able to view for long periods", "able to detect fine detail", and comfort (Qs, 1, 2, & 4), this appears to occur at focal-plane separations of 1.2 D.

3.7 Discussion

These experiments reveal that fusion time and stereoacuity were largely unaffected by focal-plane separation. That is, as assessed by these two indices, there was no appreciable difference between stereo performance when viewing real-world (i.e. single-plane) stimuli and depth-filtered approximations to the same stimuli (at least up to focal-plane separations of 1.2 D). However, stereoresolution thresholds were poorer for depth-filtered stimuli, compared to real-world stimuli, for focal-plane separations in excess of 0.9 D. Finally, the experiment examining the appearance of (2-D) depth-filtered images showed that they could be reliably discriminated from real-world stimuli even with a focal-plane separation as small as 0.3 D, and that this discrimination became increasingly easy as focal-plane separation increased. However, the questionnaire data, designed to measure subjective opinion of image-quality, revealed that subjective positive experience of the depth-filtered images decreased systematically with increasing focal-plane separation. Using this measure, depth-filtered images with focal-planes separated larger than 0.9 D may be deemed as significantly worse in terms of device satisfaction.

For Experiments 1-3, we measured stereo performance for depth-filtered images at multiple focal distances within the display, using the same focal-plane separation (0.6 D) at multiple distances. Here, we found the same null effect for this focal-plane separation, regardless of stimulus location. Taken together, these findings suggest that the effects of plane separation (in dioptres) are likely to be similar at different overall focal distances. We also considered oculomotor response type (convergent or divergent response) for all distances and focal-plane separations during these experiments. However, contrary to our hypothesis, we found no effects of response direction on TTF. This is surprising because previous research has found divergence eye movements to be slower, and less actively controlled (Semmlow, Hung, & Ciuffredo, 1986; Semmlow & Wetzel, 1976; Hung, Zhu, & Ciuffreda, 1996).

3.7.1 Effectiveness of Depth Filtering

Our results, taken together with those of previous studies examining oculomotor responses to depth-filtered stimuli, confirm that depth filtering can be highly effective, but trades-off aspects of image quality. MacKenzie et al. (2010, 2012) showed previously that depth-filtered images can effectively eliminate the vergence-accommodation conflict for focal-plane separations up to ~0.9 D. That is, viewing such depth-filtered stereoscopic stimuli is essentially equivalent to viewing real-world objects in terms of the oculomotor system. Our results extend these findings to perceptual aspects of stereo performance. Again, with 0.9 D focal-plane separation, stereo acuity, time-to-fuse, and stereoresolution are essentially the same for real-world stimuli and depth-filtered approximations to them. However, the reduction in retinal-image contrast (effectively retinal blur) that results from viewing depth-filtered images can produce noticeable differences even with focal-plane separations that are impractically small for a general-purpose display. If focal planes were separated by 0.3 D, for example, 10 planes would be required to present images at distances ranging from 36 cm (2.8 D) to infinity (0 D).

These results suggest a somewhat complex trade-off may be required in the various aspects of depth-filtered images depending on the particular requirements of a given application. Where the finest possible detail is required, depth-filtered images may be problematic, and perhaps a conventional stereoscopic display should be used. However, this must be traded-off against the fact that (i) vergence-accommodation conflicts themselves can cause reduced stereo performance (Akeley et al., 2004; Watt et al., 2005b, Hoffman et al., 2008; and see later chapters in this thesis), and (ii) inaccurate accommodation, which may cause blur in the retinal images. It should also be noted, however, that subjective ratings of image-quality to depth-filtered images were not significantly worse relative to real-world images for focal-plane separations of less than 0.9D.

Alternatively, for applications where fine detail is less important (such as entertainment), focal-plane separations of up to 0.9 D could be used, eliminating the vergence and accommodation conflict but with a somewhat blurry image. Together with the findings of MacKenzie et al. (2010; 2012), these findings suggest that a display with a range of 33 cm to around optical infinity (10 m; 0.1 D) would require only four focal planes, with separations of up to 0.9 D.

The benefits of matching the cues to vergence and accommodation, in these applications may out-weigh the perception of a slightly blurry image. This seems possible, given that there is a continuous relationship between detectability of depth-filtered images and focal-plane separation. Considerably more research is required to map out the space of these trade-offs if general guidelines are to be realised.

3.7.2 Improvements, refinements, and alternatives to depth filtering

So far we have considered the practicality and effectiveness of depth filtering as a general solution for presenting approximately correct focus cues in a stereoscopic display. By this measure, the picture is mixed: there are clear benefits for eliminating vergence-accommodation conflicts, while preserving stereo performance, but in a practical display (i.e. with practical focal-plane spacings) there is a cost in terms of image quality. Two important caveats should be attached to this conclusion. First, we have considered the case of building a display that covers essentially the whole distance range we normally interact with (~35 cm to infinity). In practice, many stereoscopic devices may need only to present images across a specific range of distances, which is known and fixed. Here, inter-plane spacings of much less than 0.9 D might therefore be practical. For example, just four image planes, spaced 0.4 D apart, could cover the range 40-77 cm (2.5 to 1.3 D), allowing problem-free stereoscopic viewing for the majority of reaching space, with very little loss of image quality.

Second, and more generally, depth filtering was developed to solve the problem of visible discontinuities in multi-plane images, and was not optimised for presenting correct focus cues, or other aspects of image quality. Indeed, Narain et al. (2015) report how, when using this rule to create complex scenes, various perceptual issues other than blur/reduced contrast arise. Haloing effects occur at occlusion boundaries, and reflections are not presented at the correct focal distance, for example. Narain et al. describe an optimisation approach to assigning image intensity to focal planes. Using a model of image formation at the eye (based on a typical observer's optics), their algorithm finds the distribution of image intensities that minimises the difference, across all focal distances, to natural viewing of the same scene. In their paper, they present examples of the images that can be created using multi-plane displays and their technique. Currently this is proof-of-concept work rather than a practical solution. It is computationally very intensive (it takes several hours to compute image intensities for a single image) and it is not yet known if, for example, the stimulus to accommodation is correct. However, it potentially offers a solution to the degrading of image quality observed in 'conventional' depth-filtered images.

3.7.3 Fixed-viewpoint Volumetric Displays

Using this technique, our research suggests that fixed-viewpoint volumetric displays could replace conventional stereo 3-D displays for single-viewing applications, particularly where head-mounted displays could be used. As mentioned in the Introduction, several approached to building fixed-viewpoint volumetric displays allow, in principle, miniaturised displays that could be built into head-mounted displays. For instance, the birefringent lens display (see General Introduction; Love at al., 2009) and other similar displays (Liu et al., 2008) are physically small and can be incorporated into head-mounted displays relatively trivially. Indeed, even the beam-splitter based multi-plane displays (such as ours) can be miniaturised with the use of a powerful badal lens, located close to the eye.

However, for larger audience applications where a multi-view component is required, such as T.V and cinema, conventional stereo 3-D displays may still be required. This adds to the growing amount of research into these multi-plane displays, which demonstrate their potential as a viable solution to improving stereoscopic displays.

3.7.4 Effect of distance on stereoacuity: A side-note.

In Experiment 3, we were able to measure how stereoacuity thresholds changed with distance. Though not related to our main research questions, we felt it was an interesting question that our experiments may shed light on.

As mentioned previously, researchers have pondered whether disparity sensitivity at far distances is limited by low-level disparity-based mechanisms (i.e. the magnitude of disparities per se), or other factors including the ability to represent depths perceived as being at far distances. For obvious practical reasons, it is difficult to measure stereoacuity (or other aspects of stereoscopic depth perception such as magnitude of perceived depth) using real stimuli at very large distances. One solution has been to simulate large distances by altering vergence, without presenting appropriate cues to accommodation (Dees, 1966; Kaufman, et al., 2006). This is problematic not only because of the effecs on stereo depth perception caused by the vergence-accommodation conflict (the main topic in this thesis), but also because the stimulus to accommodation has been shown to affect how disparities are interpreted or 'scaled' (Watt et al., 2005b; Hoffman et al., 2008). Studies by Allison, Gillam, and Vecellio, (2009), and Palmisano, Gillam, Govan, Allison and Harris, (2010) addressed this problem by using physically large spaces (disused railway tunnels and long hallways), to present far-distant, real-world (i.e. matched accommodation and vergence cues) stimuli. Allison et al. (2009) measured depth-interval estimates and discrimination performance for distances of 4.5, 9.0 and 18 m. They were able to show that binocular estimates of depth increased in an essentially linear fashion with increasing disparity at these far distances. This showed that stereopsis is an effective cue for depth discrimination and estimation for distances beyond those traditionally assumed. Palmisano et al. (2010) went on to show that, at larger distances still (up to 40m), as disparity increased, estimated stereoscopic depth also increased. It is important to note, however, that by measuring depth estimates, neither study has been able to examine how disparity *sensitivity* scales with distance, only that stereopsis is an effective cue at large distances.

Because our multi-plane display used Badal optics, we were able able to measure stereo performance for far distances (in principle up to infinity/0 D although we measured up to 10m/0.1 D) with no vergence-accommodation conflict, while controlling for all other cues to distance and depth. Importantly, in our study, we were able to use random-dot stereograms, which allow us to isolate disparity as the only cue to depth in the stimulus. We were therefore able to measure stereoacuity thresholds for a large range of distances.

Our findings show that stereoacuity thresholds appear to be relatively stable for image distances between 1.3 D and 0.7 D, after which they increase. This suggests that the amount of disparity in an image required to perceive depth increases at far distances. Therefore, this implies that disparity sensitivity is not limited solely by low-level disparity-based mechanisms. If this were the case, this would predict identical stereoacuity thresholds *in units of binocular disparity* (Bradshaw & Glennerster, 2006). Instead, it is possible that disparity sensitivity may be controlled by other factors, including the ability to represent depths perceived as being at far distances. This is consistent with the findings of Palmisano et al. (2010) and Allison, et al., (2009).

CHAPTER 4

4 EFFECTS OF VERGENCE-ACCOMMODATION CONFLICT ON STEREO PERFORMANCE.

4.1 Introduction

In this chapter we explore stereoscopic performance as a function of vergence-accommodation conflict. Previously, we examined a possible approach to eliminating the conflict using multi-focal-plane displays, together with depth filtering. This approach is limited to single viewer applications, where the display is viewed from a fixed position with respect to the head (head-mounted displays, displays with viewing apertures, and so on). As such, it cannot be used in many situations where stereoscopic presentation is commonly employed, including television and cinema viewing. Therefore, identifying the effects of vergence-accommodation conflict, and understanding the tolerance to it, remains an important task if effective and comfortable stereo content is to be created.

4.1.1 What is the problem?

As discussed in the General Introduction, one fundamental problem with conventional stereo 3-D displays is the unnatural stimulus to accommodation. In conventional stereo 3-D displays, in order to converge and accommodate accurately to image points at distances other than the screen, the viewer must accommodate at the distance of the screen while converging at another. This can be a problem as the accommodation and vergence systems are neurally coupled; changes in one system result in changes in the other (Martens and Ogle, 1959; Mays and Gamlin, 2000; Yang & Sheedy, 2011; Heron, Charman & Schor, 2001; Fincham & Walton, 1957; Schor, 1992). In natural viewing this is advantageous because the coupling between them results in faster and more effective responses. In 3-D viewing it can be problematic, however, because while the stimulus to vergence varies according to the depth structure of the portrayed scene, the stimulus to accommodation is fixed at the screen surface. Thus, the viewer is required to decouple their vergence and accommodation responses, which is effortful, and accurate responses of both systems may not be possible.

In order to have a clear, single binocular view of an object, these responses do not have to be completely accurate. The accommodation system needs to respond within the depth of focus of the eye (roughly ~0.25 to 0.3 D; Campbell, 1957; Charman and Whitefoot, 1977), or the target object will appear blurry. And the viewer needs to converge within ~15-30 arcmin of the target to avoid double vision (diplopia; Schor, Wood, & Ogawa, 1984). A kind of upper limit on the stimulus to vergence and accommodation for which stimuli can be seen clearly while maintaining binocular fusion vision is given by the zone of clear single binocular vision (ZCSBV; Fry, 1939; see General Introduction). It is clear, however, that adverse effects can occur well within the established ZCSBV, particularly when viewing dynamic stimuli such as stereo 3-D content (as opposed to fixed conflicts induced by prisms or lenses in spectacles). Due to the wide range of applications of stereo displays (including entertainment, gaming and medicine), it is important to establish the effects of conflict magnitude on stereo performance.

4.1.2 Previous Research

As discussed in the General Introduction, Hoffman et al. (2008) showed vergence-accommodation conflicts can cause visual fatigue and discomfort (using a multi-plane display that allowed them to isolate the effects of conflict, *per se*). Shibata et al. (2011) extended this work, by attempting to estimate a tolerance zone to conflict, or “zone of comfort” for stereo 3-D viewing—the range of conflicts that result in acceptably comfortable viewing. To do this, they measured discomfort ratings as a function of conflict magnitude, conflict sign (i.e. targets nearer or father than the screen) and focal distance. They again used a multi-plane display allowing them to establish the effects of conflict *per se* by comparing discomfort and fatigue when viewing the same stimulus steps in vergence with consistent changes in the stimulus to accommodation, or when the latter was fixed. Their display presented four focal planes, spaced 0.6 D apart (i.e. covering a range of 1.8 D). They used trial lenses to shift this range in different conditions, allowing them to measure the zone of comfort’ for screen distances at between 0.1 and 2.5 D, which covers the range of most applications of stereo 3-D. The results suggested that tolerance to conflict was overall reduced at farther distances compared to nearer distances. They also found that at near distances there was greater tolerance to conflicts farther than the screen, compared to nearer than the screen. This pattern was reversed at far distances, with greater tolerance for nearer-than-screen targets (see their Discussion for an exploration of the implications of this for producing stereo content).

In addition to discomfort, stereo performance has also been found to be impaired with conflicts within the ZCSBV (i.e. less than the maximal possible decoupling of accommodation and vergence). Indeed, such conflicts have been shown to lead to increased time-to-fuse, poorer stereoacuity, and poorer stereoresolution (Hoffman et al., 2008; Akeley et al., 2004; Watt et al., 2005b). This is likely to be for at least two reasons. First, cross-links between vergence and accommodation may affect the speed and accuracy of vergence responses (Cumming & Judge, 1986; Krishnan et al., 1977; Semmlow & Wetzel, 1979), leading to impairment of stereoscopic fusion (Hoffman et al., 2008). Second, the cross-links may result in slow or inaccurate accommodation (with respect to the screen), leading to a loss of high spatial frequency information in the retinal images, and a consequent reduction in the precision of disparity information (Banks et al., 2004; Odom et al., 1992).

As previously mentioned, the work to date has, examined only a limited range of conflicts, and viewing distances, and so does not provide a comprehensive description of the relationship between conflict and stereo performance. That was the purpose of the experiments in this chapter.

4.2 Experiment 6: Stereo Performance as a function of conflict magnitude.

The purpose of this experiment was to attempt to map out a ‘zone of good stereo performance’, analogous to the zone of comfort determined by Shibata et al. (2011). To do this we measured the effect of vergence–accommodation conflict magnitude on stereoacuity, both for stimuli nearer and farther than the screen, and at various viewing distances.

Because our primary interest was in producing data useful to producers of content for stereo 3-D displays, we wanted to establish the effect of conflict in a somewhat realistic context. In stereo 3-D media depicting complex scenes (movies and television, for example) the magnitude and sign of the vergence-accommodation conflict is not limited to a single value but instead changes continually, while the screen distance is fixed. In our experiment, therefore, different conflicts were interleaved throughout the experiment, with trials blocked by accommodation distance (screen distance). Interleaving conflicts in this manner is not possible in studies of discomfort because ratings of symptoms must be collected following prolonged exposure to a single conflict level, if clear interpretations are to be made.

Given the constraints of our display (three focal planes, which cannot be moved within a block of trials), we could not run equivalent real-world control conditions (vergence distance = accommodation distance) for all of the vergence distances tested in the experiment. This is potentially problematic because, as we know from previous work, as distance increases, stereoacuity gets poorer (Chapter 3; Palmisano et al., 2010; Allison, et al., 2009). Rather than measuring stereoacuity, we could instead have measured stereoresolution, which we found in Chapter 3 did not change appreciably with distance. This is more problematic. As performance worsens, and an increasingly low corrugation frequency stimulus is presented, fewer cycles of the sinusoidal waveform are presented, which itself limits performance at the task. Given our relatively small stimulus (see Method Section), this would occur at relatively high corrugation frequencies, effectively resulting in a very small corrugation-frequency ‘window’ in which thresholds could be measured (Banks et al., 2004). In contrast, measuring stereoacuity (varying peak-to-trough disparity) allows us to measure performance over a wider range of deterioration in performance (effects of vergence-accommodation conflict), limited only by the disparity gradient limit at very high disparity levels (Banks et al., 2004). Thus, although not ideal, we measured stereo performance in terms of stereoacuity, keeping in mind the effect of distance on stereoacuity thresholds. We expected that as the conflict magnitude increased both in a positive (in front of the screen) and negative (behind the screen) direction, the stereoacuity threshold would also increase. The exact nature of this relationship is unknown, however, and so this was essentially an exploratory study.

4.2.1 Methods

Observers

Five observers took part in this study: LR (24 years old), AO (20 years), AP (20 years), AQ (22 years) and AF (21 years). LR is the author; the rest were unaware of the experimental hypotheses. The observers wore any corrections required; this does not interfere with the optics of the display.

Stimuli

The stimuli were random-dot stereograms, as described in the General Methods. The corrugation frequency was 1.5 cpd. The fixation stimulus was a white Maltese cross (i.e. spatially broadband).

Design

Figure 4.1.1 shows the experiment conditions. There were three accommodation distance conditions, corresponding to screens at 1.3, 0.7 and 0.1 D. Within each, there were several vergence distances, giving rise to a zero conflict condition, and several 'positive' and 'negative' conflicts. Here, positive conflict indicates a vergence distance *nearer* than the accommodation distance, and negative conflict indicate a vergence distance farther than the accommodation distance.

Figure 4.1.1 shows that, in each condition, the smallest non-zero conflict (in either direction) was 0.6 D, with increments of 0.4 D thereafter. We did not present conflicts between zero and 0.6 D because Hoffman et al. (2008) showed that such conflicts do not cause significant decrements in stereo performance. These conflict magnitudes were measured for positive and negative vergence distances. However, we did not present stimuli at distances beyond 0 D (infinity) because they create the unnatural situation in which divergent eye movements are required to fuse the stimuli. This meant that for the conditions with focal distances at 0.7 D and 0.1 D, there were a reduced number of vergence distances relative to the 1.3 D condition.

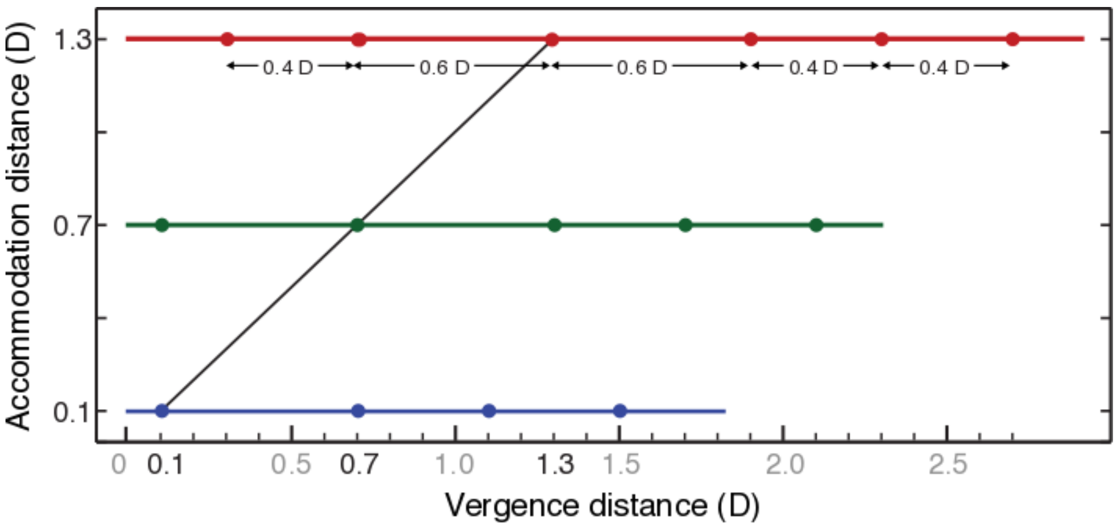


Figure 4.1.1 – Stimulus conditions for Experiment 6. The x-axis is the vergence distance (in dioptres) and the y-axis is the accommodation distance (in dioptres). The different colours denote the different accommodation distance conditions, and the symbols show the vergence distances tested at each focal distance. The diagonal black line shows the locus of zero vergence-accommodation conflict in this space.

Procedure

The observers used a bite-bar throughout the experiment. Observers initiated each trial with a button press. A Maltese cross was presented at the focal distance for a random time interval, drawn from a uniform distribution, between 650-1500 msec (to prevent anticipatory eye movements). Following the fixation, the stimulus appeared for 2 sec. Observers then indicated the orientation of the corrugation using a button press. They received no feedback.

Trials were blocked by accommodation-distance condition (0.1, 0.7, and 1.3 D). Within each accommodation-distance condition, the individual vergence distances (conflicts) tested were selected at random on each trial. Peak-to-trough disparity for the various conflict magnitudes was controlled using interleaved adaptive staircases (e.g. six interleaved staircases for the 1.3 D condition). Observers completed six blocks of each condition (randomly ordered), giving rise to six staircases per vergence/accommodation combination. In each case three of the staircases used a 2-down, 1-up reversal rule, and three used a 3-down, 1-up rule, to distribute data along the psychometric function. In this experiment, ‘down’ refers to decreasing the disparity level, and therefore increasing the difficulty of the task. The observers had to be accurate two or three consecutive times (depending on the reversal rule) before the disparity level decreased. The initial disparity level was chosen at random between 15 and 50 seconds of arc. The initial step size was 5 arc sec, halved after four reversals of the staircase. Each staircase quit after 12 reversals.

4.2.2 Results and Discussion

For each observer and for each vergence/accommodation combination, we computed a stereoacuity threshold as described in the General Methods. For some observers, at some conflict magnitudes, it was not possible to compute stereoacuity threshold values because their performance collapsed catastrophically. That is, even at large peak-to-trough disparities performance was erratic across trials, or simply near chance. This is to be expected if they were unable to achieve stereoscopic fusion on some proportion of trials. These cases were easily identified because performance was far too noisy to achieve a good fit to the psychometric function. Figure 4.1.2 shows examples of a typical, well fitting psychometric function (left graph) and an example of where the observer was largely unable to do the task (right graph). In each case, the size of the data points (blue circles) is proportional to the number of times that disparity value was measured. In the right hand plot there is no systematic relationship between the disparity in the stimulus and the observer’s performance, indicating a catastrophic breakdown in stereoscopic performance. The lack of concentration of data points at relatively few disparity values also indicates that the adaptive staircases did not converge. It was not meaningful to compute threshold values in these cases.

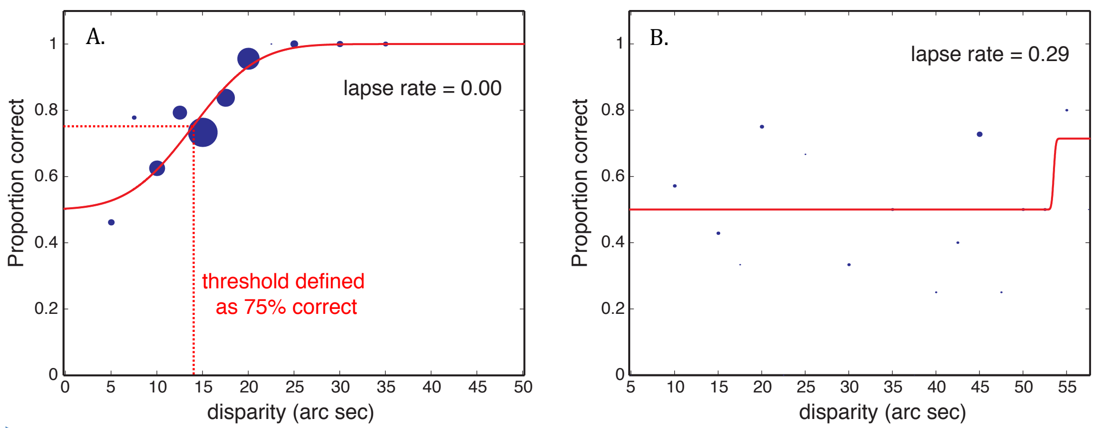


Figure 4.1.2. Examples of fitted psychometric functions. The figures plot proportion correct at identifying the corrugation orientation, as a function of its peak-to-trough disparity. A typical, good fit is shown on the left, and a case in which the performance broke down catastrophically on the right. The size of the data points (blue circles) is proportional to the amount of times the observer responded to that disparity. The red line is the best-fitting cumulative Gaussian through the data.

Figure 4.1.3 plots the stereoacuity thresholds as a function of vergence distance for each observer (graph 1-5) and the average of the measured thresholds (graph 6; bottom right). The coloured lines indicate the different focal distances: 1.3 D (red), 0.7 D (green) and 0.1 D (blue). The solid lines and filled data-points represent positive vergence distances, and the dashed lines and open circles show negative vergence distances. The large, filled circles are the data points for the no-conflict stimuli. In cases where it was not possible to compute stereoacuity threshold values, because performance collapsed catastrophically, lines extending outside the graphs represent these points. Lower threshold values denote better performance as they mean smaller peak-to-trough disparities were required to recognize the orientation of the disparity-defined corrugations.

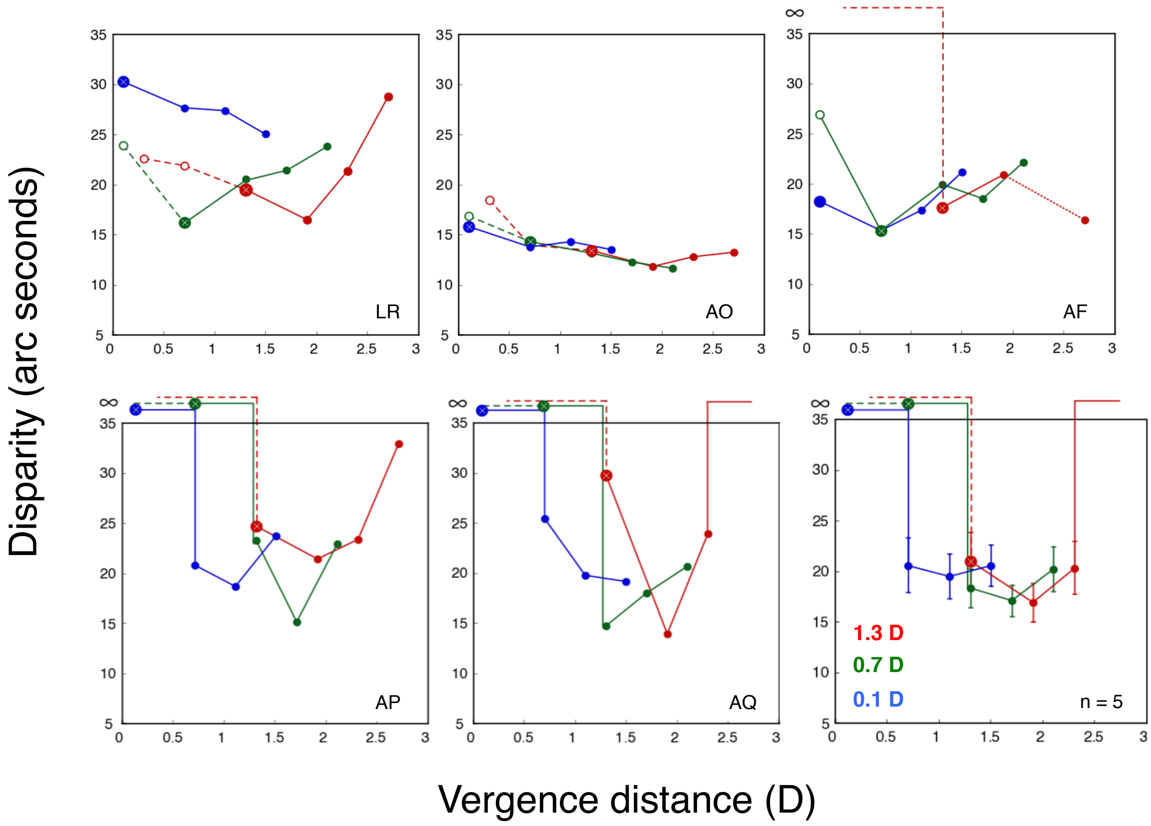


Figure 4.1.3. Results of Experiment 6. The first five panels show individual stereoacuity threshold for each observer, as a function of vergence distance. The sixth panel (bottom right) shows the average across observers. Red, green, and blue lines represent accommodation distances of 1.3, 0.7, and 0.1 D, respectively. Solid lines/filled circles represent positive conflicts (nearer than the screen). Dashed lines/open circles represent negative conflicts (farther than the screen). The larger data points, with a central 'x', indicate the no-conflict conditions. Dashed lines leading to data points outside the graphs denote conditions in which we could not determine a threshold (see main text). Error bars (average data, graph 6) represent standard error of the mean.

Summary and discussion of results:

The results show effects of vergence-accommodation conflict on stereoacuity. For most observers, the nearest and farthest vergence distances resulted in the highest (worst) stereoacuity thresholds (including cases where performance collapsed completely). In detail, however, the data depart from the clear pattern we expected in several ways. We discuss these below.

A clear overall feature of the data is the degree to which individual observers show different patterns of effects of conflict. If the data were purely the result of changes in stereoacuity thresholds with distance (c.f. Chapter 3, Expt. 4), thresholds would be lowest at the nearest distance and highest at the farthest, regardless of conflict magnitude. Indeed for observer AO, the data follow this pattern, suggesting they may have been completely unaffected by the vergence-accommodation conflict. In contrast, AP and AQ were unable to do the task at all in several conditions. For the other observers, the stereoacuity thresholds increased either side of the focal distance, although there is wide variation in the data. This heterogeneity means it is not meaningful to postulate a general 'zone of good stereo performance' from our data.

Figure 4.1.3 also shows several instances in which the no-conflict stimuli did not result in the lowest stereoacuity thresholds. Our prior assumption was that stereoacuity would be best with zero vergence-accommodation conflict, and systematically worse with stimuli either nearer or farther than the screen, based on theoretical considerations and previous data (Hoffman et al., 2008; Watt et al., 2005; Akeley et al., 2004). This is clearly not the case for several of our observers. Indeed, for two of the five (AP and AQ), we were unable to compute thresholds for the no-conflict conditions at 0.7 and 0.1 D accommodation distances. That is, across whole blocks of trials, these observers could not reliably see stereoscopic depth in a 'natural' stimulus, for which accommodation and vergence stimuli were matched.

It should be noted that we cannot rule out that lower thresholds nearer than the screen resulted from effects of distance, as seen in Chapter 3. It is possible, for instance, that there was in fact no effect of positive conflict (same underlying stereoacuity for no-conflict and small, positive-conflict stimuli) and the reduction in stereoacuity resulted from the stimulus being nearer. This is examined further in Experiment 7.

These results suggest that interleaving trials covering a relatively large range of conflicts around a single screen distance can, at least in some observers, result in large decrements in stereoscopic performance, even with no-conflict stimuli. This is despite being able to perform on the task and identify the stimulus orientation of the stimuli in pre-screening (see General Methods). We speculate that there are at least two, related contributory factors to this. Research has shown that the vergence and accommodation cross-links can adapt to conflicting demands following prolonged exposure to a given conflict (e.g. using prism glasses; Miles, Judge & Optican, 1987; Wann, Rushton & Williams, 1995). The idea is that this adaptation effectively results in a new 'resting state' for the accommodation-vergence system that corresponds to a constant conflict between the stimuli. Then, departures from this new 'zero effort' position (which could include to zero-conflict conditions) become difficult. How might this happen for our stimuli, for which demand was constantly changing? One possibility is that once it becomes difficult for observers to maintain accurate accommodation and vergence for certain stimuli, they no longer contribute to adaptation/maintenance of the cross-links, and the adaptation state of the cross-links drifts toward the most easily viewed stimuli.

One might think that the most easily viewed stimuli will always be zero-conflict stimuli. This is not necessarily the case, however, due to large variations in individual's phorias. Phoria is the relative vergence position of their eyes when viewing a stimulus at a given focal distance, and when there is no disparity stimulus to vergence (Shibata et al 2011; Kim, Granger-Donetti, Vicci and Alvarez, 2010; Walline, Mutti, Zadink and Jones, 1998; Schroeder, Rainey, Goss and Grosvenor, 1996). Exophoria describes when a person converges at a farther distance than the accommodation stimulus (when vergence is not stimulated). Esophoria describes when a person converges nearer than the accommodation stimulus. For our purposes, phoria can be thought of as the resting state, or zero-effort state of vergence when viewing a screen at a given distance. Thus, someone with a large esophoria might need to exert relatively high effort to converge on zero-conflict stimuli, but relatively little effort to converge on positive-conflict stimuli (targets nearer than the screen). Consistent with this, Shibata et al. (2012) found phoria to be related to the discomfort observers experience when viewing stereo 3-D content and the associated conflict. Therefore, they found phoria to be predictive of the symmetry of a persons' tolerance to conflict before experiencing discomfort or visual fatigue (i.e. whether they were more tolerant to positive of negative conflicts). Accordingly, we might expect that a person's phoria will predict to some extent the combination of vergence and accommodation stimuli for which their stereo performance is best. We explore this in the next Experiment.

4.3 Experiment 7: Phoria and tolerance to vergence-accommodation conflict

This experiment was broadly similar to Experiment 6, but with key differences. We again examined the effect of conflict magnitude on stereoacuity. However, we used a fixed conflict 'step' of 0.6 D, allowing us to compare the effects of different accommodation distances, for the same vergence distance. This allowed us to isolate the effect of conflict magnitude on stereoacuity threshold, holding vergence distance constant. Therefore, for cases where conflict magnitude is different but the vergence distance is the same, any changes in stereoacuity threshold must be due to the conflict, not distance changes. Moreover, we measured each observer's phoria, to examine whether there was a relationship between it and their tolerance to conflict.

4.3.1 Methods

Observers

Three observers: LR (24 years old), AP (20 years), SJW (41 years) took part in this experiment. Observers LR and SJW are members of the lab, and so were aware of the hypotheses. AP also took part in Experiment 6, but was naïve to the specific hypotheses of this experiment.

Stimuli

The stimuli were the same as those in Experiment 6, except the corrugation frequency was reduced to 1 cpd (from 1.5 cpd). This change was made to make the task slightly easier, to give more 'headroom' to measure decrements in performance.

Design

The design of this experiment was similar to Experiment 6. There were again three accommodation-distance conditions (1.3, 0.7, and 0.1 D). The vergence distances used in each case are shown in Figure 4.2.1. It can be seen that these resulted in the matched vergence distances across different accommodation distances.

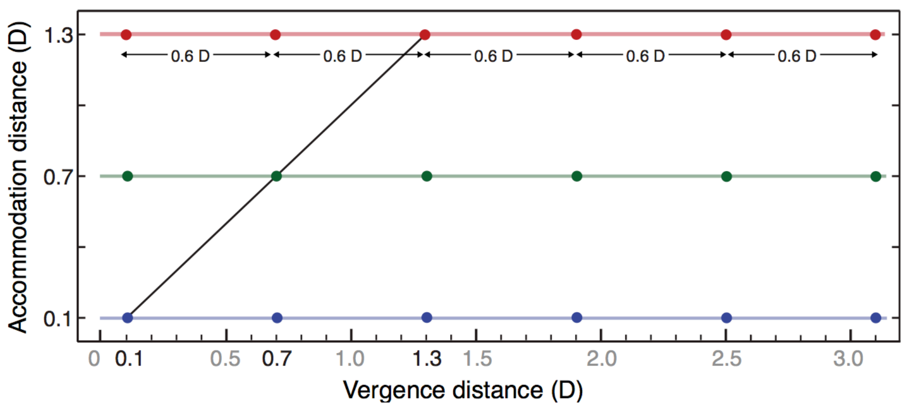


Figure 4.2.1. Stimulus conditions used in Experiment 7. The x-axis is the vergence distance (in dioptres) and the y-axis is the accommodation distance (in dioptres). The different colours denote the different accommodation distance conditions, and the symbols show the vergence distances tested at each focal distance. The diagonal black line shows the locus of zero vergence-accommodation conflict in this space.

Procedure

This was identical to Experiment 6, with the following exceptions. Observers completed a total of four staircases for each vergence/accommodation combination, instead of six, and in all cases a 2-down, 1-up staircase controlled the disparity level in the stimulus only (we found these changes reduced the duration of the experiment with no noticeable effects on the quality of the psychometric function fits). Again, 'down' refers to decreasing disparity level and thus increasing difficulty.

Measurement of phoria

We measured each observer's phoria at focal distances of 0.1, 0.7, and 1.3 D. To do this we used the multi-focal-plane display to create a virtual version of a Maddox Wing, a handheld device used by Optometrists to measure phoria in clinical settings.

The principle is to present different, unfusable images to the two eyes, at the same focal distance, so that accommodation is stimulated but vergence is not. In this way, vergence posture should be at its physiological resting state. To measure this, one eye sees a vertical arrow or pointer, and the other eye sees a horizontal measurement scale. To the observer, the arrow appears aligned with different points on the scale depending on the vergence posture. The observer is asked to report where they see the arrow, and from this the current vergence angle can be calculated. A similar technique is used to measure vertical phoria, but we measured only horizontal phoria.

In our system we presented an arrow pointing vertically down to the left eye. It was centred at the middle of the image plane, using the calibration procedure described in General Methods. A numerical scale was presented to the right eye. The zero point of the scale was also centred with respect to the image plane. The numbers were spaced 1 degree apart. The positive and negative numbers were different colours, so that the sign of the observer's response could be determined (e.g. "red 5"). At each focal distance, the display arms were rotated so that the lines-of-sight through the centre of each eye's display intersected at the tested focal distance. Zero phoria (converge distance = accommodation distance) was therefore indicated by the arrow appearing aligned with zero on the measurement scale. When viewed binocularly, the observer could see what is shown in Figure 4.2.2.

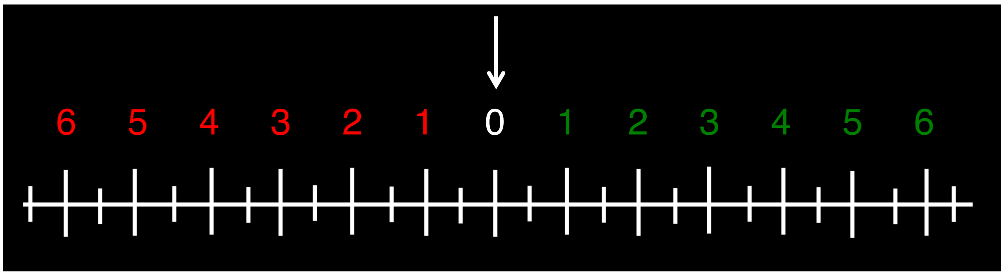


Figure 4.2.2. Observer's view of virtual Maddox Wing used for phoria measurement. This is an illustration of what was actually viewed in the experiment.

While testing phoria, the lights were turned off. The observers were instructed to report the number the arrow was pointing to and the colour of that number. They were told that the arrow may not become completely stationary, in which case they should report the values it moved between. The experimenter noted these, and took the mid-point as the phoria value. Phoria was measured three times per focal distance, and the average taken. The observer dictated the speed of measurement.

Phoria measurements in vergence angle were converted into distances in dioptres (taking into account each observer's inter-ocular distance) to allow comparison with the stereoacuity data at various distances.

4.3.2 Results

Stereoacuity thresholds for each vergence/accommodation combination are shown in Figure 4.2.3. Each column shows a single observer's data. The rows show results for the 1.3 D (top), 0.7 D (middle) and 0.1 D (bottom) focal distances. The data points and lines are the same as Experiment 6. Here, again, lower thresholds denote better performance.

Examining the results, there are a lot of similarities to what was found in Experiment 6. Firstly, there are large individual differences across observers. Despite these differences, however, stereoacuity thresholds appear to increase with increasing conflict in-front-of and behind the focal-planes. Secondly, for some observers (AP and SJW), these decreases in stereo performance with conflict lead to catastrophic failures of stereo perception. Both of these observers struggled to complete the task for some conflict magnitudes, and in some instances, for the no-conflict trials. Thirdly, where observers' 'best' (lowest) stereoacuity thresholds were measured, was not necessarily at zero conflict, but sometimes at distances away from the focal-plane location. In the previous experiment, we considered this was due, in part to the effect of stimulus distance on stereoacuity. Here we can compare results at matched vergence distances across conflict and no-conflict conditions. Examining the data, although there are large individual differences across observers, there is a consistent variation in stereoacuity threshold at matched vergence distances. This indicates that the measured stereoacuity thresholds were a result of our conflict manipulation and not stimulus distance *per se*.

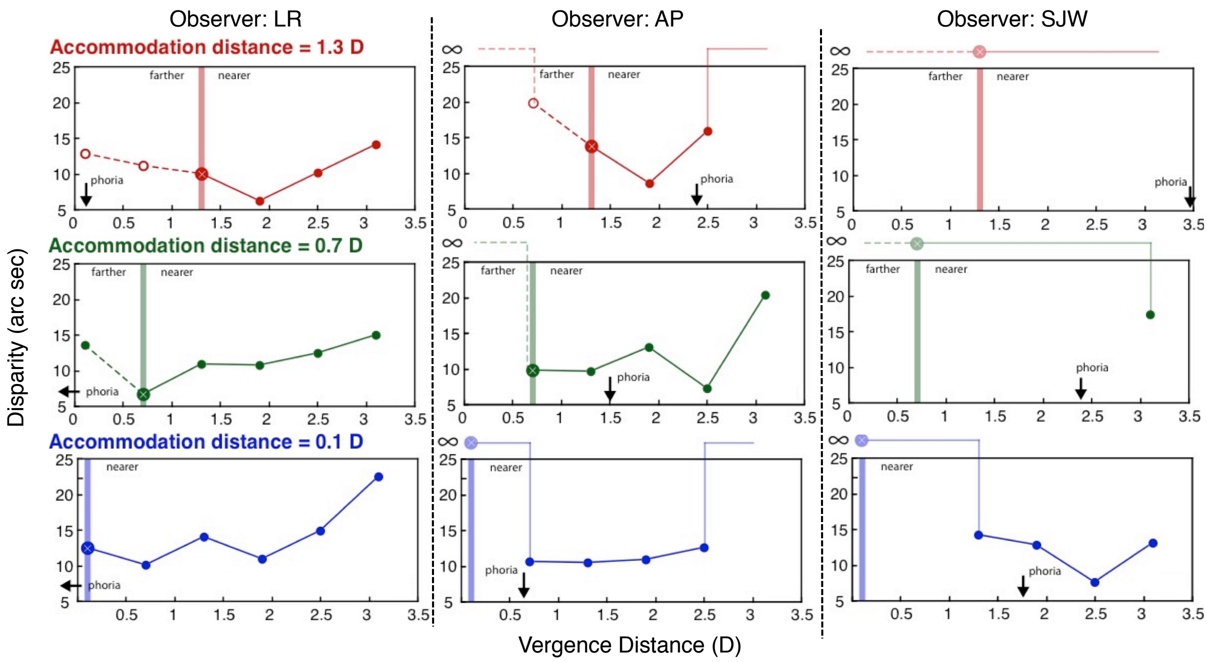


Figure 4.2.3. Results of Experiment 7. Each column represents the data from each observer, the initials of which is shown above. The rows show data from the 1.3 D (red; top row), 0.7 D (green; middle row), and 0.1 D (blue; bottom row) focal-distance conditions.. The rows show data from the 1.3 D (red; top row), 0.7 D (green; middle row), and 0.1 D (blue; bottom row) focal-distance conditions. The solid lines/filled circles represent data from positive conflict, and the dashed lines/open circles are negative conflict. The dashed lines leading to data points outside the graphs denote cases where we could not determine a threshold. The black arrow on the x-axis shows each observer's phoria at each focal distance.

We also explored whether phoria was related to the pattern of effects of conflict. Each observer's phoria at each focal distance is plotted on the appropriate graph in Figure 4.2.3. The different observer's phorias differed substantially. Observer LR was exophoric for all focal distances, while observers AP and SJW were both esophoric at all focal distances, SJW particularly so.

For observer LR, this person was exophoric, they appear to have a greater tolerance to negative conflict relative to the other observers. However, their overall tolerance was similar for positive and negative conflict. Whereas for observers AP and SJW, stereoacuity thresholds appear to be highly linked to their phoria. Observer AP is esophoric, his data show greater tolerance for positive conflicts than negative. Also, observer SJW was highly esophoric, and stereoacuity thresholds could only be measured for relatively large positive conflicts. In fact, stereoacuity thresholds could only be measured at or around SJW's phoria, otherwise stereo performance collapsed. This suggests there is a relationship between phoria and stereo performance, although the degree of this is also subject to variation.

Finally, given that when observers are unable to do the task at the no-conflict distances, they are also unable to do the task for the negative vergence distances, we believe a possible adaptation is occurring leading to an 'after-effect' of viewing conflict stimuli, which affects subsequent viewing. In the next experiment, we measure this possible 'after-effect'.

4.4 Experiment 8: Is there an Effect of Vergence-Accommodation Conflict on Subsequent Stereo Viewing?

We saw in Experiments 6 and 7 that some individuals were unable to perceive stereoscopic depth in our stimuli even in conditions where there was no conflict (i.e. real-world viewing). Since these observers all had normal stereoacuity in the absence of any conflict, this collapse of performance appears to have be a consequence of these trials being imbedded among conflict trials (see Experiment 6, Results and Discussion). This suggests that, at least in some observers, viewing vergence-accommodation conflict stimuli may cause aftereffects on stereo perception of no-conflict stimuli.

This has important implications for at least two distinct reasons. First, in 3-D content it is often assumed (implicitly, at least) that stereo perception at the screen surface will be unaffected by preceding viewing (Mendiburu, 2009). Our data suggest that is not necessarily correct, and that stereo perception may fail even for no-conflict stimuli. Second, and more importantly, it implies that stereo perception in real-world viewing may be affected following time spent using a conventional stereo 3-D environment, with vergence-accommodation conflicts. That is, immediately after leaving the cinema, or using other stereo 3D devices, stereoscopic depth perception may be impaired for some period of time, with potential public-safety implications.

Here, we specifically explored the effect of vergence-accommodation conflict on subsequent viewing of no-conflict stimuli. To determine whether conflict magnitude was predictive of the size of aftereffects, we presented observers with small (0.4 D), medium (0.7 D), and large (1.0 D) conflicts in different conditions, and measured their subsequent stereo performance. We also ran equivalent real-world control conditions in which the changes in vergence demand were identical to the experimental conditions, but the stimulus to accommodation also varied accordingly (i.e. there was no vergence-accommodation conflict). Finally, we also administered the symptom questionnaire used by Hoffman et al. (2008) to establish the degree of various visual symptoms following viewing of images with and without conflict.

4.4.1 Method

Observers

Seventeen observers participated in the experiment (6 male and 11 female, aged between 19-23 years). All had normal or corrected to normal vision. Before testing, we screened each observer for stereoacuity using the ‘circles’ part of the Randot stereo test (Stereo Optical Company, Inc.; see General Method).

Design

As we wanted to explore the effects of conflict on subsequent stereo performance, we conducted an experiment with three testing phases. The initial baseline phase measured observers’ stereoacuity under ‘real-world’ viewing conditions (i.e. with no vergence-accommodation conflict), to provide a measure of their normal stereo performance. There was then an exposure phase, during which observers viewed either a series of conflict stimuli (experimental conditions) or no-conflict stimuli (control conditions). The conflict conditions all used the same screen distance (accommodation distance) of 1.3 D. Finally, they completed a post-test phase to determine effects of the exposure phase on stereoacuity. The post-test phase was similar to the baseline but was prolonged, in order to measure the recovery time for any decrements in performance. This was a within-subject design.

In the baseline phase, stereoacuity thresholds were measured at two focal distances (1.3 D and 0.3 D) interleaved in a single session. We measured stereoacuity at two distances to determine whether aftereffects affected performance similarly at the screen distance at which conflicts were presented in the exposure phase (1.3 D) and at a different distance (0.3 D).

In the exposure phase observers carried out our normal stereo-corrugation orientation-judgement task in one of six conditions (three conflict, and three no-conflict conditions). This phase was measured using a bespoke two-focal-planes display, described below. In the three conflict conditions the focal distance (screen distance) was always 1.3 D, and the vergence distances presented were either 1.3 D ±0.4, 0.7, or 1.0 D. For the three control conditions, the vergence distances were again 1.3 D ±0.4, 0.7, or 1.0 D, but the accommodation distance also varied in a consistent manner (i.e. there was no vergence-accommodation conflict; see apparatus, below). Figure 4.3.1 shows a cartoon of the conflict and control stimuli for the ±0.4 D conditions. Thus, accommodation distance for the conflict conditions was always at the screen (1.3 D), while the accommodation distance for the control was always matched to the vergence distance.

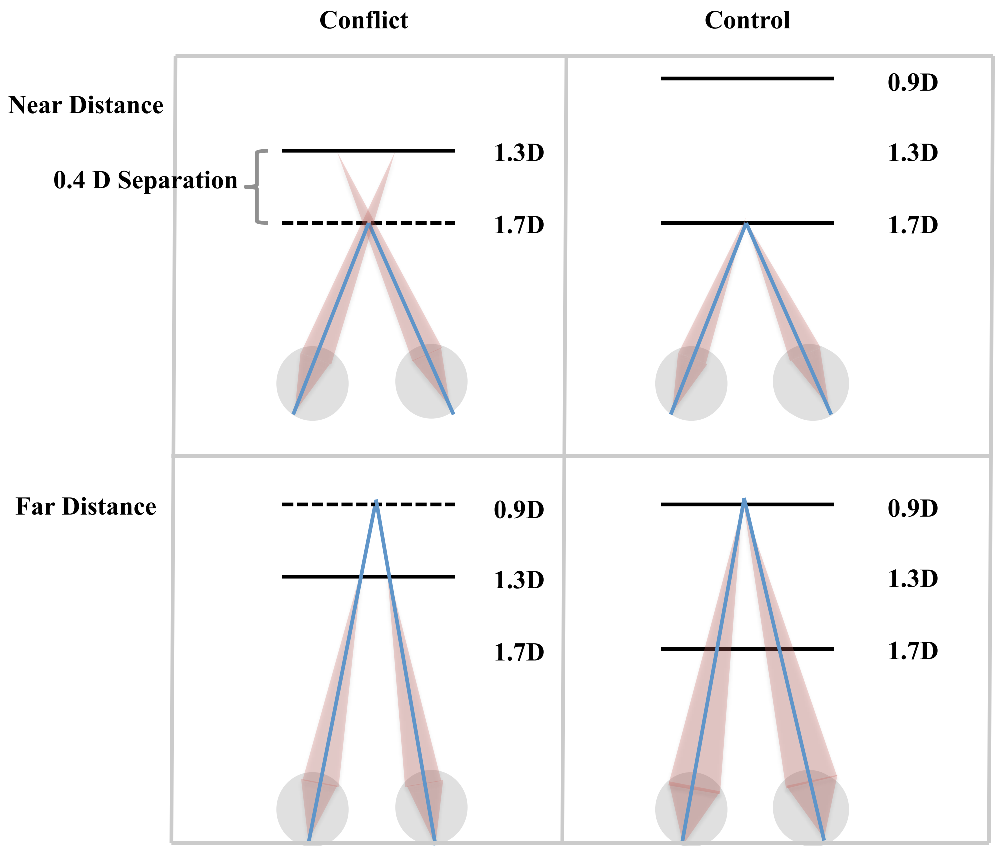


Figure 4.3.1. Example of Conflict and Control conditions in the exposure phase. The black horizontal lines show the positions of the focal planes (labels to the right). The dashed black lines are the geometrically defined vergence distance for the conflict condition. The light red regions indicate the focal distances while the blue lines represent the vergence distances.

Finally, in the post-test phase, observers again completed stereoacuity task with no conflict, at 1.3 D and 0.3 D. As noted earlier, this testing session was extended in order to measure the recovery of stereo performance in cases where there were deficits following the exposure phase.

This was a within-subjects design. As previously explained, there were six sessions in this study: three with conflict exposure phases (one at each vergence step-change), and three with no-conflict/real-world exposure phases (again, one at each vergence step-change). To determine the order the sessions were completed in, they were first grouped by the step-change in vergence distance (i.e. conflict *and* no-conflict exposure at 1.0 D vergence step-change, conflict and no-conflict at 0.7 D, and conflict and no-conflict at 0.4 D). The order of these vergence-step 'groups' was counterbalanced across observers. Then within each group, conflict and no-conflict sessions were completed on different days, in a randomised order.

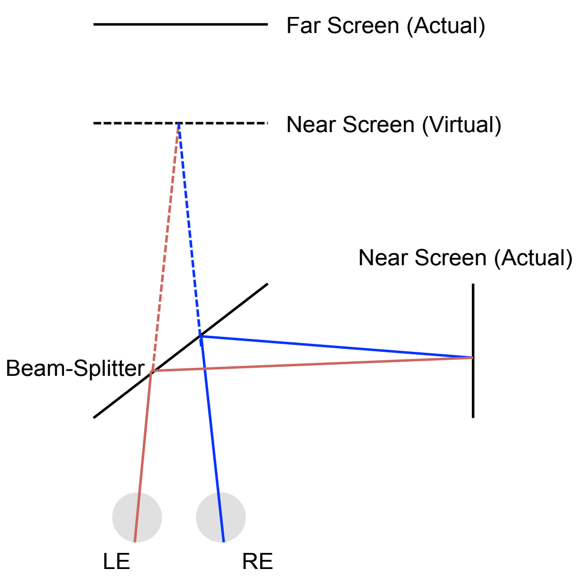
Apparatus

The baseline and post-test measures were carried out using the multi-focal-plane display. Because of the large number of observers, and the fact we did not use depth-filtered stimuli, it was neither practical nor necessary to use a bite-bar to position the observers. Instead, we used a combination of a chin rest and welding goggles (see General Methods).

Two-focal-planes Display

The control conditions in the exposure phase required presenting real-world (no-conflict stimuli) across a range of distances, in a single block of trials, that exceeded that of the multi-focal plane display (the largest range being 0.3 and 2.3 D). To achieve this, we constructed a bespoke display capable of presenting stereoscopic stimuli at two focal planes, which we refer to as the two-screens display.

A schematic of the display is shown in Figure 4.3.2. Both eyes viewed the same 'display', and we used red/blue anaglyph stereo presentation. This display used two, physically separate, screens viewed via a beam-splitter (positioned at a 45 deg angle to the line of sight) to create two focal planes. The distance to both screens could be altered to create different viewing conditions, including creating near and far real-world distances within the same block of trials. In the conflict conditions, the near screen (physically positioned to the right of the observer) was used to present all the stimuli, at 1.3 D focal distance. In the control conditions, the near and far screens were used to present the near and far stimuli, respectively.



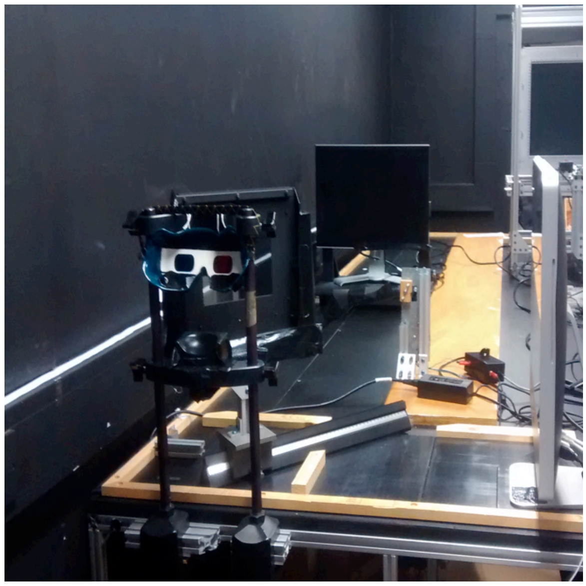


Figure 4.3.2. The display used in the exposure phase of Experiment 8. The left image is a schematic of the display. The near screen presented all images at focal distances at or nearer than 1.3 D. The light from this screen was reflected via a beam-splitter, creating a virtual screen in front of the observer. The far screen presented images at focal distances beyond 1.3 D. The right image is a photograph of the display.

The far screen moved on a wheeled platform, and the near screen was mounted on a Delrin (low-friction plastic) base, that slid on a second Delrin surface.

As explained above, we used anaglyph filter glasses to separate the two-eyes image points in this display. In the Introduction chapter (page 11) we talked about the adverse effects of viewing stereo images through anaglyph filter. These adverse effects include visible crosstalk and reductions in image luminance. The luminance of the two-screen display at the eyes was reduced relative to the multi-image-plane display. This was a result both of using red and blue filters, and having stimuli presented using fewer colour channels compared to 'white' RGB images. Also, in order to eliminate visible crosstalk of the left- and right-eye images, we used two filters (red or blue) in front of each eye, further reducing luminance. This is not ideal. However, as the 'two-screens' display was used in both the control (no-conflict) and conflict conditions, we concluded that any effect of viewing stimuli through anaglyph filters would be present in both conditions. Therefore, any differences between the conditions are a result of the conflict and not the device.

We matched the luminance for all stimuli, at all focal distances in the two-screen display using a Minolta CS-100 Chromameter (see General Methods). The red and blue dots had a luminance of 1.71 cd/m^2 in this display. The anaglyph filters were attached to welding goggles, which were in turn attached to a chin-and-forehead rest, to stabilise the observers' position.

Fixation: two-screen display

In the no-conflict control condition, near and far stimulus distances were interleaved within trial blocks. In all conditions, the fixation stimulus was always presented at the dioptric mid-point between near and far stimuli (1.3 D), so that observers had to continually make vergence (and accommodation) responses on every trial. In the control conditions, there was no screen at this distance. We therefore created an electro-mechanical fixation stimulus, which rapidly swung into position before each trial, and then swung out of view again. Although there was a screen at the 1.3 D distance for the control condition, we used the same fixation as in the control conditions. This meant that the fixation properties were the same across conditions. It also meant that the observers did not have a cue to different conditions within the Experiment.

This fixation stimulus device is shown in Figure 4.3.3. It consisted of a mechanical arm, actuated by a linear solenoid. Turning the solenoid on, via an electrical signal from a Labjack (*LabJack Corporation, Colorado, USA*) digital I/O device controlled by the experiment code, raised the arm into position. When current was turned off a return spring moved it back out of the line of sight. The fixation stimulus itself was formed by a cross-shaped aperture carved into a thin metal panel that formed the front of a small 'light box'. The light box was illuminated with an internal white light emitting diode (LED), and the cross was covered with a diffusing material, creating an emissive fixation cross that appeared similar to that displayed on the screen in the conflict conditions. The luminance of the light was matched to the screen-displayed fixation. The mechanism raised the fixation into position in less than 500 msec; the LED light was then turned on (so that it could not be seen moving). After 2 seconds, the fixation was extinguished, and the lever dropped rapidly back to its resting position.

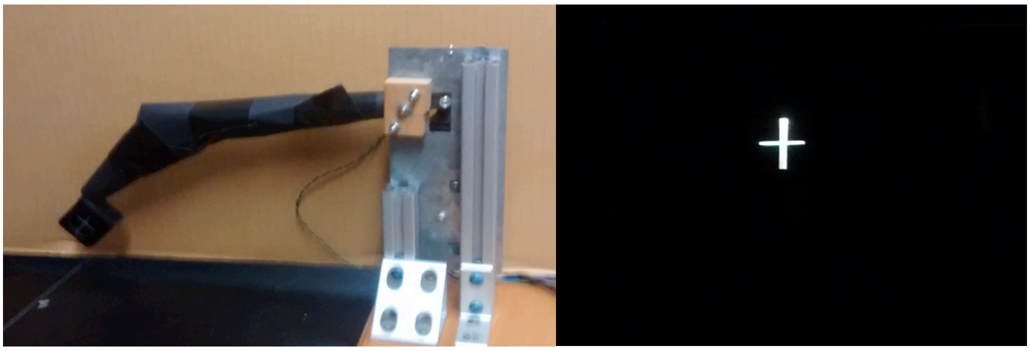


Figure 4.3.3. Electro-mechanical fixation stimulus device used in the control conditions of Experiment 8. The left image shows the apparatus (the brown cardboard is to aid clarity of the photograph and was not present in the experiment). Once engaged, the arm elevated the box to the centre of the line of sight. The image on the right is a photograph of the illuminated fixation stimulus (not taken through the anaglyph filters).

Stimuli

In all three phases of the experiment (baseline, exposure, and post-test) the stimuli were random-dot-defined corrugations in depth (see General Methods). Their diameter was 4.5 deg of visual angle.

Procedure

<u>Baseline phase</u>

In this phase, we aimed to measure the disparity level, for each observer, that corresponded to 90% accurate on the stereo task, when viewing stimuli with no conflict. This disparity level was measured for both test distances (1.3 and 0.3 D). We chose this accuracy level as a compromise between making the baseline task too difficult and too easy. Firstly, we wanted performance to start high enough that we had room to measure substantial decrements in performance (other than simply chance performance; 50% accurate). Secondly, a higher value might result in the task being too easy, in which case it may have been insensitive to the effects of conflict. This stimulus level was then used to measure performance against baseline over short time intervals (relatively few trials) in the post-test.

Observers initiated each trial with a button press. A Maltese cross was presented at the near (1.3 D) or far (0.3 D) focus and vergence specified distance. Fixation was presented for 1 second. Following fixation, the stimulus appeared for 2 sec at the same focal distance as the fixation marker. As in previous studies, the observer's task was to report the orientation of the corrugation. Two adaptive staircases (2-down, 1-up and a 3-down, 1-up) manipulated the peak-to-trough disparity for each focus and vergence specified distance. The initial peak-to-trough disparity was chosen at random between 35 and 50 seconds of arc. The stimulus had a corrugation frequency of 1.5 cpd. Disparity levels in each condition were estimated by fitting a cumulative Gaussian to all trials, using a maximum-likelihood criterion.

This threshold value was used as the peak-to-trough disparity value for the post-test measure. As such, we wanted to ensure that this disparity level did result in 90% accuracy on the task, as it is possible that small errors in threshold estimation can result in inaccurate disparity levels. Therefore, once the 90% correct points for the near and far trials were calculated we checked their accuracy by measuring percent correct over 60 trials, using this stimulus level, at each focal distance (interleaved). The different focal distances were presented in a pseudorandom order. For every pair of trials, the first focal distance was chosen at random, and the second trial was constrained to be at the other focal distance. In this way, both distances were tested with equal frequency and it was not possible to test at one distance more than twice in a row. Once complete, the percent correct performance was measured for each distance, if performance was at 90% or lower (but higher than 80% correct) the observer was then tested on the second phase. If performance was above 90% or below 80%, the disparity level was adjusted accordingly and re-tested. Observers were given a 5-minute break between the baseline and the exposure phase.

Exposure phase

During this phase, the observers viewed the two-planes display. The task was the same as in the first phase (reporting the orientation of a random-dot defined corrugation). On each trial, there was an initial pause of 500 msec, during which the fixation stimulus moved into position. The fixation was then presented for 2 seconds, after which the stimulus was presented for 3 seconds. This meant that a trial lasted a total of 5.5 seconds. We wanted to present the observers with taxing stimuli and rapid changes (4 seconds) in vergence demand when viewing conflict can cause more discomfort than slower changes (100 seconds; Kim, Kane & Banks, 2014). The stimulus had a corrugation frequency of 1.5 cpd and 240 arc seconds. This ensured that once fused, the depth in the stimulus could be easily identified.

The sign of the vergence distance change, relative to fixation, was selected at random on each trial. The magnitude of the change, and whether it was a conflict or no-conflict stimulus, was held constant throughout. The observers were instructed to focus on the fixation cross, and to try to ensure it was single and clear. They were instructed to respond to the stimulus during presentation, and not after. If they did not respond within the presentation time, the trials continued and the response was recorded as a null result. This was designed to ensure that they were constantly viewing the stimuli without breaks. Trials continued in this manner for 10 minutes. Immediately after the block of trials concluded, the observers completed the symptom questionnaire. In total, the time between testing the second phase and the post-test was no longer than 30 seconds, measured by the experimenter.

Questionnaire

The questionnaire was used to measure any adverse symptoms resulting from the exposure phase. It was the same one used by Hoffman et al. (2008). Observers rated their symptoms on a 5-point likert scale where ‘1’ indicated no adverse symptoms and ‘5’ indicated severe symptoms. The observers were asked to complete the questionnaire for all conditions. The questions were:

1. How tired are your eyes?
2. How clear is your vision?
3. How tired and sore are your neck and back?
4. How do your eyes feel?
5. How does your head feel?

Similar to Experiment 5, question 3 was intended as a control question, to assess whether observers were responding to the specific symptoms being measured by each question, or simply answering each one according to a general sense of how they were feeling, or to detectable differences in the stimuli (e.g. more blur) (see Hoffman et al., 2008). If they were responding to the level of specific symptoms present we would expect no effect of focal-plane separation on ratings of head and neck comfort, because this should be equal in all conditions. We also included reversed questions to check for observers simply marking all responses in the same manner, without attending to the specific question being asked.

Post-test phase

The observers’ stereo performance was then re-measured. The test phase lasted 10 minutes total. Presentation times were the same as the baseline phase for fixation and stimulus. Once the stimuli were presented, the observer had 2 seconds to respond. If they did not respond within that time, the next trial began and that trial would be recorded as a void. This was for two reasons, firstly it ensured that all observers had roughly the same amount of trials. Secondly, it ensured the observers did not take a break, which may affect the results. The two focal distances were again selected pseudo randomly, as per the baseline phase, because it was important to ensure that both distances were tested an equal number of times in a given time period. The stimuli had a corrugation frequency of 1.5 cpd and the same peak-to-trough disparity from the baseline measure was used.

4.4.2 Analysis

In this experiment we measured stereo performance by measuring percent correct for a stimulus with a constant peak-to-trough disparity. We used this measure, as opposed to a 'full' psychometric function, because it requires many fewer trials, allowing us to track post-test changes in stereoacuity with much better temporal resolution. Ideally, we would like to measure changes in individuals' underlying sensitivity that is comparable across observers. Unfortunately, percent correct does not provide that because percent correct is affected by an individual's sensitivity and bias to the stimulus. However, we can transform percent correct scores into performance in terms of d-prime (d'), a measure based on signal detection theory, allowing us to consider any changes in performance in terms of changes in underlying sensitivity.

Signal detection theory (SDT) attributes responses to a combination of sensitivity and bias. Sensitivity is what we are interested in, while bias is something we have to account for in order to recover sensitivity. Sensitivity is the ability to detect a signal and also a representation of how an observer decides whether a signal is present. Classical SDT is normally applied to a M-AFC task, where M represents two or more alternatives. Using this conventional method, four possibilities of detection that might happen during a discrimination task can be calculated:

	Response: Same (Yes)	Response: Different (No)
Stimuli: Yes (same)	Hits	Miss
Stimuli: No (Different)	False Alarms	Correct Rejection

Only the numbers of responses in two of the four categories shown in the table (1 per row), plus the total number of trials, are required to calculate performance (the rest can be inferred from the data available). The items of interest tend to be the Hits and False Alarms and these are then given as proportions of the row totals, which are in turn viewed as estimates of probabilities of responses. Hit rate is the proportion of ‘yes’ trials that the observer responded ‘yes' to, while the false alarm rate is the proportion of ‘no’ trials that the observer responded ‘yes’ to.

In our experiments, we are measuring using a 1-AFC task: we are measuring whether observers can accurately identify whether a single stimulus has ‘left’ or ‘right’ orientation Frederick and Kingdom, (2009). For our purposes then, we use an adaptation of the above table. Instead of considering ‘yes’ and ‘no’ trials/responses, we instead consider ‘left’ and ‘right’ (relating to orientation) trials/responses. It is arbitrary whether ‘right’ or ‘left’ orientations are used instead of ‘yes’, for example. For this analysis, we are considering the trials with right oriented stimuli as the ‘yes’ stimuli. Therefore, we can calculate the proportion of hits (pH) as the proportion of trials on which the observer responded ‘right’ for a right-orientated stimulus and, the proportion of False Alarms (pF) as the proportion of trials on which they responded ‘right’ for a left-orientated stimulus.

In this analysis, for the baseline phase, we calculated the d' value based on all the trials per distance (60 trials). However, for the post-test trials we wanted to calculate initial performance, and sample how it changed over time. We therefore needed to decide how many trials to average across to determine performance for the initial and subsequent time points. Obviously, using a smaller number of trials increases temporal precision, but also increases the noise in the estimate of d'. To better understand this trade-off, we explored the consequences of computing d' across ‘bins’ containing different numbers of trials. Example results shown in Figure 4.3.4, which plots the average post-test d' scores across observers for the ±1.0 D control condition, at the far distance (0.3 D). Two of the seventeen observers had to be removed (see below), thus this is the average of 15 observers’ data. The initial data point on each graph is calculated from the first 1-n trials, where n is the number of trials in each bin. The graphs show comparable initial d' scores (~1.6), regardless of the bin size. Examining this data, we decided that using 15 trials was a good compromise between the variance in the data and plotting recovery patterns.

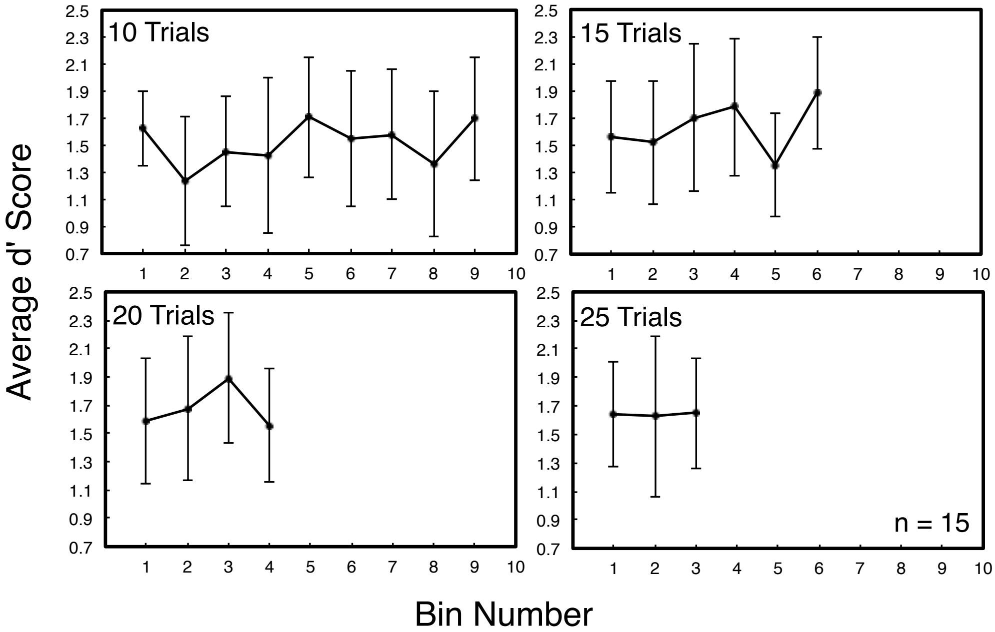


Figure 4.3.4. d' as a function of trial bin number for example post-test measurements. Each graph was generated using different bin sizes, shown on the top right of each. The data are from the ±1.0 D condition at the far screen distance (0.3 D). Each graph generates data plots using x number of trials (specified in each graph). The error bars are 95% confidence intervals.

We therefore calculated the d' scores for the baseline phases based on all 60 trials and the post-test phase based on bins of 15 trials. To calculate the pH and pF, we measured the number of times the observer reported 'right' when they saw a right-orientated stimulus and divided that by the total number of right-orientated trials they would have seen (pH). We then measured the number of times they reported 'right' when they saw a left-orientated stimulus and divided that by the total left orientated stimuli (pF). For some cases, the observers were 100% correct, or 0% correct, resulting in proportions of '1' or '0'. In this situation d' values cannot be calculated. There are two conventional methods of dealing with these extremes: the log-linear rule, and the 1/(2N) rule (Hautus, 1995). The log-linear rule involves adding 0.5 to the *number* of hits and false alarms, and adding 1 to the total right-orientated trials and left-orientated trials, making extreme values impossible to obtain. The 1/(2N) approach involves replacing extreme values of zero or 1 with 1/(2N) and 1-1/(2N), respectively, where N is the amount of trials on which the calculation is based. Hautus (1995) compared the effectiveness of both approaches; the results indicate that both applications work reasonably well. When comparing the approaches with a reduced number of trials (less than 100 in the paper), the log-linear approach appeared to be the more conservative of the two, in part because the transform is applied to all the data, not just the extreme scores. We therefore decided to apply the log-linear approach. Once this transform was applied, we could calculate d' scores for each individual's baseline and post-test measures.

4.4.3 Results and Discussion

Percent correct performance on the exposure phase of the experiment was calculated to ensure that all observers were paying attention and could do the task. Of the 17 observers, two were removed for 'poor' performance (70% accuracy or less) on at least one session. The concern here was that if these observers could not do the task (i.e. they stopped trying), they are effectively not being exposed to the conflict anymore, and so won't show after-effects even if they are susceptible to them.

We calculated the d' score for the baseline and all bins of data in post-test phases, for each observer and each condition. To examine the effect of the exposure phase on subsequent stereo performance, we calculated the difference between the baseline scores and the post-test scores in the same session. The results are plotted in Figure 4.3.5. The x-axis plots the time in seconds instead of bin number to give a better idea of recovery over time. If there is an effect of viewing conflict on subsequent viewing, we would expect a decrease in stereo performance in the post-test relative to baseline. Therefore, in the graphs, a difference score below zero means the post-test performance was poorer than baseline. Each column in the figure shows the control and the conflict conditions for each magnitude of vergence distance change. The right column shows the results for the screen distance (1.3 D). While the left column shows the results for the distances away from the screen (far distance - 0.3 D). It is important to compare performance in the control and conflict conditions because the control conditions reveal any effects of the task and display, beyond effects of conflict per se.

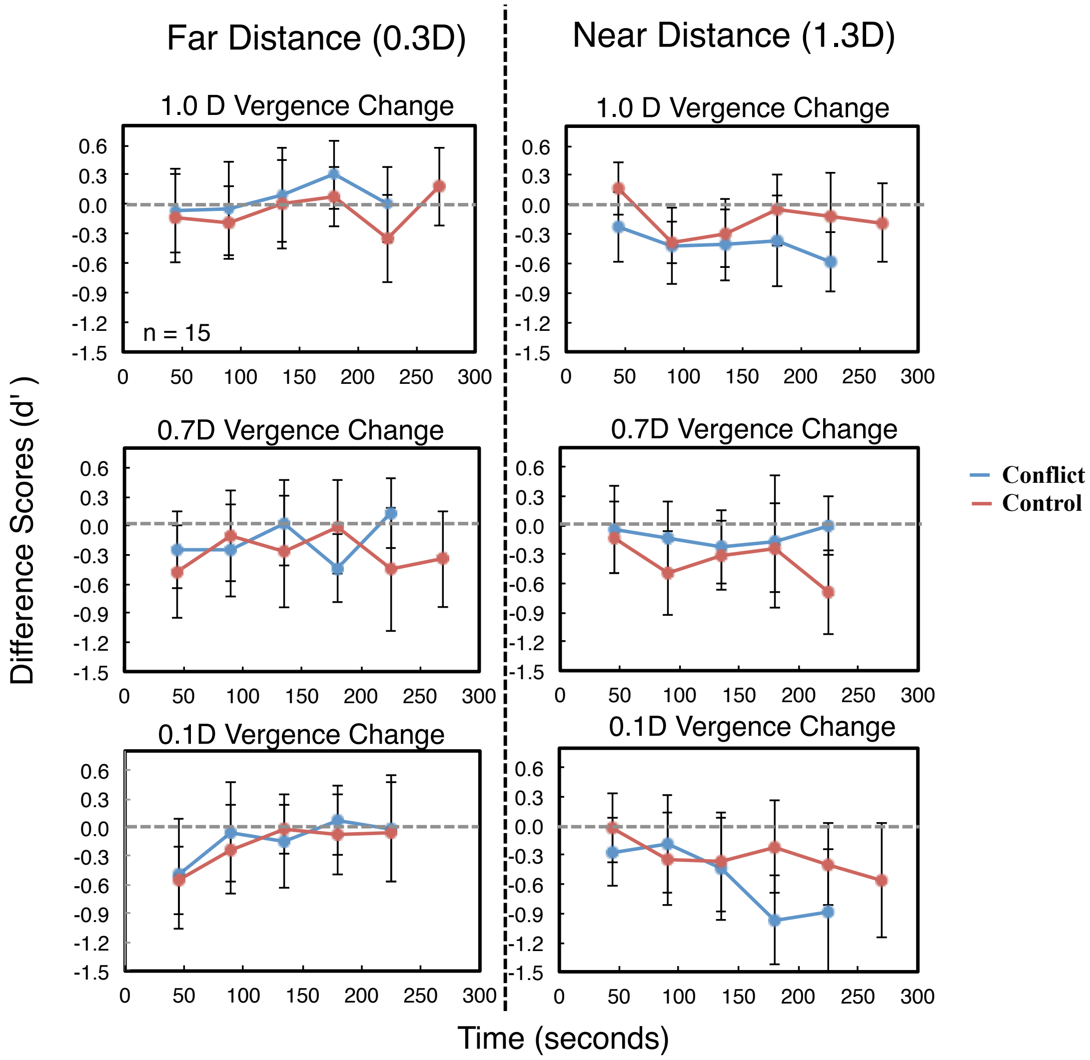


Figure 4.3.5. The results of Experiment 8. Each graph shows the difference scores (post-test minus baseline) for the conflict and control conditions as a function of time, averaged across the 15 observers whose data were included. The blue symbols/lines denote the conflict conditions and the red symbols/lines denote the control conditions. The dashed grey lines denote zero difference score (that would result if baseline and post-test performance were equal). Each row of graphs represents a different vergence change in the exposure phase; 1.0 D (top row), 0.7 D (middle row) and 0.4 D (bottom row). The columns represent performance at the test (1.3 D) or far (0.3 D) distance (baseline and post-test phases). The error bars are 95% Confidence Intervals.

The average results show no clear effects of viewing vergence-accommodation conflict stimuli on subsequent stereo performance with no conflict. There may be a difference between the control and conflict performance for the 1.0 D vergence distance change tested at the screen distance (1.3 D), but if so the effect is very small. It is interesting to note that, for some cases, initial stereo performance for the control conditions (red) fall below ‘zero’, indicating a decrement in stereo performance. This is especially important for researchers who test complex scenes using conventional stereo displays without a matched control. These results mean it is hard to make specific conclusions from the data in these cases.

Statistically, we found no significant difference between the control and conflict performance for the 1.0 D vergence distance change tested at the screen distance (1.3 D), using a Wilcoxon signed rank test. To visualise the differences between the control and conflict conditions, we took the difference between the mean scores of each. The results are shown in Figure 4.3.6. Similar to Figure 4.3.5, this figure shows the difference scores for the far distance (0.3 D) on the left and the near (1.3 D) on the right. The 1.0 D, 0.7 D and 0.4 D vergence changes are on the top, middle and bottom of the graphs respectively. Considering the initial measures (no.1 on the y-axis), we can see that there are no significant differences between the conflict and control conditions at any distance and for any vergence change.

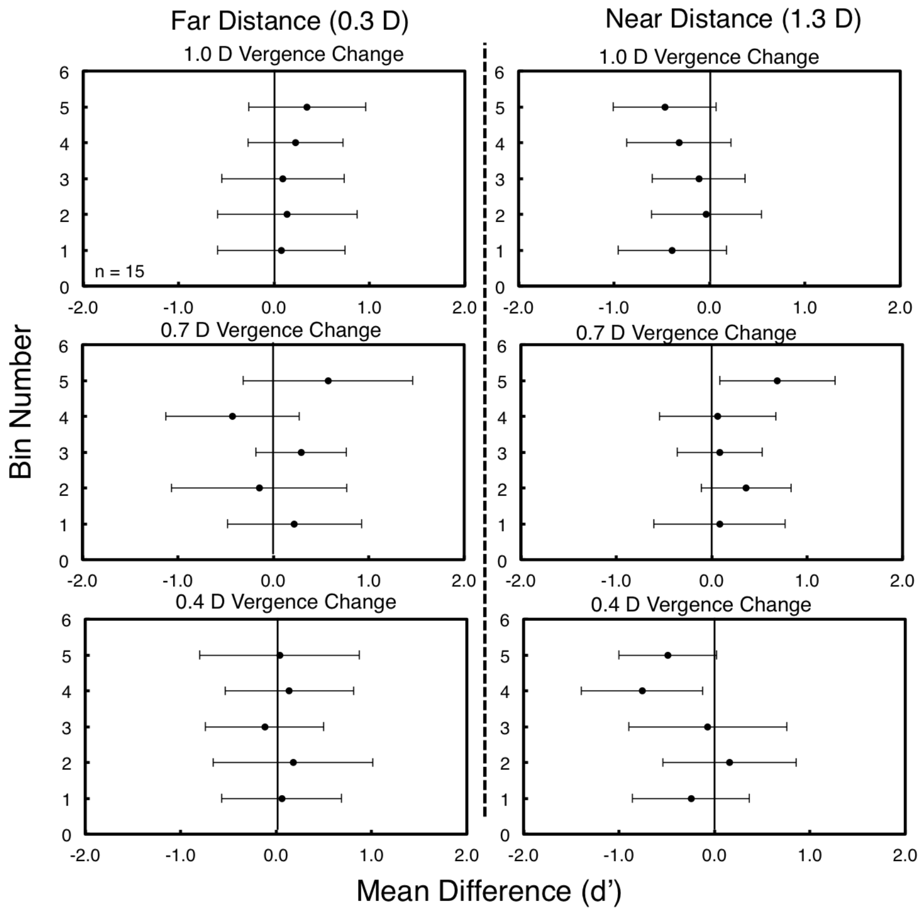


Figure 4.3.6. Difference scores between conflict and control conditions. The graphs on the left show the data for the far distance (0.3 D), the graphs on the right show the near distance (1.3 D). The top, middle and bottom graphs show the vergence change of 1.0, 0.7 and 04 D respectively. The x-axis is the mean difference in d' units, the y-axis is the bin number. Error bars are 95% confidence intervals.

In one sense, these results are not surprising; we know from the past two experiments that tolerance to conflict varies greatly across individuals. As well as this, it is likely that we excluded observers like those in Experiments 6 and 7, who are most affected by conflict. It is therefore likely that in our sample, there are people who are affected by viewing conflict, and those who are not. To determine this, we plotted the difference (between the baseline scores and the post-test scores) scores for individual's data. The results are plotted in Figure 4.3.7. The x-axis plots the participant numbers. For each participant there is a corresponding conflict (blue) and control (red) data point/bar which represents the different score for the first bin of trials. As with Figure 4.3.5, scores below zero on the y-axis represent poorer post-test performance relative to the baseline measure. If there is an effect of viewing conflict on subsequent viewing, we would expect stereo performance for the control to remain the same or get better (higher on the y-axis). Whereas we would expect a decrease in stereo performance for conflict (blue). Each column in the Figure shows the control and the conflict conditions for each magnitude of vergence distance change. The left column shows the results for the screen/test distance (1.3 D). While the right column of graphs show the results for the distances away from the screen (far distance - 0.3 D).

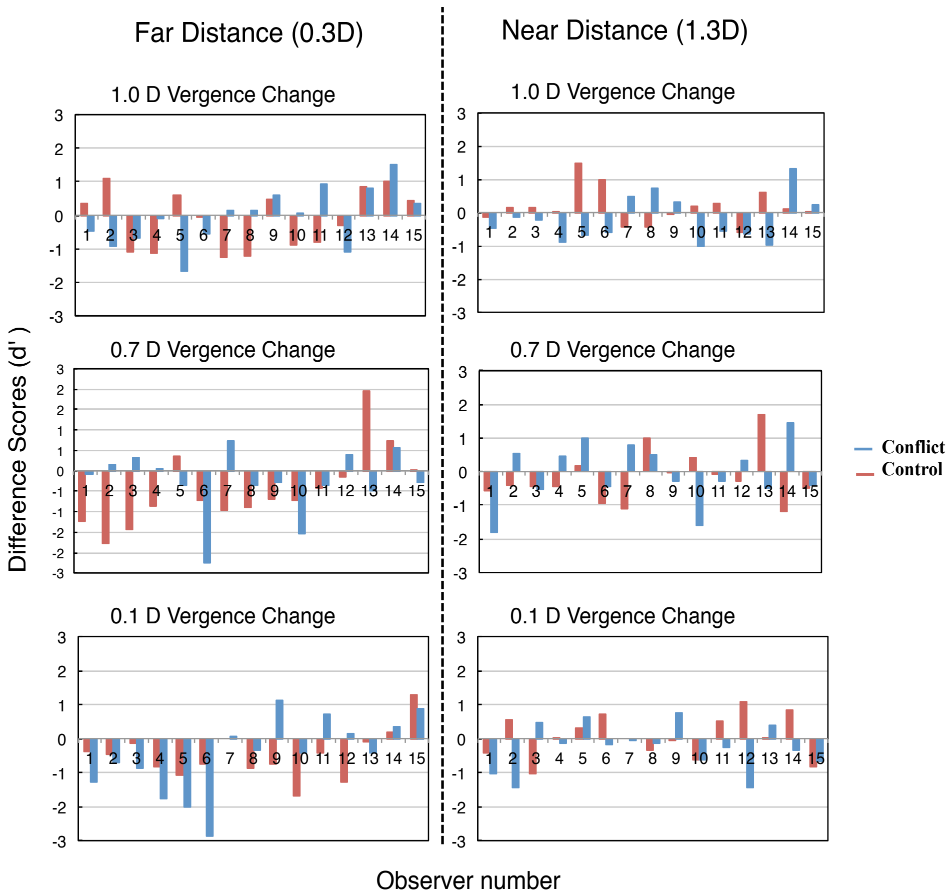


Figure 4.3.7. Individual data plots for Experiment 8. Each graph shows the difference scores (post-test minus baseline) for the conflict and control conditions for each observer. The blue bars denote the conflict conditions and the red bars denote the control conditions. Zero on the y-axis denotes zero difference score (that would result if baseline and post-test performance were equal). Each row of graphs represents a different vergence change in the exposure phase; 1.0 D (top row), 0.7 D (middle row) and 0.4 D (bottom row). The columns represent performance at the test (1.3 D) or far (0.3 D) distance (baseline and post-test phases).

Although there is no effect of exposure to conflict on stereoacuity in subsequent no-conflict viewing, individual data on figure 4.3.7 show that some observers appear to have reduced stereo performance after viewing conflict relative to after viewing the control condition. Given the results of earlier experiments in this Chapter, we expected some but not all of our observers to be affected, and so this result is not entirely surprising. We therefore wanted to examine how those affected in the initial bin of trials recovered over time. To do this, we examined the largest vergence step condition. Of the fifteen observers, four had reduced stereo performance at the far, 0.3 D post-test distance (observers 1, 2, 5, and 12) while another six observers had reduced stereo performance for the test screen distance (1.3 D; 4, 5, 6, 10 and 13).

This indicates that any effects on stereo performance as a result of conflict may vary with distance.

The results of the average performance across the observers who were affected by viewing conflict stimuli for the 1.0 D vergence-change conditions are shown in Figure 4.3.8. It can be seen that there are initial decrements in performance relative to the control conditions at both distances. In both cases, stereo performance after viewing conflict increases over time until at, or around baseline values. Sensitivity after viewing conflict recovered two minutes after exposure (30 seconds between exposure phase and post-test, plus 90 seconds of post-test). Recovery here is deemed as the point at which the sensitivity for the control and conflict measures converge. A Wilcoxon signed rank test shows that for the Far distance, initial performance in the conflict condition was significantly different to the control condition ($p < .01$) after which there were no significant differences. For the near distance (1.3 D), initial performance was again significantly different from the control condition ($p = .03$). Performance on the conflict condition approached significance at this distance at 180 seconds ($p = .06$) and 270 seconds ($p = .06$). This suggests that for every 10 minutes of relatively large conflict viewing, it takes up to 2 minutes for some individuals visual systems to recover to 'normal' stereoacuity levels when viewing stereo content.

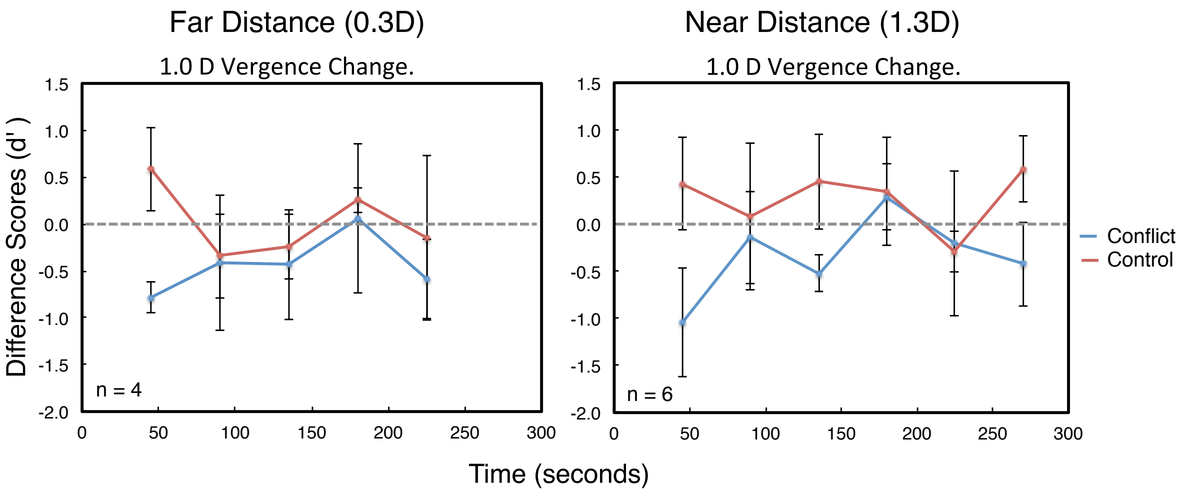


Figure 4.3.8 – Results of Experiment 8. Each graph shows the difference scores (post-test minus baseline) for the conflict and control conditions as a function of time, averaged across the observers whose data showed an effect of conflict. The blue symbols/lines denote the conflict conditions and the red symbols/lines denote the control conditions. The dashed grey lines denote zero difference score (that would result if baseline and post-test performance were equal). Each graph represents a different focal distance in the post-test phase (1.3 D; left graph), (0.3 D; right graph). The error bars represent 95% confidence intervals.

Questionnaire Results:

The questionnaire data were averaged across the fifteen observers for each session to establish whether some conditions caused more adverse symptoms than others. The results are shown in Figure 4.3.9. Each graph shows the results for the titled question, plotting severity of symptoms for the conflict and control conditions as a function of vergence change from fixation. The results show no difference between the control and conflict conditions for each measure of discomfort. This suggests that the observers, on average, did not experience excessive discomfort in the conflict conditions, relative to the control. These results are contrary to those found by Read and Bohr (2014) and Shibata et al. (2011). It should be borne in mind, however, that their observers viewed individual conflict stimuli for longer overall testing periods; 20 minutes (Shibata et al. 2011) or ~1 hour (Read & Bohr, 2014) compared to our 10 minutes. The results suggest that perhaps the observers in our experiment were not taxed in the same way, possibly contributing to our overall findings.

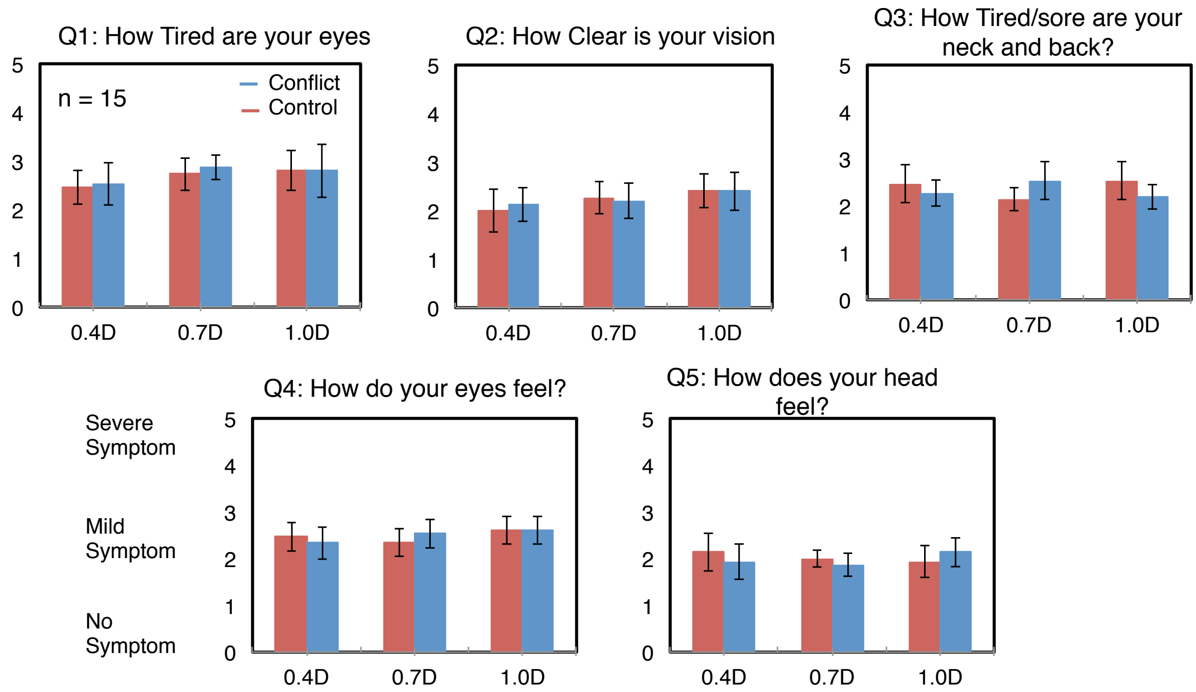


Figure 4.3.9. Results of the questionnaire data in Experiment 8. Severity of each following each exposure condition. The blue bars represent responses in the control conditions and the red bars represent the conflict conditions. . The error bars are 95% Confidence Intervals.

There are a number of outcomes from these findings. Our results repeat the findings of the previous experiments; there are individual differences with conflict tolerance. In the previous research, we found a larger proportion of the observers potentially experiencing aftereffects while viewing conflict. This could be for a few reasons. Firstly, the largest conflict magnitude used here was smaller than the largest conflicts in the previous experiments. Secondly, the observers in Experiments 6 and 7 experienced a range of different conflict magnitudes in a single block of trials, rather than fixed conflict magnitudes. Because an aim here was to examine whether aftereffects scaled with conflict magnitude, we needed to use fixed conflict magnitudes. It would clearly be worthwhile, however, to repeat our aftereffects study with more varied conflict conditions (although presenting an appropriate real-world control condition would be challenging). Finally, we had to remove some observers who couldn't do the task, potentially excluding those most affected by conflict from the study.

For those that had reduced sensitivity to the stimulus after viewing the large conflict, it took up to 2 minutes to recover from 10 minutes of viewing conflict stimuli. A possible avenue for future research is to examine whether this recovery time scales with exposure time and whether the effect is linear.

Finally, the questionnaire results suggest that the conflict stimuli were comparable to the control stimuli in terms of subjective ratings of adverse effects. This is contrary to research examining the subjective experience of viewing stereo 3-D content (Read & Bohr, 2014; Shibata et al., 2011; Yang & Sheedy, 2011; Lambooij et al. 2009; Hoffman et al., 2008). Again, this may be indicative of the study itself, suggesting that the conflict magnitude, or the exposure time, was not large enough. However, despite the caveats of our study (single conflict magnitudes and short duration time), the data imply that any effect of viewing conflict is relatively small when considering both the proportion of individuals affected and the seriousness of the effects. The results are therefore good news for conventional 3-D stereo content. These findings are consistent with similar research conducted by Read and Bohr (2014). In their study Read and Bohr measured subjective ratings of visual discomfort after viewing a feature length film in either 2-D or stereo 3-D. Their results show that only 14% of their observers reported experiencing adverse effects when viewing stereo 3-D content.

4.5 Discussion

4.5.1 Summary of results

In this chapter we aimed to measure the effects of vergence-accommodation conflict on stereo performance. In Experiment 6, we attempted to map out a 'zone of good stereo performance', analogous to the zone of comfort determined by Shibata et al. (2011). Unfortunately, given our results we were unable map out a zone. This was, in part, due to the effects of viewing multiple conflict magnitudes simultaneously, which lead to unexpected results. We found that, generally, as conflict magnitudes increased, stereoacuity thresholds also increased. We also found large individual differences in stereo depth perception with conflict. For some observers, as the conflict magnitudes increased, their stereo performance became catastrophic to the point where they could not complete the task at these vergence distances. Surprisingly, however, two of the observers were also unable to complete the task for the no-conflict trials. We suggest this could be a result of adaptation of the vergence response causing 'after-effects' on subsequent viewing.

The other unexpected finding was that the lowest stereoacuity thresholds were not always measured when the vergence and focal distances were matched. This was explored further in Experiment 7 where the effects of conflict were measured with fixed vergence and focal distances. We also measured each observer's phoria, to investigate whether phoria predicts the vergence distance with the lowest stereoacuity threshold. We replicated the findings of Experiment 6, as the conflict magnitude increased, so too did the stereoacuity thresholds. Again, we found large individual differences across observers, with phoria being related to performance on the task for one observer (SJW) and to an extent, for the other observers.

Finally, in Experiment 8, we investigated potential 'after-effects' of viewing conflict on subsequent stereo depth perception. This also found evidence of individual differences, with a smaller proportion of observers experiencing negative effects of viewing conflict. We examined the overall recovery time, averaged across the observers who were affected at the 1.0 D vergence-step conditions for both the test screen (1.3 D) and far screen (0.3 D) distances. We found that observers were affected differently at the different screen distances. This indicates that exposure to conflict might have different subsequent effects for real stimuli at the distance of the screen and real stimuli at other distances. Nonetheless, performance appears to recover after ~2 minutes of real-world viewing from ten minutes of 'exposure' to conflict for both accommodation distances.

Implications

Individual's tolerance of stereo depth perception to conflict is highly idiosyncratic. In order to create content that is comfortable and easy to view, individuals' tolerance to conflict needs to be fully understood. As already discussed, some research has been done by Shibata et al. (2011) to outline the adverse effects of viewing conflict with respect to visual discomfort and fatigue. Not enough is known, however, to create comprehensive guidelines for the production of stereo 3-D content (Banks et al., 2013). Many factors are yet to be fully researched that are currently not yet well understood, these include but are not limited to stereo performance with conflict and the factors predicting individual susceptibility to conflict (Banks et al., 2013).

It is therefore important to understand tolerance to conflict in terms of stereo performance and how it varies across individuals if meaningful guidelines are to be established. In Experiment 6, we unsuccessfully attempted to measure the 'zone of good stereo', analogous to the zone of comfort set out by Shibata et al. Part of the reason for this, was an inability to measure stereo performance for all conflict magnitudes, and for some cases, even for the no conflict trials. In Experiments 6 and 7 the data showed that viewing stereo content with conflict that is outside an individuals' tolerance range could result in subsequent decrements in stereo performance. Also, research by Kim, et al. (2014) has shown that the temporal frequency of stereo content can affect how comfortable individuals feel when viewing stereo 3-D content. Therefore, the high-frequency changes in the vergence position of the eyes may have also contributed to the poor subsequent stereo performance.

A possible explanation for the poor performance of some individuals in Experiments 6 and 7 is an inability to fuse the stimuli for either crossed (near) or uncrossed (far) disparities (Wilmer, 2008; Patterson & Fox, 1984; Jones, 1977; van Ee & Richards, 2002; Richards, 1970). Richards (1970) states that the probability of lacking the ability to respond to crossed, uncrossed or zero disparities is about 30%, which translates to 2.7% of the population missing one of these functions. However, a study by Jones (1977) found that a relatively high proportion of individuals in their study were found to have deficiencies of vergence eye movement when viewing disparities. This resulted in a marked reduction or absence of convergent or divergent eye movements. This may explain some of our findings, as those individuals who were unable to perform at the task, were often unable to do so for uncrossed disparities. It might be important to consider this for future studies (explored below).

Our results also suggest that some observers may experience an 'after-effect' of viewing conflict stimuli (Experiment 8). This effect and the proportion of individuals who experienced it was greatly reduced relative to the previous experiments. However, the poor performance measured in Experiments 6 and 7 might also be due to the dynamic change in the stimulus to vergence within a range of distances. This is very difficult to control for, but is important too, if we are to understand the effects of conflict on depth perception for all audiences. A possible solution to these issues is to vary conflict magnitude according to the observers' stereo performance instead of using fixed conflict magnitudes, which may or may not be within their tolerance range. This will ensure that during the course of the experiment, the issues associated with viewing stimuli with conflict that is beyond the individuals' tolerance and viewing multiple conflicts will not affect stereo performance. In our next experiment, we again try to measure a 'zone of good stereo', employing this revised method (see Chapter 5).

When considering individual tolerance to conflict, it is reasonable to expect that visual factors that affect these systems can predict how well a person tolerates conflict. For instance, ophthalmological variables such as 1) an individual's ability to decouple their accommodation and vergence responses, 2) their ability to accommodate, 3) as well as their phoria, are likely to affect tolerance (Banks et al. 2013). Shibata et al. (2011) found that an individual's phoria and zone of clear single binocular vision (ZCSBV) can predict their susceptibility to visual fatigue and discomfort. Our research (Experiment 7) has also shown phoria to be a possible predictor of an individual's stereo performance. In the next experiment we examine how those and other visual properties predict an individual's tolerance to conflict. This adds to the growing amount of research into tolerance to conflict.

Conclusions

The results suggest there are large individual differences for stereo performance with conflict. For some of our observers, viewing conflict beyond their tolerance resulted in catastrophic performance such that they could not perform on the task. For some, they also had catastrophic performance for the no-conflict trials. We suggest that these individuals may, therefore, be experiencing an after-effect of viewing conflict. We also found that the location of best (lowest) stereo performance is not always at the location of the screen, but can be away from it. We found that this was highly related to one observer's phoria, and slightly related to the other observers' phorias.

Given the findings in this chapter, individual differences should be taken into account. Given the possibility that individuals may not be able to fuse crossed or uncrossed disparities, it seems that when measuring tolerance to conflict, it is important measure to the limits of tolerance rather than test with conflicts that are beyond the individual's tolerance. Also, it is possible that certain ophthalmological factors, such as phoria, are related to stereo performance with conflicts.

CHAPTER FIVE

5 AGE-RELATED CHANGES IN ACCOMMODATION PREDICTS PERCEPTUAL TOLERANCE TO VERGENCE-ACCOMMODATION CONFLICTS IN STEREO DISPLAYS.

5.1 Introduction

An emergent theme of Chapter 4 was that there can be large individual differences in tolerance of stereoscopic depth perception to vergence-accommodation conflict. In this chapter, we examined various possible predictors of these differences. To do this, we again measured tolerance to conflict (in terms of its effect on stereoacuity) for different groups of people, and related them to other measures of visual function. The aim was both to understand more about the causes of different tolerances to conflict, and also to provide information that might be of use to content producers when setting parameters such as the 'depth budget' — the range of depths around the screen distance at which objects of interest are presented (Mendiburu, 2009).

Also, in Chapter 4 we found some evidence that exposing people to a variety of wide-ranging conflicts, some of which are beyond their fusion capability, can lead to subsequent failures of stereo depth perception for small conflicts, or even no-conflict stimuli, for which they would be expected to perform normally. This is problematic for measuring tolerance, because it becomes a moving target. This poses particular difficulties for this study because we wished to compare tolerance across different population groups (see below). In this chapter we therefore developed a different method: we measured tolerance by examining the conflict required to reach a criterion level of performance, while not presenting stimuli with conflicts far beyond that which a person could fuse/tolerate. This is explained in detail, below. First, we outline the various 'predictor variables' studied.

Presbyopia

The principal variable of interest in this study was the degree of participants' presbyopia. Presbyopia refers to the fact that, as we age, there is a progressive decline in the ability to vary accommodation, due to hardening of the eye's lens (Mordi and Ciuffreda, 1998). At 20 years old people can generally adjust accommodation through a range greater than that required by most everyday tasks (>10 D). This range declines at different rates in different individuals, but by 60 years of age most people have no more than a 1 D range (Ramsdale & Charman, 1989). There are two classes of thought about how presbyopia may affect tolerance to vergence-accommodation conflict. One idea is that people who are presbyopic are more susceptible to adverse effects as a result of conflicts. Mendiburu (2009) states that decoupling the accommodation and vergence cross-links can be painful and the effort involved increases with age. As a result presbyopes 'may be unable to practice it' whereas 'kids' are more robust and can practice decoupling (Mendiburu, 2009). Another idea is the exact opposite. Because presbyopes cannot vary their accommodation, and so remain focused at a fixed distance, they view the world with conflicts between vergence and accommodative responses naturally, and in fact their oculomotor responses are very similar to those required for viewing stereo media (Yang et al., 2011; Banks, et al., 2012; Read & Bohr, 2014). Neither of these ideas specifies predictions for individuals who are not yet presbyopic but are beginning to experience effects of decreased amplitude of accommodation (referred to here as intermediate presbyopes; this typically begins around the age of 40; Duane, 1912) and how they might tolerate conflicts. If there are effects of presbyopia on tolerance to conflict, however, information about the age of the audience could potentially be used to create stereo content that has as much stereoscopic depth as possible, while remaining suitable for the target audience.

In this experiment we examine how degree of presbyopia and other visual properties predict an individual's tolerance to conflict. The properties being measured are discussed below, they are; visual acuity, ZCSBV, phoria, and pupil size.

Pupil Size

Another visual factor that changes with age is pupil size. In their study, Birren, Casperson and Botwinick, (1950) found a pupil size difference of 2.74 mm between the younger (aged between 20-29 years) and older (aged between 60-69 years) observers when measured in dark conditions. It is advantageous for older individuals to have a reduced pupil size, as defects of refraction and accommodation are reduced when viewing through a pinhole aperture (Birren, et al., 1950). A smaller pupil size also leads to an increase in the depth-of-focus of the eye (Watt, Akeley, Ernst and Banks, 2005; Atchison, Fisher, Pedersen and Ridall, 2005).

Most of what we know about the depth-of-focus of the eye and tolerance to blur comes from optometry research designed to recognise the point of just-noticeable blur. However, individuals will tolerate a reasonable amount of blur before images are perceived to be 'troublesome' (Atchison et al., 2005). In their study Atchison et al. examined limits of blur that would be considered 'troublesome' and how it changes with pupil and letter sizes. They found that as pupil size increased the amount of tolerable blur decreased. This indicates that a smaller pupil size means a greater tolerance to defocus away from the stimulus to accommodation. Therefore, older observers who have smaller pupil sizes may have a greater depth-of-focus relative to younger observers and therefore might tolerate a larger conflict magnitude.

The size of pupil does not only change with age, other factors also affect pupil size when viewing images. Examples include emotional stimulation (Partala & Surakka, 2003; Bradley, Miccoli, Escrig, & Lang, 2008), processing load (Koivisto, Hyona, & Revonsuo, 2004; Laeng, Ørbo, Holmlund & Miozzo, 2011), and attention (Binda, Pereverzeva & Murray, 2013). The luminance of the stimulus being viewed also affects pupil dilation (De Groot & Gebhard, 1952; Leibowitz, 1952; Denton, 1956; Campbell & Gregory, 1960). With increasing luminance, the pupil size decreases, increasing the depth-of-focus of the eye. In as much as is possible, these factors (emotional stimulation, processing load, attention and luminance) are controlled for in this study (see method section). Finally, the near reflex of the pupil to images that are in close proximity to the observer can also influence the size of an individual's pupil (Campbell & Wetheimer, 1959; Schafer & Weale; 1969; Kasthurirangan & Glasser, 2006; Lowenfield, 1999). The near reflex refers to the pupil constriction, which accompanies the accommodation and vergence response to objects/images that are near to the observer (Kasthurirangan & Glasser, 2006). Near pupil constriction does not increase with age (Schafer & Weale, 1070; Kasthurirangan & Glasser, 2006). Therefore, overall, individuals will have a larger tolerance to conflict for near relative to far distances.

Phoria

Phoria is the relative position of the eyes (vergence posture) in the absence of a disparity stimulus to vergence. That is, when the accommodation cross-link is the only stimulus to vergence. This is typically measured using two methods. Either using a conventional clinical phoropter (Shibata et al 2011; Kim, Granger-Donetti, Vicci and Alvarez, 2010), or a practice where different images are presented to the left and right eyes of the observer so that the stimuli cannot be fused (Walline, Mutti, Zadink and Jones, 1998; Schroeder, Rainey, Goss and Grosvenor, 1996). In either case, the phoria represents the physiological resting vergence position for a given focal distance. As such, it could be argued to be a position of zero vergence effort that corresponds to a given screen distance, and so might be expected to predict tolerance to vergence-accommodation conflicts.

Esophoria and exophoria describe when the phoria state corresponds to either in front of (esophoria) or behind (exophoria) the focal distance of the object. Orthophoria is when the eyes align at the focal distance (Schroeder et al., 1996; Kim et al., 2011). Although there are individual differences (Tait, 1951), people are on average exophoric for near distances and esophoric for far. An individual's phoria could represent the amount of natural offset between vergence and accommodation responses at varying accommodation distances (Banks, et al., 2013). Therefore, zero vergence effort for a certain focal distance may not correspond with the screen distance as explained above, but instead may rest at a distance away from the screen distance. If this is the case, a stimulus that requires an individual to converge at the same location as this position of zero vergence effort (phoria) might actually be less effortful to fuse than stimuli that require a vergence response at the screen. In these instances, stereo performance might be affected accordingly. For example an individual with an exophoric response at a given focal distance may have lower stereoacuity thresholds for stimuli presented beyond the screen relative to stimuli in front of the screen. Research conducted by Shibata et al (2011) investigated subjective ratings of discomfort as a function of viewing distance as well as sign of conflict (stereo images in-front-of and behind the screen). They found an asymmetry of tolerance to conflict with distance. For near distances, ratings of discomfort were greater for images nearer than the screen and at far distances, these ratings were greater for images further than the screen. This asymmetry was linked to the observers' phoria measures. This supports the idea that stereoacuity thresholds for sign and magnitude of conflict may be predicted by phoria.

Zone of Clear Single Binocular Vision (ZCSBV)

The zone of clear single binocular fusion (ZCSBV) is the set of vergence and focal stimuli that can be seen as single and focused while maintaining binocular vision (Fry, 1939). The visual system is tolerant to inaccuracies in accommodation and vergence responses away from the object of interest. The depth-of-focus (approximately 0.25 to 0.3 D) indicates the range of tolerance to inaccurate accommodation, which still results in a clear and focused image. Panum's fusional area (approximately 0.25-0.5 deg) specifies the range of vergence responses away from the object while still perceiving a single image. The range of conflicting stimuli to vergence and accommodation for which responses are within these tolerances (i.e. for which stimuli do not appear blurry or diplopic) is referred to as the *motoric fusion limits* or the *zone of clear single binocular vision* (Fry, 1939; Shibata et al., 2011; Watt and MacKenzie, 2014; Lambooiji, et al., 2009). This differs from a zone of comfort, in that it indicates the *upper limit* of an individual's ability to tolerate differences between the stimuli to accommodation and vergence, for short periods of time, and might require high levels of effort and be highly uncomfortable. The ZCSBV is usually measured by placing equal and opposite prisms (base-in or base-out) in front of the eyes. This varies vergence demand while keeping accommodation distance constant. Therefore, this measure involves forcing the vergence and accommodation system to adjust to increasing distances between the two eyes images. Nonetheless, because it relates to the ability to decouple accommodation and vergence responses, one might expect the shape of an individual's ZCSBV to be related to both their stereoscopic zone of comfort, and their zone of good stereo performance (Shibata et al., 2011). Consistent with this, Shibata et al. (2011) found the ZCSBV to be predictive of the discomfort experienced by observers when viewing conflicts. We therefore also expect that the ZCSBV should be a predictor of when stereo depth perception becomes affected by conflict.

Summary

Previous research has estimated a zone of comfort for stereo viewing — a range of conflicts that result in comfortable viewing (Shibata et al. 2011). We know from previous research and studies presented in this thesis that vergence-accommodation conflict also affects stereo performance. Less is known, about the tolerance of perception of stereoscopic depth to conflict, however, even though this too is important for the creation of effective stereo media. Here we examine how the tolerance of perception of stereoscopic depth to conflicts is related to various other factors that might be expected to influence an individual's 'zone of good stereo performance'.

5.2 Methods

Observers

We recruited fifty-one observers for the experiment, thirty-two of which were included in the final sample. There are a few the reasons why some observers/observers data were excluded. Firstly, one observer was excluded because he was unable to fuse the stimuli at the farthest distance measured (0.1 D). Second, another three observers were excluded due to a change in method part-way through the experiment. Initially, accommodation response was measured using the multi-plane display, and autorefractor (see MacKenzie et al., 2010). Unfortunately, this proved inadequate for measuring accommodation response for observers wearing glasses, due to reflections. We therefore developed a different method to measure accommodation response (see below). From the initial observers measured, we were unable to re-test three on the new method, and so their data were omitted. Finally, a further 15 observers' data were excluded due to a procedural error. During the experiment we discovered an intermittent problem with the code controlling the adaptive staircase procedure. This problem lead to some staircases terminating early. Because we used the convergence point of the staircase as our threshold measure, data could not be used in these cases. These observers' data were removed and the code problem corrected.

We hypothesise that degree of presbyopia will be the main predictor of stereo tolerance to conflict. We therefore attempted to recruit observers with the full range of accommodation response to those with very little range in their accommodation response. As previously discussed, the age of an individual is highly correlated with their accommodation response function. We therefore attempted to recruit from three age groups: 18-25 years old (9 observers) with a mean age 20.5 years (Std Deviation ±1.3 years), 30-50yrs (15 observers) with a mean age of 39.3 years (Std Deviation of ±5.3 years) and 50-70yrs (8 observers) with a mean age of 62.87 years (Std Deviation of +5.7 years).

The observers wore their optical corrections during the experiment, including during the measures of predictive factors (two observers wore contact lenses, 14 wore glasses with a single optical power, two wore varifocals; 15 observers had no correction). Observers wore the corrections they would use under natural conditions for all focal distances in the experiment. Similarly, for the observers with varifocals, the stimuli were viewed through the part of the lenses they would normally view through, under natural conditions. Observers with stereo-blindness or significant ocular pathology or abnormality were excluded from the study.

All observers were volunteers that were reimbursed for their time. We recruited via opportunity sampling, through advertisement around the university and by email. For the older participants, we recruited using Bangor Universities 'participant panel' which is a contact list of volunteer observers. This list is maintained to provide age-matched controls for neurological-patient studies. We used this panel to recruit older observers with no known ocular pathology or abnormality.

Measures of visual function

Measurement of pupil size

We measured the pupil size of the right eye while the observer viewed the test stimulus (random dot stereogram) monocularly, at 0.7 D focal distance. We calculated pupil size in the following manner. First, we captured an image of the observer's pupil using the Autorefractor (normally used to measure the accommodation response), which is part of the right eye's display (see General Methods). The autorefractor uses infrared (i.e. not visible) light, and the image was captured while the observer viewed the stimulus, in the same conditions as the main experiment (i.e. with the room lights extinguished, same stimulus luminance etc.). Subsequently, a ruler was placed in the same plane as the pupil and a second image was captured by the autorefractor. This provided a calibrated reference, allowing us to accurately calculate the size of the pupil in the original image.

Therefore, the luminance, task and instructions were all controlled for all observers to ensure these factors (luminance, emotional stimulation and processing load) had no effect on pupil size (see Introduction). That said, we could control how attentive individuals were during the task.

Measurement of accommodation stimulus-response function

We measured each observer's ability to accommodate to stimuli at a wide range of distances, in order to characterise their degree of presbyopia. To do this, we constructed a monocular stigmascope (Figure 5.2.1 C, D), allowing us to measure amplitude of accommodation subjectively (note our multi-plane display/autorefractor setup can present stimuli only in the focal-distance range 0–1.33 D, which is insufficient to characterise the accommodation stimulus-response function fully).

The stigmascope is shown in Figure 5.2.1. The observer views a reference target, and a stigma (a point-light source), simultaneously, with their right eye, via a beam splitter positioned at 45 deg to the observer's line of sight. The principle of the device is that the point-light source is a poor accommodation stimulus, and the reference target is a good accommodative stimulus. Thus, only the target drives accommodation. The current state of accommodation, for a given target focal distance, is determined by adjusting the focal distance of the stigma until it appears maximally sharp (see below) (Ciuffreda, 1991).

Both target and stigma were viewed by separate Badal lenses (10 D, at 10 cm from the eye). This allowed presentation of the reference target and stigma without size changes or looming cues to distance. It also allowed presentation of focal distances from 0 D (infinity) to 7 D (14.3 cm) in a physically small device.

The reference target was comprised of a number of high-contrast letters (white letters on a black background), illuminated from above. The illumination was provided by a 'light box', containing a white filament bulb, covered by diffusing material, ensuring bright, even lighting across the target (Figure 5.2.1 D). The stigma was created by a second light box, again containing a white filament bulb, and with a 0.5 mm hole drilled in the front to create a point-light source (Figure 5.2.1 C). The exterior of both light boxes was painted matt black. The interior of both light boxes was coated with silver foil to maximise the emitted light. The illumination levels of each light was adjustable using a variable resistor.

Both the stigma and the reference target were mounted on optical rails, to allow their focal distance to me moved, while preserving their visual direction. A chin and forehead rest positioned the observer's head relative to the stigmascope/badal lens system. This rest was adjusted for each observer, in x and y, to ensure they were positioned correctly. The stigma could also be adjusted vertically and horizontally so that it was superimposed on a black portion of the target stimulus (and so maximally visible).

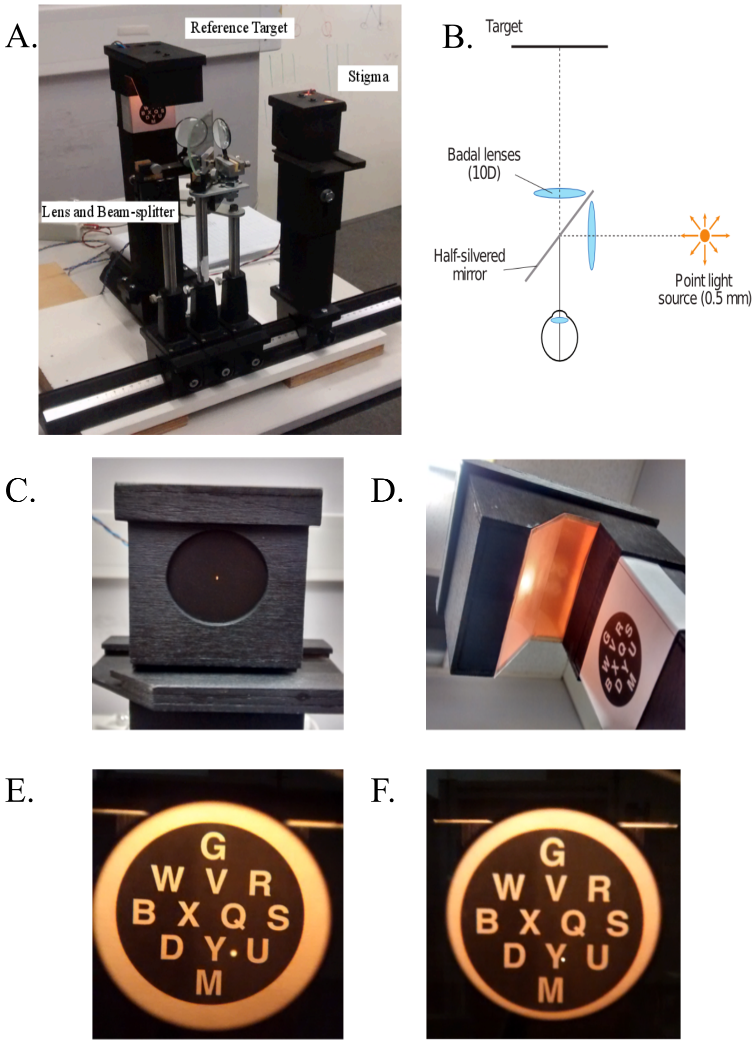


Figure 5.2.1 – Images of the stigmascope used in this experiment. Image (A) is a photo of the device while (B) is an illustration, displaying the components of the stigmascope; this arrangement includes a beam-slitter and two convex lenses (10D). Image C is the point light source (stigma), a pinhole is cut into the otherwise sealed box allowing the light to be directed along the line of sight. Image D is the stigma, with a light illuminating from above with diffuse material to ensure an even distribution of light. Images E and F show what the observer saw during the visual test, image E shows an unfocused stigma while image F shows a more focused stigma.

To measure the accommodation response the experimenter first adjusted the reference target to a given focal distance. The observer was then instructed to keep the letters clear and in focus while moving the stigma to the point where it looked sharpest. This indicated the stigma was optically conjugate to the observers' retina and thus in focus (Ciuffreda, 1991). Where the observer positioned the stigma was recorded and the next distance was then measured. To begin with, observers were allowed to practice this technique until they were comfortable with focusing on the stimulus while attending to the stigma.

The accommodation response was measured at several distances so that we could map out an accommodation stimulus-response function for each observer (0-7 D in 0.5 D increments). Each distance was tested twice. Observers were first tested at each distance, in order, from farthest to nearest (0 D to 7 D). They were then tested a second time, in reverse order (nearest to farthest; 7 D to 0 D). The direction of the required movement of the stigma can result in biased measurements of accommodation state: because the eye has a non-zero depth of focus, movements inward until it appears sharp, or outwards, can result in different settings. Prior to each trial, the stigma was therefore positioned nearer or farther than the target an equal number of times, to minimise this bias. The observer dictated the speed of the measurement session.

Phoria

Each observer's phoria was measured in the same way as in Experiment 7. Phoria was measured at the same three focal distances used for the main experiment: (0.1, 0.7, and 1.3 D).

Visual Acuity

Visual acuity was measured as an inclusion criterion to ensure that all observers had normal or corrected to normal vision. We measured visual acuity using the adaptive Freiburg Visual Acuity & Contrast Test (FrACT; Bach, 2007). This is a free, web-based program that uses anti-aliasing and psychometric methods to provide measures of visual acuity (Bach, 1996). We used FrACT because it is widely used, freely available, and has been shown to work reliably (Bach, 2007). It works by presenting 'Landolt-Cs' in one of eight orientations. The observer's task is to report the orientation (or position of the 'gap') using the numeric keypad. The size of the letter/gap is varied to determine the observers' acuity.

The test was conducted using a Macintosh computer located 3m from the observer at eye level. The observers' visual acuity was tested with and without their corrections, where applicable. Before measuring their acuity, they were allowed to have a practice session to ensure they understood the responses. The stimuli were displayed for 30 seconds, unless a response was made within that time, in which case it would immediately move to the next trial. There were a total of 24 trials. The test was repeated three times.

ZCSBV

The zone of clear, single binocular vision was measured using a procedure similar to that used in clinical assessment (and to that used by Shibata et al., 2011). Observers wore goggles that placed Risley prisms in front of each eye. Risley prisms are continuously adjustable laterally displacing prisms, which allowed the experimenter to alter the vergence stimulus as required, while keeping focal distance constant.

Observers viewed a column of high contrast letters (matched for angular size at different distances), printed on paper. Vergence demand was continually adjusted by adjusting both prisms simultaneously, in an equal and opposite manner, until the observer stated that the letters appeared blurry and/or diplopic (double vision). Inward and outward measures were taken separately. The experimenter adjusted the vergence stimulus at a rate of ~2 prism dioptres per second. In clinical practice, the prisms are adjusted until the image appears blurry, then diplopic and finally the prisms are turned towards ‘zero’ until the image recovers (Scheiman, M., & Wick, B., 2008). The prism values at all three points are recorded. We chose to record when the image appears blurry and/or diplopic to be in line with the measures of Shibata et al. (2011) who used the same method. Also, whether the observer first reported the image appearing diplopic or blurry varied, if the image appeared diplopic, continuing to increase the prisms would have no meaning.

The separation of the prisms goggles was adjusted to match each observer’s IOD. There was an initial practice to ensure the observers understood the task and to get comfortable with the task. We then made inward and outward measurements of the ZCSBV at 1.3, 0.7, and 0.1 D (measurements were made in a 12 corridor). The order in which the distances were tested was randomised.

Procedure

Before testing, we screened each observer for stereoacuity using the ‘circles’ part of the Randot stereo test (Stereo Optical Company, Inc.). Observers with stereoacuity >= 1 arc min did not proceed to the main experiment. The observers’ then completed a centring procedure, which measures the centre of the focal planes for the left and right eyes in the multi-focal-plane display. The centring technique in this experiment was slightly different to the technique described in the general methods chapter. The front focal-plane was centred in the usual manner, using the cross-hair device only here the observers moved an ‘x’ (with the response pad) to where the cross hairs overlap. The subsequent planes were centred by projecting a cross on the front focal plane and manipulating an ‘x’ on the mid focal-plane until it aligned and intersected the cross. This was repeated for the far focal-plane with the cross on the middle focal-plane and an ‘x’ on the far. As with the conventional centring technique, each task was completed three times for each focal plane, and alignment was considered acceptable if the standard deviation of the three centring values was less than 1.25 pixels.

We used this adapted centering method, as it is quicker and easier to perform on the task than the method used on earlier experiments. This was important, as we did not want the testing sessions to be longer than one hour yet we wanted to be able to conduct multiple measures within this time.

Once all these measures were complete, we measured the observers' accommodation response, phoria, visual acuity, and ZCSBV. Accommodation response was tested first as we did not want other measures (including testing in the dark) to affect the response. The other measures, visual acuity, and phoria were tested in a random order across the observers. ZCSBV was tested at the beginning of the second session; this was to prevent other tests or visual fatigue from interfering with the response. The observers were encouraged to take breaks between the visual tests as required.

Practice

Our task (identifying the orientation of a random-dot defined stereoscopic corrugation in depth) can be difficult for observers who have never taken part in a psychology experiment before, and do not know what to expect/look for. To overcome this, observers undertook various practice activities to gain familiarity with the stimuli and task.

First, each observer was shown an example of a random dot stereogram depicting a corrugation in depth (a red/blue anaglyph stimulus shown on a desktop computer). They were given unlimited viewing time, and the experimenter carefully explained the structure of the stimulus, and checked that all observers could readily identify the corrugation pattern in the stimulus, and understood what was meant by corrugation orientation.

Next, observers viewed an example of the experiment stimulus in the multi-focal-plane display. This was to ensure that the observers could do the task at the focal distances measured in the experiment. The peak to trough disparity used for this stimulus was 50 seconds of arc; this magnitude was used, as all observers were accurate on the stereoacuity test for at least this value of crossed disparity. The observer viewed two stimuli at each focal distance. They reported to the experimenter the orientation of the stimuli. For those who could not perceive depth in the stimuli, the peak-to-trough disparity in the stimulus was adjusted until they could repeatedly report the correct orientation of the stimulus.

Finally, observers completed a practice test of the actual experiment task. During this practice test, they received feedback about the correctness of their responses (via a high- or low-pitched tone; not given in the main experiment). As well as training the observer on the task, this also allowed the experimenter to monitor performance and intervene as necessary with additional questions and instructions.

Measuring Tolerance to Conflict

In this experiment we measured the tolerance of stereo performance to conflict at different focal distances, in order to characterise each observer's zone of good stereo. The overall approach of the experiment was rather different to those in Chapter 4. In the earlier experiments we measured stereoacuity at a range of fixed conflicts, interleaved randomly in blocks of trials. This is not efficient, because data points are spread out throughout the space, some of which contribute little or nothing to establishing tolerance (too easy or too hard). Moreover, we found in Chapter 4 that exposure to conflicts beyond a person's tolerance appeared sometimes to result in a 'non-linear' collapse of 'normal' stereo performance, even for stimuli with no conflict, making measurement of a zone impossible.

For this experiment, we therefore developed an alternative method to characterise tolerance to conflict that was both quick, avoided grossly exceeding an observer's tolerance, and would provide data comparable across different observers, whatever their individual tolerance. To do this, we first established a stereo stimulus for each observer that resulted in a criterion level of performance (90% correct). We then measured the 'zone of good stereo', at a single focal distance at a time, by manipulating vergence-accommodation conflict using a staircase procedure (and with fixed disparity level), to identify at what conflict stereo performance degraded to a criterion level. Said another way, the experiment began at zero conflict, and then conflict was gradually adjusted (positive and negative conflicts were interleaved) until the performance decrement reached the intended level. In this way, we could measure a zone in all observers, in a way that was comparable.

Procedure: Initial performance

In order to measure stereoacuity thresholds that corresponded 90% accuracy, peak-to-trough disparity was varied for non-conflict images. Similar to previous Experiments, the disparities were varied using three staircases: one repetition of a 1-down, 2-up staircase, and two repetitions of a 1-down, 3-up staircase. The initial peak-to-trough disparity was the same as used in the practice session for that observer (typically 50 seconds of arc, but see *Practice* above). The peak-to-trough disparity in the stimulus was initially changed in steps of 5 seconds of arc. This step size was halved after the first four reversals, then halved again after another four reversals.

A trial consisted of an audible beep, followed by a Maltese fixation cross, presented for a random time between 1 and 1.5 seconds. This was followed by the stimulus, presented for 2.5 seconds after which the observers made their response. As soon as they responded, the next trial began. Trials finished after 12 reversals of all staircases. The data were then immediately analysed, to find 90% threshold. This stereoacuity threshold was then used as the peak-to-trough disparity value for the main experiment in which the vergence distance was manipulated. All random dot stimuli had a corrugation frequency of 1 cpd.

Procedure

To measure tolerance to conflict, we manipulated conflict magnitude, for both positive and negative conflicts (in-front-of and behind the screen) while fixing the stimulus to accommodation. The stimuli were the same as above, except the peak-to-trough disparity in the image was fixed according to the individual observer's 90% correct point. We tested negative and positive vergence distances in the same block, chosen at random on each trial. Conflict magnitude was controlled by two separate, interleaved staircases; one for each sign of conflict. Both staircases used a 1-up, 2-down rule where up refers to making the task easier (in this case reducing conflict, irrespective of sign). These staircases would be expected to converge at ~75% correct, yielding a zone width that corresponds to a comparable decrement in performance across conditions and observers. The staircase step-size was equivalent to 0.1 D, which was halved after four reversals of the staircase. A block of trials continued for twelve reversals of both staircases. There were three repetitions of this measure for each session (and therefore each distance).

On each trial, the fixation cross was presented at the focal distance. Other stimulus parameters were the same as in *Initial Performance.*

5.3 Results

Analysis

For some of the analysis we examine the relationship between two variables by applying a linear trend line, the equation of the line and the correlation coefficient (r^2) value for the fit. This value shows the strength of the relationship between the two variables. The closer the r^2 value is to 1.00, the stronger the fit.

Degree of presbyopia

We characterised the degree of presbyopia (DoP) for each observer on the basis of his or her accommodation stimulus-response function. Figure 5.3.1A shows examples for three observers. The left plot shows an observer whose responses were highly accurate (and linear) across the whole range tested. This is typical of younger, non-presbyopic observers (Ciuffreda, 1991). The middle plot shows an observer whose responses varied linearly with changes in stimulus distance up to ~5 D, but beyond that could not accommodate nearer. This is also a typical pattern of responses, where accommodation has a clear 'near-point' (Ciuffreda, 1991). Note, however, that the slope of the linear portion of the response function was substantially less than the previous example, indicating that this observer was unable to vary accommodation by the same amount. We describe such observers as 'intermediate', with respect to DoP: there is still a systematic response with respect to the stimulus, but its range is reduced compared to a non-presbyope. Finally, the right plot shows an example of an observer who is almost completely unable to vary his or her accommodation response with changes in stimulus distance, and is therefore presbyopic. We derived a single number to describe the DoP for each observer using the slope of the linear-regression fit to the linear portion of his or her accommodation stimulus response function (the solid lines in each plot).

Figure 5.3.1B shows the degree of presbyopia for all the observers in our sample, characterised in the manner described above. The data are ranked from lowest to highest degree of presbyopia (high to low slope of the linear portion of the stimulus-response function). For the purposes of some analyses we then divided the observers into three groups, on the basis of their degree of presbyopia. Nine observers had an accommodation response slope of near zero. We classified these observers as presbyopic. We classified the seven observers with a slope between 0.2 and 0.7 as 'intermediate', and the remaining 16 observers (slope >0.7) as non-presbyopes. There was a clear step in slope between the intermediates and presbyopes. However, above a slope of 0.2 there was an essentially continuous variation in slope, and so the upper distinction between intermediates and non-presbyopes is more arbitrary. We therefore chose a slope of 0.7 and higher to indicate a non-presbyope because above this accommodation responses were reasonably accurate, and it placed the divide between non-presbyopes and intermediates at the biggest drop in slope value.

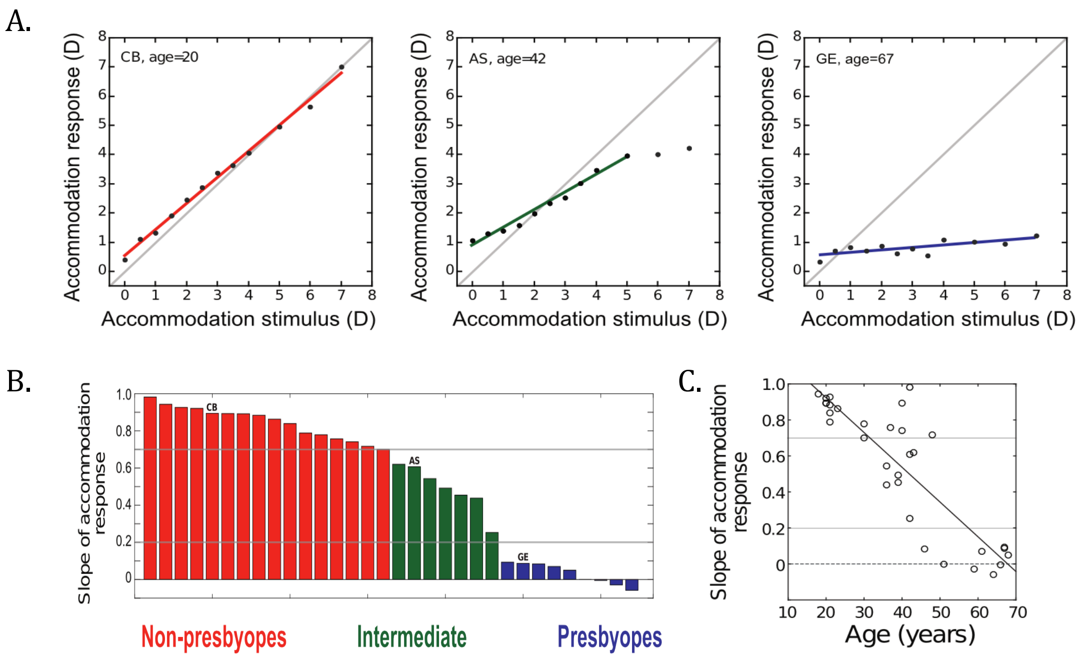


Figure 5.3.1 – Classifying the degree of presbyopia. A) Three observers' accommodation stimulus-response functions. The x-axis is the stimulus to accommodation in dioptres; the y-axis is the accommodation response in dioptres. The grey diagonal line indicates perfect accuracy. The solid lines in each case are linear regression fits to the linear portion of the response function, used to classify the degree of presbyopia. The initials and age of the observers is shown at the top left of each graph. B) The degree of presbyopia for all observers in the study. Each bar represents an individual's accommodation response slope (y-axis values), arranged in magnitude order. The thin horizontal lines show the dividing lines between categories of degree of presbyopia, also indicated by different coloured bars. C) The relationship between age and degree of presbyopia for each observer. The x-axis is the observers' ages in years, and the y-axis is the slope of accommodation response. The thin horizontal lines again show the dividing lines between degree-of-presbyopia categories. The solid diagonal line is a linear regression fit to the data.

Figure 5.3.1C shows the relationship between the age of the observers and DoP. As expected, there was a strong reduction in accommodation response slope (increased presbyopia) with increasing age. It is interesting to note, however, that while there is clear clustering at either end of the graph—observers under 25 years old all have very low DoP and observers aged 50 and above all had very high DoP—in the middle-age range (25-50 years old) the relationship was very unsystematic. Indeed, slopes for observers in their 40s ranged from 0.1 to almost 1.0. This shows that although age is a good indicator of degree of presbyopia, there is a lot of variability in its onset.

Visual Acuity

Figure 5.3.2. shows mean visual acuity for each DoP group. It can be seen that visual acuity was very similar in the three groups, and so should not have been a limiting factor in viewing the stereoscopic images.

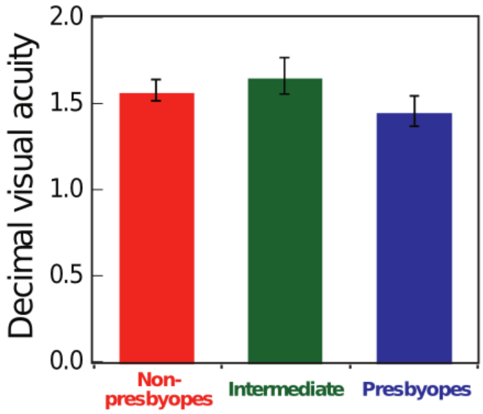


Figure 5.3.2 – Visual acuity for each degree-of-presbyopia group. The bars denote the average decimal visual acuity for non-presbyopes (red), intermediate (green) and presbyopes (blue). The error bars denote standard error of the mean.

Zone of clear, single, binocular vision (ZCSBV)

For each set of prism values at which observers reported the images as blurry and/or double, we calculated the corresponding vergence distance (in dioptres). As the ZCSBV was measured three times, the average of the three repetitions at each focal distance was calculated. Which generates inward and outward ZCSBV values for each observer, at each focal distance. The average ZCSBV values for each DoP group are plotted in Figure 5.3.4. The results show that all three groups had a similarly sized ZCSBV farther than the focal distance. Nearer than the focal distance, however, while the non-presbyopes and intermediates had a similarly sized ZCSBV, the presbyopes had a larger ZCSBV. This suggests that presbyopes can tolerate greater positive conflict magnitudes and still maintain clear vision relative to non-presbyopes and intermediates.

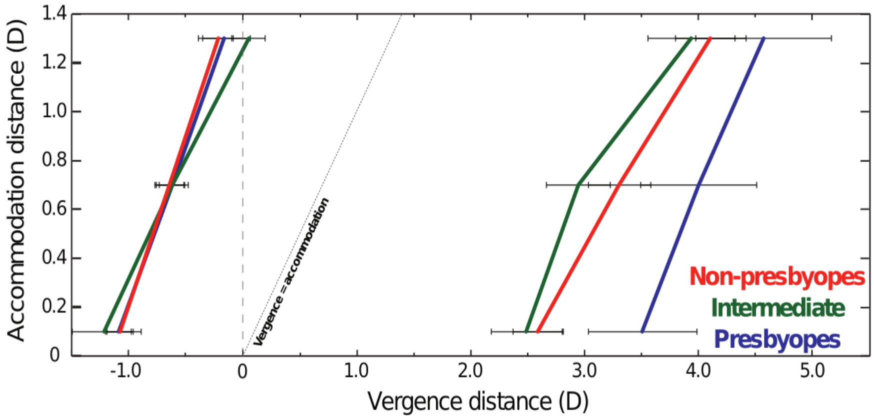


Figure 5.3.4 – The measured zone of clear single binocular vision. The x-axis is the vergence response in dioptres; the y-axis is the focal distance in dioptres. The coloured lines show the average positive and negative ZCSBV measurements for non-presbyopes (red), intermediates (green) and presbyopes (blue). The thin grey line is the natural-viewing line where vergence and focal stimuli are matched. The vertical dashed line shows optical infinity in terms of vergence (0 D); data beyond this corresponds to divergent responses. The error bars are standard error of the mean.

Pupil Size

Unfortunately, some of pupil size images were difficult to discern. To ensure the correct pupil sizes were measured, the experimenter and a lab assistant measured the size of each pupil image. Only the pupil images that both researchers agreed the size of were used in the sample. A total of 17 observer's pupils were used; five presbyopes, three intermediates and nine non-presbyopes. An example of the pupil images and ruler are shown in Figure 5.3.5. Panel A shows a pupil outline. Panel B is an example of a photograph of the ruler, which was used to scale the pupil image.

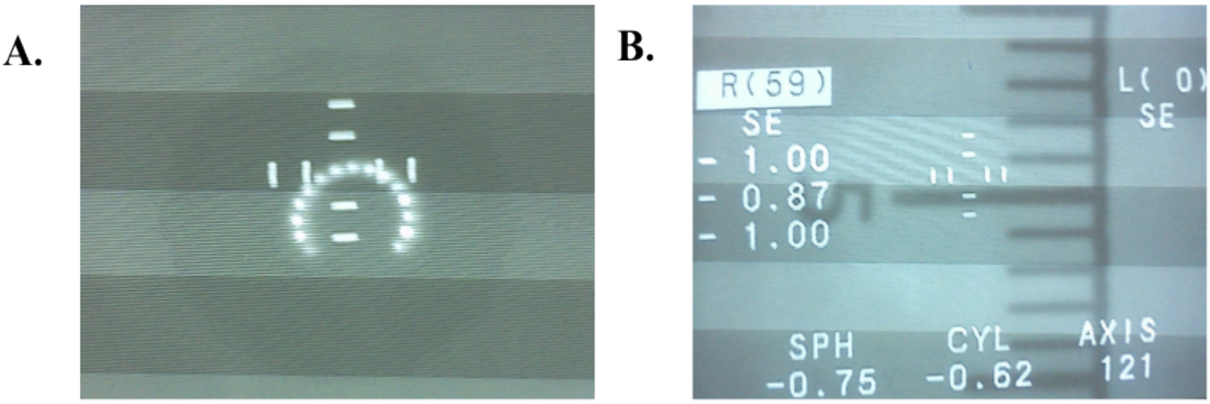


Figure 5.3.5 – Example image of a pupil (A) and corresponding image of the ruler (B). A. An observer's pupil. The white dots are a reflection of the infrared measurement light emitted by the autorefractor, and used to measure accommodation. The outer, dark-grey circle is the pupil itself. B. An example of the ruler pictured at the same location as the observer's eye. The numbers are readings from the autorefractor.

Figure 5.3.6 plots each individual's pupil size (in mm) as a function of degree of presbyopia (a) and age (b). These values indicate no relationship between pupil size and degree of presbyopia and a negative trend between age of the individual and pupil size. This suggests age is a related factor to pupil size and not degree of presbyopia. However, this result is minimal (with an average change of ~1mm across age). This may be due to small sample sizes.

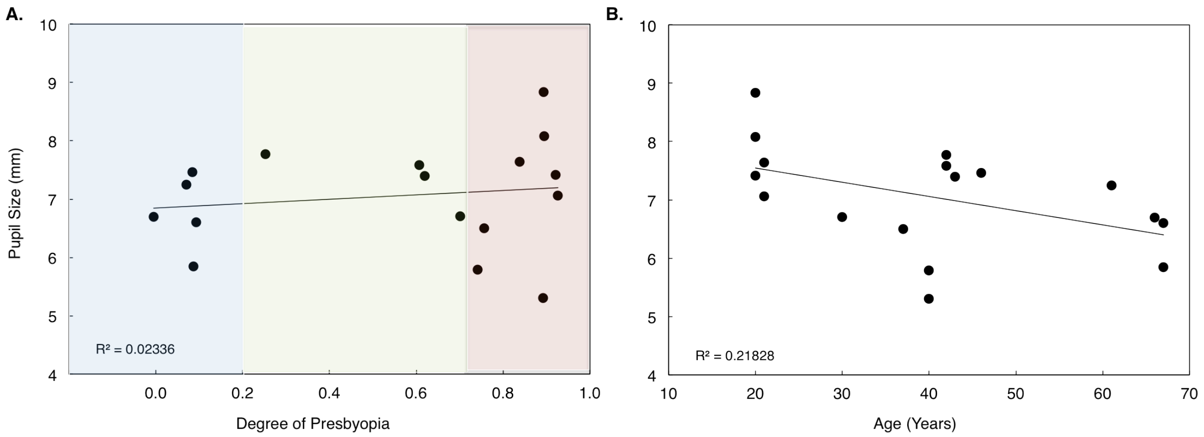


Figure 5.3.6 Pupil size as a function of the slope of the accommodation response (A) and age (B) for the 17 observers for which data were available. The coloured bars on A denote whether the observers were classed as presbyopic (blue), intermediate (green) or non-presbyopic (red). The solid lines in both plots are the best fitting linear regression to the data..

Phoria

As discussed, an observer's phoria represents the physiological position of resting vergence for a given focal distance. It typically varies with differences in accommodation, presumably due to the accommodative convergence cross-link (Martens & Ogle, 1959; Mays

& Gamlin, 2000; Yang & Sheedy, 2011; Heron, Charman & Schor, 2001; Schor, 1992). As such, variation in phoria should be affected by the observers' degree of presbyopia: if the accommodation response does not change with target focal distance, phoria will not be expected to change either.

We therefore calculated the average phoria results for each DoP group. The results are shown in Figure 5.3.7. Panel A shows the average phoria per group at each focal distance. The data show a general trend across groups; there is an exophoric response at the near distance (1.3 D) and a large esophoric response at the far distance (0.1 D). This asymmetry in response is consistent with previous research (Tait, 1951; Shibata et al., 2011). However, unexpectedly, the average phoria response at all focal distances was similar across groups. For the presbyopes (blue line), phoria varies only slightly, resulting in an almost linear data line, regardless of focal distance. For the intermediates (green lines), again the vergence response appears to be monotonic, also regardless of focal distance. Finally, for the non-presbyopes, although the phoria seemed to scale with focal distance for the near and mid focal distances (07 D), there was a particularly large esophoric response at the far distance. These phoria data are atypical for the non-presbyopic and also intermediate groups when compared with previous research. Both Shibata et al. (2011) and Tait (1951) report only a slight esophoric response at far distances. However, their measurements were at nearer focal distances (0.33 and 0.17 D respectively) and are therefore not directly comparable to our far focal distances (0.1 D). It is possible, although unlikely, that the data for the intermediate and non-presbyope groups are consistent with an inaccurate accommodative response, resulting in data patterns similar to the presbyope group.

We also analysed the distribution of the phoria response for all observers. The histograms in Figure 5.3.7B show the distribution of the phoria response for the near (1.3 D; top graph) mid (0.7 D; middle graph) and far (0.1 D; bottom graph) focal distances. The data show that the majority of observers' responses for the near and mid focal distances were only slightly exophoric or esophoric respectively. However, there is a large esophoric response to the far distance that is consistent with panel A.

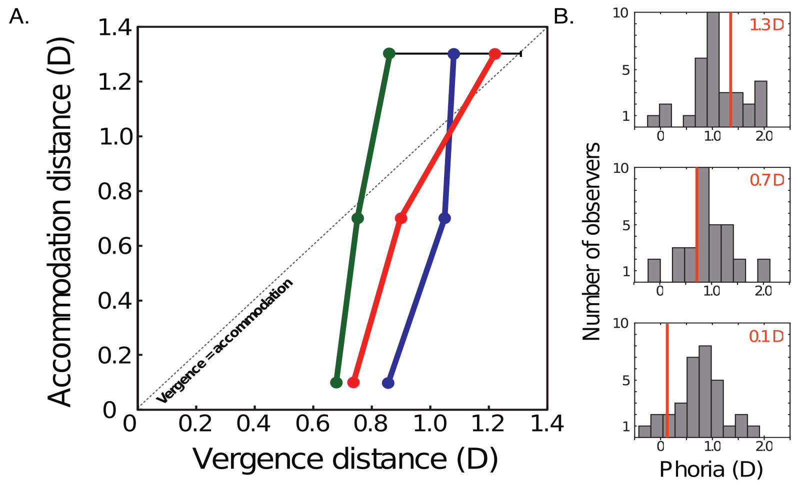


Figure 5.3.7 – Average phoria for each degree-of-presbyopia group. Panel A shows the vergence distance in dioptres on the x-axis and the focal distance in dioptres on the y-axis. The colours of the lines represent the non-presbyopes (red), intermediates (green) and presbyopes (blue). The error bars (plus/minus standard error of the mean) overlapped closely at all focal distances, and so we have plotted one representative bar, for clarity. The histograms plot the frequency of each phoria value, at each distance (shown on the top right corner of each graph). The vertical orange lines also show the focal distance in each case.

A possible explanation is the effect of observers' visual correction on their phoria measure. We chose to measure phoria with observer's wearing the corrections they would use for the focal distance measured. As mentioned previously, the observers wore the corresponding corrections for all measures in the experiment. This allows us to make accurate comparisons across measures and importantly here meant our phoria measure was accurate for a given stereo focal distance condition. However, an individuals corrective lenses change the focal distance and therefore the possible accommodative response. This would possibly affect their phoria response to that distance. Figure 3.5.7 shows the phoria responses for observers when wearing their corrective lenses. However, we also (where possible) measured phoria response while the same individuals were not wearing any corrections. Therefore, to investigate whether the observers' corrections influenced the average phoria responses, we analysed the phoria data without the corrections. There were a total of eight presbyopes, four intermediates and 15 non-presbyopes in each group that either didn't need corrections or were able to be tested with and without them. Figure 5.3.8 shows the result of the phoria measures, without corrections. This shows a less-steep slope of responses for the non-presbyopes and intermediates, still with a large esophoria at the far distance. For the presbyopes, the slope of the response across distance does not change but the responses are more esophoric. This suggests that observers' corrections may have contributed to the results but other factors such as an inadequate accommodative response might be an issue.

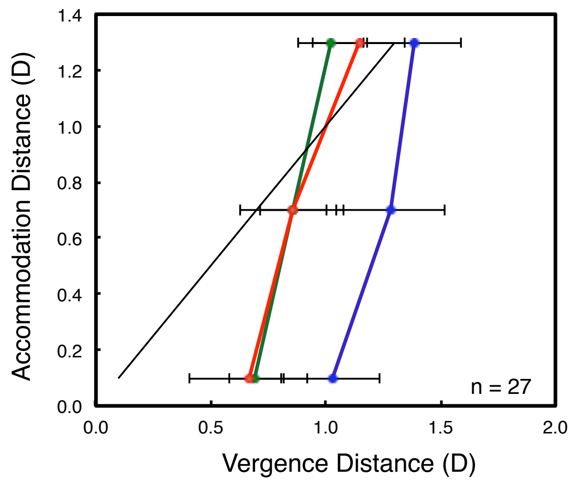


Figure 5.3.8 – Average phoria for each degree-of-presbyopia group measured with no corrections. The vergence distance in dioptres on the x-axis and the focal distance in dioptres on the y-axis. The colours of the lines represent the non-presbyopes (red), intermediates (green) and presbyopes (blue). The error bars are plus/minus standard error of the mean.

The 'Zone of Good Stereo'

The vergence distance (i.e. vergence-accommodation conflict) that corresponded to ~75% correct on the stereoacuity task was determined for each staircase by averaging the last four reversals points. We then computed an average vergence distance across the three repetitions of each staircase, yielding a single vergence distance for each observer and each distance. This provided a measure of 'zone width' corresponding to the same decrement in stereoacuity at different screen distances, and across different people. The absolute width of the zone is arbitrary, because it is defined by the criterion levels we selected (90% correct to 75% correct). However, because it represents the conflict required to evoke the same performance change, it is comparable across conditions, and observers.

Relationship between the zone of good stereo and DoP

Figure 5.4.1A shows the average 'zone of good stereo' for each classification of presbyopia. To understand how our zone, and the 'zone of comfort' estimated by Shibata et al. (2011) are related, we have also plotted their 'zone of comfort' on this Figure. It is important to note that the 'limits' of both zones have been arbitrarily chosen depending on what might be deemed uncomfortable, or 'poor' stereo performance. Therefore, the 'zone of comfort' has been included for interest and not for systematic comparison.

Panel A plots the 'zone of good stereo' for each group (non-presbyopes, intermediates and presbyopes). The data show that all groups have a comparable negative zone (beyond the screen distance). However, for the mid (0.7 D) and far (0.1 D) focal distances, the non-presbyope group had a larger zone relative to the intermediate and presbyope groups. The results are more varied for zone of good stereo in front of the screen. At 0.1 D focal distance, the positive zone of good stereo (nearer than the screen) for both the presbyopes and non-presbyopes was substantially larger (~0.75 D) than for the intermediates. This result is similar at the 1.3 D focal distance with the positive zone being larger (albeit by a smaller amount) for the presbyopes and non-presbyopes relative to the intermediates.

These results suggest that individuals with an intermediate ability to change their focusing response will have a reduced zone of good stereo relative to those with a greater and poorer DoP. Therefore, perhaps it is those who are experiencing a reduction in focusing response that are more susceptible to conflict than those who are either non-presbyopic or presbyopic. This is interesting, as researchers to date have not specified predictions for individuals who are not yet presbyopic and instead discuss the possible effects for those who are (Mendiburu, 2009; Yang et al., 2011; Banks et al., 2012; Read & Bohr, 2014).

In their paper, Shibata et al. (2011) state that their estimated zone of comfort can be used to determine the 'likely discomfort with a given onscreen disparity for a young adult viewer'. Although they suggest further work is required to establish clear guidelines, some researchers have used it to help determine the conflicts to use in their studies (e.g. Miles, Pop, Watt, Lawrence & John, 2012; Winkler & Min, 2013). Therefore, although not useful to directly compare the two zones, it is interesting to consider the differences between the estimated zone of comfort and zone of good stereo. For example, if researchers use the zone of comfort as a guideline for maximum conflict, for most individuals, stereo performance would be substantially reduced. This is especially true for the negative values of the zone of comfort. However, it is also true for those with an intermediate DoP at the positive zone of comfort. Thus, it might be advisable to consider both zones when considering conflict magnitude. Although it must be noted that further studies are required to create definitive guidelines for both the zone of comfort and the zone of good stereo.

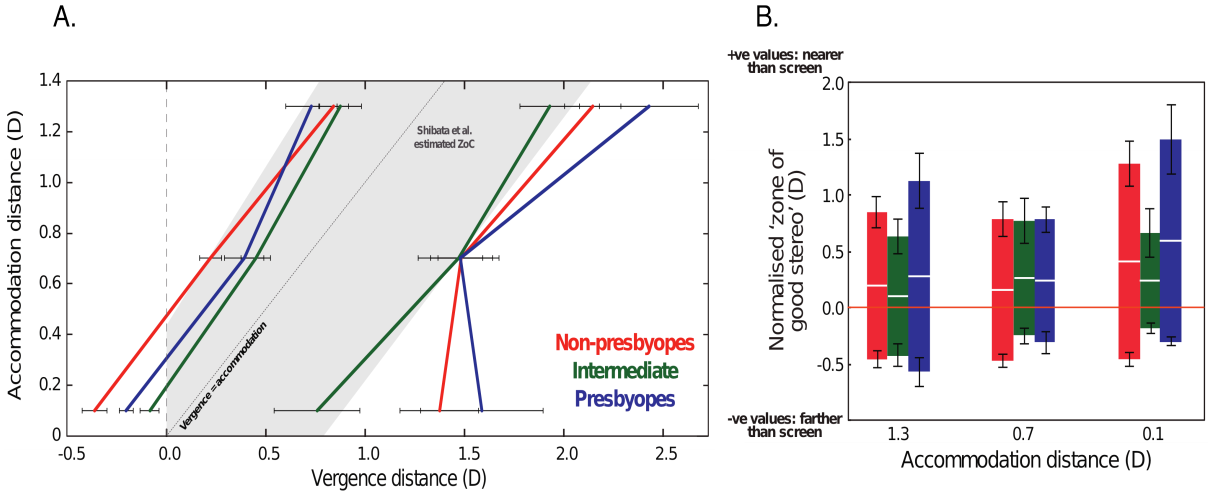


Figure 5.4.1 – The zone of good stereo. Part A is the zone of good stereo according to degree of presbyopia. The x-axis is the vergence distance in dioptres and the y-axis is the accommodation distance in dioptres. The colours of the lines represent the non-presbyopes (red), intermediates (green) and presbyopes (blue) which form the average negative and positive vergence distances that form the threshold limit of good stereo. The thin grey line is the natural-viewing line where vergence and focal stimuli are at equal distances. The zone of comfort according to Shibata et al (2011) is the grey shaded region. (B) The histograms plot the normalised zone of good stereo, illustrating the symmetry of the zone across groups and as a function of distance. The x-axis is the accommodation distance in dioptres and the y-axis is the values for the normalised zone. The thin red line is the screen distance; negative values are farther and positive values are nearer than the screen. The red, green and blue coloured bars are data for the pre-, inter- and presbyopic groups respectively. The white lines on the data illustrate the middle of the zone. Error bars on both graphs denote plus/minus standard error of the mean.

Figure 5.4.1B shows the symmetry of the zone of good stereo at each focal distance, broken down by presbyopia group. The data are normalised for focal distance (i.e. the y-axis shows the magnitude of vergence-accommodation conflict in each case, not absolute vergence distance). The white lines show the mid-point of the zone in each case. The symmetry of the graphs shows that the zone of good stereo was larger for positive conflicts (nearer than the screen) than for negative conflicts at all focal distances. As the focal distance increased, the negative tolerance limits generally decreased while the tolerance to positive conflicts increased. However, the results show a difference in symmetry across the groups, which is evident in Figure 5.4.1A. Finally, Figure 5.4.1B shows the changes in zone width across distances: the zone narrows at the middle distance for the pre- and presbyopic groups and widens at the near and far distances. However, the symmetry of the zone for the intermediates is generally consistent across distances.

Figure 5.4.2 explores the relationship between degree of presbyopia and zone size in more detail, by plotting individual observer's zone of good stereo data as a function of the slope of the accommodation response. Each graph shows the data for a different focal distance: 1.3 D (left graph), 0.7 D (middle), and 0.1 D (right). The curves are second-order polynomial fits to the data. There are a few emergent points. First, there is evidence from the continuous data that the general effect shown in the group data—larger positive zones for presbyopes and non-presbyopes, and smaller zones for intermediates, is evident in all conditions. In all cases the fit to the data shows that the smallest zones occur for the intermediate accommodation responses. Moreover, this pattern is driven not by the presence of larger zones for all presbyopes/non-presbyopes, but by increased variance in those groups. That is, some observers with low and high accommodation response slopes have small zones, and some have rather large zones. In contrast, almost all observers with intermediate accommodation response functions have small zones of good stereo.

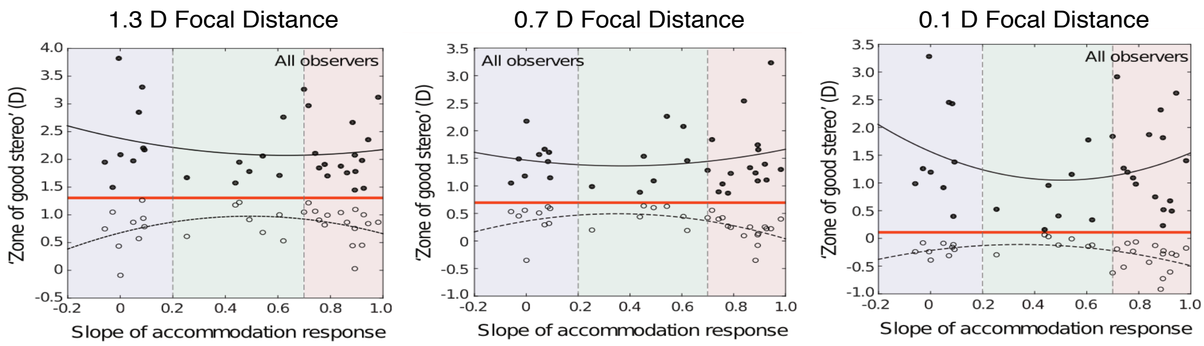


Figure 5.4.2. Individual observer's zone of good stereo data, as a function of the slope of their accommodation response at each screen/focal distance. The left, middle and right plots show the data for the 1.3, 0.7 and 0.1 D screen distances, respectively. The positive and negative 'zones of good stereo' are plotted as a function of slope of accommodation response. The filled circles are the observers' positive zone limits, and the open circles are their negative zone limits. The curves are second-order polynomial fits to the data for the positive (solid line) and negative (dashed line) conflicts. The red horizontal line indicates the screen distance in each case. The colours in the background show the presbyopia categorisation: presbyopic (blue), intermediate (green) or non-presbyopic (red).

Relationship between age and zone of good stereo

We also considered whether the zone of good stereo was systematically related to observer's age. Degree of presbyopia is the underlying variable expected to predict the size of the zone of good stereo. It is reasonable to assume, however, that content producers will not know the accommodation response slopes of their target audience (indeed individual viewers will not know their own slopes), but they can know their ages. Obviously age and the degree of presbyopia are strongly related (Figure 5.3.1C), but we have seen in our data that there can be great variability in middle age. It is worthwhile therefore examining whether age is a useful variable for specifying guidelines for stereo content.

We therefore grouped the observers by age; observers aged 18-25 years were classed as 'younger', the 30-50 year olds were classed as middle aged and the 51-70 year olds were classed as older. Figure 5.4.3 shows the zone of good stereo for different age groups - younger (dashed red), middle aged (dashed green) and older (dashed blue). The younger group having a larger negative zone (beyond the screen) than the middle and older groups. There are comparable negative zones at the nearer focal distances for all groups. For the positive zone at the 1.3 D and 0.1 D focal distances, the middle and older groups have larger positive zones than the younger group. This effect is larger at the 1.3 D focal distance than the 0.1 D distance.

Figure 5.4.3B compares how the zone of good stereo grouped by age relates to the zone of good stereo grouped by degree of presbyopia. The left plot shows the young group compared with the non-presbyopes (solid lines), the middle plot compares the middle-aged group with the intermediates, and the right plot compares the older group with the presbyopes. The plots in Figure 5.4.3B show a difference in tolerance between the middle-aged and intermediate groups for the near and far positive vergence distances. There is also a difference between the younger and the non-presbyopes for the near positive vergence distances. However, there is no difference between the older age group and the presbyopic group. These results are most likely due to the variance in degree of presbyopia in the middle-aged group. Figure 5.3.1C shows that some in the middle-aged group would be classed as non-presbyopes. Their data may be inflating the zone of good stereo for that age group. Therefore, if creators of stereo content were to use age as a guide for conflict magnitude, they may over-estimate the tolerance of middle-aged individuals.

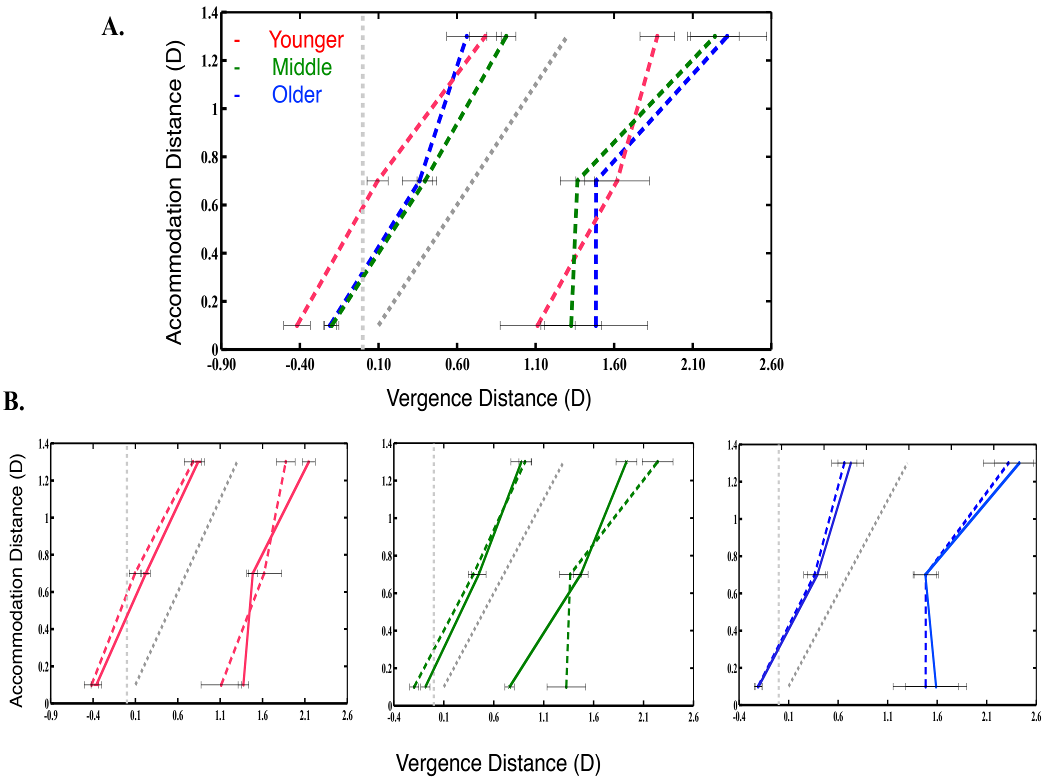


Figure 5.4.3. Zone of good stereo with age. Part A is the zone of good stereo for young (18-25, dashed red), middle aged (30-50, dashed green) and older (50-70, dashed blue) observers. Part B shows three graphs, each comparing the age groups to the degree of presbyopia groups; the young age with the non-presbyopes (left graph, solid lines), the middle age with the intermediates (middle graph, solid lines) and the older group with the presbyopes (right graph, solid lines). The sloped grey line is the natural-viewing line where vergence and focal stimuli are at equal distances. The vertical grey line is at zero dioptres. The x-axis is vergence distance in dioptres; the y-axis is the accommodation distance in dioptres. Error bars on both graphs denote plus/minus standard error of the mean.

Finally, we investigated whether there was a relationship between the age of the viewer and the zone of good stereo. We examined this for the 0.1 D focal distance, where the largest zone was measured when grouped by DoP. Figure 5.4.4 shows the correlation plots for the two groups (age and DoP). The plot on the left shows the relationship between the age of the individual and their zone of good stereo. The right plot shows the relationship between the DoP of the individual and their zone of good stereo (taken from Figure 5.4.2). This data shows no clear relationship between the age of the individual and the measured zone of good stereo when compared with the zone and DoP. This suggests age is overall a poor predictor of tolerance to conflict.

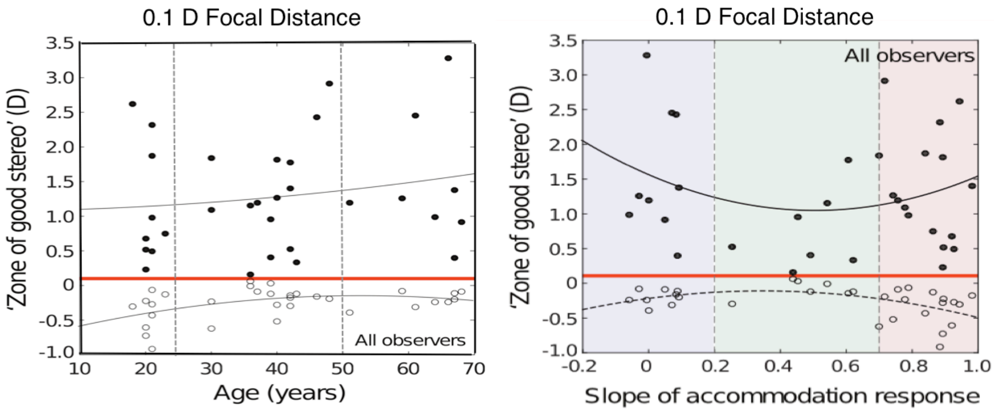


Figure 5.4.4. Individual observer's zone of good stereo data. The positive and negative 'zones of good stereo' are plotted as a function of age (left plot) and slope of accommodation response (right plot). The filled circles are the observers' positive zone limits, and the open circles are their negative zone limits. The curves are second-order polynomial fits to the data for the positive (top line) and negative (lower line) conflicts. The red horizontal line indicates the screen distance. The vertical dashed lines indicate the separate groups in both graphs. The colours in the background for the right graph show the data categorisation: presbyopic (blue), intermediate (green) or non-presbyopic (red).

Relationship between stereoacuity and zone of good stereo

As before, psychometric functions were fit to each individual's data for each distance (1.3, 0.7 and 0.1 D) by fitting a cumulative Gaussian to all trials, using a maximum-likelihood criterion (Wichmann and Hill, 2001). In each case we identified the peak-to-trough disparity at which the orientation of the depth corrugation was correctly identified on 90% of trials (the value that was used to specify the starting stimulus parameters in the main experiment, measuring the zone of good stereo performance).

Although our main experiment was designed to produce comparable zone measurements across people, it is informative to examine whether there were any large differences in stereoacuity (90% correct) across different observers/groups that could potentially explain any differences in zone width. The change in stereoacuity threshold with distance and as a function of degree of presbyopia is shown in Figure 5.4.5. The line graph (Figure 5.4.5A) shows the variance in stereoacuity threshold with viewing distance (in dioptres) across the three groups, presbyopes (blue lines), intermediates (green lines) and non-presbyopes (red lines). When examining each group individually, we see only slight variations in the average stereoacuity thresholds with distance. They also show that the groups follow the same pattern of results at all distances, with non-presbyopes having the smallest stereoacuity thresholds, and presbyopes having substantially larger thresholds.

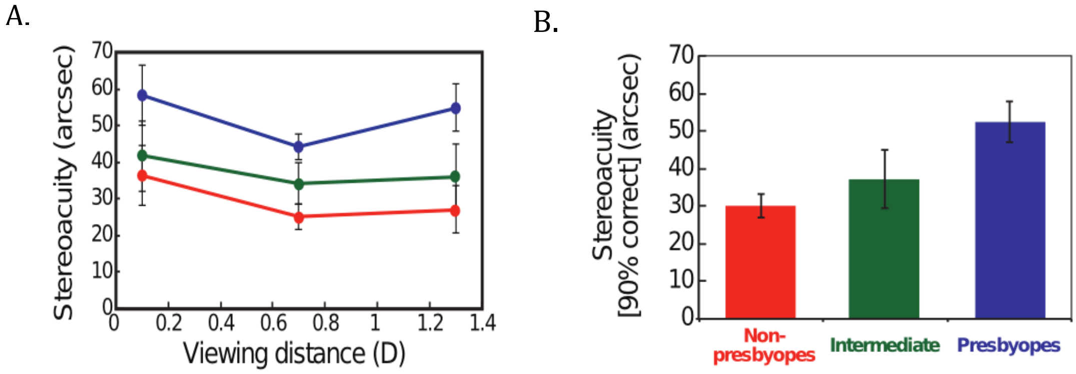


Figure 5.4.5. Stereoacuity 90% correct values as a function of distance (A) and with Degree of presbyopia (B). The line chart (A) shows the variation of 90% correct with distance for the different groups by degree of presbyopia. Each colour line represents a different group; non-presbyopes (red), intermediates (green) and presbyopes (blue). The x-axis is the viewing distance in dioptres; the y-axis is the stereoacuity 90% correct point. The bar chart (B) shows the average of the stereoacuity 90% correct values for each group. The x-axis is the corresponding group, the y-axis is the stereoacuity value which corresponds to 90% accuracy. All error bars are plus/minus standard error of the mean.

We examined whether stereoacuity threshold predicted the zone of good stereo to conflict by calculating the ‘width’ of each individual’s zone. Figure 5.4.6 shows these zone-width measurements for each observer plotted as a function of stereoacuity, for each distance. These results show very small positive trends for the near (1.3 D) and far (0.1 D) distances, indicating that as the stereoacuity gets poorer, the width of the zone increases. The converse is true for the middle distance (0.7 D). However, the linear trend line (black line) as well as the r^2 values for each distance shows a poor fit. This indicates that there is a weak relationship between stereoacuity and the width of an individual’s zone of good stereo.

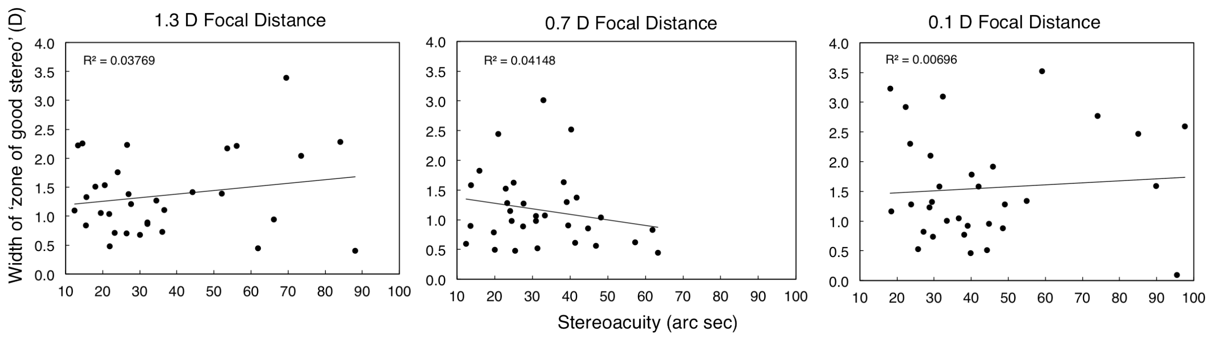


Figure 5.4.6. Relationship between stereoacuity threshold and the width of the zone of good stereo. Each graph shows the width of the zone of good stereo, as a function of stereoacuity 90% correct point, at different focal distances; 1.3 D (left), 0.7 D (middle) and 0.1 D (right). The data points are individual observers scores. The x – axis of all graphs is the stereoacuity threshold in seconds of arc (arc sec) and the y-axis is the width of the zone of good stereo in dioptres. The black line is a linear trend line of the data.

Relationship between zone of clear single binocular vision and zone of good stereo

Figure 5.4.7 is a comparison between our measured ZCSBV and zone of good stereo, grouped by presbyopia classification. The dashed lines are the ZCSBV for the non-presbyopes (red), intermediates (green) and presbyopes (blue). The ZCSBV lies between these dashed lines. The solid lines are the zones of good stereo for the same groups. The results of the ZCSBV show that all groups had similar negative ZCSBV (beyond the focal distance). The presbyopes have a consistently larger positive ZCSBV compared with the intermediates and non-presbyopes; this is especially true at the 0.1 D focal distance. Overall the size of the ZCSBV is significantly larger than the zone of good stereo. This is unsurprising given the work by Percival (1982) who proposed that although there is a range of possible vergence and accommodation tolerance limits, comfortable viewing lies within that range. Thus indicating that although it is possible converge at one distance and accommodate to another, the effort involved might not be comfortable. We therefore expected the zone of good stereo to be within the ZCSBV. However, we also expected the ZCSBV would be a predictor of when stereo depth perception becomes affected by conflict.

Generally, the ZCSBV does predict the zone of good stereo but there are differences. While the groups with intermediate and presbyopic observers generally mimic the ZCSBV, the non-presbyopes appear to have a larger tolerance to positive and negative conflicts than expected from their ZCSBV. As previously suggested, it is possible that non-presbyopes may be able to adjust their accommodation response, opting to tolerate some blur without a large performance decrement on the task.

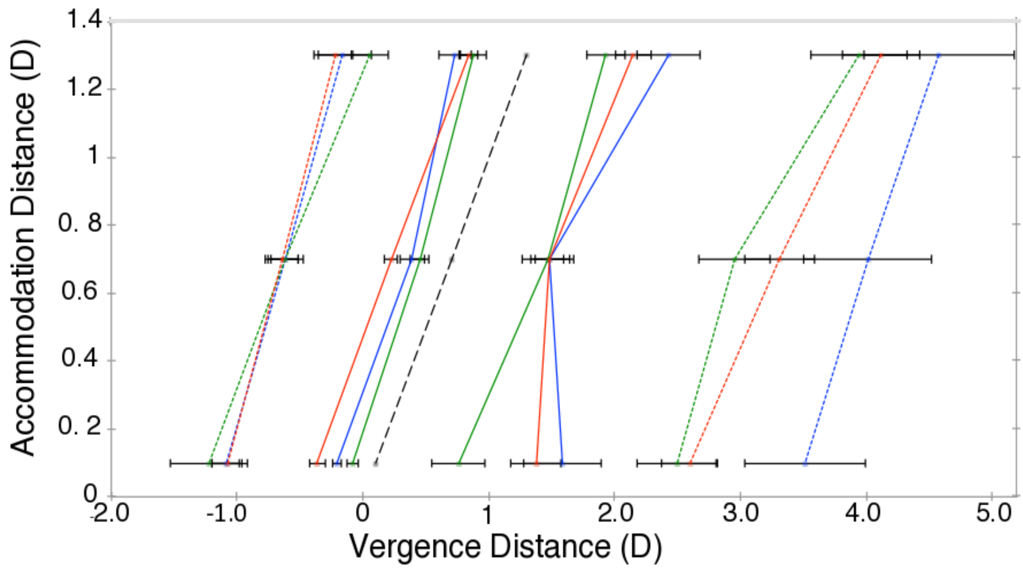


Figure 5.4.7. The estimated ZCSBV and zone of good stereo. The x-axis is the vergence distance in dioptres; the y-axis is the accommodation distance in dioptres. The colours of the lines represent the non-presbyopes (red), intermediates (green) and presbyopes (blue), which form the average negative and positive vergence distances that form the ZCSBV (dashed lines) and the zone good stereo (solid lines). The thin grey line is the natural-viewing line where vergence and focal stimuli are at equal distances. The grey dashed line is the 'natural viewing' line where accommodation and vergence response match. The error bars are plus/minus standard error of the mean.

Phoria

We expected that phoria would influence the symmetry of the zone of good stereo. For instance, we would expect the centre of an esophore's zone of good stereo to have a greater positive value relative to that of an exophore. We explored this in two ways. First, we examined the relationship between the centre of each observer's zone of good stereo and his or her phoria. Figure 5.4.8 shows these data for each focal distance. For the 1.3 D and 0.1 D focal distances, there was a positive relationship between the observers' phoria and the location of the centre of the zone of good stereo. At the 0.7 D focal distance, there appears to be no effect of phoria on the centre of the zone of good stereo. Despite this general trend for these focal distances, the data are very noisy.

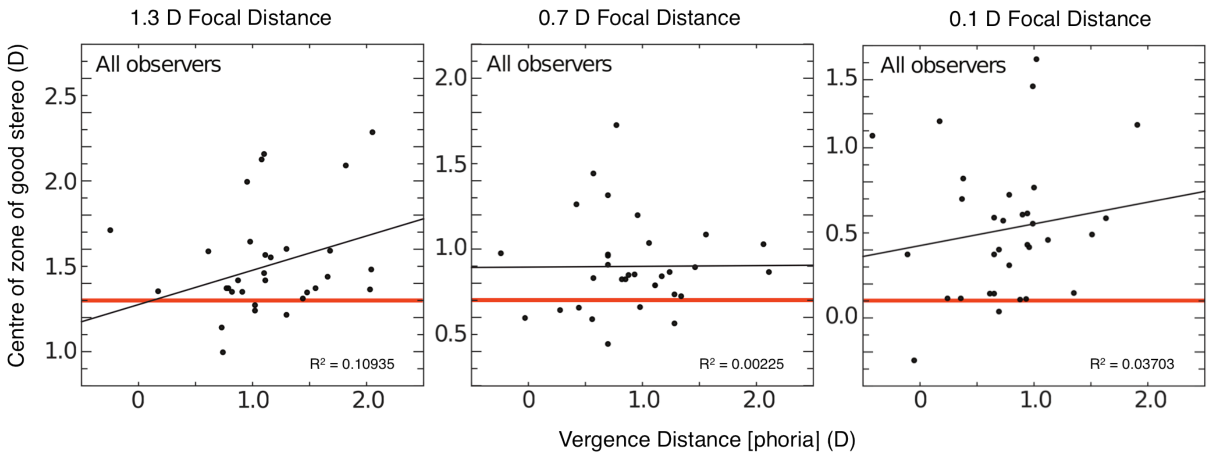


Figure 5.4.8. Individual observer's centre of zone of good stereo data, as a function of their phoria response at each screen/focal distance. The left, middle and right plots show the data for the 1.3, 0.7 and 0.1 D screen distances, respectively. The centre of the zones of good stereo are plotted as a function of phoria in terms of vergence distance (D). The lines are linear fits to the data. The red horizontal line indicates the screen distance in each case.

Second, we performed a 'group analysis', whereby we compared the zone of good stereo for the observers with the most positive phorias, and the most negative phorias. To classify each individual we calculated a single phoria values by averaging across their phoria at the three distances. We then took the distribution of all observers' average phorias, and identified the observers in the upper and lower quartiles of this range, and classified them as groups with more positive and less positive phoria, respectively. Figure 5.4.9 plots the zone of good stereo for these groups. It can be seen that the observers with less positive phorias have a somewhat smaller positive zone at near and far distances and a more clearly larger negative zone at all distances relative to those with a more positive phoria. This suggests that phoria affects the symmetry of the zone of good stereo.

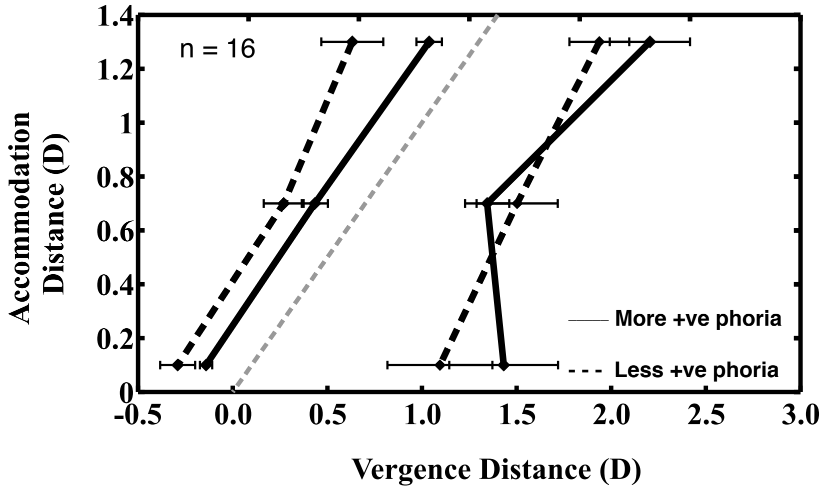


Figure 5.4.9– Zones of good stereo compared for upper and lower quartiles of observers with respect to average phoria. The x-axis is the vergence distance in dioptres and the y-axis is the accommodation distance in dioptres. The dashed lines are for the less positive average phoria while the solid lines are the more positive average phoria. The grey dashed line is the natural-viewing line where accommodation and vergence stimulus are matched. The error bars are plus/minus standard error of the mean.

5.4 Discussion

Summary

In this experiment, we aimed to measure the zone of good stereo. To do this, we measured tolerance to conflict (in terms of its effects on stereoacuity) as well as the possible predictors of that tolerance. The aim was both to better understand the causes of different tolerances to conflict, and also to provide information that might be of use to content producers when setting parameters such as the 'depth budget'.

Our main predictor for tolerance to conflict was degree of presbyopia. We therefore plotted the zone of good stereo for each classification of presbyopia (non-presbyopes, intermediates and presbyopes). The results showed that the non-presbyopic and presbyopic groups had consistently larger zones relative to the intermediate group, at all focal distances.

The tolerance for the non-presbyopic and presbyopic groups to positive conflict at the far focal distance is a lot larger than the tolerance for the intermediates. There are two possible explanations. The first is a possible adaptation effect where the viewer's vergence response is adapting to increasing vergence demand. To try and prevent this, we presented the negative and positive conflicts in the same blocks of trials. The dynamic change of the vergence response to positive and negative distances, as well as at the fixation distance should prevent this adaptation effect. The other possible explanation is the non-presbyopes adjusted their accommodation response away from the focal distance, tolerating some blur in order to perform well on the task. To test this, individuals with a large amplitude of accommodation could be cyclopleged; the paralysis of the ciliary muscle in the eye leading to loss of accommodation response. Through the use of lenses, we could fix their accommodation to the surface of the screen. The experiment would be conducted in the same way, by varying the stimulus to vergence using an adaptive staircase. This would prevent the observers from varying their response. If the resulting zone of good stereo were narrower than the current zone, this would indicate that the observers did alter their accommodation response, resulting in a greater tolerance to conflict.

It would be most useful for establishing guidelines if the size of the zone of good stereo could be based on simple ‘demographics’, such as age, which are likely to be predictable for content makers, and known precisely by individuals. However, although there was a strong relationship between DoP and the age of the individual there was high variability in degree of presbyopia for the individuals aged between 30 and 50 years. This meant that within this age group there was a mix of non-presbyopic or presbyopic individuals with the majority being intermediates. We compared the zone of good stereo for observers grouped by age and degree of presbyopia. The results, when compared to the zone of comfort as a function of degree of presbyopia, showed zone for the middle age group (ages 30-50) was underestimated for the positive zone of good stereo relative to estimates of the intermediate group. Therefore if the zone was characterised by age, it is possible that the zone of good stereo could be over-estimated for some of the target audience. This indicates that detailed recommendations for stereo 3D content creation based on simple demographics (such as age) may not be particularly useful.

We also measured the predictive power of other visual factors such as phoria, and ZCSBV. We expected phoria to be predictive of the zone of tolerance as this measure can indicate the natural difference between vergence and accommodation responses at varying accommodation distances (Banks, et al., 2012). Our analysis of phoria showed some unexpected results (see below). However, our results found that phoria was predictive of the zone of good stereo. They revealed that on average, individuals with a less positive phoria have a somewhat smaller positive zone at near and far distances and a more clearly larger negative zone at all distances relative to those with a more positive phoria. This suggested that phoria affects the symmetry of the zone of good stereo.

The ZCSBV is a measure of the upper limit of an individual's ability to decouple their accommodative and vergence responses. As a result, we hypothesised that the zone of good stereo would be a reflection of the ZCSBV. The results of the ZCSBV measure suggests that ability to decouple the responses were monotonic across the groups, with the presbyopic group having a greater tolerance compared to the non-presbyopes and intermediates who had very similar tolerance range. In the zone of good stereo, however, the non-presbyopes have a greater tolerance relative to the intermediates. This may be due to the instructions given in both tasks. For the ZCSBV, the observers were instructed to report when the stimulus appeared blurry and/or diplopic. Whereas the non-presbyopes may have changed their accommodation response on the stereo task, allowing them to fuse stimuli with large disparities between the two-eye's images, causing a tolerable amount of blur. This could result in the larger zone of good stereo relative to the intermediates who had a reduced capacity to vary their accommodative response.

Finally, we found that although stereoacuity thresholds scaled with the degree of presbyopia, there was no clear relationship between stereoacuity threshold and zone width. Therefore, any differences between the zones of good stereo with changes in degree of presbyopia are not related to the observers' stereoacuity threshold.

Taken together, these results suggest that there is a measurable zone of good stereo and that there are predictors to individuals' tolerance to conflict. These predictors can inform content creators and researchers about likely stereo performance as a function of distance across various visual functions including degree of presbyopia, phoria and ZCSBV. Factors including age of the individual can also be used but estimates using this criterion can over-estimate the tolerance for certain age groups.

Implications

These results are relevant for optimising stereo content for different target audiences. The aim is to create content that is easily viewable by as many viewers as possible. However, given the variance in individual zone of good stereo, and in factors that influence it, this appears to be a challenging (if not impossible) task. Establishing techniques to overcome this is advisable. For example, if we know how the visual system affects our tolerance to conflicts, we can adapt glasses and head-mounted displays according to that individual's visual properties. Therefore, instead of adapting content to suit a wide range of individuals, instead we could adjust (using lenses and prisms) an individual's conflict tolerance to suit the stereo content. Thus allowing the creation of effective stereo content that the audience will be able to view comfortably and effectively.

In order to establish a detailed set of guidelines which accounts for tolerance to conflict, a large-scale population study needs to be conducted. The results presented here provide an estimation of the effects of individual properties on the zone of good stereo and what that zone might look like, as well as suggesting some techniques for its measurement.

Future Research

Similar to the zone of comfort estimated by Shibata et al., (2011), the zone of good stereo is arbitrary: the zones can be different according to the criterion used to establish them. Therefore, what we have reported here is the range of conflicts, which result in adequate stereo performance (better than ~75% accurate). However, this can change according to the level of performance required. As more research is conducted on the stereo performance with conflict, we hope that this continuous representation of the zone of good stereo can be better defined and understood.

We should state some caveats to our measures. Our analysis of phoria for each DoP group showed some surprising results. The results were largely unvarying for the presbyope and intermediate groups, regardless of focal distance. For the non-presbyopes, the results were similar but with some variation at the 1.3 D and 0.7 D focal distances. The results for the presbyopic group were expected, as their reduced accommodative response function would affect their phoria. However, the results were surprising for the other groups. There are two possible reasons for these findings. First, this could have been a result of measuring phoria with corrective lenses. When measured without corrections, the results show a less-steep slope of responses for the non-presbyopes and intermediates, still with a large esophoria at the far distance. This suggests that observers' corrections may have contributed to the results but other factors might be an issue. Secondly, the non-presbyopes and intermediates may not have accurately adjusted their accommodation response. It is therefore possible that the non-presbyopic and intermediate groups were able to do the phoria measurement task without accurate accommodation. For future testing, we would therefore recommend using a conventional clinical phoropter as used by Shibata et al., (2011).

It should also be noted that instead of (or as well as) measuring phoria, we could also measure fixation disparity. Fixation disparity, defined as vergence errors of binocular fixation

(Schor, 1978), occurs in the presence of binocular feedback, so is a closed-loop error. Whereas phoria refers to the physiological resting vergence position for a given focal distance. This is typically measured open loop. Measuring fixation disparity, as well as associated phoria (the amount of prism needed to reduce fixation disparity to zero minutes of arc) has been shown to be a reliable indicator of binocular stress and visual symptoms (Pickwell, D., Jenkins, T., & Yekta, A.A., 1987). It would, therefore, be a good measure to include for future research. However, in the current experiment, for the reasons previously mentioned (see Chapter 5 Introduction) and because we wanted visual measures that indicate the relationship between the vergence and accommodation response systems, we measured phoria.

The analysis of pupil size resulted in images that were difficult to discern. This was largely due to recording pupil size through the autorefractor. It was not possible for us to record the observers' pupil size directly while they were viewing the stimulus in the device. Our recommendation would be to capture the pupil size directly, while the observer views stimuli of the same properties as in the experiment (luminance, size etc).

Finally, we used some observers who had varifocal corrections. Varifocal corrections are not ideal for these measures. This is because the observers are required to look through different parts of their corrective lenses in order to view images at various focal distances. Therefore, if they changed where they looked through their lenses to view the image, the correction may not be accurate for that focal distance. Despite having the head and chin-rest, which was aimed at reducing these errors, we cannot rule out that this happened. Therefore, we would suggest either using trial lenses using the correction they require for the focal distance, or not testing observers with varifocal corrections.

However, with slight adjustments to the visual measures and as more data is collected on individual's tolerance to conflict in terms of stereoacuity and discomfort, with degree of presbyopia and other visual factors, the generation of detailed recommendations for stereo 3D content creation is possible. We hope this will lead to optimised stereo content for a range of different audiences.

Future Research: Additional Measures

We would suggest some additional visual properties are measured, specifically the AC/A and CA/C ratios. The accommodation and vergence systems are neurally linked. As such, when one system responds to a stimulus, the other system also changes to a corresponding degree that is never fully accurate (Tait, 1951). The AC/A and CA/C ratios are a reflection of that. Accommodative convergence (AC) is the convergence resulting from the accommodative effort. Convergence accommodation (CA) is the accommodation resulting from the convergence response. A ratio of one means that the response of the non-stimulated system matched the response of the stimulated system exactly. Therefore, these ratios are a reflection of the strength of the connection between these systems. This interaction creates a combined response of the two systems that is different to their individual responses (Bruce, Atchison and Bhoola, 1995). Therefore, a strong connection could indicate a reduced tolerance to conflicts, as either the accommodation or vergence response would be influenced by the response of the other. Unfortunately, we could not measure these ratios as ideally they would be measured using an autorefractor and eye-tracker device and we could not use either device in our study.

It should be noted, however, that our method of measuring phoria is very similar to methods of measuring AC with some noticeable differences. In our phoria measure, the observers viewed the stimuli for as long as they needed for their vergence to return to its resting state for that focal distance. The AC response is measured by changing the accommodative task and measuring the consequent changes in accommodation and convergence (Bruce, et al., 1995). This is a relatively quick measure. Ideally, the AC/A ratio would be measured using an autorefractor and eye-tracker device, which would measure the accommodation and vergence response for the same distance simultaneously. However, it is possible to use our phoria measure to record the AC response (but not AC/A ratio). The difference being that the stimuli could be initially presented at a distance away from the focal distance of interest (e.g. far distance) and changed to the focal distance of interest (near distance). Instead of allowing the vergence response to return to its resting state, the measure would be taken immediately. Thus, once the stimulus moved the participant reports where the arrow is pointing as soon as they identified the number (and colour) on the scale. Remember, this is an open loop task and any change in vergence position would be a result of the accommodation response.

CHAPTER 6

6 GENERAL DISCUSSION

6.1 Overview

With the ongoing proliferation of stereo 3D displays, their use is becoming more commonplace in applications like medicine (including laparoscopic surgery and imaging), scientific visualization, education, video conferencing and entertainment (TV, cinema, and gaming). It is therefore important to address the potential adverse side effects of viewing conflict in order to develop ways to manage, reduce or eliminate them. With this in mind, we have attempted to investigate one of the more fundamental issues in conventional S3D displays. Specifically, the vergence-accommodation conflicts caused by incorrect focus cues in conventional stereo 3-D displays. To achieve this, we used two approaches. The first is to attempt to eliminate the conflict, by investigating possible display solutions to the conflict that match the stimulus to vergence and accommodation distances. We did this using a beam-splitter based multi-plane display. For the instances where the multi-plane display solution is not feasible (e.g. multi-viewing situations) the second option was to better understand the causes and effects of conflict. The aim being to optimise content to provide an enhanced stereo 3-D experience, while minimising adverse visual effects. In this general discussion we will report the results of several experiments in Chapter 3 through to Chapter 5. These experiments focus on eliminating vergence-accommodation conflict and focusing on the effects of conflict on stereo performance as well as understanding individuals' tolerance to conflict.

6.2 Multiple-Focal-Plane Displays

6.2.1 The practicality of the multiple-focal-plane display as a solution to conflict

In chapter three, we explored the relationship between focal-plane separation and stereo performance, to examine whether reasonable performance can be achieved in a practical fixed viewpoint, multi-plane display. There is a fundamental limitation associated with these multi-plane displays; specifically the number of focal-planes used ultimately determines the luminance of the resultant images. This limitation, applies to any current implementation of these displays. Therefore, it is optimal to use as few focal-planes as possible, while still producing images that are equivalent to real-world images (viewed naturally). Thus, there is a trade-off between the number of focal-planes used and the range of accommodation distances available. However, with decreasing focal-planes, the distance between the focal-planes necessarily increases and therefore the range of accommodation distances decreases. To overcome this, we use a depth-filtering technique—in which the luminance ratio of image points is varied across the focal planes, to simulate focal distances between them. Importantly, accommodation (and vergence) responses are accurate to spatially broadband depth-filtered images for focal-plane separations up to 0.9 D (MacKenzie et al., 2010; MacKenzie et al., 2012). However, depth-filtered images cannot present the same level of fine detail (high spatial frequencies) as images on a single focal plane. This decrement in retinal-image contrast increases with increasing focal-plane separations (MacKenzie et al., 2010). Thus, we might expect that depth-filtered images will result in reduced precision of stereoscopic vision (Legge & Gu, 1989; Simons 1984; Smallman & McKee, 1995) as a function of image-plane separation.

In terms of stereo performance, Experiments 2 and 3 have shown that some aspects of stereo performance—stereoacuity and fusion time—are relatively unaffected by the spacing between image planes in depth-filtered images. Experiments 1 and 4 show there is a systematic relationship between focal-plane separation and stereo resolution. This relationship makes sense as retinal-image contrast for depth-filtered images reduces with increased focal-plane separation. However, our results show that for focal-plane separations of up to 0.9 D, stereo resolution thresholds for depth-filtered images were comparable to real-world images. This is promising for these volumetric displays and the depth-filtering technique, meaning that accommodation and vergence responses (MacKenzie et al., 2010; 2012), as well as stereoacuity, fusion time and stereo resolution are all comparable to real-world images at this focal-plane separation. We concluded our investigations looking at the quality of depth-filtered images in terms of blur. The depth-filtered images could be reliably discriminated from real-world stimuli even with focal-plane separations as small as 0.3 D. This discrimination became easier with increasing distances between the focal-planes. However, subjective ratings of image-quality from depth-filtered images show that perceived quality in these images is again comparable to real-world equivalents for focal-plane separations of up to 0.9 D. Finally, we postulate that accurate discrimination for the focal-planes separated by less than 0.6 D might have been a result of miss-alignment of the depth-filtered images. However, typically for these types of images, anti-aliasing techniques like applying a Gaussian blur are used to remove any pixellation of the stimulus edges. This effectively adds blur to the image. Therefore, as both the depth-filtered and real-world images would have a Gaussian blur applied, any discrimination based on blurred edges in the depth-filtered images would be eliminated.

These findings are 'good-news' for the combined multi-plane display and depth-filtering approach. They suggest that a multi-plane display with a range of 33 cm to around optical infinity (10 m; 0.1 D) would require only four focal planes, with separations of up to 0.9 D. Although the actual number of focal planes (and therefore focal-plane separation) required could change according to the users criterion. These results suggest a somewhat complex trade-off may be required in the various aspects of depth-filtered images, depending on the particular requirements of a given application. Where the finest possible detail is required, depth-filtered images may be problematic. However, for these instances, just four image planes, spaced 0.4 D apart, could cover the range 40-77 cm (2.5 to 1.3 D), allowing problem-free stereoscopic viewing for the majority of reaching space, with very little loss of image quality.

6.2.2 Depth-filtering

The main aim with the depth-filtering approach was to solve the problem of visible discontinuities in multi-plane images, and was not optimised for presenting correct focus cues, or other aspects of image quality. We have measured the depth filtering technique using simple scenes. But it is possible to produce complex depth-filtered scenes with geometric monocular cues such as occlusion and reflections. However, haloing effects occur at occlusion boundaries, and reflections are not presented at the correct focal distance, for example. Instead, Narain et al., (2015) have proposed creating these complex scenes using a novel optimisation technique and not the depth-filtered technique presented here. This optimisation technique is aimed at creating the closest approximation to the image that would be viewed when looking at the scene naturally. Using a model of image formation at the eye (based on a typical observer's optics), their algorithm finds the distribution of image intensities that minimises the difference, across all focal distances, to natural viewing of the same scene. Put another way, they are attempting to establish a way of distributing image intensity that gives the closest match to the intended retinal image. In their paper, they demonstrate the technique for synthesised and natural scenes. They also analyse their approach, relative to the depth-filtering technique. They found that the artefacts produced around occluded images are eliminated with their technique. This approach makes it possible to create near-correct focus cues, similar to the depth-filtering approach, only with no artefacts in complex scenes. Unfortunately, their algorithm is not fast enough and requires more research to address the real-time performance issues (Narain et al., 2015). They have also yet to measure whether the images act as an effective stimulus to accommodation. Therefore, this is proof-of-concept work rather than a practical solution. However, it offers a potential solution to image-quality degradation with increased focal-plane separations.

6.2.3 Conclusions

Thus, our research suggests that the multi-plane display along with the depth-filtering technique could replace conventional stereo displays for single-viewing applications. Many of the volumetric multi-plane displays can be reduced in size to fit into head-mounted displays. Also, with the recent advances in this technology, and with continued research the quality and complexity of images used by these displays could make them a conventional, useable product. However, for now, such displays are still only practical for specialised applications.

6.3 Understanding Adverse Effects of Vergence-Accommodation Conflict

6.3.1 Summary of findings

Given the specialised nature of the multi-focal-plane display and depth-filtering approach, possible adverse effects of viewing conflict is still problematic for many situations where stereoscopic presentation is commonly employed, including television and cinema viewing. Both Chapters 4 and 5 aimed to gain a better understanding of tolerance to conflict in terms of stereo performance and the underlying factors that predict it. We therefore begin by discussing these Chapter results together before investigating the outcomes of both.

Results: Chapter 4

In chapter 4, we sought to parameterise the problem of viewing conflict, in terms of stereo performance. The research that has been done to date on stereo performance with conflict *per se* has found adverse effects on stereoacuity (Hoffman et al. 2008) and fusion time (Akeley et al. 2004; Hoffman et al., 2008). However, this research has examined only a limited range of conflicts and viewing distances, and so just proves there is a problem. Little is known about the relationship of stereo performance with changes in conflict magnitude and focal distance.

In Experiment 6, we aimed to measure a 'zone of good stereo performance', analogous to the zone of comfort determined by Shibata et al. (2011). To do this we measured the effect of vergence–accommodation conflict magnitude on stereoacuity, both for stimuli nearer and farther than the screen, and at various viewing distances. We generally found that as conflicts increased, so too does the peak-to-trough disparity required to recognise depth in the stimulus. We also measured, the effects of viewing multiple conflict magnitudes at positive and negative vergence distances within the same block of trials. This resulted in surprising findings, in which we could not calculate stereoacuity thresholds for some of the observers due to catastrophic stereo performance, regardless of peak-to-trough disparity. For some individuals, this was the case even for the no-conflict trials where the vergence and accommodation distances were matched. We therefore, could not calculate a 'zone of good stereo performance'. We also found that stereoacuity thresholds could be lowest (indicating better performance) for vergence distances away from the focal distance. We investigated these results further in Experiments 7 and 8.

In Experiment 7 we measured the phoria of the participants and then measured their stereoacuity thresholds for the same focal distances and matched vergence distances. Again, we found large individual differences across observers, with phoria being related to performance on the task for one observer and to an extent, for the other observers. We therefore suggest there may be more ophthalmological factors that explain the individual differences seen in stereoacuity thresholds with vergence distance.

In both Experiments 6 and 7, some of the observers were unable to do the task at all conflicts (and even when there were no conflicts). In Experiment 8, we examined the 'after-effects' of viewing stimuli for a brief exposure time (10 minutes). There were large individual differences again across the observers. However, for those who appeared to be affected after viewing conflicts, it took on average 2 minutes for their stereoacuity threshold to return to baseline levels. Interestingly, we found that some of the observers appeared to be affected at the test screen distance (1.3 D) while others appeared to be affected for distances away from this (0.3 D). This indicates that exposure to conflict might have different subsequent effects for real stimuli at the distance of the screen and real stimuli at other distances.

Results: Chapter 5

An emergent theme from the results of Chapter 4 was that there could be large individual differences in tolerance to stereoscopic depth perception with conflict. This experiment looked at these individual differences, measuring tolerance to conflict as a function of focal distance and sign of conflict. We also measured possible visual functions that might predict tolerance to conflict. Little is known about how individual differences (such as age and degree of presbyopia) affect a person's stereoacuity with conflict. The estimated 'zone of comfort' by Shibata et al. (2011), goes someway to outlining the limits of conflict tolerance. However, if comprehensive guidelines are to be produced, more specific knowledge regarding stereo performance both during and after viewing needs to be established, as well as the factors that affect it.

Our main predictor for tolerance to conflict was degree of presbyopia. We therefore plotted the zone of good stereo for each classification of presbyopia (non-presbyopes, intermediates and presbyopes). Overall we found evidence of a non-monotonic relationship between the ability to vary accommodation and tolerance of stereoscopic performance to accommodation-vergence conflicts. Tolerance was greatest for the non-presbyopes, and presbyopes, while the intermediates (i.e. beginning to become presbyopic) were the least tolerant to conflict. This is somewhat contrary to what is believed by the industry (Mendiburu, 2009) and by academics (Yang et al., 2011; Banks, et al., 2012; Read & Bohr, 2014); it is not the older observers who are most affected but the middle-aged individuals. That said, the age of the individual might not be useful for the creation of detailed recommendations for stereo 3d content due to the variability in degree of presbyopia across the individuals aged 30-50 years old. Stereoacuity threshold and degree of presbyopia was found to be strongly related, however, this did not correlate with width of the zone of good stereo. Finally, our results found that individual differences appear to be related to phoria, ZCSBV and degree of presbyopia.

6.3.2 Implications for stereo content

In Experiments 6 and 7 (Chapter 4), we found that viewing conflicts that are beyond an individual's tolerance can result in catastrophic stereo performance. This was even true for stimuli presented at the distance of the screen (with no conflict). It is often assumed (implicitly, at least) that stereo perception at the screen surface will be unaffected by preceding viewing (Mendiburu, 2009). Our data suggest that is not necessarily correct, and that stereo perception may fail even for no-conflict stimuli if the 'Depth Budget' (i.e. the conflict magnitude) is beyond what is tolerable by members of the audience. These findings suggest that there are large individual differences in tolerance to conflict. Therefore, it is extremely difficult (and probably impossible) to produce stereo content that would suit all tolerance ranges. However, our results (Experiment 9) and that of Shibata et al. (2011) have shown that it is possible to estimate a zone of good stereo/ zone of comfort according to different criteria, within which individuals can view stereo content that is comfortable and easy to view. Our experiment and that of Shibata et al. have also found that these zones can be predicted by various ophthalmological factors, such as degree of presbyopia, ZCSBV and phoria.

The creators of stereo content are aware of the issues surrounding conflict. Although they try to address them, there are currently no fixed guidelines. There is widespread use of what is termed the 'percentage rule' (Mendiburu, 2009; Shibata et al., 2011; Banks et al., 2013) in which the maximum parallax (screen offset of image points), as a proportion of screen width, is limited to 2-3% for crossed disparities (nearer then the screen) and 1-2% for uncrossed disparities (farther than the screen). This rule is fundamentally flawed, in principle, because it does not account for the angular size of the display being viewed. The disparities at the viewer's eyes, and thus the conflict, depend on the angular size of the disparities at the two eyes. When the content is considered in pixels, or as a proportion of screen size, the disparities change with viewing distance and the size of the display (Banks et al., 2013). Also, the positive vs. negative conflict zone width we see in both the zone of comfort and zone of good stereo changes as a function of focal distance. Therefore, when creating content, producers need to create content that is optimized for screen distance. However, despite this over-simplification, the percentage rule works quite well in practice. Shibata et al. suggest that this is because larger screens tend to be viewed at greater distances, going some way to normalising the screen parallax, and so the projected disparities/conflict. They examined this percentage rule, against the estimates zone of comfort for near and far distances. They conclude that the percentage rule, coupled with the assumptions about viewing distance is a reasonable guideline for creating comfortable viewing.

That said, it would be optimal if content producers could actually base the conflicts in their content on the tolerance range of their target audience. As previously mentioned, our results and the findings by Shibata et al. (2011) suggest there is some relationship between an individuals visual properties (including ZCSBV, phoria and degree of presbyopia) and their tolerance to conflict. Although interesting scientifically, in order for a content producer to utilise this information, they would need to know information about the target audience that is difficult obtain. Indeed, many individuals would not know their own phoria or ZCSBV, for example. We have suggested that, if this information could be obtained, the individuals' phoria (for example) could be adjusted (with the use of prisms/lenses). Thus, an exophore could change their phoria so that they tolerate more positive conflict than negative, according to the stereo content they are viewing. In principle, an individual's phoria response could be crudely measured using a display and red/green glasses, allowing relatively easy access to this information.

However, a large-scale population study needs to be conducted to establish guidelines based on the relationship between ophthalmological variables and adverse effects during stereo viewing. Although informative, our results fall short of what is required for these guidelines to be established. Experiment 9 should act as a pilot-study for this large-scale population experiment, highlighting the potential methodological pitfalls for some of the measures used (see Section 5.4, page 160).

6.3.3 Implications for research

Thus far we have considered the implications of viewing stereo content on conventional displays and our findings for entertainment/medical applications. However, it is also important to consider our findings with regards to other uses of stereo 3-D imagery, including scientific research. There are many fields that use stereo 3-D stimuli which include, but are not limited to, attention (Anderson, 1990; He & Nakayama, 1995; Anderson & Kramer, 1993), object recognition (Edelman & Bulthoff, 1992; Patterson, Cristino, Hayward & Leek, 2012; Cristino, Davitt, Hayward & Leek, 2015), visuo-motor control (Saunders & Knill, 2003; Scarfe & Hibbard, 2011; Knill, 2005) depth estimation (Knill & Saunders, 2003;

Bulthoff & Mallot, 1988) and shape estimation (Hougen & Ahuja, 1993; Johnston, Cumming, & Landy, 1994). Here, we briefly discuss why researchers using conventional stereo displays to study visual processing in 3-D space should consider the results reported here (and elsewhere) when considering their methodologies.

Research has shown that a conflict between the stimulus to accommodation and vergence leads to increased fusion time (Watt, et al., 2005; Hoffman et al., 2008). Slow fusion could be a confound in reaction time studies as a result of the position in space of the perceived stimulus. Therefore, these issues could affect any research measuring reaction time that use stereo 3-D stimuli. This includes attention as well as object recognition studies, for example.

Another potential issue is conflicts that result in unfusable stimuli. We have seen in Experiments 6 and 7 that there are large individual differences in what conflicts observers can and cannot tolerate. If individuals cannot, or struggle to fuse the images, then the results may be artificially altered. This would affect every experiment that is designed to measure visual performance using stereo 3-D, including attention, object recognition, viso-motor, basic depth/distance/shape perception experiments.

Importantly for object-recognition studies, there is evidence of reduced stereo performance (current research; Hoffman, et al., 2008; Watt, et al., 2005). The effect of this differs across individuals. Generally, this indicates that, for some individuals, the stereo 3-D image may appear altered (flattened or distorted) relative to its actual shape.

However, it is not just performance issues regarding fusion times, poor fusion, and stereoacuity. There are also examples of after-effects of viewing conflict affecting some individuals (Experiment 8), fatigue and discomfort after viewing stereo 3-D images (Shibata et al, 2011; Hoffman et al, 2008) and distortions in perceived depth (Hoffman et al., 2008; Watt et al., 2005).

There are a number of ways to alleviate these issues. The first option is to use fixed viewpoint volumetric displays to present stereo 3-D images with matched cues to vergence and accommodation. Another possible option is to measure individuals' stereoacuity for the focal and vergence distances being measured and apply it to the image-point disparities. This should ensure that the image is perceived the same across participants. Finally, reaction time experiments may use displays that can present images with and without conflict at the same vergence distance, similar to the '2-screen display' used in Experiment 8.

6.3.4 The zone of good stereo and zone of comfort

In terms of establishing guidelines for the vergence distances in stereo content at certain focal distances, our results suggest that both the zone of comfort and the zone of good stereo should be considered together. As previously mentioned, if researchers use the zone of comfort as a guideline for maximum conflict, for most individuals, stereo performance would be substantially reduced at the maximum negative conflict. Therefore, stereo performance is, in some cases, a limiting factor when considering the conflict a viewer can tolerate with focal distance (even within the established zone of comfort). Although it must be noted that the 'limits' of both zones have been arbitrarily chosen depending on what might be deemed uncomfortable, or 'poor' stereo performance. Therefore further studies are required to create definitive guidelines for both the zone of comfort and the zone of good stereo according to changing criteria.

6.3.5 Future Research

Some studies to date have investigated the effects of conflict using a single conflict and often a single sign of conflict (Akeley et al., 2004; Hoffman, et al., 2008; Shibata et al., 2011; Emoto et al., 2005). If clear interpretations are to be made about the effects of viewing conflict, interleaving conflicts is not possible. However, here we measured a more realistic example of stereo viewing in which the observers viewed a combination of conflicts. Importantly, in our studies we were able to present a known level of interleaved conflicts. This allowed us to control what conflict magnitudes the viewers observed, unlike in other 'natural' stimuli in which observers view feature films. For example, Read and Bohr, (2015) used the animated film "Toy Story" (1995, produced by Pixar Animation Studios) as their stimulus to measure visual discomfort. Although reflective of 'normal' stereo 3-D viewing, the 'Depth Budget' in these movies tends to be unknown. Therefore the researchers have no control over the magnitude of conflict, which therefore might change substantially across experiments. Hence, it might be advisable (where possible) to test using interleaving, but controlled conflict magnitudes. Thus, more research needs to be conducted to establish the effects of viewing multiple conflict magnitudes to better understand how audiences are affected.

It also remains to be established whether there are any short- or long-term effects of viewing stereo images with conflict. Here, we have gone some way to answering this question. The results in Experiment 8 are good news for the short-term effects of viewing conflict. Our results imply that any subsequent effect of viewing conflict is relatively small when considering both the proportion of individuals affected and the seriousness of the effects. These findings are consistent with similar research conducted by Read and Bohr (2014), who found that only 14% of their observers reported experiencing adverse effects when viewing stereo 3-D content. Similarly, in a recent study, Read et al. (2015) investigated the short-term effects of viewing stereo 3-D on a series of visuomotor tasks. These tasks were aimed at measuring the effect viewing conflict has on balance and coordination. They concluded there were no differences in performance between 2-D and stereo 3-D viewing. However, they did not examine individual effects of viewing conflict. Therefore, it is possible that some of their observers experienced an effect but not all. In our studies, we found large individual differences across observers. Thus, it might be necessary to consider individual data when examining the effects of viewing conflict.

Also, for instances where individuals are affected by viewing conflict, there is currently no research into whether recovery time scales with exposure time to the stereo content. For example, we found that after 10 minutes of exposure to conflict, on average, the observers affected took up to 2 minutes to recover. The question is, does recovery time scale with exposure time? Had they been exposed to conflict for 20 minutes, would it take 4 minutes to recover? This is difficult to measure for two main reasons. Firstly, there are a number of stimulus properties that might affect the results, including the temporal and spatial properties of the stimuli (Kim et al., 2014). Secondly, as we found, the amount of conflict viewed by the observer can be difficult to control for.

Even less is known about the long-term effects of viewing conflict. For ethical reasons, long-term studies into the effects of viewing stereo 3-D cannot be conducted. However, with the continued use of conventional stereo 3-D displays not only for entertainment (including gaming where the content can be viewed for significant amounts of time) but also in work and educational environments, this is an important topic and ideally, should be better understood. Arguably, of particular importance are the long-term effects on children, whose visual systems are still developing (Rushton & Riddell, 1999; Lambooij et al., 2009). In their paper, Ruston and Riddell (1999) suggest avenues of research that might help provide some insight into this topic, including computational studies, animal studies, and research on adult populations (in terms of comparing the visual properties of those who use virtual reality headsets to those who do not).

6.4 Conclusion

We have considered the practicality and effectiveness of using a multi-plane display along with depth filtering as a general solution for presenting approximately correct focus cues in a stereoscopic display. By this measure, the picture is mixed: there are clear benefits for eliminating vergence-accommodation conflicts, while preserving stereo performance. But a somewhat complex trade-off may be required in the various aspects of depth-filtered images, depending on the particular requirements of a given application. However, the proposed optimisation technique (Narain et al., 2015) offers a potential solution to image-quality degradation with increased focal-plane separations. Therefore, our research suggests that fixed-viewpoint volumetric displays, similar to the one reported here, could replace conventional stereo 3-D displays for single-viewing applications, particularly where head-mounted displays could be used.

However, for larger audience applications where a multi-view component is required, such as T.V and cinema, conventional stereo 3-D displays may still be necessary. Our research into the effect of viewing conflict has shown large individual differences. We have found that some, but not all observers experience after-effects of viewing conflict stimuli. However, the proportion of observers affected and the size of the effects, along with the previous research into short-term effects of viewing conflict (Read et al., 2015; Read & Bohr, 2014) is good news for conventional stereo displays. Suggesting there are minimal after-effects of viewing conflict. We have also found that, similar to the zone of comfort (Shibata et al., 2011), a zone of good stereo can be established. This outlines the range of conflicts an individual can tolerate with reasonable stereo performance. We have found that certain ophthalmological factors (such as degree of presbyopia, ZCSBV and phoria) can predict an individual’s zone of good stereo. This research is the first attempt at framing the zone of good stereo and we concede that more research in the form of a large-scale population study is required to create comprehensive guidelines regarding individuals’ tolerance to conflict.

7 REFERENCES

Akeley, K., Watt, S. J., Girshick, A. R., & Banks, M. S. (2004). A stereo display prototype with multiple focal distances. ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH, 23(3), 804–813. doi:10.1145/1015706.1015804

Allison, R. S., Gillam, B. J., & Vecellio, E. (2009). Binocular depth discrimination and estimation beyond interaction space. *Journal of Vision*, *9*, 10.1–14. doi:10.1167/9.1.10

Alpern, M. (1958). Variability of accommodation during steady fixation at various levels of illuminance. *Journal of the Optical Society of America*, 48, 193–197. doi: 10.1364/josa.48.000193

Andersen, G. J. (1990). Focused attention in three-dimensional space. *Perception & Psychophysics, 47*(2), 112-120

Andersen, G. J., & Kramer, A. F. (1993). Limits of focused attention in three-dimensional space. *Perception & Psychophysics, 53*(6), 658-667.

Atchison, D., Fisher, S. W., Pedersen, C., & Ridall, P. G. (2005). Noticeable, troublesome and objectionable limits of blur. *Vision Research*, *45*, 1967–1974. doi:10.1016/j.visres.2005.01.022

Bach, M. (2007) The Freiburg Visual Acuity Test – Variability unchanged by post-hoc re-analysis. *Graefe's Arch Clin Exp Ophthalmol*, *245*(7): 965-971

Banks, M. S., Akeley, K., Hoffman, D. M., & Girshick, A. R. (2008). Consequences of incorrect focus cues in stereo displays. *Information Display*, *24*(7), 10–14.

Banks, M. S., Gepshtein, S., & Landy, M. S. (2004). Why is spatial stereoresolution so low? *The Journal of Neuroscience*, 24(9), 2077–2089. doi:10.1523/JNEUROSCI.3852-02.2004

Banks, M. S., Read, J. C., Allison, R. S., & Watt, S. J. (2012). Stereoscopy and the Human Visual system. *SMPTE motion imaging journal*, 121(4), 24-43.

doi:10.5594/j18173.

Binda, P., Pereverzeva, M., & Murray, S. O. (2013). Attention to bright surfaces enhances the pupillary light reflex. *The Journal of Neuroscience, 33*(5), 2199-2204.

Birren, J. E., Casperson, R. C., & Botwinick, J. (1950a). Age changes in pupil size. *Journal of Gerontology*, *5*(3), 216–221. doi:10.1093/geronj/5.3.216

Blundell, B. G. (2012). *On Exemplar 3D Display Techniques: Summary Discussion*. Auckland.

Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology, 45*(4), 602-607.

Bradshaw, M. F., & Glennerster, A. (2006). Stereoscopic acuity and observation distance. *Spatial Vision*, 19(1), 21–36. doi:10.1163/156856806775009250

Bradshaw, M.F. & Rogers, B.J. (1999). Sensitivity to horizontal and vertical corrugations defined by binocular disparity. Vision Research, 39, 3049-3056. doi:10.1016/S0042-6989(99)00015-2

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10: 433 – 436.

Brewster, D. (1849). Account of a new stereoscope. *Report of the British Association, Transactions of the Sections.* pp. 6–7.

Bruce, V., Green, P. R., & Georgeson, M. A. (2003). Visual perception: Physiology, psychology, & ecology. Psychology Press.

Bruce, A. S., Atchison, D. A., & Bhoola, H. (1995). Accommodation-Convergence Relationships and Age. *Investigative Ophthalmology & Visual Science*, *36*(2), 406–413.

Bülthoff, H. H., & Mallot, H. A. (1988). Integration of depth modules: stereo and shading. *Josa a, 5*(10), 1749-1758.

Burt, P., & Julesz, B. (1980). A disparity gradient limit for binocular fusion. *Science* (New York, N.Y.), 208, 615–617. doi:10.1126/science.7367885

Campell, F. W. (1957). The depth of field of the human eye. Optica Acta: *International Journal of optics*, *4*, 157-164.

Campbell, F. W., & Gregory, A. H. (1960). Effect of size of pupil on visual acuity. *Nature, 187*, 1121 – 1123.

Campbell, F. W., & Westheimer, G. (1960). Dynamics of accommodation responses of the human eye. *The Journal of physiology, 151*(2), 285-295.

Charman, W. N., & Whitefoot, H. (1977). Pupil Diameter and the Depth-of-field of the Human Eye as Measured by Laser Speckle. Optica Acta: *International Journal of optics*, *24*(12), 1211-1216.

Ciuffreda, K. J. (1991). Accommodation and its anomalies. *Vision and visual dysfunction*, *1*, 231-279.

Costa, M. F., Moreira, S. M. C. F., Hamer, R. D., & Ventura, D. F. (2010). Effects of age and optical blur on real depth stereoacuity. *Ophthalmic and Physiological Optics*, *30*(5), 660–666. doi:10.1111/j.1475-1313.2010.00750.

Cristino, F., Davitt, L., Hayward, W. G., & Leek, E. C. (2015). Stereo disparity facilitates view generalization during shape recognition for solid multipart objects. *The Quarterly Journal of Experimental Psychology, 68*(12), 2419-2436.

De Groot, S. G., & Gebhard, J. W. (1952). Pupil Size as Determined by Adapting Luminance. *The Journal of the Optical Society of America, 42*, 492-495.

Dees, J. W. (1966). Accuracy of absolute visual distance and size estimation in space as a function of stereopsis and motion parallax. *Journal of Experimental Psychology*, 72, 466–476.

Denton, E. J. (1956). The responses of the pupil of Gekko gekko to external light stimulus. *The Journal of general physiology, 40*(2), 201.

Duane, A. (1912). Normal Values of the Accommodation at all Ages. *Journal of the American Medical Association*, LIX(12), 1010. doi:10.1001/jama.1912.04270090254042

Duchowski, A. T., House, D. H., Gestring, J., Wang, R. I., Krejtz, K., Krejtz, I., Mantiul, R., & Bazyluk, B. (2014). Reducing visual discomfort of 3d stereoscopic displays with gaze-contingent depth-of-field. In *Proceedings of the ACM Symposium on Applied Perception* (pp. 39-46). doi: 10.1145/2628257.2628259

Easdown, R. (2013, August 15). Future looks flat for 3D TV. The Sydney Morning Herald. Retrieved from http://www.smh.com.au/digital-life/digital-life-news/future-looks-flat-for-3d-tv-20130813-2rv1q.html

Edelman, S., & Bülthoff, H. H. (1992). Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. *Vision research, 32*(12), 2385-2400.

Emoto, M., Niida, T., & Okano, F. (2005). Repeated vergence adaptation causes the decline of visual functions in watching stereoscopic television. *NHK Laboratories Note*, *1*(501), 328–340. doi:10.1109/JDT.2005.858938

Falmagne, J. C. (1986). Psychophysical measurement and theory. In Boff, K. R., Kaufman, L., & Thomas, J. P. (Eds), *Handbook of perception and human performance.* New York: John Wiley and Sons.

Fattal, D., Peng, Z., Tran, T., Vo, S., Fiorentino, M., Brug, J., & Beausoleil, R. G. (2013). A multi-directional backlight for a wide-angle, glasses-free 3D display. *Nature*, 24–25. doi:10.1109/IPCon.2013.6656348

Favalora, G. E., Napoli, J., Hall, D. M., Dorval, R. K., Giovinco, M. G., Richmond, M. J., & Chun, W. S. (2002). 100-Million-Voxel Volumetric Display. In *SPIE's 16th Annual International Symposium on Aerospace/Defense Sensing, Simulation, and Controls*. doi:10.1117/12.480930

Fender, D. H. (1964). Control mechanisms of the eye. *Scientific American*, 211, 24–33. doi: 10.1016/0014-4835(92)90789-u

Fincham, B. Y. E. F., & Walton, J. (1957). The reciprocal actions of accommodation and convergence. *Journal of Phsyiology*, *137*(1956), 488–508.

Fry, G. A. (1939). Further experiments on the accommodation convergence relationship. *American Journal of Optometry*, *16*, 125.

Fukushima, T., Torii, M., Ukai, K., Wolffsohn, J. S., & Gilmartin, B. (2009). The relationship between CA/C ratio and individual differences in dynamic accommodative responses while viewing stereoscopic images. *Journal of Vision*, *9*(13), 21.1–13. doi:10.1167/9.13.21

Garding, J., Porrill, J., Mayhew, J. E. W., & Frisby, J. P. (1995). Stereopsis, vertical disparity and relief transformations. Vision research, 35(5), 703-722.

Gescheider , G. A. (1997) . Psychophysics: The Fundamentals. . Mahwah, New Jersey : Lawrence Erlbaum Associates .Goldstein, E. B. (1996). *Sensation and perception* (4th ed.). New York: Brooks Cole

Grosvenor, T. P. (2007). *Primary Care Optometry* (5th ed.). Sr. Louis, Mo: Butterworth-Heinemann/Elsevier, 257.

Häkkinen, J., Polonen, M., Takatalo, J., & Nyman, G. (2006). Simulator Sickness in Virtual Display Gaming : A Comparison of Stereoscopic and Non-stereoscopic. *Proceedings of the 8th Conference on Human-Computer Interaction with Mobile Devices and Services*, 227–229. doi:10.1145/1152215.1152263

Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values of d′. *Behavior Research Methods, Instruments, & Computers*, *27*(1), 46–51. doi:10.3758/BF03203619

He, Z. J., & Nakayama, K. (1995). Visual attention to surfaces in three-dimensional space. *Proceedings of the National Academy of Sciences, 92*(24), 11155-11159.

Heath, G. G. (1956). Components of accommodation. *Optometry and Vision Science*, 33(11), 569–579. doi: 10.1097/00006324-195611000-00001

Heckbert, P. (1982). Color image quantization for frame buffer display (Vol. 16, No. 3, pp. 297-307). ACM.

Held, R. T., Cooper, E. A., & Banks, M. S. (2012). Blur and disparity are complementary cues to depth. *Current Biology*, 22(5), 426-431.

Heron, G., Charman, W. N., & Schor, C. M. (2001). Age changes in the interactions between the accommodation and vergence systems. *Optometry and Vision Science : Official Publication of the American Academy of Optometry*, *78*(10), 754–762. doi:10.1097/00006324-200110000-00015

Higashi, T., & Nakamizo, S. (2003). Perceived depth and direction as a function of ocular vergence. *The Japanese Journal of Psychonomic Science*, 21(2), 103-111.

Hillis, J. M., & Banks, M. S. (2001). Are corresponding points fixed? *Vision Research*, *41*(19), 2457–2473. doi:10.1016/S0042-6989(01)00137-7

Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture and disparity cues: Optimal cue combination. *Journal of vision*, 4(12), 1. doi:10.1167/4.12.1

Hoffman, D. M., Girshick, A. R., Akeley, K., & Banks, M. S. (2008). Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision*, *8*, 33.1–30. doi:10.1167/8.3.33

Hoffman, D. M., Karasev, V. I., & Banks, M. S. (2011). Temporal Presentation Protocols in Stereo Displays: Flicker Visibility, Perceived Motion, and Perceived Depth. *J Soc Inf Disp*, *19*(3), 271-297.

Hougen, D. R., & Ahuja, N. (1993, May). Estimation of the light source distribution and its use in integrated shape recovery from stereo and shading. In Computer Vision, 1993. Proceedings., Fourth International Conference on(pp. 148-155). IEEE.

Howard, I. P. (2002). Seeing in depth, Vol. 1: Basic mechanisms. University of Toronto Press.

Howard, I. P. (2012). Perceiving in depth, Vol. 1: Basic mechanisms. New York: Oxford University Press.

Howard, I. P., & Rogers, B. J. (2002). Binocular correspondence and the horopter. Binocular Vision and Stereopsis, 35-36.

Howarth, P. A. (2011). Potential hazards of viewing 3-D stereoscopic television, cinema and computer games: A review. *Ophthalmic and Physiological Optics*. doi:10.1111/j.1475-1313.2011.00822.x

Howarth, P. A., & Costello, P. J. (1997). The occurrence of virtual simulation sickness symptoms when an HMD was used as a personal viewing system. *Displays*, 18(2), 107–116. doi: 10.1016/s0141-9382(97)00011-5

Hung, G. K., Zhu, H., & Ciuffreda, K. J. (1996). Convergence and divergence exhibit different response characteristics to symmetric stimuli. *Vision Research*, *37*, 1197–1205.

Howard, I. P. (2002) Seeing in Depth: Vol. 1. Basic Mechanisms. New York: Porteous.

Johnston, E. B., Cumming, B. G., & Landy, M. S. (1994). Integration of stereopsis and motion shape cues. *Vision research, 34*(17), 2259-2275.

Jones, R. (1977). Anomalies of disparity detection in the human visual system. *The Journal of physiology, 264*(3), 621.

Jin, Z. X., Zhang, Y. J., Wang, X., & Plocher, T. (2007). Evaluating the usability of an auto-stereoscopic display. *Lecture Notes in Computer Science*, 4551, 605. doi: 10.1007/978-3-540-73107-8_68

Julesz, B. (1960). Binocular Depth Perception of Computer□Generated Patterns. *Bell System Technical Journal*, 39(5), 1125-1162.

Kasthurirangan, S., & Glasser, A. (2006). Age related changes in the characteristics of the near pupil response. *Vision research, 46*(8), 1393-1403.

Kaufman, L., Kaufman, J. H., Noble, R., Edlund, S., Bai, S., & King, T. (2006). Perceptual distance and the constancy of size and stereoscopic depth. *Spatial Vision*, *19*, 439–457.

Kim, E. H., Granger-Donetti, B., Vicci, V. R., & Alvarez, T. L. (2010a). The relationship between phoria and the ratio of convergence peak velocity to divergence peak velocity. *Investigative Ophthalmology and Visual Science*, *51*, 4017–4027. doi:10.1167/iovs.09-4560

Kim, E. H., Granger-Donetti, B., Vicci, V. R., & Alvarez, T. L. (2010b). The relationship between phoria and the ratio of convergence peak velocity to divergence peak velocity. *Investigative Ophthalmology and Visual Science*, *51*, 4017–4027. doi:10.1167/iovs.09-4560

Kim, J., Kane, D., & Banks, M. S. (2014). The rate of change of vergence-accommodation conflict affects visual discomfort. *Vision Research*, *105*, 159-165. doi:10.1016/j.visres.2014.10.021

Kim, J., Kane, D., & Banks, M. S. (2013). Visual discomfort and the temporal properties of the vergence-accommodation Conflict. *Proc. SPIE* 8288*,* Stereoscopic Displays and Applications XXIII, 1–15. doi:10.1117/12.912223.

Kingdom F.A.A., Prins (2009) Psychophysics: a practical introduction. London: Academic Press. p28.

Klein, S. (2001). Measuring, estimating, and understanding the psychometric function: a commentary. *Perception & Psychophysics*, *63*(8), 1421–1455. doi:10.3758/BF03194552

Knill, D. C. (2005). Reaching for visual cues to depth: The brain combines depth cues differently for motor control and perception. Journal of Vision,5(2), 2-2.

Knill, D. C., & Saunders, J. A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant?. Vision research,43(24), 2539-2558.

Koivisto, M., Hyönä, J., & Revonsuo, A. (2004). The effects of eye movements, spatial attention, and stimulus features on inattentional blindness. *Vision Research*, *44*(27), 3211-3221.

Kooi, F. L., & Toet, A. (2004). Visual comfort of binocular and 3D displays. *Displays*, *25*(2-3), 99–108. doi:10.1016/j.displa.2004.07.004

Kotulak, J. C., & Schor, C. M. (1986a). A computational model of the error detector of human visual accom- modation. Biological Cybernetics, 54, 189–194.

Laeng, B., Ørbo, M., Holmlund, T., & Miozzo, M. (2011). Pupillary stroop effects. *Cognitive processing, 12*(1), 13-21.

Lambooij, M., IJsselsteijn, W., Fortuin, M., & Heynderickx, I. (2009). Visual Discomfort and Visual Fatigue of Stereoscopic Displays: A Review. *Journal of Imaging Science and Technology*, *53*(3), 030201. doi:10.2352/J.ImagingSci.Technol.2009.53.3.030201

Legge, G. E., & Gu, Y. C. (1989). Stereopsis and contrast. *Vision Research*, *29*(8), 989–1004. doi:10.1016/0042-6989(89)90114-4

Leibowitz, H. (1952). The Effect of Pupil Size on Visual Acuity for Photometrically Equated Test Fields at Various Levels of Luminance*. *JOSA,42*(6), 416-422.

Liu, S., Cheng, D., & Hua, H. (2008, September). An optical see-through head mounted display with addressable focal planes. *In Mixed and Augmented Reality*, 2008. ISMAR 2008. 7th IEEE/ACM International Symposium on (pp. 33-42). IEEE.

Liu S, & Hua H. (2009) Time-multiplexed dual-focal plane head-mounted display with a liquid lens. *Optical Lett*ers, *34*(11):1642–1644.

Loewenfeld, I. E. (1993). The reaction to near vision. *The pupil: Anatomy, physiology, and clinical applications, 1*, 295-317.

Love, G.D.; Hoffman, D.M.; Hands, P.J.W.; Gao, J.; Kirby, A.K.; Banks, M. S. (2009). High-Speed switchable lens enables the development of a volumetric display. *Optics Express*, *28*(18), 4439–4448. doi:10.1037/a0013262.Open

MacKenzie, K. J., Dickson, R. a., & Watt, S. J. (2012). Vergence and accommodation to multiple-image-plane stereoscopic displays: "real world" responses with practical image-plane separations? *Journal of Electronic Imaging*, *21*(1), 011002. doi:10.1117/1.JEI.21.1.011002

MacKenzie, K. J., Hoffman, D. M., & Watt, S. J. (2010). Accommodation to multiple-focal-plane displays: Implications for improving stereoscopic displays and for accommodation control. *Journal of Vision*, *10*, 22. doi:10.1167/10.8.22

Maiello, G., Chessa, M., Solari, F., & Bex, P. J. (2014). Simulated disparity and peripheral blur interact during binocular fusion. *Journal of vision*, 14(8), 13. doi:10.1167/14.8.13

Manny, R. E., & Banks, M. S. (1984). A model of steady- state accommodation: II. Effects of luminance. Inves- tigative Ophthalmology and Visual Science, 25, 182.

Mansson, J. (1997). Stereovision: A model of human stereopsis. *Technical Report, Lund University Cognitive Studies,* Sweden, 1–22.

Martens T. G., & Ogle K. N. (1959). Observations on accommodative convergence; especially its nonlinear relationships. *American Journal of Ophthalmology*, 47, 455–463.

Mather, G., & Smith, D. R. (2000). Depth cue integration: stereopsis and image blur. *Vision Research*, *40*, 3501–3506. http://doi.org/10.1016/S0042-6989(00)00178-4

Mathews, S., & Kruger, P. B. (1994). Spatiotemporal transfer function of human accommodation. *Vision Research*, 34(15), 1965-1980. doi:10.1016/0042-6989(94)90026-4.

Mays LE, Gamlin PDR. (2000). Neuronal circuits for accommodation and vergence in the primate. In: FranzénO, RichterH, StarkL, eds. *Accommodation and Vergence Mechanisms in the Visual System*. Basel: Birkhäuser Verlag, 1–9.

Mendiburu, B. (2009). 3D Movie Making: Stereoscopic Digital Cinema from Script to Screen. United Kingdom: Elsevier Science.

Menozzi, M. (2000). Visual ergonomics of head-mounted displays. *Japanese Psychological Research*, 42(4), 213–221. doi: 10.1111/1468-5884.00148

Miles, F., Judge, S. J., & Optican, L. M. (1987). Optically Induced Accommodation in the Couplings Between Vergence and Accommodation. *The Journal of Neuroscience*, *7*(8), 2576–2589.

Miles, H. C., Pop, S. R., Watt, S. J., Lawrence, G. P., & John, N. W. (2012). A review of virtual environments for training in ball sports. *Computers & Graphics*, 36(6), 714-726. doi:10.1016/j.cag.2012.04.007

Mon-Williams, M., & Tresilian, J. R. (2000). Ordinal depth information from accommodation? *Ergonomics*, 43, 391–404. doi: 10.1007/s002210000410

Mordi, J. a., & Ciuffreda, K. J. (1998). Dynamic aspects of accommodation: Age and presbyopia. *Vision Research*, *38*, 1643–1653. doi:10.1016/j.visres.2003.07.014

Narain, R., Albert, R. A., Bulbul, A., Ward, G. J., Banks, M. S., & Brien, J. F. O. (2015). Optimal Presentation of Imagery with Focus Cues on Multi-Plane Displays. *ACM Transactions on Graphics*, *34*(4). doi:10.1145/2766909.

O'leary, A., & Wallach, H. (1980). Familiar size and linear perspective as distance cues in stereoscopic depth constancy. *Perception & Psychophysics*, 27(2), 131-135.

Owens, D. A. (1980). A comparison of accommodative responsiveness and contrast sensitivity for sinusoidal gratings. *Vision Research*, 20(2), 159–167. doi: 10.1016/0042-6989(80)90158-3

Palmer, S. E. (1999). Vision Science: Photons to Phenomenology. *Vision Science. MIT Press*.

Palmisano, S., Gillam, B., Govan, D. G., Allison, R. S., & Harris, J. M. (2010). Stereoscopic perception of real depths at large distances. *Journal of Vision*, *10*, 19. doi:10.1167/10.6.19

Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International journal of human-computer studies, 59*(1), 185-198.

Patterson, R. (2015). Human Factors of Stereoscopic 3D Displays. London: Springer. .

Patterson, R. (2009). Review Paper: Human factors of stereo displays: An update. *Journal of the Society for Information Display*, *17*(12), 987. doi: 10.1889/JSID17.12.987

Patterson, C., Cristino, F., Hayward, W., & Leek, C. (2012). Stereo information benefits view generalization in object recognition. *Journal of Vision, 12*(9), 1044-1044

Patterson, R., & Fox, R. (1984). The effect of testing method on stereoanomaly. *Vision Research, 24*(5), 403-408.

Pastoor, S. (1995) Human factors of 3D imaging: results of recent research at Heinrich-Hertz-Institut Berlin. *Proceedings of the International Display Workshop*

Pastoor, S., & Wöpking, M. (1997). 3-D displays: A review of current technologies. *Displays*, *17*(2), 100–110. doi:10.1016/S0141-9382(96)01040-2

Peli, E. (1998). The visual effects of head-mounted display (HMD) are not distinguishable from those of desk-top computer display. *Vision Research*, *38*, 2053–2066. doi:10.1016/S0042-6989(97)00397-0

Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437 – 442.

Percival, A. S. (1892). The relation of convergence to accommodation and its practical bearing. *Ophthalmological Review*. 11:313–328.

Pickwell, D., Jenkins, T., & Yekta, A. A. (1987). The effect of fixation disparity and associated heterophoria of reading at an abnormally close distance. *Ophthalmic and Physiological Optics*,*7*(4), 345-347. doi: 10.1111/j.1475-1313.1987.tb00758.x

Poggio, G. F., Motter, B. C., Squatrito, S., & Trotter, Y. (1985). Responses of neurons in visual cortex (V1 and V2) of the alert macaque to dynamic random-dot stereograms. *Vision Research*, *25*(3), 397–406. doi:10.1016/0042-6989(85)90065-3

Ramsdale, C., & Charman, W. N. (1989). A longitudinal study of the changes in the static accommodation response. *Ophthalmic and Physiological Optics*, 9(3), 255-263. doi: 10.1111/j.1475-1313.1989.tb00903.x

Raymond, J. E., Lindblad, I. M., & Leibowitz, H. W. (1984). The effect of contrast on sustained detection. *Vision Research*, 24, 183–188. doi: 10.1016/0042-6989(84)90120-2

Read, J. C. a, & Bohr, I. (2014). User experience while viewing stereoscopic 3D television. *Ergonomics*, *57*(8), 1140–53. doi:10.1080/00140139.2014.914581

Read, J. C. A., Simonotto, J., Bohr, I., Godfrey, A., Galna, B., Rochester, L., & Smulders, T. V. (2015). Balance and coordination after viewing stereoscopic 3D television. *Royal Society Open Science*, 2(7), 140522. doi:10.1098/rsos.140522

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

Rogers, B., & Graham, M. (1982). Similarities between motion parallax and stereopsis in human depth perception. *Vision research*, 22(2), 261-270. doi:10.1016/0042-6989(82)90126-2

Rolland, J. P., Krueger, M. W., & Goon, A. (1999). Dynamic focusing in head-mounted displays, *Proc. SPIE* 3639*,* Stereoscopic Displays and Virtual Reality Systems VI, 463–470.

Rose, D (1995). Psychophysical Methods. In Breakwell, G. M., Hammond, S. & Fife-Schaw, C. (Eds). *Research Methods in Psychology* (pp. 140 – 159). London: Sage Publications.

Rushton, S. K., & Riddell, P. M. (1999). Developing visual systems and exposure to virtual reality and stereo displays: some concerns and speculations about the demands on accommodation and vergence. *Applied Ergonomics*, *30*(1), 69–78. doi:10.1016/S0003-6870(98)00044-1

Saunders, J. A., & Knill, D. C. (2003). Humans use continuous visual feedback from the hand to control fast reaching movements. *Experimental Brain Research, 152*(3), 341-352.

Scarfe, P., & Hibbard, P. B. (2011). Statistically optimal integration of biased sensory estimates. *Journal of Vision, 11*(7), 12-12.

Schäfer, W. D., & Weale, R. A. (1969). [Pupillary near reflex in various age groups]. Bericht uber die Zusammenkunft. *Deutsche Ophthalmologische Gesellschaft, 70*, 72-75.

Schor, C. M. (1992). A Dynamic Model of Cross-Coupling Between Accommodation and Convergence: Simulations of Step and Frequency Responses. *Optometry and Vision Science*, *69*(4), 258–269.

Schor, C. M. (1979) The relationship between fusional vergence eye movements and fixation disparity. *Vision Research*, 19, 1359 – 1367.

Schor, C. M., Wood, I., & Ogawa, J. (1984). Binocular Sensory Fusion is Limited by Spatial Resolution. *Vision Research*, *24*(7), 661–665.

Schor, C.M., & Tsuetaki, T.K. (1987). Fatigue of accommodation and vergence modifies their mutual interactions. *Invest Ophthalmol Vis Sc, 28,* 1250-1259.

Schor, C. M., & Tyler, C. W. (1981). Spatio-temporal properties of Panum's fusional area. *Vision Research*, 21(5), 683–692. doi:10.1016/0042-6989(81)90076-6

Schowengerdt, B. T., & Seibel, E. J. (2006). True 3-D scanned voxel displays using single or multiple light sources. *Journal of the Society for Information Dis- play*, 14, 135–143.

Scheiman, M., & Wick, B. (2008). Clinical management of binocular vision: heterophoric, accommodative, and eye movement disorders. Lippincott Williams & Wilkins.

Schroeder, T. L., Rainey, B. B., Goss, D. A., & Grosvenor, T. P. (1996). Reliability of and Comparisons Among Methods of Phoria. *Optometry and Vision Science*, *73*(6), 389–397.

Semmlow, J. L., Hung, G. K., & Ciuffreda, K. J. (1986). Quantitative assessment of disparity vergence components. *Investigative Ophthalmology and Visual Science*, *27*, 558–564.

Semmlow, J., & Wetzel, P. (1979). Dynamic contributions of the components of binocular vergence. *Journal of the Optical Society of America*, 69, 639–645.

Sexton, I., & Surman, P. (1999). Stereoscopic and autostereoscopic display systems. *IEEE Signal Processing Magazine*, *16*, 85–99. doi:10.1109/79.768575

Shibata, T., Kawai, T., Ohta, K., Otsuki, M., Miyake, N., Yoshihara, Y., & Iwasaki, T. (2005). Stereoscopic 3-D display with optical correction for the reduction of the discrepancy between accommodation and convergence. *Journal of the Society for Information Display*, *13*(8), 665. doi:10.1889/1.2039295

Shibata, T., Kim, J., Hoffman, D. M., & Banks, M. S. (2011). The zone of comfort: Predicting visual discomfort with stereo displays. *Journal of Vision*, *29*(6), 997–1003. doi:10.1016/j.biotechadv.2011.08.021.

Shiwa A, Omura K, Kishino F, (1996). Proposal for a 3-D display with accommodative compensation: 3DDAC. *J. Soc. Inf. Disp*, 4(4):255–261.

Simons, K. (1984). Effects on stereopsis of monocular versus binocular degradaton of image contrast. *Investigative Ophthalmology and Visual Science*, *25*(8), 987–989.

Smallman, H.S. and McKee, S.P. (1995). A contrast ratio constraint on stereo matching. *Proc. R. Soc. London Ser. B.*, *260*(39), 265–271.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers : A Journal of the Psychonomic Society, Inc*, *31*(1), 137–149.

Sullivan, A. (2005). 3 Deep: new displays render images you can almost reach out and touch. *IEEE Spectrum*, *42*(4), 30–35.

Suyama, S., Date, M., & Takada, H. (2000) Three-dimensional display system with dual frequency liquid crystal varifocal lens. *Japanese Journal of Applied* Physics, *39*(Part 1, No. 2A), 480–484.

Switkes, E., Bradley, A., & Schor, C. (1990). Readily visible changes in color contrast are insufficient to stimulate accommodation. *Vision Research*, 30(9), 1367–1376. doi: 10.1016/0042-6989(90)90010-i

Tait, E. (1951). Accommodative Convergence. *American Journal of Ophthalmology*, *34*, 1093-1107.

Theatrical Market Statistics (2014) Retrieved from http://www.mpaa.org/wp-content/uploads/2015/03/MPAA-Theatrical-Market-Statistics-2014.pdf

Toates, F. M. (1972). Accommodation function of the human eye. *Physiological Review*, 52, 828–863.

Tucker, J., Charman, W. N., & Ward, P. A. (1986). Modulation dependence of the accommodation response to sinusoidal gratings. *Vision Research*, 26, 1693–1707. doi:10.1016/0042-6989(86)90056-8

Tyler, C. W. (1975). Spatial organization of binocular disparity sensitivity. *Vision Research*, 15(5), 583-590. doi:10.1016/0042-6989(75)90306-5

Ukai, K., & Howarth, P. a. (2008). Visual fatigue caused by viewing stereoscopic motion images: Background, theories, and observations. *Displays*, 29, 106–116. doi:10.1016/j.displa.2007.09.004

Ukai, K. (2007). Visual fatigue caused by viewing stereoscopic images and mechanism of accommodation. *Proceedings of the First International Symposium on University Communication*, 1, 176-179.

Universale, C. O., & Committee, V. F. C. (1988). Visual Acuity Measurement Standard. *Italian Journal of Ophthalmology*, *2*, 1–18.

van Ee, R., & Richards, W. (2002). A planar and a volumetric test for stereoanomaly. *Perception, 31*(1), 51-64.

Vinnikov, M., & Allison, R. S. (2014, March). Gaze-contingent depth of field in realistic scenes: The user experience. In *Proceedings of the Symposium on Eye Tracking Research and Applications* (pp. 119-126). ACM. doi: 10.1145/2578153.2578170

Wallach, H., & Norris, C. M. (1963). Accommodation as a distance-cue. *American Journal of Psychology*, 76, 659–664. doi: 10.2307/1419717

Walline, J. J., Mutti, O. D., Zadink, K., & Jones, A. L. (1998). Development of Phoria in Children. *Optometry and Vision Science*, *75*(8), 605–610.

Wandell, B. A. (1995). Foundations of vision. Sinauer Associates; Sunderland, MA, US.

Wang, B., & Ciuffreda, K. J. (2006). Depth-of-focus of the human eye: Theory and clinical implications. *Survey of Ophthalmology*, *51*(1), 75–85. doi: 10.1016/j.survophthal.2005.11.003

Wann, J.P., & Mon-Williams, M. (2002) Measurement of visual after effects following virtual environment exposure. In: Stanney, KM., editor. *Handbook of virtual environments: Design, implementation, and applications.* Lawrence Erlbaum Associates; Hillsdale, NJ: p. 731-749.

Wann, J. P., Rushton, S., & Mon-Williams, M. (1995). Natural problems for stereoscopic depth perception in virtual environments. *Vision Research*, *35*(19), 2731–2736. doi:10.1016/0042-6989

Ward, P. A. (1987). The effect of spatial frequency on steady-state accommodation. *Ophthalmic & Physiological Optics*, 7, 211–217. doi: 10.1111/j.1475-1313.1987.tb00735.x

Watt, S. J., Akeley, K., Ernst, M. O., & Banks, M. S. (2005). Focus cues affect perceived depth. *Journal of Vision*, *5*, 834–862. doi:10.1167/5.10.7

Watt, S. J., Akeley, K., Girshick, a. R., & Banks, M. S. (2005). Achieving near-correct focus cues in a 3-D display using multiple image planes. *SPIE 5666,* Human Vision and Electronic Imaging X, *393.* doi:10.1167/4.8.463

Watt, S. J., & MacKenzie, K. J. (2013). 3D Media and the Human Visual System. *In Emerging Technologies for 3D Video: Creation, Coding, Transmission and Rendering*, 833 – 898.

Webber, J. E. (2014, December 11). Virtual reality headsets: How Oculus Rift has started a games revolution. The Guardian. The Guardian. Retrieved from http://www.theguardian.com/technology/2014/dec/11/virtual-reality-oculus-rift-gaming-vr-development

Wetherill , G. B. , & Levitt , H. (1965) . Sequential estimation of points on a psychometric function . *The British Journal of Mathematical and Statistical Psychology*, *18*, 1-10

Wheatstone, C. (1838) Contributions to the physiology of vision - Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London,* 128, 371 – 394.

Wheatstone, C. (1852). Contributions to the physiology of vision - Part the second. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London,* 142, 1 – 17.

Wichmann, F. a, & Hill, N. J. (2001). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception & Psychophysics*, *63*(8), 1314–1329. doi:10.3758/BF03194545

Wilmer, J. B. (2008). How to use individual differences to isolate functional organization, biology, and utility of visual functions; with illustrative proposals for stereopsis. Spatial vision, 21(6), 561-579.

Winkler, S., & Min, D. (2013). Stereo/multiview picture quality: Overview and recent advances. *Signal Processing: Image Communication*, 28(10), 1358-1373. doi:10.1016/j.image.2013.07.008

Vishwanath, D., & Blaser, E. (2010). Retinal blur and the perception of egocentric distance. *Journal of Vision*, 10(10), 26. doi:10.1167/10.10.26

Yamazaki, T., Kamijo, K., & Fukuzumi, S. (1990). Quantitative evaluation of visual fatigue encountered in viewing stereoscopic 3 D displays: near-point distance and visual evoked potential study. *Proceedings of the Society for Information Display*, 31(3), 245-247.

Yang, S., Schlieski, T., Selmins, B., Cooper, S. C., Doherty, R. A., Corriveau, P. J., & Sheedy, J. E. (2012). Stereoscopic Viewing and Reported Perceived Immersion and Symptoms. *Optometry and Vision Science*, 89(7), 1068–1080.

Yang, S., & Sheedy, J. E. (2011). Effects of Vergence and Accommodative Responses on Viewer's Comfort in Viewing 3D Stimuli. *IS&T/SPIE Electronic Imaging. International Society for Optics and Photonics*. doi:10.1117/12.872546

Yano, S., Ide, S., Mitsuhashi, T., & Thwaites, H. (2002). A study of visual fatigue and visual comfort for 3D HDTV/HDTV images. *Displays*, 23(4), 191–201. doi:10.1016/S0141-9382(02)00038-0