The spatial averaging of disparities in brief, static random-dot stereograms

Ariella Vered Popple

Popple, Ariella Vered (1999) The spatial averaging of disparities in brief, static random-dot stereograms. Doctoral thesis, Durham University.

The Spatial Averaging of Disparities in Brief, Static Random-dot Stereograms

by

Ariella Vered Popple

Submitted to the University of Durham

Department of Psychology

for the degree of

Doctor of Philosophy

1999

Ariella Vered Popple

The spatial averaging of disparities in brief, static random-dot stereograms

PhD. Thesis submitted to the University of Durham Psychology Department 1999

Abstract

Visual images from the two eyes are transmitted to the brain. Because the eyes are horizontally separated, there is a horizontal disparity between the two images. The amount of disparity between the images of a given point depends on the distance of that point from the viewer's point of fixation. A natural visual environment contains surfaces at many different depths. Therefore, the brain must process a spatial distribution of disparities.

How are these disparities spatially put together? Brief (about 200 msec) static cyclopean random-dot stereograms were used as stimuli for vergence and depth discrimination to answer this question. The results indicated a large averaging region for vergence, and a smaller pooling region for depth discrimination. Vergence responded to the mean disparity of two transparent planes. When a disparate target was present in a fixation plane surround, vergence improved as target size was increased, with a saturation at 3-6 degrees. Depth discrimination thresholds improved with target size, reaching a minimum at 1-3 degrees, but increased for larger targets. Depth discrimination showed a dependence on the extent of a disparity pedestal surrounding the target, consistent with vergence facilitation. Vergence might, therefore, implement a coarse-to-fine reduction in binocular matching noise. Interocular decorrelation can be considered as multiple chance matches at different disparities. The spatial pooling limits found for disparity were replicated when interocular decorrelation was discriminated. The disparity of the random dots also influenced the apparent horizontal alignment of neighbouring monocular lines. This finding suggests that disparity averaging takes place at an early stage of visual processing.

The following possible explanations were considered: 1) Disparities are detected in different spatial frequency channels (Marr and Poggio, 1979). 2) Second-order luminance patterns are matched between the two eyes using non-linear channels. 3) Secondary disparity filters process disparities extracted from linear filters.

For my grandfather Jack Pepys,

my father Aviam Heidecker,

and my brother Yannai Heidecker.

They did not live to see this finished.

Acknowledgments

I wish to thank my husband, Mark, and my three children, Ulysses, Dante and Absalom for their enduring patience and support over the last three years. My grandmother, Rhoda Pepys, is owed a debt of gratitude for the financial assistance without which this would not have been possible.

I am grateful to my supervisor, Prof. John Findlay, for his insight and enthusiasm during our frequent and often lengthy discussions. Past and present members of the psychology department are thanked for their assistance and encouragement, especially Dr. Harvey Smallman, my collaborator on chapter 2, who joined the meetings with Prof. Findlay and contributed knowledge and expertise as well as the use of his laboratory, and Dr. Iain Gilchrist for practical help with the eye-trackers. This work was done with the support of a postgraduate studentship from the University of Durham Department of Psychology.

Finally, many thanks to my tireless fellow postgraduates and other volunteer subjects who participated in the experiments.

Contents

Figures and Tables

Declaration

No part of the material contained in this thesis has previously been submitted for a degree in this or any other university. A number of joint publications have arisen from work presented in the various chapters. These are listed below.

Chapter 2:

Popple, A. V., Smallman, H. S., and Findlay, J. M. (1997). Spatial integration region for initial vergence. *Investigative Opthalmology and Visual Science, 38, 4,* s4225.

Popple, A. V., Smallman, H. S., and Findlay, J. M. (1998). The area of spatial integration for initial horizontal disparity vergence. *Vision Research, 38, 2,* 319-326.

Chapter 3:

Popple, A. V., and Findlay, J. M. (1998). Coarse to fine cyclopean processing. *Perception, 26, 10,* s22.

Popple, A. V., and Findlay, J. M. (submitted). Coarse to fine cyclopean processing. *Perception, special issue following AVA conference on Depth Perception.*

Chapter 5:

Popple, A. V. and Findlay, J. M. (1998). Is monocular alignment computed by binocular neurones? *Investigative Opthalmology and Visual Science, 39, 4,* s2889.

Popple, A. V. and Findlay, J. M. (submitted). Nonius alignment can be used to measure vergence. *Perception, ECVP Abstracts.*

Statement of copyright

0 Introduction

How the brain sees with two eyes

0.0 Summary and objectives

The aim of this PhD. thesis is to determine how horizontal binocular disparities are spatially averaged in static cyclopean random-dot stereograms to stimulate horizontal vergence eye-movements and depth perception. Mathematically, averaging and integration are closely related. The average of a function $f(x)$ in the interval from a to b is equal to its integral from a to b, divided by the size of the interval ($b-a$); i.e. the average is proportional to the integral. Although disparity integration for stereopsis has been studied (e.g. Rogers and Graham, 1982), little is known about the way in which disparities are normally integrated to control vergence. A further objective of this research is to describe the relationship between vergence angle and perceived depth in these visual images, in an attempt to determine whether and how they interact. Eye movements such as vergence constrain our visual perceptions, and also provide a measure of the underlying visual processing (Findlay and Kapoula, 1991).

The small vergence eye-movements stimulated by the stereograms investigated in this research would be difficult to measure objectively, by tracking the positions of the two eyes. Therefore, vergence has been estimated where possible using the technique of subjective dichoptic nonius alignment. The accuracy of this technique, in comparison with objective vergence measurement, is examined in section 5.2. The psychophysical techniques used to estimate both nonius/vernier alignment (Chs. 1,2,4,5) and stereoacuity/detection thresholds (Chs. 3 and 4) were adapted from the method of constant stimuli (Falmagne, 1985).

I decided to focus on responses stimulated visually, with minimal attentional mediation. Therefore, on the whole, I measured initial vergence and perceived depth, using brief (usually about 200 ms) stimulus intervals. Previous research (e.g. Semmlow and Carpenter, 1996) has indicated that this schedule corresponds roughly to the first burst of disparity vergence, prior to the initiation of further eye-movements. Subsequent vergence is more likely to be under voluntary influence, and might also affect the viewers' perceptions and encoding of the stereograms.

In this introduction I will give a brief account of human image processing, and the geometry of binocular disparity, highlighting theoretical issues which provide the necessary background for my thesis. The literature relating to vergence and depth perception from static cyclopean random-dot stereograms will be reviewed.

This thesis sheds new light on very old problems. Berkeley (1709) was concerned with the philosophical conundrum of how we are able to see depth, when the distance of an object cannot be inferred from its projection on the retina. In his own words: "For distance being a line directed end-wise to the eye, it projects only one point in the fund of the eye, which point remains invariably the same, whether the distance be longer or shorter." In solving this apparent contradiction between the geometry of vision and visual experience, Berkeley was determined to avoid what he saw as the errors of the "mathematicians" such as Descartes (Atherton, 1990). Descartes, in his Dioptrics (in Berkeley, 1709), had described how we might compute distance from the geometry of binocular vision, a theory which will be described in the first section of this introduction (0.1). Berkeley, while not denying this underlying geometry, was determined that we are able to perceive depth only from sensible qualities such as muscle strain resulting from eye-movements, with no need for geometrical computation. Vergence, the physical orientation of the two eyes in depth, will be discussed in the second section (0.2). Berkeley held that the apparent immediacy of depth perception is, like the inevitability of assigning meaning to words heard in the mother tongue, a learnt phenomenon. Stereopsis, the perception of depth from binocular disparity, is reviewed in section 0.3. The fourth section (0.4) reviews the literature on visual direction. The main motivation behind Berkeley's 'New Theory of Vision' was, according to Atherton (1990), to dispense with the need to invoke an abstract cross-modal internal representation of external space. "If the visible figure and extension be not the same with the tangible figure and extension, we are not to infer that one and the same thing has divers extensions. The true consequence is that the objects of sight and touch are two distinct things." (Berkeley, 'New Theory of Vision', 49). In section 0.5 the neurophysiological background of binocular vision is summarised. In the sixth and final section (0.6), this notion is discussed in relation to more recent theories of vision and ideas on how sensory information from the two eyes is combined to result in what we see.

0.1 The geometry of binocular vision

Many mammals, e.g. rabbits, have laterally placed eyes, giving them panoramic vision. Such an arrangement is common in animals liable to predation. Primates, however, have frontally placed eyes with a large degree of binocular overlap. A human eye has a field of view of about 150° (Bruce and Green, 1990). As well as ensuring that little field loss would result from damage to one eye, this arrangement gives the advantages of binocular vision.

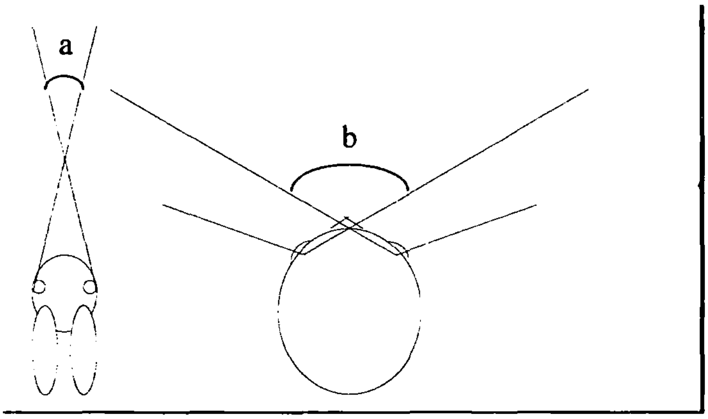


Figure 1. Schematic diagram of the degree of binocular overlap in a) a rabbit and b) a human.

Descartes (17th C; in Atherton, 1990) described how the angle an object subtends between the two eyes can be used to calculate the distance of that object from the viewer, given the distance between the eyes. He used the example of a blind man holding two sticks in front of him, to feel the distance of an object by the angle at which the sticks cross at the object. This passage translated from Descartes is quoted in an appendix to Berkeley's 'New Theory of Vision':

> "We apprehend distance, moreover, through a sort of joint activity of the eyes. For in the same way as our blind man, holding two sticks of indeterminate length, AE and CE, and knowing only the distance between his hands, A and C, together with the size of the angles ACE and CAE, can thence determine the position of E by a sort of innate geometrical knowledge shared by all men, so, when both our eyes, RST and rst, are focused on X, the length of the line Ss and the size of the angles XSs and XsS let us know the position of the point X. We can also discover that position by means of either one of our eyes alone, by changing its location. If we keep the eye fixed on X and hold it first at the point S and then immediately afterwards at the point s, that will be enough for the length of the line Ss and the size of the angles XSs and XsS to be present together in the imagination and thus to

inform us of the distance of the point X. They do so in virtue of an act of the mind which, while it may seem to be a simple judgment, nevertheless involves a kind of complicated reasoning process like that by which geometers calculate inaccessible positions from two separate given points." (From Descartes' Dioptrics VI 13).

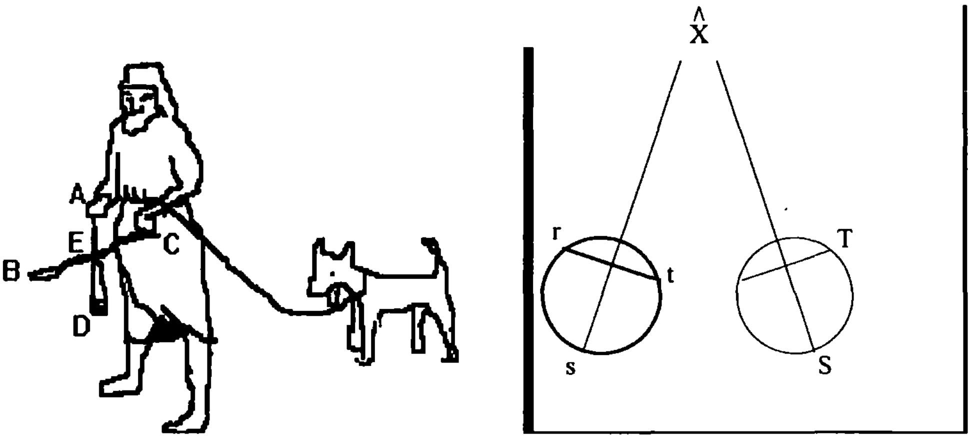


Figure 2. Descartes' analogy of a blind man with sticks

The ancient Greeks estimated the distance of astronomical objects such as the moon, using the principle of calculating distance by triangulation based on Euclidean geometry. This is the quality of physical space that underlies our ability to utilise binocular overlap in perceiving depth. Descartes did not, however, specify how we gain knowledge of the necessary visual angles. Berkeley supposed our depth perception relied on feedback from the muscles supporting the eyes, in a direct analogy to the blind man feeling the position of his hands as he directs the sticks to a target location.

Vergence eye-movements, which do exactly that, do not, however, provide good depth perception. Collewijn and Erkelens (1990), in a review of psychophysical studies addressing this question, were generally disparaging. Foley (1978) found that observers were able to correctly order separately presented small vergence targets according to depth, but tended to underestimate the absolute distance of the targets. It appears that vergence, or absolute horizontal disparity alone, is a poor depth cue. This conclusion reflects the results of Erkelens and Collewijn's (1985ab) experiments. Using a large (30 °) field of random dots, the images in the left and right eye were shifted in opposite directions following a temporal sinusoid. In this condition, vergence was imperfect but neither the vergence eye-movements nor the residual disparity led to any perception of motion in depth. (When only one eye's image was viewed monocularly, the shifting dots were seen in lateral motion).

The vergence angle is not, however, the only source of geometrical information which could be used to compute depth. It is easily apparent, by covering first one eye and then the other, that the view from the two eyes is slightly different. Euclid (4th C. BC; in Howard and Rogers, 1995) discussed the differing views of an object obtained by the two eyes. Alhazen (11th C; in Howard and Rogers, 1995) was the first to formulate rules for binocular visual direction. The precise rules that govern the apparent direction of objects under monocular and binocular viewing conditions were later delineated by Wells (1792; cited in Ono, 1991) and Hering (1879|1942; cited in Ono, 1991) (see section 0.4).

Leonardo da Vinci (16th C; in Howard and Rogers, 1995) noted that this difference between the two eyes' views gave a sensation of depth absent from monocular viewing. In other words, it could be used as a source of information about depth. The depth of an object relative to the plane of fixation, where the two eyes are pointing, can be specified by the disparity between its images in the two eyes' views. Imagine a point P some distance in front of the point of binocular fixation, O. The images of an object at O will fall on the centre of the fovea in both eyes (O_L and O_R). P_L and P_R denote the images of P at the two retinae. If the left and right retinae were overlaid, with the points O_L and O_R on corresponding locations, the points P_L and P_R would fall on slightly different locations. This distance between P_L and P_R is the horizontal retinal disparity of P. This disparity can be expressed as a visual angle, as shown in figure 3.

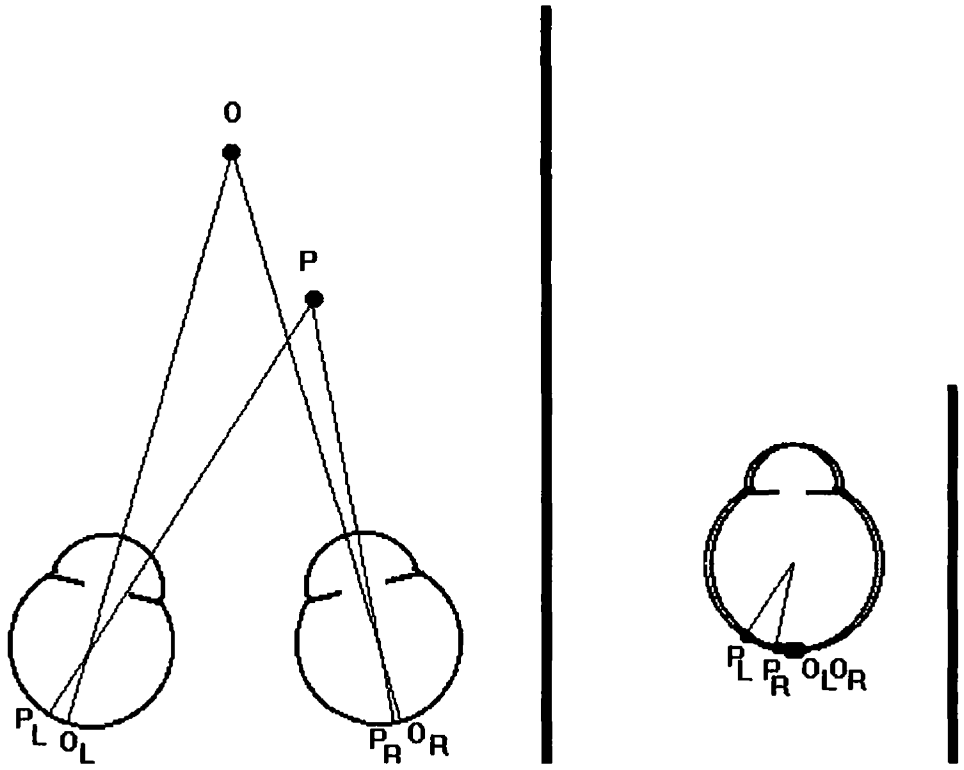


Figure 3. Retinal disparity (see text for details).

The disparity described above can be termed 'absolute' horizontal disparity (Collewijn et al., 1991). It specifies absolute depth, given the distance of the fixation plane. 'Relative' horizontal disparity is the difference in disparity between two points, neither of which necessarily lies in the plane of fixation. In the example above, imagine that O_L and O_R did not fall on corresponding points on the two retinae. By shifting the retinae until O_L and O_R were in line, the distance between P_L and P_R would then give the relative horizontal disparity of points O and P. Relative horizontal disparity remains unchanged during changes in fixation, and therefore provides information about the relative depths in the environment which is invariant under a variety of eye movements (Van Ee and Erkelens, 1996).

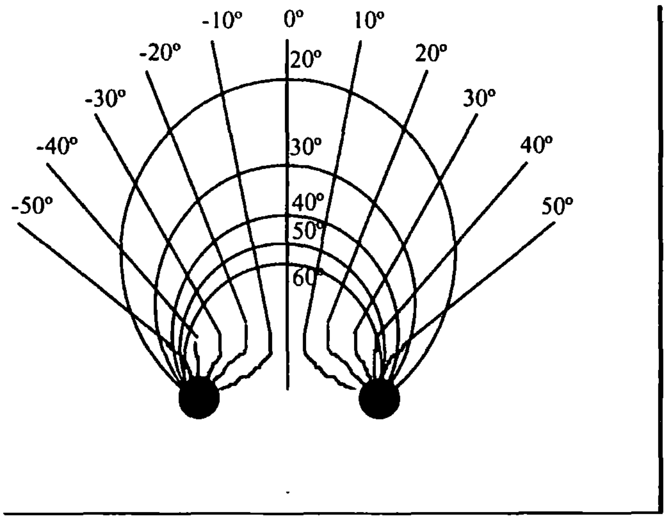


Figure 4. Vieth-Müller circles, or isophores denoting points of equal absolute horizontal disparity, and Hillebrand hyperbolae, or isotropes, marking points of equal mean binocular visual direction (after Carpenter, 1988 after Lunenberg, 1948).

All the points subtending the same horizontal disparity must geometrically lie on a circle including the fundus points of the two eyes. These Vieth-Müller circles are illustrated in figure 4. Also shown are Hillebrand hyperbolae, or lines of equal mean binocular visual direction. The circles are the orthogonal cross-section of cylinders that describe surfaces of equal horizontal disparity in three-dimensional space. In fact, however, points above or below the horizontal plane, and which are not in the vertical plane midway between the two eyes and orthogonal to their axis, will have vertical as well as horizontal binocular disparity.

Vertical disparities arise because of the different perspective of the two eyes. Imagine looking directly at a large chessboard. The right-hand edge of the chessboard will be closer to the right eye than the left-hand edge of the chessboard, and hence it will subtend a larger visual angle. Likewise, the left-hand edge will be closer to the left eye, and this will subtend a larger visual angle in the left eye as a result. These distortions in the two eyes' images of the board are shown in figure 5. Figure 5 also shows the disparity vector field, first described mathematically by Koenderink and van Doorn

(1976). The pattern of vertical disparities alone can, in principle, be used to gauge fixation distance (Helmholz, 1866; in Tyler, 1991). More recently, Mayhew and Longuet-Higgins (1982) showed that the vector disparity difference between a small number of points is sufficient to specify absolute depth differences.

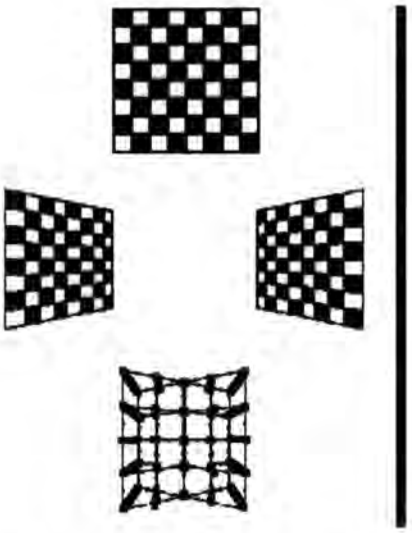


Figure 5. The pattern of vertical and horizontal disparities created by an object extending beyond the plane of the horizon. The chessboard is projected differently onto each retina. To obtain the disparity vector field, imagine overlaying the two monocular images. Heavy arrows are drawn between left and right image points arising from the same physical intersection. After van Ee (1995).

The geometrical horopter is the point or points where disparity is zero relative to the point of fixation. The form of the horopter depends on its geometrical definition (Tyler, 1991). Zero horizontal disparity occurs on the surface of a cylinder orthogonal to the Vieth-Müller circle, which is its cross-section. On the other hand, zero horizontal and vertical disparity only occur where this cylinder intersects the vertical plane midway between the two eyes, and where it intersects the horizontal plane containing the two eyes (that is along the Vieth-Müller circle). This is the case when gaze is directed toward points where vertical disparity is zero, i.e. points that lie in the plane of the horizon or in the vertical plane between the two eyes. When looking at a general point which has vertical as well as horizontal disparity, the horopter forms a single line which loops along the cylinder of equal horizontal disparity. The line deviates from the vertical midline to include the point of fixation, then goes along the Vieth-Müller circle through the eyes. It includes the reflection of the fixation point in the opposite quadrant (lower left for top right, and vice versa) and then curves back towards the vertical midline. Helmholtz (1866; in Tyler, 1991) mathematically described the shape of the horopter with oblique gaze, that is with both horizontal and vertical version. Tyler (1991) and Howard and Rogers (1995) give detailed descriptions of the horopter.

The empirical horopter, where visual stimuli appear to be at equal disparity, differs somewhat from the geometrical horopter. The main difference, as shown in figure 6, is that the empirical vertical midline tips back about the longitudinal horopter. The longitudinal horopter also differs slightly from the Vieth-Müller circle. This small difference, known as the Hering-Hillebrand deviation, can be described as the addition of a fixed angular disparity as the fixation distance is varied (Hering, 1864; Hillebrand, 1893; cited in Tyler, 1991). Ogle (1932, 1950) showed that the empirical horopter in the visual plane is always one of the conic sections (circle, ellipse, straight line, parabola or hyperbola). The form of this horopter is consistent with nasal eccentricity being larger than temporal eccentricity for each pair of corresponding points, such that this difference increases with eccentricity. In addition, when gaze is elevated from the horizontal plane, some degree of cyclovergence occurs. Cyclovergence is the difference between the torsional movements of the two eyes. Torsion is rotation about some axis between the pupil and the retina. As a result of cyclovergence, the shape of the horopter is distorted.

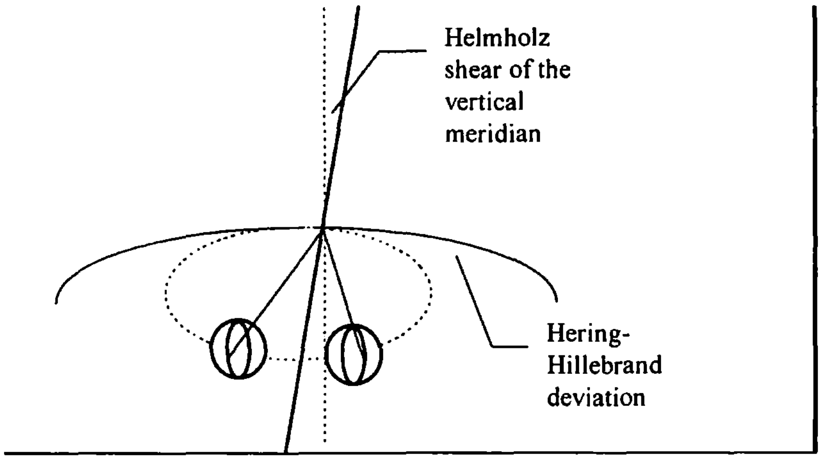


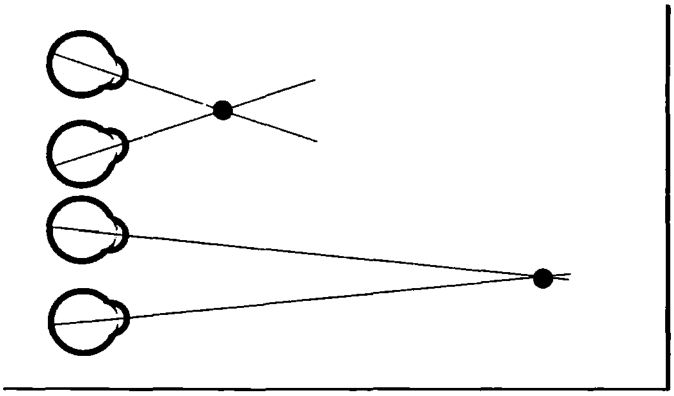
Figure 6. Deviation of empirical from geometrical horopters

The geometry of binocular vision provides a number of different sources of information about depth. However, these 'binocular' depth cues are not essential for perceiving the third dimension. This can be readily observed by occluding one eye. The three dimensionality of the visual world is essentially unchanged. Depth can also be computed from blur, given the focal distance of accommodation of the lens. Accommodation, and its interaction with vergence, will be discussed in the next section (0.2). Other, so-called 'monocular' depth cues allow us to recognise the 3D world in two dimensional images such as paintings, photographs and film. These monocular depth cues arise from the projective geometry of image formation. A number of recent theories have addressed the issue of how these monocular depth cues are combined with depth derived from stereopsis (e.g. Bulthoff and Mallot, 1988; Young et al., 1993). The remaining sections of the introduction (0.2-0.4) concentrate, however, on how the geometry of binocular vision is utilised in practice to control vergence eye-movements, stimulate depth perception, and determine visual direction.

0.2 Vergence

Vergence eye movements are movements of the two eyes in depth. The vergence angle is the angle between the lines of gaze of the two eyes. Convergence onto a nearby point makes this angle larger, divergence onto a distant point makes it smaller (see figure 7). A change in vergence is the difference between the angular deviations of the two eyes, a convergent change is positive (making the vergence angle larger) and a divergent change is negative (making the angle smaller).

Figure 7. Convergence and divergence



Although vergence could be achieved if each eye were able to target the object of vision independently, this is in practice not the case with the human visual system (for exceptions, see Enright, 1996b). Hering (1868; cited in Alpern, 1962) observed that the movements of the two eyes are normally yoked together, because corresponding muscles of each eye receive equal innervation (Alpern, 1962). Version eye-movements or saccades are conjugate, the two eyes move equally in the same direction. Hering's 'principle of equal innervation' has been taken to mean that version and vergence eye-movements are effected independently (e.g. Carpenter, 1988). Thus, any possible eye-movement can be said to have a (conjugate) versional component, which is the mean of the movements made by the two eyes, and a vergence (disjunctive) component, which is the difference between the movements made by the two eyes. Alpern (1957) demonstrated that a yoked saccade occurred when a prism was introduced in front of one eye (see figure 8). The base-out prism in front of the right eye caused the image of the target to shift laterally in that eye. Although only movement of the right eye was required to correct for this shift, both eyes began a symmetrical vergence movement. This was followed by a saccade to the left before the completion of the vergence movement. It is not clear whether the origin of such saccades is near the effector site, as would be implied by Hering's observation. The probability of such saccades depends on the stimulus layout (Ono and Nakamizo, 1978) and the saccades themselves are often asymmetrical (Enright, 1984; 1986). The weight of the evidence suggests a partial interaction between version and vergence requirements in the ongoing computation of eye-movements (for review see Ono, 1983). Neural mechanisms of saccade-vergence interaction have been proposed (Zee et al., 1992; Mays and Gamlin, 1995).

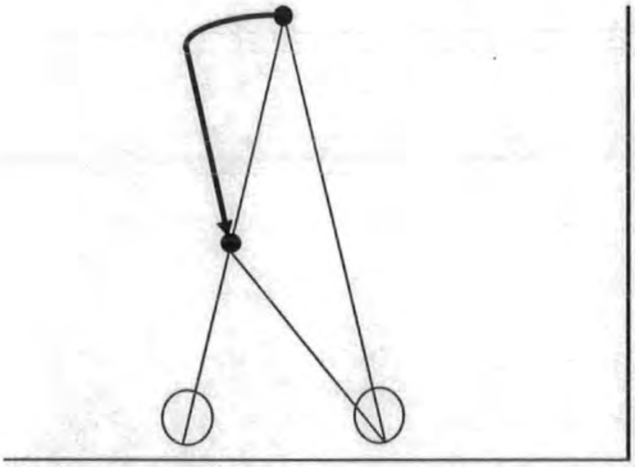


Figure 8. Alpern's (1957) demonstration of a yoked saccade. The position of an object (filled dot) has been shifted suddenly by introducing a base-out prism in front of the right eye. Both eyes saccade to the left, before a slow vergence movement is undertaken. Adapted from Alpern (1962).

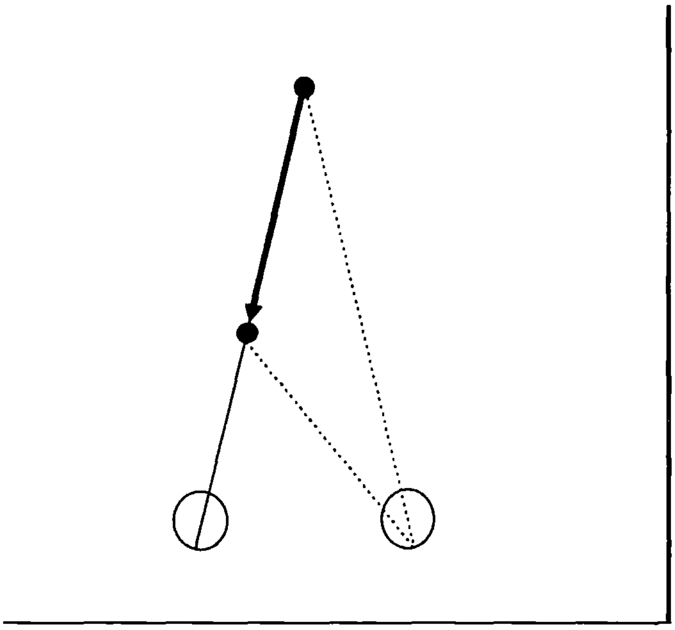


Figure 9. Alpern and Ellen's (1956) reproduction of Müller's (1826) experiment. A lens introduced in front of the left eye, causes the occluded right eye to perform a vergence movement.

The main reason for believing that versional saccades and vergence are controlled separately is that the dynamics of the two kinds of eye-movements differ. Whereas saccades are rapid, vergence (although it has a shorter latency) is slow and can take up to a second to reach completion (Carpenter, 1988).

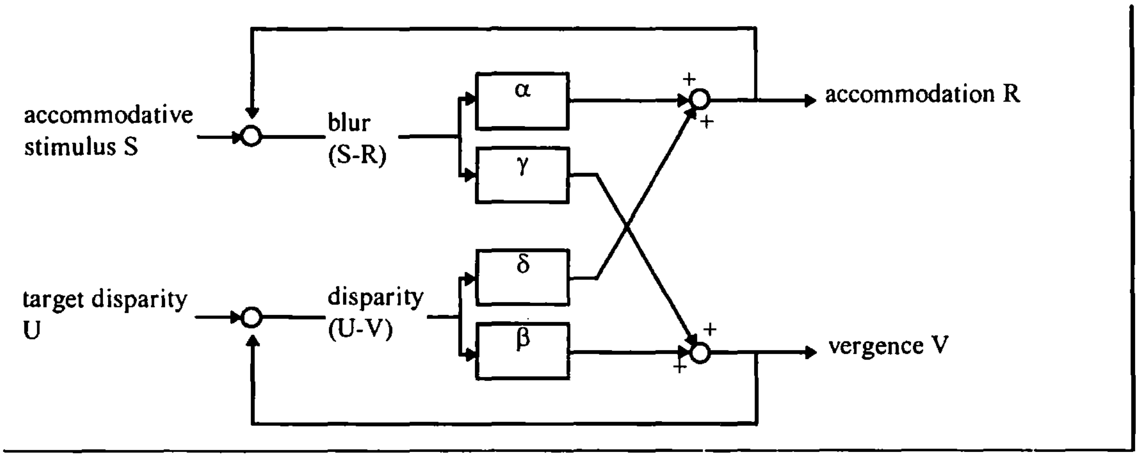
What, then, is the stimulus for vergence eye-movements? Maddox (1893; cited in Stark, 1983) attempted to quantify the component causes of the deviation of vergence angle from zero. The four components discussed by Maddox were tonic vergence, accommodative vergence, voluntary or proximal vergence, and reflex fusional vergence (Stark, 1983). Tonic vergence is the angle at which the eyes are held in the absence of visual stimulation, an angle that normally deviates in the convergent direction from parallel lines of gaze. Tonic vergence varies with adaptation to different optical apparatus (e.g. prisms) and task, but normally relaxes to an average of about 3°, equivalent to a distance of about 120 cm (Owens and Leibowitz, 1980).

The largest component of vergence was, according to Maddox, accommodative. The human eye contains a lens which serves to focus incoming light on the retina. In individuals under the age of about 55, the thickness of this lens can be varied by the ciliary muscles to vary the focal distance (Alpern, 1962). The direct stimulus for accommodation is blur. Blur and accommodative focus give an indication of viewing distance, which can be used to direct vergence eye-movements. Müller (1826|1842; cited in Alpern, 1962) showed that an accommodative stimulus alone can cause vergence change. Alpern and Ellen (1956) reproduced his experiment. With one eye occluded, a lens was introduced in front of the other eye. The target in the viewing eye would then be blurred. A convergent movement of the occluded eye was the result. (See figure 9).

Maddox termed the fourth component of vergence 'reflex' or fusional. This is what we now refer to as disparity vergence. Accepting that the eyes move together rather than independently, the difference in target location between the two eyes can be used to determine their relative motion. Absolute or retinal disparity (see figure 3) and vergence form part of a closed feedback loop, in the same way as accommodation compensates for image blur. The interaction between the two systems is summarised in figure 10 (Carpenter, 1988). For a target subtending a given angle between the two eyes, reflex vergence is a function of retinal disparity. Retinal disparity is the difference between target angle and vergence angle. The new vergence angle is fed back in the re-registration of retinal disparity. Retinal disparity cross-links with accommodation, which in turn feeds back in the registration of blur and vice versa. These two systems also interact with the pupillary reflex, forming the 'near triad' of automatic responses to visual stimuli (e.g. Semmlow and Jaegar, 1972; cited in Semmlow and Hung, 1983).

Figure 10. The interaction of vergence and accommodation systems, adapted from Carpenter (1988). Target focus (S), and visual angle (U) are compared with accommodative focus (R) and vergence angle (V).

$$R = \alpha(S - R) + \delta(U - V)$$
$$V = \beta(U - V) + \gamma(S - R)$$



The exact way in which disparity determines vergence has been studied under open loop conditions. With the image in each eye stabilised to the position of that eye, Rashbass and Westheimer (1961) found that vergence velocity is proportional to the magnitude of disparity, for disparities up to about 4°. When a step disparity is applied under open loop conditions, after a 160 ms latency vergence increases at a constant rate. This suggests the signal is being integrated with respect to time in the control of vergence, as with smooth pursuit eye-movements (Carpenter, 1988). Under normal conditions, when feedback is available, an integrator would slow down and stop the response when target vergence had been attained. The dynamics of vergence are suggestive of a linear system, however even under open loop conditions there is some evidence of predictive behaviour, for example to a temporal ramp-step combination (Rashbass and Westheimer, 1961). A step disparity of under 1° was followed by a slow ramp lasting about 1 second, concluding with a step return to the starting disparity. In this case, vergence velocity reached zero before the ramp crossed the zero disparity point. Under normal, closed-loop conditions, similar anticipatory vergence responses occur following repeated stimulation. Vergence is also markedly non-linear under normal viewing conditions; convergence is faster than divergence (Zuber and Stark, 1968). A dual-mode model of vergence was proposed (Hung et al., 1986; Semmlow et al., 1994) which distinguished between a delayed slow component driven with continuous feedback, and a fast component incorporating a disparity predictor based on sampling target velocity.

Little attention has been given to the spatial constraints operating in the control of vergence. In a natural viewing environment, objects are seen at many different depths, and we are able to shift vergence between them effectively (Erkelens and Collewijn, 1989). The way in which the disparities of two transparent planes are integrated in the control of initial, stimulus driven vergence was determined by Mallot et al. (1996), replicated in chapter 1. Further discussion of this topic is in chapter 2, which looks at the spatial parameters of disparity integration in the control of initial vergence. Vergence is thought to be directed by attention (Collewijn and Erkelens, 1990). Indeed, trained observers are able to gate between targets at different disparities voluntarily (Erkelens and Collewijn, 1995) even when these are presented parafoveally, and in competition with a stabilized foveal target (Collewijn and Erkelens, 1991). However, inexperienced subjects are likely to respond automatically to a step change in disparity, even when trying to maintain fixation on a target (Stevenson et al. 1997).

Although vergence is of little direct relevance in the perception of depth (0.1), it is used to scale the retinal image. Conversely, the apparent size of objects can determine Maddox's third component of vergence, proximal vergence (Ittleson and Ames, 1950;

McLin et al., 1988). Wheatstone (1852; cited in Howard and Rogers, 1995) first noted that increased convergence leads to a decrease in perceived size. Heinemann et al. (1959) confirmed that this effect was due to vergence, and not the accompanying accommodative and pupillary reflexes. It may be this tendency to scale images according to vergence, in the direction of size constancy, such that nearer objects of the same size appear smaller, which accounts for the difficulty in estimating depth explicitly from vergence. Here is an analogy to explain the way we use vergence to indicate depth. Patients with cerebral achromatopsia are able to segment objects from colour boundaries, without being able to see the different colours themselves (Heywood et al., 1991). In the same way, we may be able to scale objects from vergence, while not being able to perceive the vergence distance itself.

Others have demonstrated, however, that vergence and/or absolute disparity can be used to give an accurate, direct estimate of relative depth - sometimes as good as stereoacuity. Enright (1996a) found viewers were able to perform 'sequential stereopsis' to determine the relative depth of peripheral patches of sandpaper, which could not be resolved simultaneously. Frisby et al. (1997) replicated this result using filtered textures, although they found the depth thresholds obtained depended on spatial frequency, and were poorer for high-frequency images. This may have been because the high frequencies could not be resolved in the periphery, or because they were less effective in driving vergence or the underlying disparity coding.

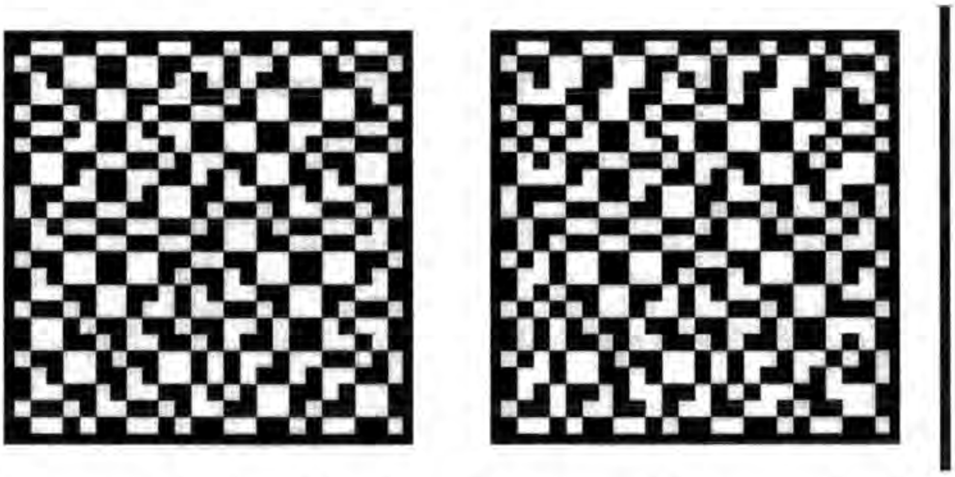
In a natural image, size constancy functions along the z-axis as well as in the fronto-parallel plane. This is why stereograms containing a fixed horizontal disparity appear to contain a greater depth interval the further away they are from the viewer. Vergence scaling results in poor size constancy along this dimension, a finding which has been attributed to our poor ability to estimate depth from vergence (Foley, 1980; Tyler, 1983). However, under certain stimulus conditions, when the target is relatively small (10°), vergence can give good disparity scaling (Bradshaw et al., 1996). Scaling, rather than direct depth perception, appears to be the primary function of distance estimates based on absolute disparity or vergence.

0.3 Stereopsis

When two 2D images depicting the view of an object or scene from the vantage points of the two eyes are observed, each by its own eye, a profound sensation of depth is experienced. Binocular stereopsis is the perception of depth from binocular disparities. The mirror stereoscope, a device which directs each half of a stereogram to the appropriate eye, was first invented by Wheatstone in 1832 (Wheatstone, 1838; in Howard and Rogers, 1995). At around the same time, Wheatstone devised the prism stereoscope which uses different optical apparatus to perform the same function. Brewster (1849; in Howard and Rogers, 1995) later publicised the prism stereoscope and initiated its commercial production. There followed a fashion for stereoscopic photographs, or stereographs. These were generated using the techniques of early photography, by a special camera which simultaneously registered images from two different vantage points, equivalent to the positions of the two eyes.

To obtain such an impression of depth, the images of each component must be matched in the two eyes' views. Early speculation about matching elements favoured considerable pre-processing of each image, such that edges or even whole, recognised objects were matched (Wheatstone, 1838). However, in 1960 Julesz invented the random-dot stereogram, a stereogram in which sharp depth contours emerged, despite the uniformity of the two monocular half-images (figure 11) (Julesz, 1960; 1971). The original random dot stereogram was generated using two identical copies of a random dot pattern. A square cut from the centre of the right copy was shifted to the left, to give the impression of a raised square when viewed binocularly. The gap created by shifting the square was filled with more random dots, to eliminate any monocular contours. Depth in such random-dot stereograms is termed cyclopean, because it can only be seen by combining the images from the two eyes. The allusion to the one-eyed giant from Odysseus' travels refers to an imaginary central eye, where the images from the two eyes are combined to give the singular visual world we perceive. For further discussion of this topic, see section 0.4. Random-dot stereograms indicated that the visual system was able to use primitive matching elements to achieve stereopsis. Binocular combination can occur early in the visual processing stream, before objects or even edges have been identified.

Figure 11. Random-dot stereogram after Julesz (1960; 1971).



Early computational models of stereopsis were mainly concerned with addressing the correspondence problem, the problem of dealing with false matches, which occurs with random dot stereograms. Two main strategies were used to address this problem. The first strategy used constraints on the possible outcome (Marr and Poggio, 1976;

Pollard, Mayhew and Frisby, 1985), implemented by cooperation between neighbouring units. This is termed a cooperative algorithm. The first of these algorithms (Marr and Poggio, 1976) was among the earliest examples of a neural network, as it employed parallel processing. Figure 12 shows the possible matches between four elements. Each point in this cross correlation matrix is represented by a single unit. The constraints used by Marr and Poggio were those of uniqueness and continuity. Uniqueness implies that each point in the right eye is matched with only one point in the left. This is achieved by inhibitory connections along the solid lines. Continuity comes from the relative smoothness of objects. Sharp disparity boundaries are present only at corners and edges, which form a small part of the image area. The algorithm was designed to solve simple random-dot stereograms such as figure 11. In this case, continuity could be equated with excitatory connections across fronto-parallel planes, as shown by the dotted lines in figure 12. The second model of this sort (Pollard, Mayhew and Frisby, 1985) was able to deal with more complex depth profiles, including natural images after edges had been extracted. In this case, the continuity or 'smoothness' constraint was expressed by a limit on the disparity gradient rather than a preference for flat planes.

Figure 12. Marr and Poggio (1976) Filled circles show stimuli and 'true' matches, unfilled circles show 'false' matches. The white arrows represent axes of interaction between units.

The second approach to the correspondence problem has been to limit false matches by constraining the input, rather than the output. Marr and Poggio (1979) used this approach, taking as a starting point Richards' (1971) suggestion concerning the existence of two 'pools' of neurons representing large and near-zero disparities. This suggestion was based on work with stereoanomalous subjects, whose impairments in stereoacuity were often restricted to one or another of these pools (Richards, 1970, 1971). In Marr and Poggio's (1979) theory, three pools (near, far and zero) are represented at a number of different spatial frequencies or filter sizes. For each spatial frequency (rate of luminance modulation) only three disparity bands are considered, making the correspondence problem almost irrelevant. The difficulty with this approach is that fine disparities can only be represented near the plane of current fixation. This gives vergence an important role in scanning the range of depths present in the visual scene, driven entirely by low spatial frequency channels as only these can register sufficiently large disparities. The role of vergence in determining stereoacuity, and the consequent influence of the spatial parameters constraining vergence on stereoacuity, are examined in chapter 3. According to Marr and Poggio, the sampled disparities (both small and large) are temporarily stored in a visual buffer termed the 2½ -D sketch.

Another way of constraining the input to binocular correlation is to combine information from the different spatial frequency channels monocularly, to obtain a rich representation of each eye's image so that false matches are unlikely. Such an approach was used by Mayhew and Frisby (1980) who filtered the stereo image halves with a broad band filter (a circularly symmetric Laplacian centre-surround operator) and located zero-crossings forming monocular contours before binocular combination.

Broad band filtering is equivalent to convolving the inputs from a number of different spatial-frequency and orientation tuned channels. In conclusion, the main difference between these two non-cooperative solutions to the correspondence problem is the order in which information is combined binocularly and from different spatial-frequency channels. Mayhew and Frisby (1980) criticised Marr and Poggio (1979) for the unrealistically limited range of disparities their model could represent. In fact, the visual system does impose a limit on the range of disparities about the horopter which can be fused. This limit, known as Panum's fusional area, will be discussed in the next section (0.4). Mayhew and Frisby's model also requires range limitation, but this is based on psychophysical estimates of Panum's area.

Cross-correlation of the images in the two eyes can be used to generate an estimate of depth without, explicitly, addressing the correspondence problem (e.g. Sperling, 1970). A cross-correlation matrix of the stimulus, such as shown in figure 12, is used to derive a cross-correlation profile (see figure 13). It is plain that such a profile will peak at the correct disparity, since matches in this plane outnumber matches in any other plane. Cross correlation can give a local estimate of depth within the window used to perform the correlation, however the extraction of depth boundaries would depend on the size and density of such windows, as with any spatial frequency account. This approach is nonetheless useful because it can provide a coarse disparity estimate, as is required for vergence, without the need for prefiltering. The detection of interocular correlation can be predicted from differences in the cross-correlation functions between correlated and uncorrelated random-dot stereograms. Cormack (Cormack et al., 1994) found that preformance in this task depended on the number of elements in the display. Cross-correlation has been used to model disparity integration for vergence (Mallot et al., 1996, see also chapter 1). The outcomes of processing image correspondence for vergence and correlation detection are compared in chapter 4.

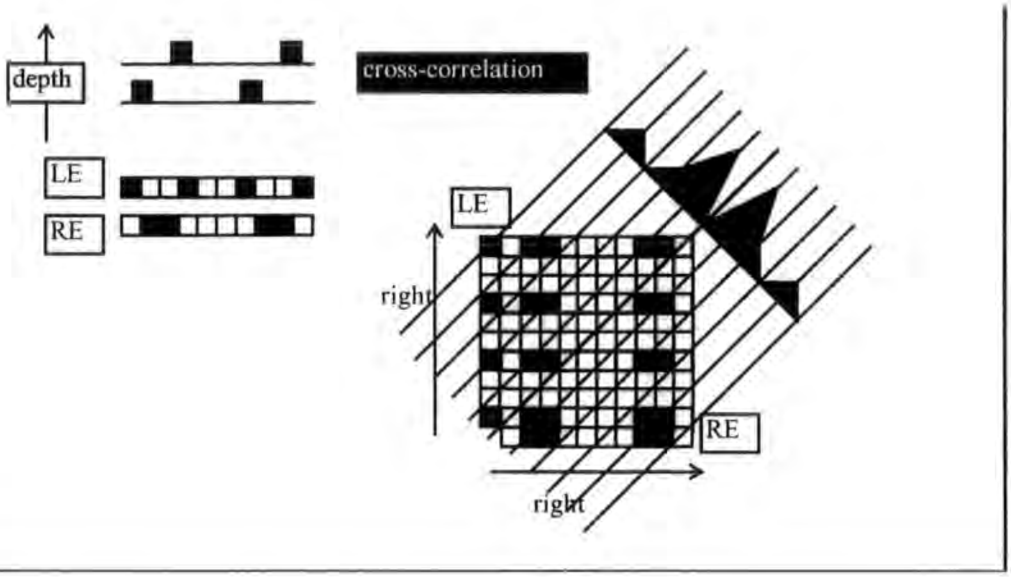


Figure 13. A cross-correlation matrix and profile.

Neurophysiological evidence is, at present, insufficient to select between these different matching strategies. This is because all the different kinds of disparity sensitivity required by the different models appear to coexist among the binocular neurones of the primary visual cortex (for review, see Howard and Rogers, 1995). Further discussion of the neurophysiological basis of binocular combination is in section 0.5. Psychophysical evidence is also ambiguous, for example recent experiments have supported both spatial-frequency based (Smallman and MacLeod, 1997) and cooperative (Glennerster and McKee, 1997) accounts. Cross-correlation also explains a number of stereoscopic phenomena (Stevenson et al., 1994). Perhaps all three strategies are used by the human visual system, depending on the stimulus and the task. (This approach to stereoalgorithms draws from the account given by Poggio and Poggio, 1984).

Is binocular depth perception learnt, as Berkeley had suggested? About 5% of the adult population lack good stereoscopic vision (Richards, 1971). Binocularity is rare in strabismics, who frequently suppress information from one eye. In other words, binocular depth perception is associated with accurate binocular fixation, and vergence. Strabismus is often the result of poor vision in one eye, for example anisometropia (a difference in focus between the two eyes), giving rise to amblyopia (impaired visual resolution in one eye). It appears that accurate binocular fixation can only develop when disparity information is available, that is when the images in the two eyes are sufficiently similar to be matched. Binocular vision is unlikely to be restored by surgical correction of the strabismus, showing instead a critical period of development. Early treatment of monocular aberrations, for example by removing a cataract in one eye, results in the development of normal binocular vision. (The aetiology of strabismus and its relation to binocularity is reviewed by Flax (1983)). Monkeys reared with an esotropia surgically induced at an early age, failed to develop normal stereoacuity after eye alignment recovered spontaneously, although their fusional vergence eye-movements were accurate (Harwerth et al., 1997). It appears that binocular depth perception is learnt from early visual experience. Like first language acquisition (Brown, 1958), it follows a maturational pattern and can only be achieved during a critical period, as cases of monocular visual deprivation indicate. Berkeley (1709) concluded his essay with the following observation: "... we cannot without great pains cleverly separate and disentangle in our thoughts the proper objects of sight from those of touch which are connected with them. ... consider how hard it is for anyone to hear the words of his native language pronounced in his ears without understanding them. Though he endeavour to disunite the meaning from the sound, it will nevertheless intrude into his thoughts ...". Berkeley's analogy between visual depth perception and language acquisition is more apt than can have been foreseen. It remains to enquire whether this reliance on learning in (at least binocular) depth perception also reflects the absence of an internal representation of abstract space, as Berkeley believed.

0.4 Visual direction

We are able to distinguish visual directions that differ by only arcseconds of visual angle, less than the width of a single photoreceptor. Such hyperacuity can be demonstrated by our ability to align vernier lines, a method which will be described more fully in the experimental chapters of this thesis. Similarly, stereoacuity enables us to make equally accurate comparisons of depth based on disparity of a few arcseconds (McKee et al., 1990).

How do we integrate the different directions of an object seen in the two eyes? The general geometry of binocular vision was described by Alhazen in the 11th Century. A set of rules governing apparent visual direction was formulated by Wells (1792). Later, Hering (1879|1942) delineated the 'Laws of visual direction' in a theoretical framework including the notion of the 'cyclopean eye'. These are summarised in figure 14.

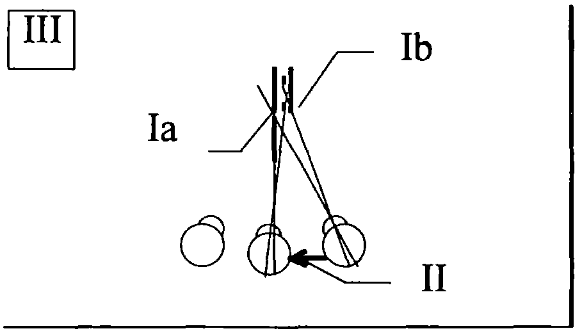
According to Ono (1991), 'information about the visual line is provided by the afferent signal from a retinal image location (local sign)', and 'information about the common axis is provided by the efferent signals sent to the eye muscles and/or by the afferent signals from the eye muscles.' Gregory (1958, 1966) proposed two subsystems to determine visual direction, a retinal-image system and an eye-head system. Ono (1974,1975; in Ono, 1991) suggested that, in line with the laws of cyclopean projection (figure 14, IV), the joint binocular local signals combine with the joint eye position signals. Panum's fusional area is the region delimiting the boundaries of V(b). In this area, which is normally broader horizontally than vertically (Panum, 1858; cited in Howard and Rogers, 1995), the disparate images of an object in the two eyes can be fused and therefore appear to have the same visual direction, which is the mean of their directions in the two eyes. Beyond this area, the images are seen in two different visual directions (diplopia). The size of this region depends on a large number of factors. These include eccentricity (Hampton and Kertesz, 1983) and spatial frequency (Schor et al., 1984), but not contrast or phase (Schor et al., 1989). Panum's area is often represented as a thickening of the empirical horopter compared with the geometrical horopter (e.g. Tyler 1983,1991). However, this is confusing as objects within Panum's area, although fused, appear at different depths. Ogle (1952) described a region of patent stereopsis, where veridical depth impressions could be obtained with diplopic stimuli. Although monocular half-images may appear at different locations and cannot be fused, they must be matched in some way to obtain a disparity value for this region. Similarly, disparities beyond Panum's area must be processed to stimulate appropriate vergence eye-movements as demonstrated in chapter 5 (5.2) and discussed in chapter 3.

Figure 14. Herings' Laws of Visual Direction.
(overleaf)

I. The laws of oculocentric direction.

a) Objects on a given visual line appear to be aligned, or superimposed. That is, they appear to have the same oculocentric direction.

b) Non-coincident retinal images give rise to a judgment of spatial separateness.



II. The law of the cyclopean eye.

All visual objects are judged as if from the cyclopean eye located midway between the two eyes.

III. The laws of monocular visual direction.

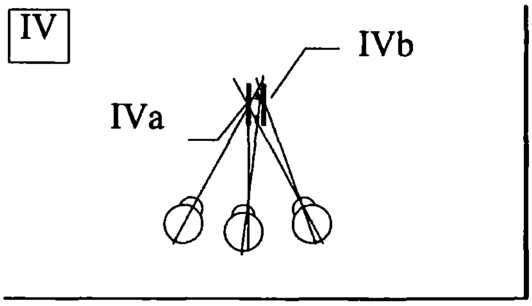
a) In monocular viewing, objects on the same visual line are judged to be in the same visual direction, which is unique to that visual line.

b) All visual lines appear to point to the same cyclopean eye.

IV. The laws of cyclopean projection.

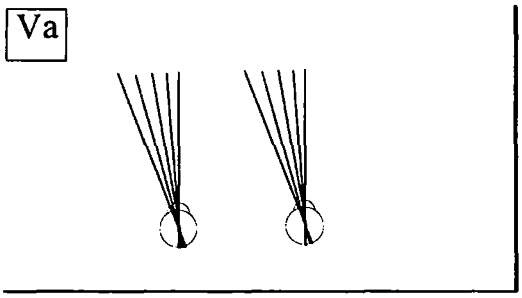
a) Objects on the visual axes of the two eyes are judged to be in the common axis (passing through the cyclopean eye and the intersection of the two visual axes).

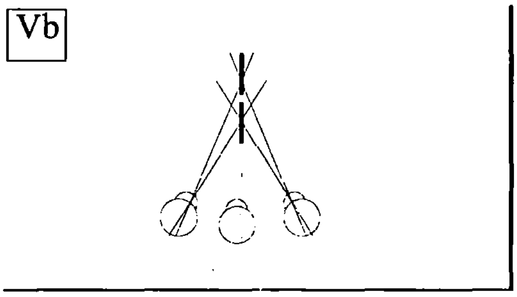
b) An unfused monocular object on a visual line is judged to deviate from the common axis with the same angle subtended by the visual axis and the visual line that contains the object.



V. The laws of binocular visual directions.

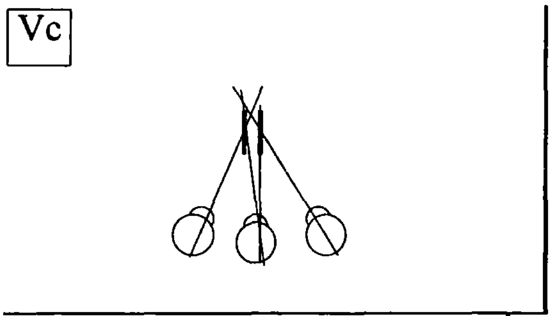
a) Every visual line in the binocular field of one eye has a corresponding visual line in the other eye, with an identical apparent visual direction.

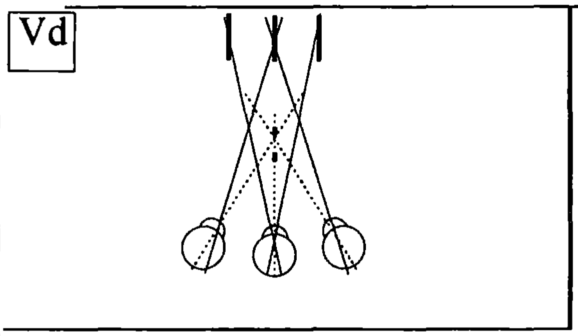




b) The visual direction of slightly disparate, fused images is the average of the visual directions of the monocular components.

c)The visual direction of rivalling images in the two eyes is that of the dominant image.





d) Widely disparate images are seen in two different visual directions.

Recently, these laws of binocular visual direction have come under some criticism. Erkelens and Van de Grind (1994) showed that the laws of cyclopean projection and mean visual direction made paradoxical predictions in the case of half-occlusions. When one object is positioned in front of another, larger object parts of the second object will be visible only monocularly, and hence follow the law of cyclopean projection. In effect, this means that two points of the partially occluded object will appear to have the same visual direction, a prediction of the laws which is not borne out in practice. Erkelens and Van de Grind showed that, instead, when a binocular and a monocular object are aligned under such conditions, alignment between the monocular object and the same-eye half image of the binocular object takes place. Subsequently, Erkelens and van Ee (1997ab) went on to show that the apparent location of a monocular line was 'captured' by a nearby binocular surround. Similarly, McKee and Harrad (1993) demonstrated that fusional suppression restricted access to monocular alignment information except in the dominant eye of stereoanomalous observers. In the field of binocular rivalry, Logothetis et al. (1996) found that, when different images were presented to the two eyes, instead of perceptions alternating between the two half-images, global constraints were taken into account, to cause alternation between grouped objects or coherent scenes containing neighbouring regions from either eye.

Another area of debate has been the visual direction assigned to stimuli differing in contrast between the two eyes. Mansfield and Legge (1996) found that gabor luminance patches fused in a visual direction weighted by their contrast, and closer to the higher-contrast half-image. This, they proposed, could be accounted for if the position of the 'cyclopean eye' shifted depending on the stimulus. As the cyclopean eye is a theoretical construct which describes the reference frame in which we see visual 'where' has been proposed (Ungerleider and Mishkin, 1982) which is associated with two relatively independent 'dorsal' and 'ventral' pathways (Baizer et al., 1991). These themselves perhaps originate from the M and P ganglion cells in the retina (Livingston and Hubel, 1988). An alternative scheme, espoused by Zeki (1993), contains specialised visual areas for colour (V4), motion (MT or V5) and dynamic form (V3). What all these theories have in common is a reliance on V1, the primary visual area, as a relay station for different properties of the visual scene.

The striate visual cortex, V1, is also significant in the integration of the images from the two eyes. The optic fibres from the two eyes partially decussate (cross over) at the optic chiasma, so the right LGN (lateral geniculate nucleus) receives its input primarily from the left visual field of both eyes, and the left LGN from the right visual field. However, the two monocular inputs are shunted to different, alternating layers of the LGN. From the LGN, these unintegrated monocular inputs are carried up to V1 where binocular combination takes place. Only the input layer of V1, layer 4b, contains monocular neurones. In visual areas beyond V1, all neurones are binocular, responding to stimulation from either eye or both. Therefore, it is in V1 that we must look for physiological evidence of disparity processing and this is the area I will concentrate on here.

Cells in V1 are specialised along a number of stimulus dimensions. These include spatial frequency, orientation, phase, ocular dominance, and disparity. The way in which such preferences are distributed among cells in V1 suggests that this stage of visual processing is more than a sorting system for subsequent stages. The organisation of orientation and phase preferences can be described by topological equations, and predicted from a self-organising neural network (Tanaka, 1995). Further, physiological evidence of tangential connections between cells of the same orientation suggests a role for V1 in image segmentation following Gestalt grouping processes such as collinearity (Gilbert and Wiesel, 1989; Hirsch and Gilbert, 1991; Schmidt et al., 1997). If similar connections were found between neurones tuned to the same disparity, this would provide support for cooperative models of disparity processing.

A receptive field is the area of visual space in which a visual neuron responds to stimulation. Receptive fields can be classical, responding with decreasing strength to stimulation further from their centre following a gaussian receptive field profile. Alternatively, they may have a centre-surround organisation where stimulation in the surround prohibits a response. Such centre-surround cells may be linear, such that an inactive balance can be achieved between static stimulation in the centre and surround areas. Non-linear cells in V1 are described as complex, since they may be particularly sensitive to moving stimuli and have specialisations along the temporal dimension. Simple cells respond equally to static stimuli. The concept of a receptive field is used for visual neurones from the retina and up to higher visual areas such as MT and MST.

In V1, receptive field profiles to oriented bar stimuli outline the spatial frequency, orientation and phase selectivity of a cell. Spatial frequency tunings from 2 to 8 c/deg have been found foveally, and lower in the periphery (DeValois et al., 1982). This arrangement could be used to perform a patch-wise fourier analysis of the retinal image, to extract useful boundary information necessary to reconstruct both objects and the spatial layout of the environment. It is also a useful starting point for any theory of visual processing as such linear filtering precedes subsequent operations.

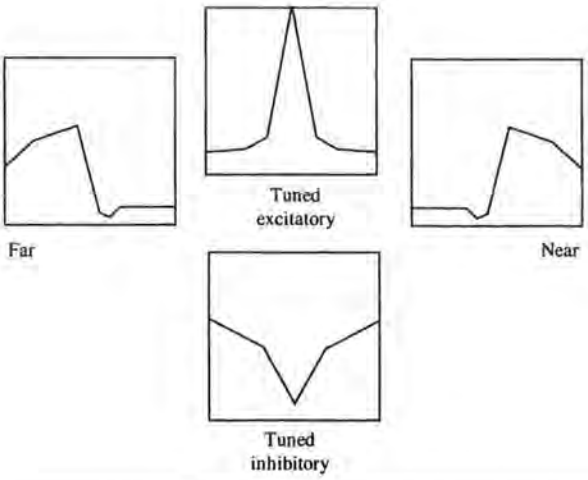


Figure 15. Disparity tuning profiles of neurones recorded by Poggio (1984). Schematic illustration, showing neural response plotted against disparity.

About two thirds of cells in V1 are selective for binocular disparity. Responses along this dimension were obtained in alert monkeys by Poggio and Fischer (1977) and Poggio and Talbot (1981). The cells were originally categorised along two dimensions - tuned or untuned, and near vs. far. Tuned excitatory cells peaked at zero disparity, whereas tuned inhibitory cells troughed at this value. Near and far cells both exhibited a broader range of disparity selectivity, responding most strongly to disparities in front of, or behind, the fixation plane respectively. The disparity tuning functions of these different cells are shown in figure 15. Later (Poggio, 1991) two further categories were added to this scheme: tuned near, and far, detectors.

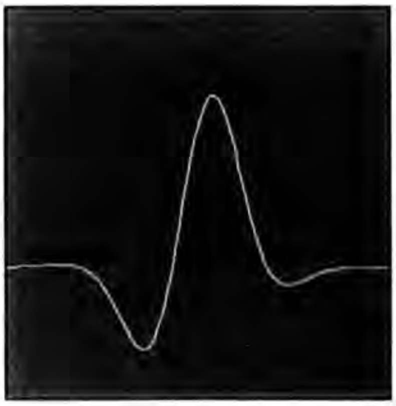


Figure 16. Left (black) and right (white) receptive fields of neurones modelled by Freeman and Ohzawa (DeAngelis, 1991). Neural activity is plotted against position. Phase shift of left and right gabor response profiles is used to obtain disparity selectivity.

One problem with the scheme proposed by Poggio was that disparity selectivity was isolated from other stimulus dimensions. Freeman and Ohzawa (1990) (see figure 16) developed a model where disparity is selected by a phase shift between the two monocular receptive field profiles of a simple cell, rather than a shift in the centre location of the monocular fields. They found physiological evidence of this (in the anaesthetized cat, DeAngelis et al. 1991, 1993; Ohzawa et al. 1996) compatible with the spatial frequency and orientation preferences of the simple cells they describe. In the original Freeman and Ohzawa model, these simple cells specialised for both disparity and location were summed in quadrature by a complex binocular disparity neuron, relatively insensitive to location (also in Ohzawa et al., 1997b). This latter part of the model appears an oversimplification in the light of further research on complex cells, which cannot easily be described in terms of phase shifts (Ohzawa et al., 1997a). This kind of model is nevertheless better suited to current understanding of V1 than the disparity profiles provided by Poggio. The two are fully compatible, as 'near' and 'far' disparity tuning curves may arise from low spatial frequency selectivity, whereas the tuned excitatory profiles could be the result of high spatial frequency selection, with cells having binocular receptive fields shifted by 180° corresponding to the tuned inhibitory neurones of Poggio et al.. These phase-shifted neurones easily fit in with Marr and Poggio's (1979) spatial frequency theory of disparity processing (DeAngelis, 1991). What might appear to be a Keplerian representation of different disparities (Lehky and Sjenowski, 1990) could actually be a sparse representation repeated at different spatial frequencies (Smallman and McLeod, 1997).

0.6 How does the brain see with two eyes?

Is there evidence for an internal representation of external, three-dimensional space? The 17th century debate between Berkeley and Descartes has, in my view, a modern equivalent in the conflict between connectionist and symbolic theories of mental processing. Symbolists, like Descartes, describe the rules or algorithms which govern perception. Although connectionists often make no claim to emulate or simulate brain systems, their models are based on the parallel links between units which, in my opinion, resemble the parallel organization of neurones in the brain. These links can be subject to learning reminiscent of Berkeley's notion of associating vergence angles with tactile distances. Marr (1982), who led the computational approach to vision, identified three levels of representation in an information-processing system such as the brain or a computer. The first level is the computational level, which describes the input-output function. The second is the algorithmic level, the transformational rules by which the input becomes the output. The third level is that of hardware implementation. Marr held that these levels are relatively independent. If this is the case, there is no need to couch theories of visual processing in neural terms (whether biological or connectionist) providing the input output function is maintained. However, more recent work on computer information processing has shown the importance of evolutionary constraints in algorithm development, even down to the level of using the material properties of silicon chips (Thompson, 1997), and indeed physicists have recently emphasized the physical nature of information itself (Zurek, 1989).

The aim of this thesis is to unravel the processing of visual information in a specific biological substrate (the human brain) continuously with other physical processes. Although our knowledge of disparity processing in the brain is limited, various organisational principles such as the existence of interconnected, parallel networks, the use of receptive fields, and spatial frequency analysis, can be used to constrain the theories proposed. An outline of the theory supported by the thesis would look very much like Marr and Poggio's (1979) stereoalgorithm; low spatial frequencies processed rapidly drive vergence (chapters 1 and 2), which brings into register the finer disparities that determine stereoacuity (chapter 3). The processing of interocular correlation can be modelled using the same pooling mechanisms as with correlated binocular stimuli (chapter 4). Finally, however, cooperative interactions between neighbouring units are invoked to explain perceived monocular visual direction in a binocular context (ch. 5). This places the integration of disparity firmly in the disparity domain, rather than the spatial domain as originally proposed by Marr and Poggio.

In chapter 1 initial vergence was measured as the proportion of random dots in two transparent disparity-defined planes was varied, in replication of Mallot et al. (1996). Vergence was estimated using the psychophysical technique of dichoptic nonius alignment, a method corroborated by simultaneous binocular eye-movement recording in chapter 5. Vergence was a function of the mean disparity of the two planes, weighted by the number of dots in each plane. This vergence response was dissociated from the perceived depth of the stimulus, which could be seen in transparency. Coarse filtering of the input could account for the findings in this broad-band spatial frequency stimulus. Mallot preferred a model filtering (or averaging) disparity peaks obtained by quantized cross-correlation. Initial coarse-filtering was based on the ideas of Marr and Poggio (1979).

Following from the local averaging found in chapter 1, chapter 2 established that even a global stereo-figure is integrated with its surround by the pooling mechanism which determines initial vergence. The integration of neighbouring disparities extended a region of about 6°. Chapter 3 related the spatial tuning properties of vergence in the disparity domain, to those of stereoacuity. Stereoacuity was better for a smaller (1-2°) target, as can be predicted from spatial-frequency tuning to disparity gratings (Rogers and Graham, 1982). When the area of a depth pedestal was increased, making initial vergence more accurate, stereoacuity to a test-patch on the pedestal also increased. This demonstrates the role of vergence in the ongoing processing of disparities. Larger pedestals, however, produced better results even for very brief (40 msec) stimuli, suggesting that vergence itself relies on the early stages of disparity processing, as does transient stereopsis. But stereoacuity improved considerably when the stimulus interval was increased to 500 msec, demonstrating the dependence of binocular vision on vergence eye-movements to implement coarse to fine disambiguation.

The perception of interocular decorrelation, and the vergence response to it, were explored in chapter 4. Interocular correlation is commonly used to explore the underlying processing of disparities (e.g. Cormack et al., 1994). Surprisingly, and in contrast with previous studies (Tyler and Julesz, 1978; Cormack et al., 1994), the spatial tuning for changes in correlation was identical to that reported for changes in disparity. Decorrelation, therefore, may be represented by the visual system as a range of disparities and processed by the same mechanisms responsible for disparity processing. Some individuals verged in response to changes in interocular correlation. This may have been due to the prevalence of 'near' or 'far' disparity pools (Richards, 1970,1971), later integrated in the determination of a target vergence disparity.

One finding which did not fit neatly with the filter model is reported in chapter 5. Monocular alignment was tested across a disparity step (5.1). The results could easily be explained by interactions between congruent disparity units, or feedback from a later disparity processing stage. However, dichoptic lines presented subsequently to the binocular stimulus reflected the vergence angle between the two eyes, and not the preceding disparity (5.2).

The motivation behind most of these studies was the empirical exploration of the spatial parameters governing disparity integration, both for vergence and for depth perception. Although each chapter follows logically from the preceding one, the selection of a particular experiment undertaken from the range of possible studies was by no means exclusive. However, by combining the findings obtained in all five experimental chapters a novel theory of disparity integration emerges in the last chapter. This final coherence of ideas sublimes the sometimes haphazard progression of the doctoral studies presented in this thesis.

Bringing together the study of vergence eye-movements and perception highlights the role of the visual system as actively exploring the environment, not simply a hierarchy of selective filters operating on a constantly refreshed sensory input. The problem of integrating the two eyes' views can be seen as one of sampling from three-dimensional surroundings, using the learned constraints of Euclidean space, rather than reconstructing a 3D nomenal world.

Chapter 1

Initial horizontal vergence is stimulated by the weighted mean disparity of two transparent planes

1.0 Abstract

Mallot et al. (1995,1996) reported that initial vergence is directed to the mean disparity of two transparent, random-dot planes, weighted by the dot density and contrast in the planes. Here, one of these findings was replicated in a sample of 10 viewers, using a slightly different method. Brief (230 msec) rectangular stereograms were flashed up on a modified Wheatstone stereoscope, instead of the shutter-goggles used by Mallot. Each was followed by a 160 msec dichoptic nonius vernier. The position of perceived alignment was determined using a forced-choice procedure followed by Probit analysis, rather than the method of adjustment Mallot had employed. As in Mallot's study, vergence (as estimated by the dichoptic nonius procedure) was found to be directed to the weighted mean disparity when the number of dots in the two planes was varied. Further, in experiment 2 the planes were shifted laterally so that they overlapped, with the result that the weight of each plane in the disparity pooling for vergence was determined by the number of dots in that plane. Dot density, which remained constant while dot number was altered, had no effect. In conclusion, disparities of all the dots in the stimulus area were pooled in the calculation of initial vergence, regardless of their spatial contiguity.

1.1 Selective replication of Mallot et al. (1995,1996)

1.1.1 Introduction

Mallot, Roll and Arndt (1995,1996) showed that disparity evoked vergence is directed towards an average of two depth planes. When viewers were shown a stereogram consisting of 100 dots distributed between two planes, ±18 arcmin from fixation, their vergence response took an intermediate value, weighted by the proportion of dots in each plane. Thus, when all the dots were at a crossed disparity of 18 arcmin, vergence was directed approximately to this plane, depending on individual biases in fixation disparity and asymmetry of responses to crossed and uncrossed disparities. When 50% of the dots were in each plane, vergence remained in the fixation plane, subject to the same individual constraints.

Mallot et al. used a psychophysical method of estimating vergence, based on the alignment of a dichoptic nonius vernier. One vertical line is presented to each eye, so that they lie one above the other. Following HeringDs Laws of monocular visual direction and cyclopean projection (see section 0.4), the two lines will appear vertically aligned when their horizontal visual angles in the two eyes are identical. Therefore, the angle of vergence between the eyes is given by the angular lateral displacement between the two lines added to the vergence angle of the plane on which they are shown. If a fixation stimulus on the same plane is used, the angular displacement between the two eyes can be termed a fixation disparity. Conventionally, a positive fixation disparity is convergent (increase in vergence angle) and a negative fixation disparity is divergent (decrease in vergence angle).

The internal validity of this method was demonstrated by Mallot et al., who recorded vergence as a function of disparity and found responses consistent with data obtained by other methods. Subsequently, a close correspondence with simultaneous eye-movement recordings was observed (5.2). In fact the differences between the two (subjective and objective) methods of vergence estimation, within the range of disparities investigated in this chapter and chapters 2 and 3, proved to be insignificant.

At the time of conducting this experiment, only a single dual-Purkinje tracker was available, so this could not be used to measure vergence directly. Although the purchase of a second tracker had been planned since before the author began her doctoral studies, this only arrived during her final year of study. When delivered, the second instrument proved unreliable which led to protracted negotiations with the suppliers. Binocular eye-tracking systems available were a head-mounted Skalar Iris tracker, and a coil system. These were less accurate than the dual-Purkinje trackers, which was eventually used to validate the nonius procedure employed here (chapter 5).

The nonius stimulus was displayed, following Mallot, directly after the offset of a brief stimulus, which was preceded by a fixation interval. One major drawback of the nonius method is that, with a single presentation time in the sequence of each trial, it cannot be used to investigate the dynamics of the vergence response. Instead, it is used to sample the vergence position of the eyes at a given time. The position of the eyes may reflect a vergence eye-movement made in response to the target. Alternatively, if trials are blocked in sequences of similar or identical disparity stimuli the nonius estimate might reflect an adapted position of the fixation disparity, ie. one which is maintained even during the fixation interval. A further possibility is that the motor response to a stimulus, presented repeatedly in each block, becomes conditioned and is repeated on subsequent trials without further visual processing.

Mallot et al. conducted trials in blocks of 30-40 adjustments. Each trial consisted of a 750 ms fixation, followed by a 230 ms stimulus and finally a 160 ms vernier interval (see figure 1.1). After each trial, the viewer adjusted the relative positions of the central nonius lines until, after a number of adjustments, they appeared in line. For each condition, 20 such blocks were averaged to obtain a reliable estimate of the vergence response in that condition and of the error associated with it.

It was necessary to repeat Mallot□s experiment with a larger number of subjects, as his six observers showed considerable individual variation, so much so that an additional term had to be included in the mathematical model proposed to account for one subject□s data. A further objective was to find out whether the results could be replicated on a Wheatstone stereoscope. In Mallot□s experiment, stimuli were displayed by temporally interlacing left and right half images in synchrony with liquid crystal shutter goggles, at a frequency of 60 Hz. The risk of crosstalk between the two eyes is a major drawback of this apparatus. The psychophysical technique used in the present study differed from Mallot, the method of constant stimuli was employed in place of a staircase procedure. If his results proved reliable, this would be a way of accessing the automatic disparity processing which determines the initial vergence output.

1.1.2 Method

a Subjects

Ten observers aged between 20 and 30 participated in this experiment. With the exception of the author, the participants were naive of the theory involved. All had normal or corrected to normal vision.

b Apparatus and stimuli

(All dimensions are given in Height x Width). Stimuli were generated on a Cambridge Research Systems VSG graphics board driven by a Logix Systems Processor PC. They were displayed on two EIZO Flex-6500 21ŋ flat monitors viewed through a modified Wheatstone stereoscope. Head movements were minimised using a chin and forehead rest. The stereoscope consisted of two 5 x 7 cm front silvered mirrors mounted at a right angle. The total viewing distance was approximately 90 cm (10 cm to the mirrors and a further 80 cm to each screen). The layout of the apparatus is illustrated in figure 2.1. The monitors were set at a resolution of 768 x 1024 pixels on a screen 25.5 x 34 cm, subtending a total visual angle of about 16 x 21 degrees, 1.25 arcmins per pixel.

The two monitors were positioned at an identical distance from the mirrors, and perpendicular (in both horizontal and vertical planes) to the line between the centres of the two monitors. This line passed, at its midpoint, through the apex of the mirrors, at 45 degrees to each mirror in the horizontal plane, and perpendicular to the mirrors in the vertical plane. Rough positioning was achieved using rulers, set-squares, and spirit levels. Finer adjustments were made by monocularly aligning first the outside of the monitors, then a grid presented on both screens, in the two mirrors. Aligning the outside of the monitors ensures that the accommodative distance is equal to the vergence distance when stimuli are positioned in the same location on both monitors. Aligning the grid displayed on both screens ensures that stimuli can be accurately positioned to coincide on the two monitors. As basic optics ensure that the angle of incidence equals the angle of reflection, and the distance of the image equals the distance of the object from the mirror, providing both monitors are the same distance from the mirrors, and properly aligned, and the mirrors accurately set at 90 degrees, then the virtual images of the two monitors will fall on the same physical, spatial location. The monocular adjustment makes use of the exquisite human vernier hyperacuity, up to seconds of arc (McKee et al., 1990). At a distance of 1 m, 20 arcsec acuity is sufficient to detect an alignment error of 0.1 mm. Such a tiny displacement is much harder to measure with a ruler! No steps were taken to ensure that vergence was appropriate for the viewing distance of the stereoscopic image. Indeed, later results suggest that, for most subjects, it was not, as most subjects were found to have a fixation disparity of several arcmins (see 2.1.3).

The fixation stimulus was a pixel-wide 25 x 25 arcmin square. Target stimuli were given a fixed, arbitrary disparity of ±12.5 arcmin. As this was the author' s first experiment using the specified apparatus, a disparity was chosen at which the author could perceive transparency and thereby ensure the desired stimulus was indeed shown. Additionally, 12.5 arcmin was, it was hoped, sufficient to drive vergence. Stimuli were generated by placing 100 2.5 x 2.5 arcmin non-adjacent random dots over a 5° x 3.75° region, then giving a certain proportion of the dots an uncrossed disparity of 12.5 arcmin, and the remainder a crossed disparity of 12.5 arcmin. A nonius vernier stimulus was generated by displaying three 160 x 2.5 arcmin vertical bars 25 arcmin apart on one monitor and a single bar above them on the other monitor, at one of 21 possible horizontal locations corresponding to vergence angles separated by 2.5 arcmin.

c Design

Elicited vergence was measured, with 100 random dots distributed between two planes. The distributions tested were as follows: 0:100, 20:80, 40:60, 60:40, 80:20 and 100:0 (number of dots in far plane: number of dots in near plane).

d Procedure

Eighty-eight trials were presented in a block for each dot distribution. Trials were blocked in this way following Mallot (1985), however see the introduction to this chapter for a discussion of the disadvantages of blocking. See also experiment 2.3, where trials were randomised.

Every trial consisted of a 1 second fixation while the random dots were calculated off-line followed by a 230 ms stimulus that was, in turn, replaced immediately by a 160 ms dichoptic nonius vernier target. The nonius stimulus consisted of a single vertical bar in one eye above three vertical bars in the other eye. Subjects were instructed to attend to the fixation target, and following the briefly flashed stereogram to respond to the nonius vernier target with a key-press left or right, depending on whether the top line was left or right of the central lower line. The vernier target was presented randomly eight times at each of 11 chosen angles separated by 2.5 arcmins. The nonius half images were allocated at random to the left and right screens. Vergence angle was determined by finding the nonius displacement angle which produced 50 % far responses using probit analysis (Finney, 1947).

The rationale behind this method of vergence estimation is as follows: if the top line (seen in the right eye) appears to align with the central lower line (seen in the left eye) the physical displacement between the two lines indicates the degree of vergence from the plane of the screen (see figure 1.1). This follows from Hering's laws of visual direction, cited in the introduction (0.4). However, since the use of nonius lines has been brought into question recently, in chapter 5 this method was validated as applies to our experiments (5.2).

For the present chapter, the nonius method supplied an estimate of where, on average, the two eyes were looking during a given interval in time following the stimulus onset and offset. The estimate obtained by this method can be presumed to constitute an accurate indicator of the size of initial vergence. The only assumption made about the dynamics of vergence, was that during the fixation interval, vergence returned to its original value. This assumption seems reasonable, as viewers were instructed to keep the long (1 second) fixation stimulus fixated, and then attend to the nonius task. They were not instructed to attend the brief (230 msec) 'target' random-dot stimulus. Subjects were asked after each session whether the fixation stimulus had appeared diplopic, and if they replied positively the session was repeated. The possibility remains, however, that, as trials were blocked in sessions of equal stimuli, the viewers, in some cases, developed a fixation disparity appropriate to the stimulus, rather than respond to each trial. In retrospect, the procedure may have been improved by monitoring pre-stimulus vergence using an additional nonius interval. No assumptions were made concerning the relation between this subjective measure of vergence and oculomotor vergence, which can be recorded objectively by tracking the horizontal position of the two eyes. However, for ease of discussion the estimate obtained will henceforth be referred to as vergence.

Dichoptic nonius vernier acuity can be in the order of seconds of arc (McKee and Levi, 1987). Rather large (2.5 arcmin) vernier steps were used in the present experiment. Nevertheless, these provided sufficient sampling to produce acceptable psychometric functions in the sample of untrained observers. The stability of the results obtained by this method can be seen from the example illustrated in figure 1.1.

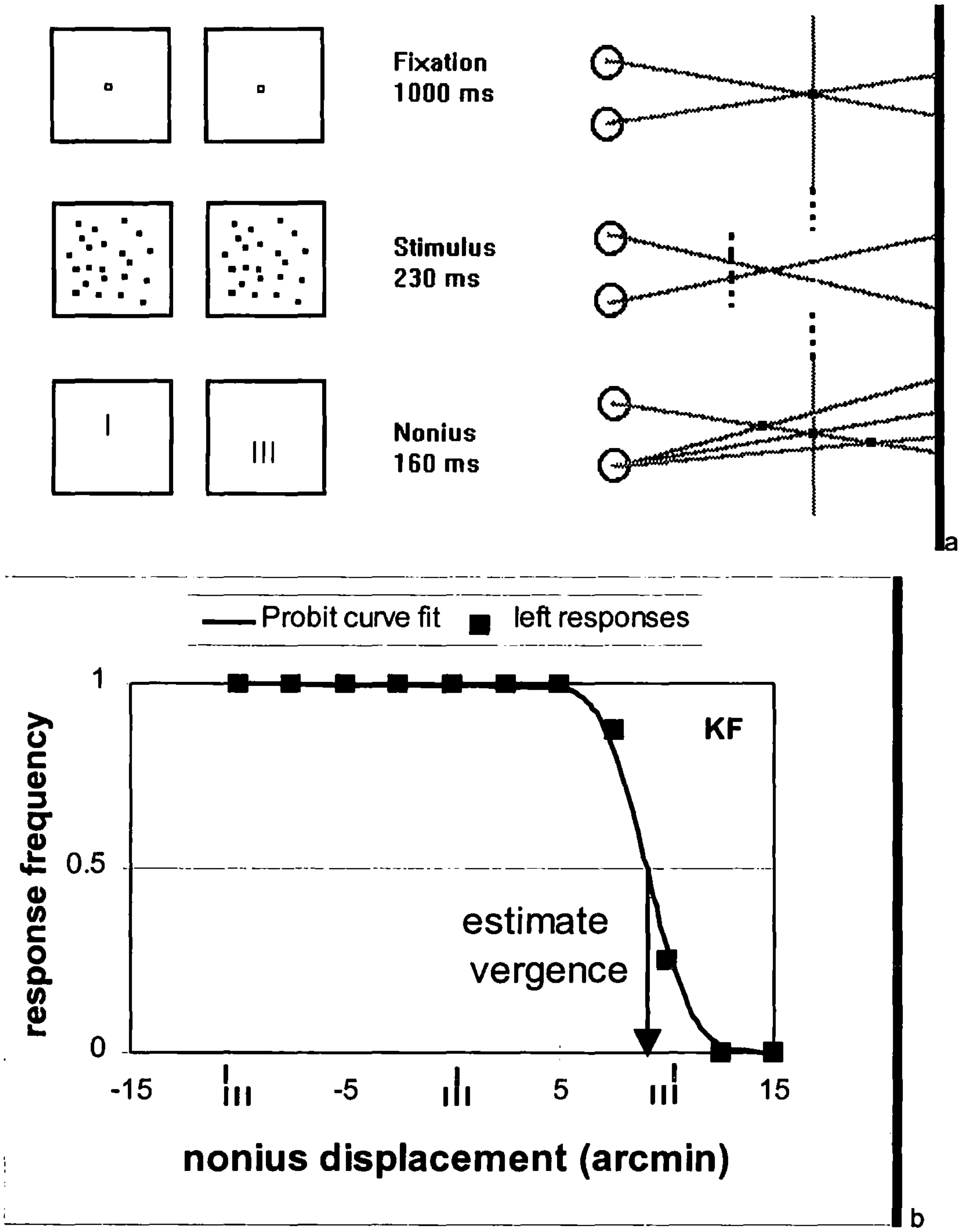
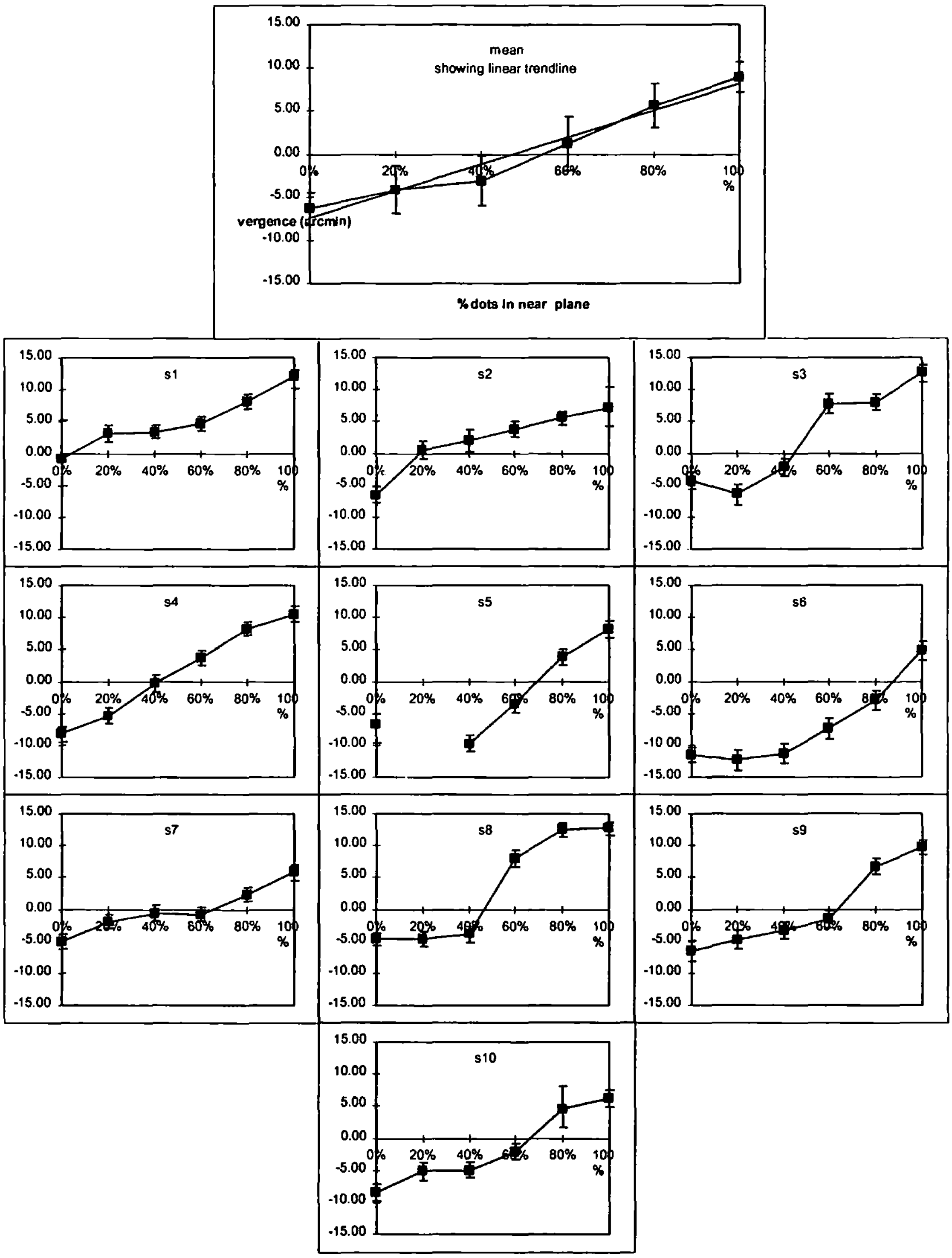


Figure 1.1. a) Procedure on each trial. A 1000 ms fixation consisting of an outlined white square on a black screen was followed by a 230 ms stimulus which preceded a dichoptic nonius vernier (160 ms). The viewer's task was to decide whether the top line was left or right of the central lower line, the vergence angles implied by alignment with each of the three lines are shown. b) Example of 'top left' response frequency. The physical displacement of the top nonius line is shown. The displacement at 50% left is a measure of vergence angle. (In the experiment, the allocation of top and bottom nonius half-images to the left and right monitors was randomised, so the left-or-right distinction was translated into far-or-near). (after Mallot et al, 1996)

1.1.3 Results

Figure 1.2. Results of experiment 1.1. All ten subjects verged in proportion to the distribution of dots between the two planes. Error bars show 95% confidence intervals.



All ten viewers responded to the disparity stimulus, although they were instructed to attend only the subsequent nonius task. The effect of dot distribution was found statistically significant in a repeated measures ANOVA taking data from the 9 subjects for whom results were available for all seven conditions ($F(5,40)=35.3$, $p<0.001$). A linear trendline fitted to the means from the entire sample closely approximated model predictions of a linear function from -10 to +10 arcmin. This linear trend was also significant ($F(1,8)=56.0$, $p<0.001$).

In seven subjects, the increase in vergence as dots were moved from the far to the near plane was monotonic. The remaining three (s3, s5 and s6) had an anomalous response only when all the dots were in the far plane. Data from one subject (s5) were too noisy to be fitted with a psychometric function in the condition where 20% of dots were in the near plane, although she repeated this condition several times. Additionally, s5 was the only observer who reported a difficulty in maintaining the fixation stimulus in single vision. In every case where diplopia was reported, she repeated the session until single vision was achieved. Due to the missing data point, her data have been removed from subsequent analyses. Inspection of this subject□s data nonetheless reveals a clear linear trend.

Looking at the individual plots, error bars are generally consistent within subjects and across conditions, indicating that when the distribution of dots was intermediate between the planes, viewers were indeed adopting an intermediate vergence position rather than alternating between the two stimulus depths. In addition, the function of vergence with dot distribution is noticeably stepped only for s3 and s8, with the remainder of the subjects showing a linear increase in vergence as more dots are moved to the near plane.

1.1.4 Discussion

As in Mallot's (1996) paper, the results confirm that initial vergence follows the weighted mean disparity of the two depth planes presented. However, like Mallot's results, the results presented here must be treated with caution. There is a potential artefact in measuring vergence using the nonius procedure. A number of recent studies have warned of the pitfalls of monitoring vergence state using dichoptic alignment (eg. Shimono et al., 1998). The issue of the relation between nonius and objective vergence estimates is addressed in chapter 5. Additionally, as in Mallot's procedure, the nonius here was used to sample vergence only during a particular interval. Therefore, no conclusions can be drawn as to whether the results represent a final state of vergence in the different conditions, or merely different dynamics of the response depending on the stimulus variables. Further, as trials were blocked according to condition, it is possible that some viewers developed a fixation disparity appropriate to that condition and did not, in fact, change their vergence state in response to the stimulus on each trial. To prevent this eventuality, viewers were strictly directed to attend the fixation stimulus and report any diplopia at the end of each session (see 1.1.2 Procedure). Nevertheless, adaptation to the conditions of each session through fixation disparity remains feasible. Subsequent experiments reported in this thesis provide similar results with blocked and randomised trials (2.1 and 2.3), but this is insufficient to preclude blocking effects in the present experiment. With hindsight, firmer conclusions might have been drawn had a randomised trial sequence been used.

In this current experiment, the two planes were separated only by 25', a separation at which transparency can be perceived in the stimulus (Anderson, 1992). The contrast between the unimodal distribution of vergence responses (each response intermediate between the two planes, not sometimes to the near plane and sometimes to the far plane), and the bimodal depth percept (both planes clearly visible, not an average between the two) supports the independence of the two on a global as well as a local scale. In confirmation of this conjecture, both the author and one other subject reported observing transparency in the stimuli. However, on debriefing, the remaining participants reported no such observation. As the participants were instructed only to fuse the fixation target, and attend to the nonius task, their lack of subtle observations regarding the unattended disparity stimulus is not surprising. It would be interesting to repeat the experiment either to test transparency alone, or, using a dual task, simultaneously estimate vergence. By varying the disparity, the upper and lower limits of both disparity averaging (for vergence) and perceived transparency could be compared, in order to establish their independence.

This conclusion could not have been drawn from Mallot's study, because the large disparity between the two planes would be unlikely to give a percept of transparency. Locally, a disparity difference of 36', can give rise to a depth step. However, on a more global scale such a large separation results in diplopia, and the perception is of 'lacy depth' (Akerstrom and Todd, 1988).

Mallot suggested that vergence is determined by pooling the responses of a large number of disparity tuned units over the stimulus area, using a population code for disparity. This model also accounts for his finding that, when the contrast of the dots rather than the number of dots in the two planes was varied, the intermediate vergence response was similarly weighted by contrast.

Mallot discussed a cross-correlation theory of disparity determination. The weight given to contrast in his mathematical model was consistent with this idea. He suggested that visual system performs a cross-correlation between the two eyes, and the vergence response is weighted by the strength of the two peak disparities in the cross-correlation profile. This is similar to saying that the cross-correlation profile is filtered to obtain a unimodal distribution. If so, it would be simpler to propose that vergence responds to the peak of the cross-correlation profile taken through coarse input filters which would blur the distinction between the two disparities (illustrated in figure 1.3). In fact, Arndt et al. (1995) showed, in a related paper, that the predictions from the two models (cross correlation and population coding) would be identical in the case of low frequency stimuli. One of the goals of Mallot's experiment was to distinguish between the two, however unless the random-dot stereogram were filtered at a high spatial frequency, it is not possible to exclude the possibility that a low frequency filter precedes cross-correlation. As another test of this hypothesis, Mallot's experiment could be repeated with increasing disparity between the two planes, to find out whether a bimodal response can be obtained, as would be predicted from a coarse pre-filter.

The foregoing formulation of disparity processing for vergence removes the need to solve the correspondence problem. False matches, if they occur, will have a mean disparity value of zero and an even distribution across the disparity range, and therefore will not affect the output of an averaging mechanism.

Such an averaging mechanism may be unique to transparent planes, or operate also for adjacent stimuli. There are two possible ways in which the distribution of dots between two planes could contribute to signal strength. The mean dot density for each disparity plane might add to signal strength, as with contrast, or the overall level of activity at that disparity might simply be pooled over the entire area. In other words, the increased weighting may be due to increased dot density, or dot number. In the next experiment the aim was to distinguish between these different options.

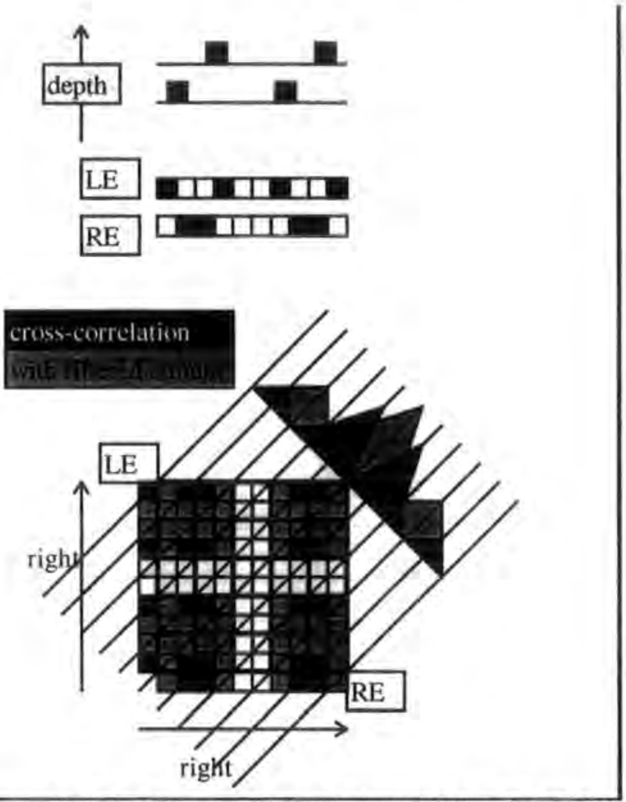


Figure 1.3. The cross-correlation between left and right half-images of a 2-disparity stereogram, showing a bimodal distribution (black) which is transformed into a unimodal distribution (grey) by filtering the image through 2-unit wide filters.

1.2 Disparity integration of overlapping planes

1.2.1 Introduction

In 1.1 Mallot's (1995, 1997) finding, that initial vergence responds to the mean disparity of two transparent planes, weighted by the number of dots in each, was replicated. In that experiment, the *density* of dots in each plane covaried with the *number* of dots so either may have determined its weight in the disparity averaging process. In addition, the averaging may be unique to transparent planes, or operate also in the case of adjacent disparity regions. Experiment 1.1 was repeated with the two disparity planes shifted one in relation to the other. The planes were shifted diagonally to obtain two overlapping rectangles. Further, the region of overlap either remained transparent (as in 1.1) or contained dots only from one or other of the two planes. If averaging occurs when the overlap region is not transparent, it might depend on the number of dots in each plane, or their density, which remained constant when one plane was occluded.

1.2.2 Method

a Design

As in 1.1, the proportion of dots in the near and far planes was varied. However, this time only four conditions were presented; 20:80, 40:60, or vice versa. In addition, the region of overlap was either 'transparent', in which case all dots were visible, or occluded. There were two occlusion conditions. The occlusion cue was either 'consistent' with an interpretation based on the assumption of two rectangular fields, or ⊐inconsistent□ with this interpretation. In effect the 'inconsistent' condition consisted of an L-shaped region in front of a neighbouring rectangular area. (As a result of occlusion, the actual distribution of dots between the two planes differed from their relative density, which remained constant in each of the three conditions).

b Apparatus and stimuli

The apparatus and stimuli were as in experiment 1.1, except that the two planes were shifted by 2° horizontally and 2° vertically one in relation to the other.

c Subjects

Only the author, AP, and one naive subject (MB) participated in this experiment. Both had uncorrected, normal vision.

d Procedure

The procedure used was identical to 1.1, but this time with separate blocks for the different occlusion conditions.

1.2.3 Results

The results from the two viewers were very different. MB diverged, but only when the dot density in the far plane was large. AP responded poorly to the uncrossed disparity, and maintained a large crossed fixation disparity throughout the experiment. Nevertheless, when AP's results were replotted as a function of actual dot distribution, a clear linear trend emerged. A similar linear trend can be found in MB's data, taking only the 80:20 and 20:80 dot distributions.

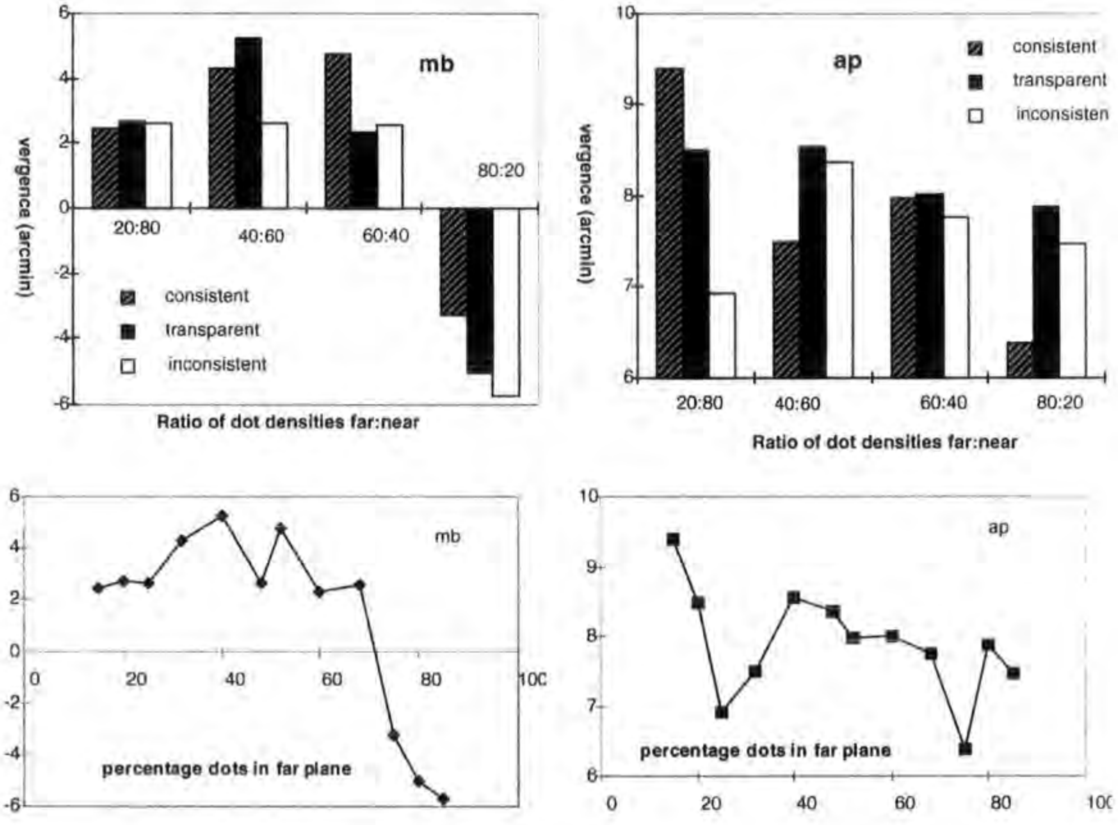


Figure 1.4. Results of experiment 1.2 - initial vergence is plotted by condition and also as a function of the dot distribution. The plots look similar, as the proportion of dots in the far plane is always slightly smaller in the 'consistent' occlusion condition (where dots in the far plane were deleted) than in the 'transparent' condition, and larger in the 'inconsistent' occlusion condition (where some of the dots in the near plane have been deleted). Note the different scales used for the two subjects, to best accommodate their results.

1.2.4 Discussion

This experiment is included in the thesis merely to illustrate the authorÕs motivation in proceeding to the experiments of the next chapter. In addition to the drawbacks of the procedure (already discussed in 1.1.4), the data from the two subjects are very different. No conclusions can be drawn in this section.

The results are, however, consistent with the spatial averaging of disparities for initial vergence, although in the case of MB a sigmoid operator must be added to the averaging function, as Mallot (1995,1996) was forced to do for one of his subjects. This idea of spatial averaging, which emerged from this experiment led to the experiments of the next chapter, which deal with the size of the region over which such averaging takes place.

Chapter 2

The area of spatial averaging for initial horizontal disparity vergence

2.0 Abstract

In experiment 2.1, we investigated over what central area disparity in a random dot stereogram is averaged to stimulate an initial vergence response. Vergence was measured subjectively, with a forced choice dichoptic nonius vernier task following a brief (230 ms) stimulus presentation. Stimuli were random-dot stereograms showing a central circular disc of 12.5 arcmins crossed retinal disparity in front of, and occluding, a same-density fixation plane surround. (The author also completed the experiment in a 12.5 arcmin uncrossed disparity condition, and as there were no major differences between the results in the two conditions, the volunteer subjects only carried out the crossed disparity condition). The size of the disc was varied. All ten observers responded to the brief stimulus. For six out of the seven observers shown discs smaller than 3° diameter, surround and target disparities were averaged together in this condition. In nine out of ten observers, the response saturated with larger discs. Initial horizontal vergence responds automatically to a cyclopean target presented in the centre of gaze by pooling disparities within a limited area.

Further to this study, we conducted a series of experiments to determine the constraints on this pooling mechanism. Experiment 2.2 was concerned with the effect of varying dot size, which was found to be neglibible. However, the results of 2.1 were replicated in 2.2 using an improved method with randomised trials, and more data was gathered on the rising portion of the vergence-size function. In experiment 2.3, we

increased the stimulus disparity to 19 arcmin, and used a larger display to allow us to rule out any effects of peripheral stimulation on the asymptotic portion of the curve.

2.1 The pooling region for initial vergence

2.1.1 Introduction

Because the two eyes are laterally displaced, the image of an object off the plane of fixation falls on slightly different, or disparate, points on the left and right retinas. The binocular disparities that arise from such situations can be used by the visual system, not only to reconstruct the relative depths of objects, but also to make astonishingly precise discriminations (Berry, 1948; Westheimer and McKee, 1978). We are best able to discriminate the binocular disparities of objects in, or near, the fixation plane (Ogle, 1953; Blakemore, 1970; Smallman and MacLeod, 1997). Vergence eye movements serve to bring objects into the plane of fixation partly to bring this exquisite stereo sensitivity to bear. In this chapter we investigate what information is processed from the monocular retinal images to mediate the successful programming of vergence eye movements. We show that the information for initial horizontal disparity vergence for foveal targets is spatially integrated over a surprisingly large region (up to 6°).

Vergence eye-movements are stimulated by changes in the disparity of the entire visual field (Erkelens and Collewijn, 1985b). However, in a real visual environment, surfaces at many depths are present. Humans are very flexible in their ability to redirect their vergence to different objects in cluttered displays (Erkelens and Collewijn, 1989). Some type of attentional control must presumably mediate the programming of such eye movements (Erkelens and Collewijn, 1991). Although Erkelens and Collewijn (1991) showed that the vergence system can be subject to attentional control, it seems likely that the system often operates efficiently without the need of such influence.

Given automatic processing, one possibility is that disparities are initially pooled over a large but discrete region. Vergence would respond as well to disparity in this region as it responds to disparity over the entire visual field, without the need for prior target selection. If disparities were averaged over an extended, but limited, central region of the visual field, no active attentional selection of a target need take place. In a natural environment we are more likely to require optimal vergence for a particular target than for the entire visual field. A large but discrete region of spatial averaging would also account for the differences between vergence and stereopsis, because initial vergence would be essentially blind to the fine grain relative local disparities that determine stereoacuity (Harris et al., 1997). Appropriate vergence without attention could help pick up salient objects rapidly. Vergence to disparities in the centre of gaze might be considered part of the orienting reflex.

We tried to estimate the proposed spatial pooling region in the absence of monocular spatial integration. We accomplished this through the use of cyclopean presentation. Attention was directed to the task of aligning a nonius vernier, to provide a subjective measure of initial vergence to the briefly presented stereogram. The size of the central cyclopean target depicted in the stereogram was varied. If disparities are averaged over a limited area, the target size at which the vergence response asymptotes should give an indication of the size of this area.

2.1.2 Method

Mallot et al. (1996) described a subjective method for measuring elicited vergence following the brief presentation of a stereogram, and this was the method used in the previous chapter (1.1.2). We adapted this method for the purposes of the present study. We estimated vergence when the size of a central occluding disc in front of a fixation plane surround was varied. The stimuli were cyclopean random dot stereograms (Julesz, 1971).

a Subjects

Eight female and two male volunteer subjects participated in this experiment as observers. All had normal or corrected to normal vision, and were aged between 20 and 40. All except two (HSS and AVP) were naive as to the purpose of the study.

b Apparatus and Stimuli

The apparatus was described in 1.1.2b. Stimuli were generated on a Cambridge Research Systems VSG graphics board driven by a Logix Systems Processor PC. They were displayed on two EIZO Flex-6500 21" flat monitors viewed through a modified Wheatstone stereoscope. Head movements were minimised using a chin and forehead rest. The stereoscope consisted of two 5 x 7 cm front silvered mirrors mounted at a right angle. The total viewing distance was approximately 90 cm (10 cm to the mirrors and a further 80 cm to each screen). The monitors were set at a resolution of 768 x 1024 pixels on a screen 25.5 x 34 cm, subtending a total visual angle of about 16 x 21 degrees, 1.25 arcmins per pixel. The fixation stimulus was a pixel-wide 25 x 25 arcmin square. Target stimuli were generated by placing 2000 2.5 x 2.5 arcmin non-adjacent random dots over the screen, then shifting a circular region on one of the two monitors and later filling the remaining crescent region to create a cyclopean Julesz stereogram of a central disc in front of a fixation plane surround. A nonius vernier stimulus was generated by displaying three 160 x 2.5 arcmin vertical bars 25 arcmin apart on one monitor and a single bar above them on the other monitor, at one of 21 possible horizontal locations corresponding to vergence angles separated by 2.5 arcmin. (All dimensions are given in Height x Width).

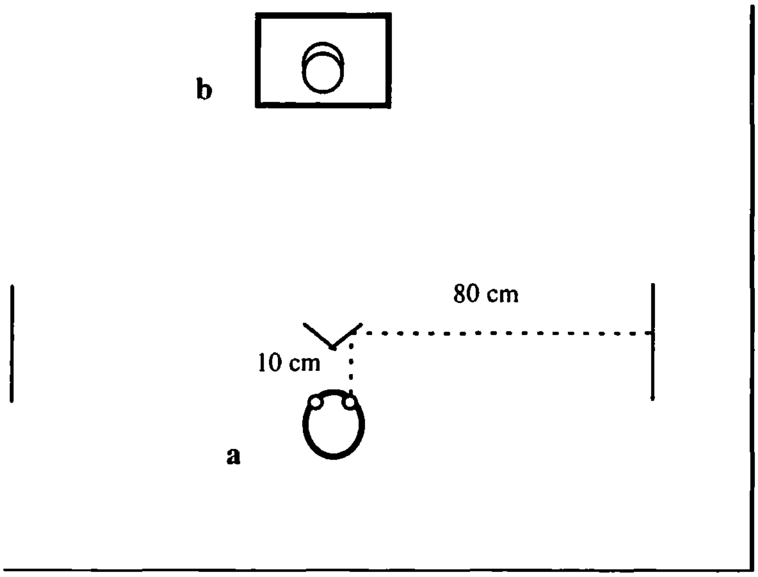


Figure 2.1. a) The setup of our mirror stereoscope. Mirrors were positioned at right angles, 10 cm from the viewer's face. The monitors were a further 80 cm away on either side, facing each other and parallel. The stereogram was seen at a distance of 90 cm in front of the viewer, disparity causing the central disc to protrude as illustrated. b) The stimulus profile depicted - a central, protruding disc of variable size but constant (12.5 arcmin) disparity to the surround.

c Design

The size of a central disc, protruding from a fixation plane surround in a Julesz random dot stereogram, was varied up to 16 degrees of visual angle. The disparity of the disc was fixed at 12.5 arcmin crossed disparity, except for the author, who also completed a 12.5 arcmin uncrossed disparity condition. The value of 12.5 arcmin was selected arbitrarily (but see 1.1.2). The (unpaid) volunteer subjects completed only the crossed disparity condition, as data from the author (AP) revealed no striking differences between the crossed and uncrossed conditions (see figure 2.2). There was a no-target (background-only) control, to provide a baseline for fixation disparity. This is shown in the results as a zero-diameter disc condition. Elicited vergence was measured subjectively following a brief stimulus interval.

d Procedure

The procedure was as delineated in 1.1.2d. Eighty-eight trials were presented in a block for each central circle size. As discussed in 1.1.2d, the blocking of the trials in this way introduces the potential for adaptation to the stimulus by way of a fixation disparity. As in the previous experiments, the observers were instructed to direct their fixation towards the attended fixation target. Nevertheless, as trials were blocked in this way, the results may reflect such adaptation rather than vergence change. The vergence change in each trial was not directly measured, only inferred from the estimated vergence state at the end of each trial. In experiment 2.3, the results of 2.1 are, however, closely replicated using a randomised trial sequence.

Every trial consisted of a 1 second fixation while the random dots were calculated off-line followed by a 230 ms stimulus that was, in turn, replaced immediately by a 160 ms dichoptic nonius vernier target. The nonius stimulus consisted of a single vertical bar in one eye above three vertical bars in the other eye. Subjects were instructed to attend to the fixation target, and following the briefly flashed stereogram to respond to the nonius vernier target with a key-press left or right, depending on whether the top line was left or right of the central lower line. The vernier target was presented randomly eight times at each of 11 chosen angles separated by 2.5 arcmins. The nonius half images were allocated at random to the left and right screens. Vergence angle was determined by finding that nonius displacement angle which produced 50 % far responses using probit analysis (Finney, 1947).

The rationale behind this method of vergence estimation is as follows: if the top line (seen in the right eye) appears to align with the central lower line (seen in the left eye) the physical displacement between the two lines indicates the degree of vergence from the plane of the screen (see figure 1.1).

The nonius method supplied an estimate of where the two eyes were looking at a given point in time following the stimulus onset and offset. The estimate obtained by this method can be presumed to constitute an accurate indicator of the size of initial vergence. No assumptions were made about the dynamics of vergence, and whether the response was sustained or transient. Because of the blocking procedure (like trials were grouped together) two kinds of adaptation to the stimulus may have occured. The vergence response may have been conditioned, and repeated without further analysis of the stimulus on each given trial. The fixation disparity may have become adapted to the stimulus, and not returned to baseline on subsequent fixation intervals. Further discussion of these drawbacks is in the introduction to Chapter 1. However, similar results were obtained later in this chapter (experiment 2.3) where the order of trials was randomised and not blocked.

No assumptions were made concerning the relation between this subjective measure of vergence and oculomotor vergence, which can be recorded objectively by tracking the horizontal position of the two eyes. However, for ease of discussion the estimate obtained will henceforth be referred to as vergence. This proposal was later corroborated by an empirical comparison between objectively measured vergence, and the nonius procedure used in this study (experiment 5.2).

Although dichoptic nonius vernier acuity can be in the order of seconds of arc (McKee and Levi, 1987) and we used rather large vernier steps (2.5 arcmin) nevertheless these provided sufficient sampling to produce acceptable psychometric functions. One possible explanation for the poor vernier acuity we obtained is the use of naive subjects rather than trained psychophysical observers. Another reason is that any vergence eye-movement made in response to the stimulus was probably still ongoing during the nonius interval. (This is confirmed by the final experiment of the thesis, 5.2, where eye-tracker recordings during the nonius interval showed vergence change). Although when asked, no subjects reported that the nonius lines were seen in apparent motion, any residual vergence may have interfered with the acuity of alignment. Nevertheless, the stability of the results obtained by this method can be seen from the example illustrated in figure 1.2b.

2.1.3 Results

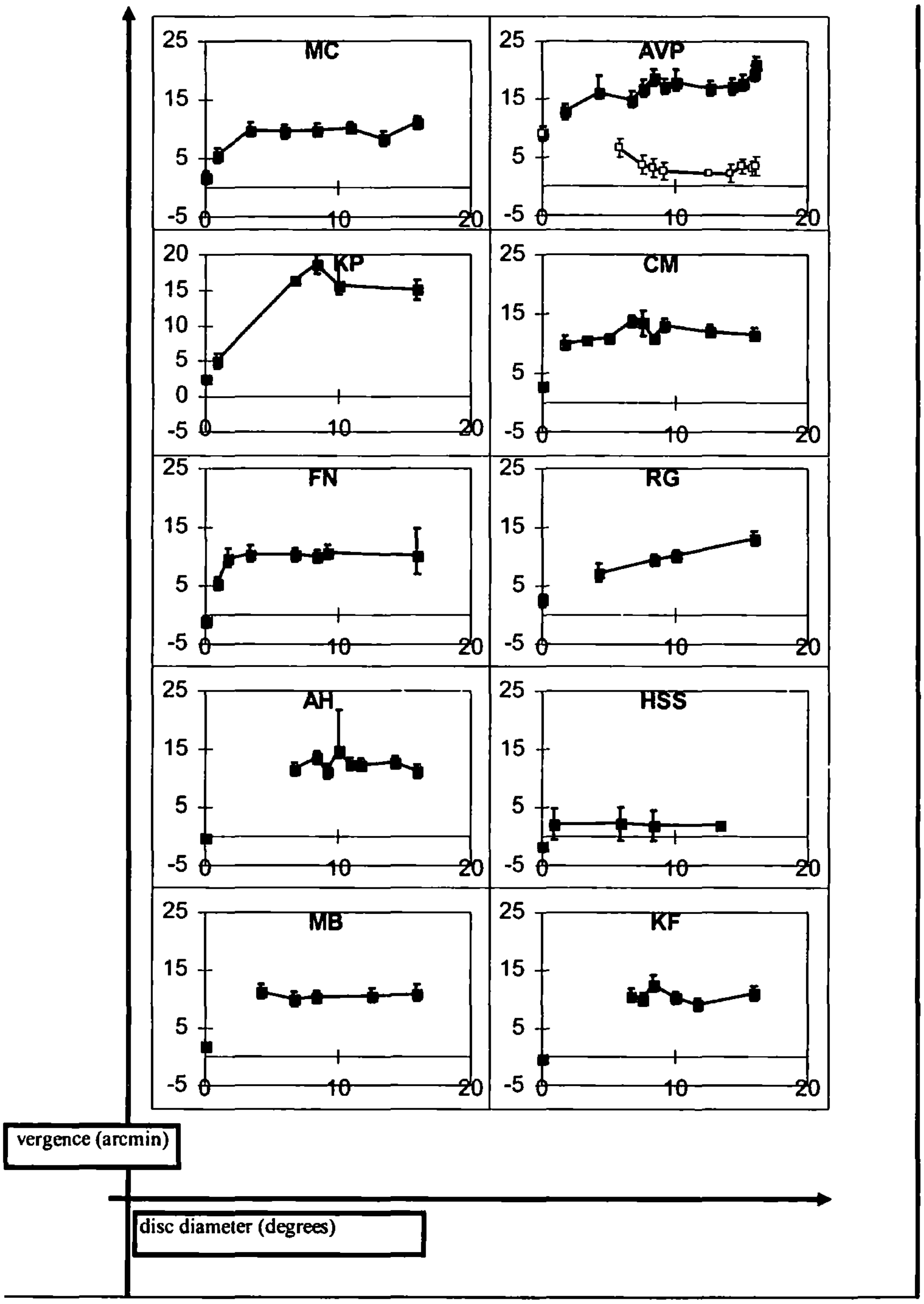


Figure 2.2. Elicited vergence with increasing target size for all ten subjects. Error bars show 95% confidence intervals. The upper six graphs provide evidence for disparity averaging. All except RG show saturation beyond a target size of about 3-5 degrees. The no-disc fixation-disparity control is plotted on x=0.

Although subjects were not instructed to attend to a particular region of the random-dot stimulus, all ten subjects clearly verged to the disc that was presented in it. Inspection of figure 2.2 shows that this response asymptoted near the target disparity for discs larger than about 3-5 degrees, with the exception of RG. Disparity averaging is evident in the data of six subjects (MC, AP, KP, CM, FN, and RG) who were tested on targets smaller than 4 degrees. The vergence response to these small discs is intermediate between the response in the baseline (no disc) condition, and the response to a large (16 degree) disc. This suggests that the disparities of the disc and its surround might be averaged. The region over which vergence reaches an asymptote potentially reveals the spatial limits of this averaging, as the response no longer increases when more positive disparities are added outside this region. HSS responded equally to small and large targets, and this response was not significantly different from zero but slightly higher than his no-target control vergence.

Vergence to small targets (<5°) was intermediate between the cyclopean target and surround disparities. We wanted to determine whether this may have been due to subjects verging to the target on some trials, but to the surround on other trials. Had this been the case, the variability of responses to the dichoptic nonius would be increased in comparison with that for the no-target condition (plotted as target size=0°), or for larger disc sizes, where presumable vergence is directed accurately to the target. However, confidence intervals on the vergence estimate for small targets (<5°) appear no larger than those for the other cases. This indicates that disparity averaging rather than target selection is responsible for the depth-averaged response we found.

To specify the suggested averaging of disparities, the following two mathematical models were fitted to the data. a) A cortical magnification model, in which the lesser weight given to peripheral disparities is accounted for by their diminished cortical representation, and b) a gaussian model which pools disparities within a central integration area (for averaging). The no-target condition was included in the data for the model fits, to provide an estimate of fixation disparity (as this was found, on average, to be non-zero). However, all the models were forced to go through this point, so the no-target condition was effectively only used to estimate a constant term in the equations.

a) Foveal magnification means that the centre of the visual field is neurally over-represented or informationally enlarged when compared with the periphery. Cortical magnification as a function of retinal eccentricity is conventionally modelled by the following formula:

$$M_c = M_f \Big/ \left(1 + {}^{E}\!/_{E_2}\right) \quad (1.1)$$

M_c is the cortical magnification factor, E is the retinal eccentricity, M_f is a foveal magnification constant and E_2 is the half-width or the angle at which magnification is halved (Wilson et al., 1990). This function was used to derive a model which integrates disparities over the entire visual field (equation 1.2).

$$\alpha_{(r)} = \alpha_0 + C_m\left[\ln\left(1 + r/E_2\right) + \frac{1}{1 + r/E_2} - 1\right] \quad (1.2)$$

Where is the vergence response as a function of r, the radius of the target disc. α_0 is a constant of integration that allows for fixation disparity, and C_m is a magnification constant.

This function was fitted to the data as a non-linear regression using SPSS for Windows (Release 6, SPSS Inc.). The value of α_0 was estimated from the no-target control condition data, which was therefore included in the analysis. E_2 was estimated visually from the data at 5 arcmin. The only comparable value from previous research is

E_2 for stereoacuity, estimated at about 20 arcmin (McKee et al., 1990). The estimated values of C_m and α_0 from the regression were 3.54 and 2.5 arcmin respectively. This model accounted for 49.9% of the variance in the pooled data. However, because cortical magnification falls exponentially with eccentricity, the model could never asymptote, merely slope more gently as target size increases. Thus, although the 'knee-point' in the data can be modelled by this kind of function, the saturation obtained prior to full-field stimulation cannot.

b) Alternatively, initial vergence might integrate disparities over a limited central area. To model this possibility, a cumulative normal centering on zero target size was fitted to the pooled data.

$$\alpha_{(r)} = \alpha_0 + C_m/2 + C_m \int_0^r e^{-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2} dx \qquad (1.3)$$

As before, α_0 is a constant added to allow for fixation disparity, and C_m is a magnification constant. σ is the standard deviation of the distribution. Again, data from the no-target control were used to estimate the fixation disparity, α_0 which was valued at 1.68 arcmin. C_m was estimated by regression at 21 arcmin. σ was estimated at 0.84 degrees. From sigma, the radius at which 95% of the response is accounted for can be estimated to be 1.39, which is equivalent to a disc diameter of about 2.8 degrees. (At 5 degrees, 99.9% of vergence is accounted for).

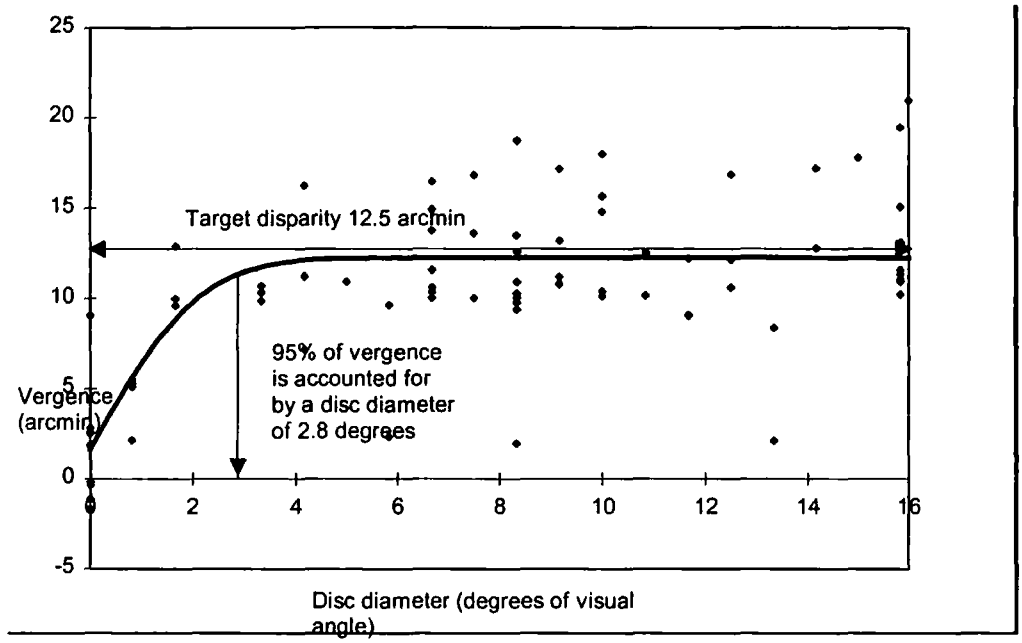


Figure 2.3. Vergence pooled across all ten subjects. The model is a cumulative normal predicted from a Gaussian pooling region, and accounts for more than half the variance across subjects.

This model accounted for 50.4% of the variance in the pooled data. Figure 2.3 shows that model vergence saturated at a vergence of 12.23 arcmin (to 2 d. p.), close to target vergence (12.5 arcmin). This would be the mean of individual responses, showing that (regardless of fixation disparity) the disc, when sufficiently large, was verged to, on average, with considerable accuracy, even after only a brief (230 msec) interval.

Looking at the asymptotic portion of the curves, there appears to be a slight fall in vergence as disc diameter increased from 8-12 degrees for five subjects with more than two relevant data points (MC, AH, AVP, CM and KF in figure 2.2). Three more subjects (KP, FN and HSS) had some fall in vergence beyond 8°. Only one subject (RG) had vergence continuing to rise with target size for targets larger than 8°.

However, the main finding remains that initial vergence increased with the increasing size of a cyclopean target, but disparate regions more eccentric than the central 3-5 degrees of gaze provided a negligible contribution to initial vergence. Due to the sparsity of data points inside this area, and the small number of subjects who were tested on these disc sizes, the estimate obtained must be treated as a first guess to be qualified by later results. To anticipate, the value of sigma is closely replicated in 2.3, where more appropriate data was gathered.

2.1.4 Discussion

The results show that the automatic initial vergence response to a brief stimulus pools disparities over a limited central region. Disparities within this area are averaged despite the presence of a smaller central cyclopean target, ie. there is no evidence of trial-by-trial target selection favouring either the disc or the surround. Instead, vergence can be modelled by taking a gaussian weighted mean of the two disparities within the pooling region. The weighting function fitted to the data had a standard deviation of about a degree, such that 95% of vergence is accounted for by a target 3 degrees wide. It should be noted that, in the present experiment, there was insufficient data on the rising portion of the function to be confident of these estimates. These values are replicated in experiment 2.3, where more care was taken to test within the appropriate stimulus domain and trial order was randomised.

This finding has recently been replicated also for dynamic vergence stimuli (Fang et al., 1998). A drawback of both that study and the present experiment is that it was always the central figure, rather than the surround, which was given a non-zero disparity. To control for a saturation of vergence velocity (despite the influence of eccentric disparities) the result must be repeated also for a disparity annulus. In chapter 4, the effect of uncorrelated dots (modelled as multiple disparities) will be shown negligible at an eccentricity of 4 degrees. Thus, for this purpose of the thesis, the necessary control is provided.

Others have already demonstrated the ability to make fast shifts in vergence between small, attended targets in different regions of the display (Erkelens and Collewijn, 1991; Collewijn et al., 1995). These shifts occurred following instructions to attend voluntarily to specific foveal or parafoveal targets. However, to reiterate, in our procedure no attention to the disparity stimulus was required. Instead, subjects were instructed only to attend the prior fixation and complete the subsequent dichoptic nonius alignment task.

Simply weighting disparities across the visual field by cortical magnification to derive a signal to drive vergence fails to account for the saturation of the data at 5°-10°. It is well know that sensitivity to higher spatial frequencies also declines in the periphery (c.f. Wilson et al., 1990). The saturation of vergence with increasing central target size could be accounted for by this fact, because of the greater energy of the small random dots in our stimuli at high spatial frequencies. Further experiments with low-pass filtered stimuli would be needed to test this possibility. Nevertheless, on the basis of present findings, a limited pooling region as exemplified by the Gaussian model offers the best description of the data.

What, if any, physiological mechanism might account for the proposed averaging of neighbouring disparities over a central region? Short-latency disparity vergence has been linked directly to the properties of tuned disparity-sensitive cells. Busettini et al. (1996) showed that short-latency vergence breaks down for disparities larger than about 1°, the limit placed on the disparity sensitivity of tuned cells (Poggio and Talbot, 1981). Indeed, Poggio and Talbot (1981) outlined a crude way in which the output of 'near' detectors might be subtracted from 'far' detectors to determine the direction of vergence, while the tuned receptors are pooled to calculate the amplitude of eye-movement required. However, our averaging region is far larger than the pooling region envisaged by Poggio and Talbot in 1981, which was presumably based on receptive fields found in early visual areas.

There is evidence for large integration regions in the control of other oculomotor functions. Analogies have often been drawn between motion parallax and binocular parallax (e.g. Rogers and Graham, 1982). As well as leading to a sensation of depth, motion parallax and binocular parallax also stimulate oculomotor responses. Optokinetic following, which is thought to stabilise gaze during motion, like vergence, responds well to full field stimulation. However, Miles et al. (1986) found that the relative motion of a 20-40 degree central disc in a random dot stimulus produced more gain in monkeys' short-latency ocular following than absolute full-field motion. They argue that maximal response to a whole moving field would be less ecologically beneficial in the animal's attempts to stabilise the retinal image during motion than tracking a target of limited size. The same argument can be applied to disparity vergence. An extended central disparity region would offer the best target to correct for vergence error following saccades, by allowing observers to focus rapidly on the object to which they have turned their gaze.

The analogy between initial disparity vergence and short-latency ocular tracking is given empirical support by the finding that both responses are enhanced following a saccade (Busettini et al., 1996; Kawano and Miles, 1986). The tenuous but suggestive downturn in vergence for larger disc sizes (>8°) could be indicative of cyclopean lateral inhibition. That is, the data hints at a centre-surround receptive field model for vergence as has been successful for motion tracking (Tanaka et al., 1986).

Cyclopean receptive fields have been proposed to account for lateral inhibition in depth perception from stereopsis (Anstis et al., 1978; Lunn and Morgan, 1995). The size of receptive fields processing disparity gratings can be estimated at 1°-2° of visual angle, from the modulation transfer function to such stimuli (Rogers and Graham, 1982; Tyler, 1983). It may be noted that this estimate, half of a single sine-wave cycle at 0.25-0.5 cycles per degree, is also confirmed in chapter 3. There may be differences between vergence and stereopsis at the input end (size of receptive fields) as well as at the output end (stereoacuity as opposed to limited oculomotor muscle control). The spatial frequency functions for vergence and depth sensitivity will be compared directly in chapter 3.

Is the function of initial vergence simply to bring the images from the left and right eyes into approximate correspondence, to facilitate the more refined mechanisms of stereopsis? Absolute disparity and vergence are believed to contribute little to depth perception (Erkelens and Collewijn, 1985a), which is based almost exclusively on relative disparity (Collewijn and Erkelens, 1990). Relative disparity is fixed regardless of vergence angle (Collewijn et al., 1991; Van Ee and Erkelens, 1996). However, relative horizontal disparities alone are insufficient to determine the thickness of objects in our visual environment. They must first be scaled according to viewing distance (Kaufman, 1974).

In the absence of monocular indicators of distance (such as texture gradients), only two sources of information are available to scale relative horizontal disparities. These are vergence, which presumably reflects the mean absolute horizontal disparity of a surface, and vertical disparity, which varies over a surface because of the differential perspective of the two eyes. Foley (1980) suggested that vergence angle might be used to obtain an estimate of viewing distance to scale relative disparities. Indeed, the association between vergence angle and perceived size was first noted by Wheatstone (1852). Vertical disparity information alone, however, is also, in principle, sufficient to recover viewing distance (Mayhew and Longuet-Higgins, 1982).

How are these two sources of information combined to scale relative disparities? Bradshaw et al. (1996) compared the influence of vergence and vertical disparity on disparity scaling for different target sizes. They found that vergence alone led to effective scaling for a target 9° of visual angle, whereas vertical disparity only affected the scaling of larger targets. Vertical disparity is pooled over a region of 14° (Adams et al., 1996) to 20° (Kaneko and Howard, 1995), perhaps because the geometry of differential perspective dictates that only large targets give rise to appreciable amounts of vertical disparity. It makes sense that vergence, which influences depth judgments for relatively small targets, should use horizontal disparities pooled over a smaller area. However, whether the dependence of vergence on target size found in our results extends to longer inspection times is something we still need to investigate.

The main contribution of this chapter has been to show that the default initial vergence response does not simply follow the disparity of the entire visual field, or of a cyclopean target presented in that field. Neither is it determined completely by disparities at the fovea, although 95% of the response is accounted for by an eccentricity of approximately 1.5 degrees. In conclusion, it appears that more initial vergence is elicited for a target about 3° of visual angle than for smaller targets. Most everyday objects we encounter in our proximal environment subtend 3° or more. Binocular vision is mostly used for tasks in the near environment (Sheedy et al., 1986) . The mechanism discussed here might be used to compensate for vergence error following a saccade. It complements the input of vertical disparity in estimating viewing distance to resolve stereoscopic depth by operating best for smaller targets.

2.2 Disparity and display size have little effect on the disparity averaging area for initial vergence

2.2.1 Introduction

It has been suggested that the area over which disparities are averaged depends on the disparity presented. This would account for the difference between the integration regions for stereoacuity, controlling threshold disparities of under 1 arcmin (1.5°, Rogers and Graham, 1982) and the averaging region we found for vergence to a disparity of 12 arcmin (3°, experiment 2.1). Indeed, a preference for lower corrugation frequencies of a disparity grating is found at supra-threshold disparities (Iounnou et al., 1993; cited in Howard and Rogers, 1995). If this were the case, there would be no need to postulate different mechanisms controlling the two responses.

Another difference between disparity processing for vergence and depth perception is the sensitivity of the latter only to relative disparity, which might be the result of lateral interactions between disparities. Although a vergence response is obtained even in the absence of relative disparity, it is possible that the response could be enhanced by lateral interactions, as suggested in the discussion section 2.1.4.

In experiment 2.1, we found that horizontal disparities were averaged over an area of 6° to determine initial vergence, much smaller than the area over which vertical disparities are pooled in the global computation of viewing distance (e.g. Kaneko and Howard, 1996). One possible reason for this discrepancy was that we had used only a small disparity of 12 arcmin, a disparity which under our experimental conditions was within Panum's area so no vergence eye-movement was required for fusion. Another possibility was that we were unable to pick up any slow increase in vergence due to cortical magnification beyond 6°, because of the limits of our display. Additionally, in the previous section we suggested there may be a downturn beyond 8° but there was insufficient data in this range to test the suggestion. To examine these possibilities in further detail, we increased the size and disparity of the stimuli.

2.2.2 Method

a Subjects

Four female adult observers with normal or corrected to normal vision participated in this extension of the first study. Data from a fifth (male) observer had to be discarded because his inconsistent results indicated that he was he was unable to follow the instructions, as indeed he reported verbally. All viewers except AP were naive about the exact aims of the study, although MB and CM had participated in 2.1 and been debriefed.

b Apparatus and stimuli

The apparatus was identical to experiment 2.2. The mirrors were positioned directly in front of the viewer's nose (an estimated viewing distance of 5-7 cm) to allow a horizontal binocular viewing angle of about 45°-60° (assuming an interocular separation of approximately 5.8-6 cm, as measured by the author in her (unpublished) undergraduate dissertation study in a sample of 2 male and 8 female observers including herself). The monitors were positioned at a total viewing distance of 22 cm, 17 cm from the mirrors. The dots were now 1 pixel wide, at this viewing distance an angle of 4.7 arcmin. Again, 2000 dots were distributed over the screen. But this time the central disc was given an arbitrary fixed disparity of 4 pixels, or 19 arcmin.

c Design

As in 2.1, the diameter of a central crossed-disparity disc was varied. The dependent variable was the initial vergence response. For AP, the diameter of the disc was varied in multiples of 50 pixels (3.9°). The remaining viewers responded to multiples of 100 pixels (7.8° approx.). For all observers, a zero-diameter condition was used as a measure of fixation disparity.

d Procedure

As before, trials were blocked in same-diameter sessions. See 1.1.2 d for procedure. The method of constant stimuli was used to determine the position of the nonius lines, and left response frequencies were fitted with a cumulative normal to find the position of dichoptic alignment and hence vergence.

2.2.3 Results

Again, observers responded to the presence of a crossed-disparity disc (see figure 2.5). For AP and MB, the response to a disc smaller than 16° was intermediate between the response when no disc was present and the response to larger discs. LG and CM simply converged steadily when the disc was present, suggesting any averaging had taken place within the smallest (8°) disc diameter. Beyond the knee-point in the data, there was no hint of any further steady increase in vergence, as might be expected from cortical magnification, or any clear downturn, as would be predicted from lateral interaction. Because of the small number of subjects, no statistical tests were carried out on these data.

Some further details are worth noting in this results section. The author AP, who features as a subject in all the experiments of the thesis, normally has a steady convergent fixation disparity of approximately 10 arcmin. Here, however, because of the short viewing distance, this is replaced by a small (1 arcmin but not significant) divergent fixation disparity. Such discrepancies between near-viewing and far-viewing fixation disparities are typical of the type-I FVFD (forced-vergence fixation disparity; Ogle, 1950) curve recorded for AP in experiment 5.2. Another observation is that despite the brief (230 msec) stimulus interval, one subject (LG) was able to complete the 19 arcmin vergence movement demanded by the stimulus disparity, unless this was an artefact of the blocking procedure, and this subject in fact simply developed a convergent fixation disparity during the course of the experiment.

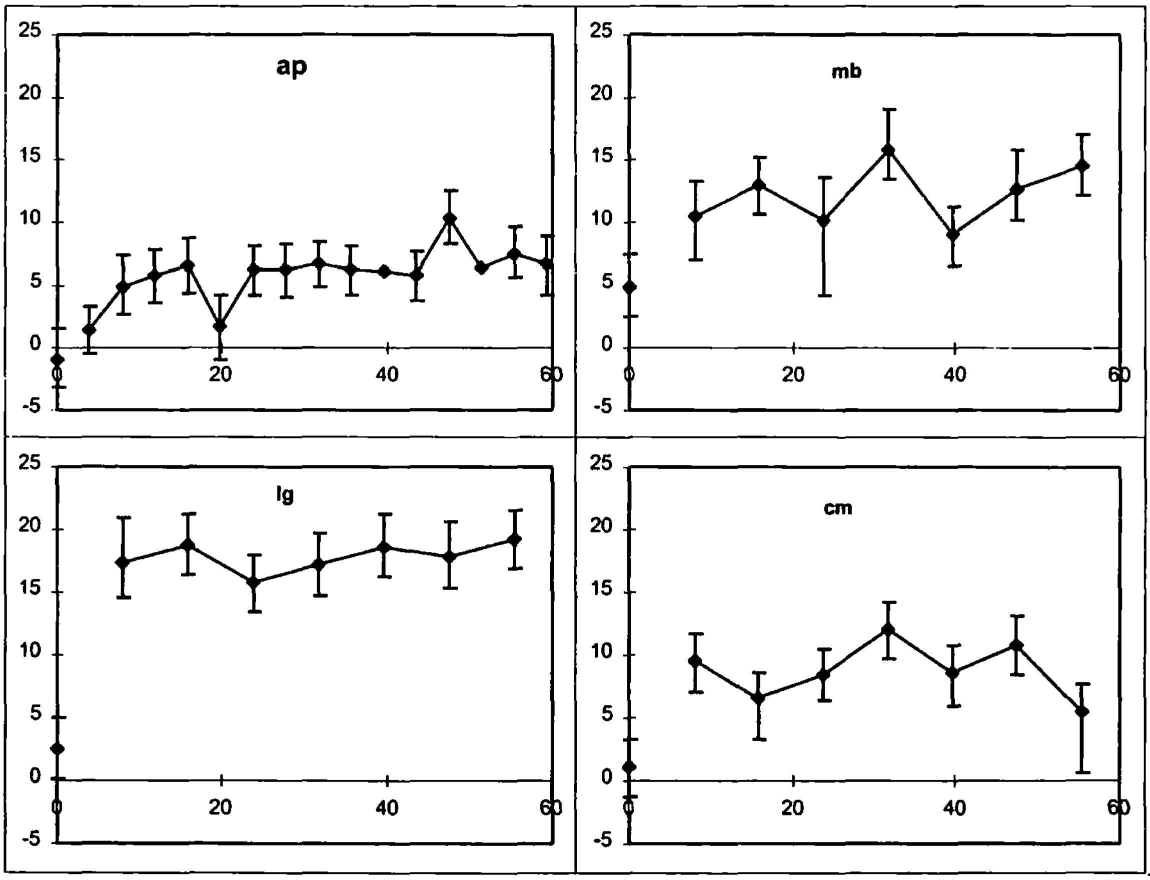


Figure 2.4. Vergence (arcmin) is plotted against target size (degrees) for a 19 arcmin crossed disparity disc protruding from a fixation plane surround. Vergence saturated well before the maximum size of over 50 degrees was reached. For AP and MB the point of saturation appears to be between 10 and 20 degrees, whereas both LG and CM verged no further to targets beyond 10 degrees in diameter, LG obtaining near target vergence at this size. Note the direction of fixation disparity in AP is reversed for the shorter viewing distance used in this study.

2.2.4 Discussion

From these results, there is no consistent indication of a continued increase in vergence beyond the knee-point in the data, as would be predicted if the results of experiment 2.1 were due to cortical magnification rather than a limited averaging region. Neither is there any sign of a downturn, which would be expected if lateral interactions in the disparity domain were involved in the processing of vergence.

There is insufficient data here to prove or disprove the suggestion of a fixed 5° region of disparity integration for initial vergence. However, the functions of vergence against target size for AP appears to saturate at a slightly larger diameter here than in 2.1 (compare figures 2.4 and 2.2). A number of reasons for this difference can be suggested. The pooling region observed in 2.1 may be specific to small disparities, with larger scale pooling operating for greater disparities. There may be integration regions of different sizes depending on absolute vergence angle, or viewing distance. The larger stimulus dots in the present experiment will have stimulated lower spatial-frequency tuned mechanisms, which may be averaged over larger regions in the determination of vergence.

Two kinds of pooling are possible – disparity averaging and disparity integration. Disparity integration (without dividing by the area) would lead to an increase in response when the area of a stimulus (without a surround) is simply increased. Such disparity integration is negligible for horizontal vergence (Fang et al., 1998). Averaging, however, where the integral is subsequently divided by the area, does occur when horizontal disparities are spatially varied, i.e. the size of a target area is varied in the presence of a disparity surround (also Fang, et al., 1998). This second kind of spatial integration or pooling, which is really averaging, is the one intended here.

It is the first suggestion, of a size-disparity correlation in the cyclopean domain, which is most consistent with the model we were beginning to develop in 2.1. This was based on Marr and Poggio's (1979) theory, but replacing luminance spatial frequency with the spatial frequency of disparities. The third suggestion, however, is compatible with an explanation based on luminance spatial frequency distributions. Therefore, in the next section we examined the effect of dot size alone.

2.3 Dot size has little effect on the disparity averaging area for initial vergence

2.3.1 Introduction

Experiment 2.1 was repeated with different sizes of stimulus dots. Increasing the size of the dots correspondingly lowers the peak spatial frequency (Morgan and Fahle, 1992). Following the results of 2.1, data points were chosen on a logarithmic scale so the rising portion of the curve could be determined more effectively. The method in 2.1 was improved by randomly interleaving trials with different target disc diameters (see 2.1.2 for a discussion of the drawbacks of blocking by condition).

2.3.2 Method

a Subjects

Five observers, including AP, participated in this study. All had normal or corrected to normal vision. CM and KP had been debriefed after experiment 2.1.

b Apparatus and stimuli

The apparatus and stimuli were similar to those used in 2.1. However, although the EIZO Flex-6500 monochrome monitors were still used in this experiment, the VSG card was now mounted on an Ambra 486 PC. Additionally, large (25 x 20.5 cm) mirrors were used to permit a wider viewing angle. The viewing distance was 65 cm. At this distance, each pixel subtended 1.72 arcmin. As the number of dots was not changed, mean luminance increased with dot size.

c Design

Once more, the diameter of a central 6 pixel crossed-disparity disc was varied. The disparity of the disc was again arbitrary, here set to 10.3 arcmin. There was a no-target control. The disc sizes presented were chosen from a set which increased on a logarithmic scale, from 0.6°, 1.1°, 2.3°, 4.6° to 9.2°. The size of the dots from which the stereogram was composed also formed a set that increased logarithmically. The dots were 1.72 arcmin, 3.44 arcmin or 6.88 arcmin squares. Initial vergence was estimated in each condition of the two variables.

d Procedure

The nonius method was used to estimate vergence, as before. However, here the nonius stimulus consisted simply of two dichoptic lines. A staircase procedure was introduced to better fit the nonius offset to perceived alignment, but as before the response frequencies at each offset were analysed using Probit to obtain the vergence estimate. An additional difference in the procedure was that the trials were blocked into only three sessions, by dot size. During each session, the presentation order of trials in the different diameters was fully randomised, with separate nonius staircases operating under the different conditions. These measures were taken so that the number of trials taken for each data point could be reduced to 50, and more data could be gathered from each subject in a shorter time.

2.3.3 Results

The pattern of results in 2.1 was replicated despite the change in procedure (compare figures 2.2 and 2.5). All five observers showed intermediate vergence to targets smaller than about 3°, suggesting an averaging of target and ground disparities. The data from new subjects EM and NC were noisy, and, like RG in experiment 2.1, show no clear saturation as disc size increases. Only CM's response varied consistently with dot size.

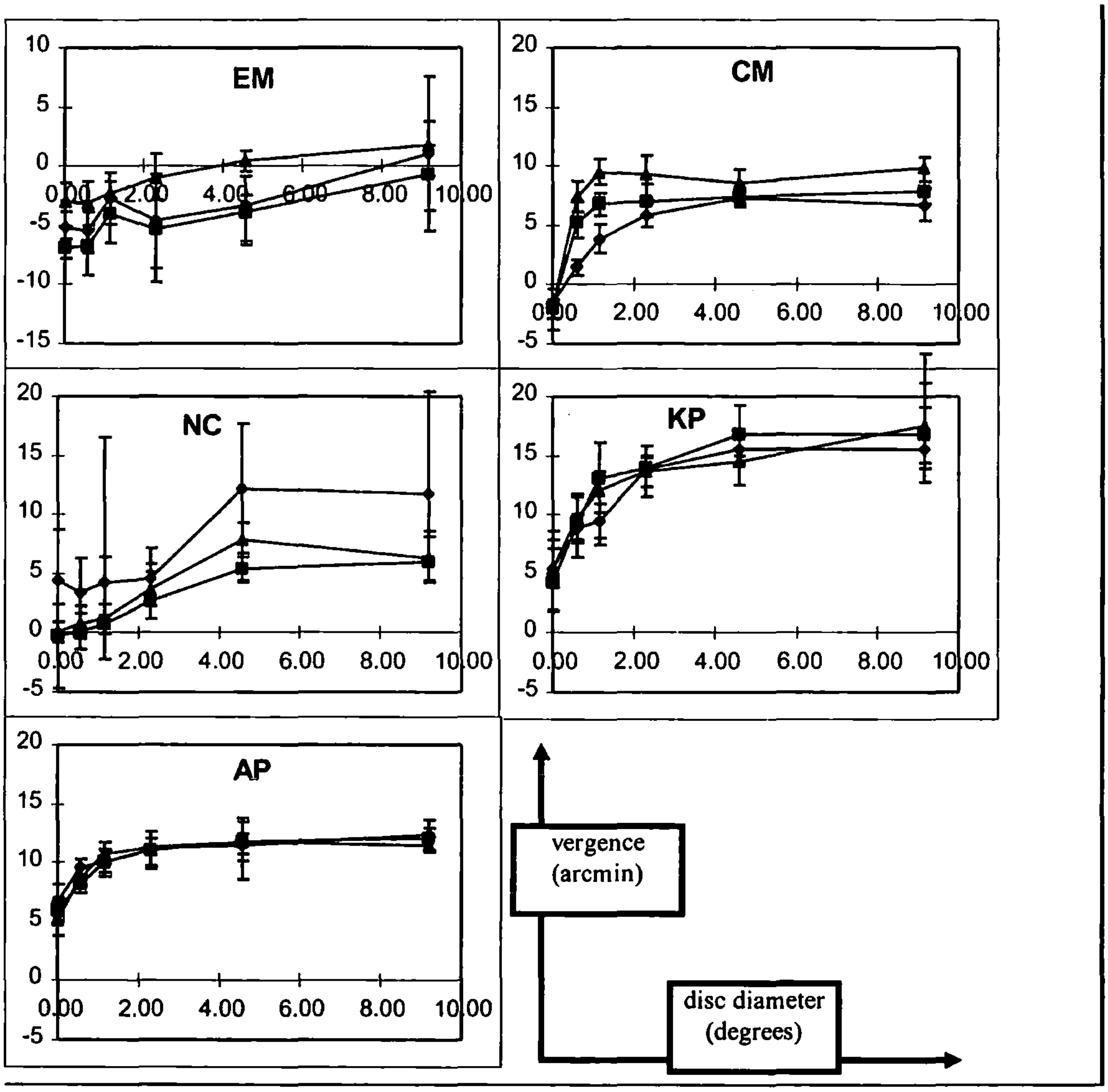


Figure 2.5. Vergence (arcmin) is plotted against target size (degrees) for three different dot sizes: 2 arcmin (diamonds), 4 arcmin (squares) and 8 arcmin (triangles). Across observers, dot size makes no consistent difference to the function of vergence with target size. Target disparity was always 10.3 arcmin. Note different axes for EM because of a large fixation disparity.

As in 2.1.3, the data were fitted with a cumulative normal, this time using probit fitted to proportions of full response. Because the purpose of this modelling was to estimate the width of the pooling region, only data from AP, CM and KP were included as the functions of EM and NC did not asymptote. This time, σ obtained was 0.86 (compared with 0.84 in experiment 2.1), giving a 95% pooling diameter of 2.84, and a 99.9% region of 5.16 degree diameter. (For comparison with experiment 2.1, see figures 2.3 and 2.6).

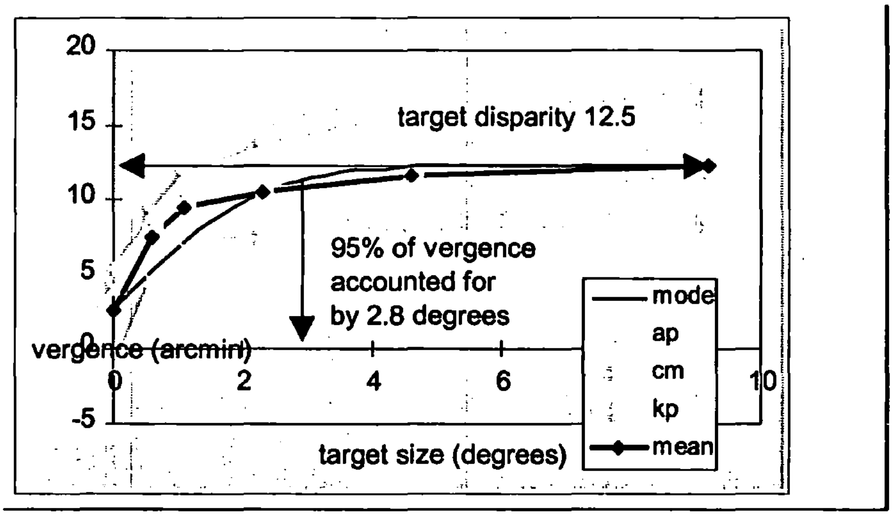


Figure 2.6. Means for AP, CM and KP were averaged across dot sizes. The grand mean across these three observers was fitted with a cumulative normal, as in figure 2.3.

2.3.4 Discussion

The aim of this experiment was to determine the effect of dot size. Across subjects, we found no difference in saturation point (or any other aspect of the vergence: size function) between the different dot sizes. If anything, one subject (CM) had a lower saturation point for the larger dots.

The possibility remains that the input filters which control initial vergence are larger than any of the dot sizes we used, as Morgan and Fahle (1992) have argued for motion. This is consistent with the short latencies of initial vergence (50-100 msec, Masson et al., 1997). It is compatible with a response driven by the early stages of visual processing and controlled by coarse, 'transient' spatial frequency channels (Campbell and Robson, 1968).

The main conclusion from this section is that the 3 degree pooling region suggested by the modelling in 2.1 was a genuine parameter of disparity processing, not an artefact of the sparse sampling for small disparity discs in that experiment, nor of the blocking procedure employed. However, three of the subjects tested so far (EM and NC from 2.2, and RG from 2.1) showed no vergence saturation in the disc-size domain tested. Therefore, the area of disparity averaging appears to be subject to individual variation and should be examined in a larger sample of observers.

Chapter 3

Stereoacuity - spatial integration and the effects of vergence

3.0 Abstract

Experiment 3.1 confirms the effects of disparity modulation scale on stereoacuity thresholds, described by Tyler (1975), Rogers and Graham (1982, 1985), and others. Using a disc target, as in chapter 2, stereoacuity was optimal at a target diameter of 1-3° in a sample of 6 observers. Although there was considerable variation between individuals, this was much smaller than the integration region we found for initial vergence. Additionally, stereoacuity, unlike vergence, declined as the disc size was increased beyond the optimal diameter.

Previously (Chapter 2; Popple et al., 1998) we found, using random-dot stereograms, that initial vergence increases with the size of a cyclopean disc. A corresponding improvement in stereoacuity within the disc was predicted, because disparities in the disc would be brought closer to the plane of current fixation.

In experiment 3.2, we looked at the effect of the spatial extent of a briefly presented (≤500 msec) cyclopean depth pedestal on stereoacuity thresholds. Observers were required to judge the depth of a small 1.7° central disc relative to a larger surrounding disc in a random-pattern stereogram. The larger disc was set, initially, at a pedestal disparity of ±24' against a fixation-plane surround. The size of the larger disc was varied from 2.6°-8.0°. As predicted, stereoacuity thresholds fell significantly with increasing pedestal disc size. Next, the disparity of the pedestal disc was varied. When pedestal disparity was reduced to ±2.4□ (experiment 3.3), a disparity too small to

demand vergence, the size effect disappeared except when the pedestal boundary was within 30' of the test disc boundary. This shows the effect was largely due to vergence and not cyclopean integration alone.

However, the effect of pedestal size was found to persist with stimuli too brief to permit vergence (≤100 msec) although at such durations the depth discrimination thresholds were large and outside the stereoacuity range.

3.1 The effect of random-dot disc diameter on stereoacuity threshold

3.1.1 Introduction

Contrast sensitivity threshold, the minimum amount of luminance contrast required to detect a stimulus, varies with the spatial frequency of contrast modulation (ref.). Similarly, stereoacuity threshold, the minimum disparity difference required to detect a depth step (Fahle et al., 1994), varies with the spatial frequency of disparity modulation. Tyler (1975) showed that lowest thresholds of under 0.3 arcmin (20 arcsec) were obtained for modulations of about 0.4 c/deg in the depth of a vertical line, with thresholds increasing for both lower and higher frequencies. Rogers and Graham (1982) replicated this finding using cyclopean, random-dot disparity gratings. They found an optimal spatial frequency of disparity modulation in the region of 0.2-0.5 c/deg. Later this result was confirmed using difference-of-gaussians disparity profiles where the number of cycles visible did not vary between spatial frequencies (Rogers and Graham, 1985). In chapter 2 we argued that our finding of vergence saturation at a cyclopean disc of diameter 6° indicated that disparities were integrated over a larger area than for stereopsis. However, to make a more direct comparison, stereoacuity would have to be measured using a similarly broad-band stimulus. Although the fundamental spatial frequency of a square disparity profile decreases as the breadth of the square is increased, a range of spatial frequencies of depth modulation is present in the stimulus. In the present experiment, we measured stereoacuity thresholds directly in such a broad-band disparity stimulus. As in experiment 2.1, the size of a cyclopean disc was varied. From the aforementioned studies, we predicted a minimum disparity threshold at a disc diameter of 1-2.5°, which is half of one cycle at 0.2-0.5 c/deg.

3.1.2 Method

a Subjects

Three male and two female adult observers aged 25-35, and the 6-year-old authorÕs son (UP) participated in the experiment. All had normal or corrected to normal visual acuity, and stereoacuity as demonstrated by the results. Only AP had precise knowledge of the purpose of the study, although the remaining subjects except for UP and MO were experienced psychophysical observers and vision scientists who were aware of the theory involved.

b Apparatus and stimuli

As before, stimuli were generated on an Ambra 486 PC using a Cambridge Research VSG graphics card. They were displayed on two EIZO Flex-6500 28 x 36 cm monochrome monitors. The monitors were placed at a viewing distance of approximately 160 cm. To form a modified Wheatstone stereoscope, two large (25 x 20.5 cm) mirrors fixed at right angles to each other were positioned 5 cm directly in front of the viewer, with the monitors 150 cm away on either side. (All measurements are height x width). A chin rest was used to maintain viewing position. The stimuli we used were random-pattern stereograms of a central test disc in front or behind a larger disc set at a pedestal disparity against a fixation plane surround. The apparatus is illustrated in figure 3.1.

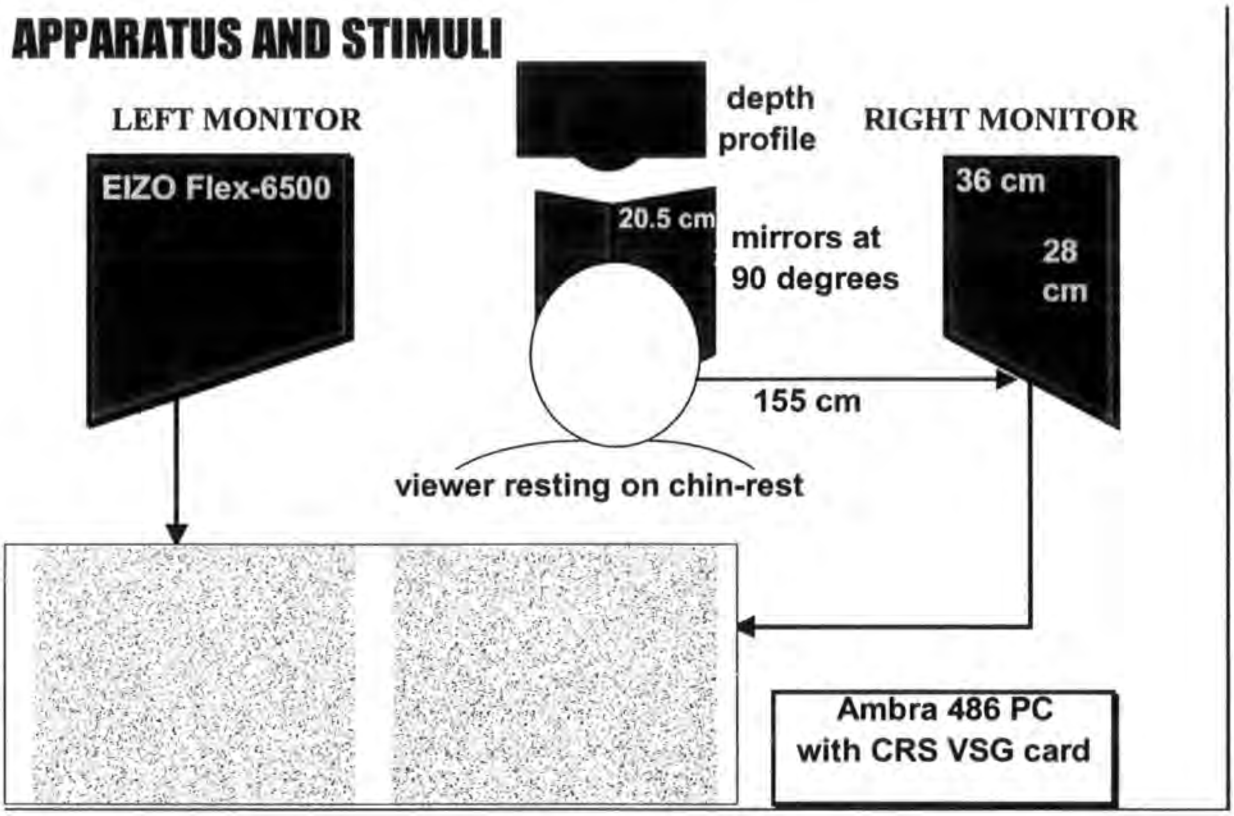


Figure 3.1. Apparatus and stimuli. We used a modified Wheatstone stereoscope consisting of two 25 x 20.5 cm mirrors mounted at a right angle. Stimuli (random-pattern stereograms, see text for details) were displayed on two 28 x 36 cm monochrome EIZO Flex-6500 monitors at a viewing distance of 155 cm. The stimuli, depicting a small disc embedded in a larger disc, were generated using a CRS VSG card on an Ambra 486 PC. A chin rest was used so the viewer could comfortably maintain head position. The stereogram illustrates one of a number of possible depth profiles, but is not based on the actual stimuli used in our experiment.

The stimuli were stereograms of a variable diameter test-disc in front of, or behind, a 10.3° x 10.3° square fixation plane surround.

Small (3"-45") disparities of the test disc were generated using sub-pixel shifting. To represent a disparity smaller than a pixel width (48") from a given pixel on one monitor, the brightness of the corresponding pixel on the other monitor and its horizontal neighbour were modified in proportion with the desired shift (Morgan & Aiba, 1986).

To reduce memory requirements and speed stimulus generation, random-pattern stereograms were used (Tyler (1979) in Tyler and Clarke, 1990). The stereograms consisted of a repeated pattern of vertical strips of 50% black and white 1.6' x 1.6' random dots. (We must emphasize that although a repeated pattern was used, the stereograms were viewed using a stereoscope and not free-fused as autostereograms). Fifteen different strips were stored in memory. For each strip, 15 different disparities produced by sub-pixel shifting were also stored, for rapid stimulus generation. Because the patterns were repeated, this was sufficient to produce a range of possible stereograms by simply cutting and pasting appropriately. Stereograms were generated on-line from these stored patterns.

Viewers reported no difficulties with the repeated pattern, or with possible screen luminance non-linearities resulting from the sub-pixel shifting.

A 16' x 16' white 0.8' outline central square on a mid-grey screen was displayed both as a fixation and following the stimulus interval. PC generated auditory signals were used for feedback to mouse-button press responses.

c Design

Stereoacuity thresholds were estimated as disc diameter was increased logarithmically from 0.4° to 6.8° (0.4°, 0.9°, 1.7°, 3.4° and 6.8°).

d Procedure

Each trial began with a 1.3 second fixation period during which the stereogram was calculated off-screen. This was followed by the stimulus, presented for 500 msec, which in turn was replaced with the fixation stimulus. Viewers were instructed to respond with a right mouse button-press if the test disc was in front of the surround, and left if behind. Errors were signalled by a tone to provide continuous feedback throughout.

There were 4 blocks of 200 trials. Disc size varied randomly, with equal numbers of presentations in each of the 5 sizes displayed. The disparity sign of the test disc was crossed or uncrossed on half the trials, each chosen randomly on every trial. Test disparity was determined by a staircase procedure for each diameter of the test disc. The five staircases for the different disc sizes operated simultaneously.

Results, in terms of number correct out of total number of presentations, were fitted with a cumulative normal to determine the 75% stereoacuity threshold. Data from crossed and uncrossed disparities the disc were pooled.

3.1.3 Results

As can be seen from figure 3.2, stereoacuity varied with disc diameter with minimum stereoacuity thresholds generally observed at a diameter of approximately 1-2°. The exception is AP, who performed best at a disc diameter of 3.4°. There was, however, considerable variation between subjects in the thresholds obtained for the largest and smallest disc sizes. EM had stereoacuity of under 1 arcmin (60 arcsec) for all disc diameters, whereas CM and MO could only detect a disparity of several arcmins when the size of the disc was greater than 6°.

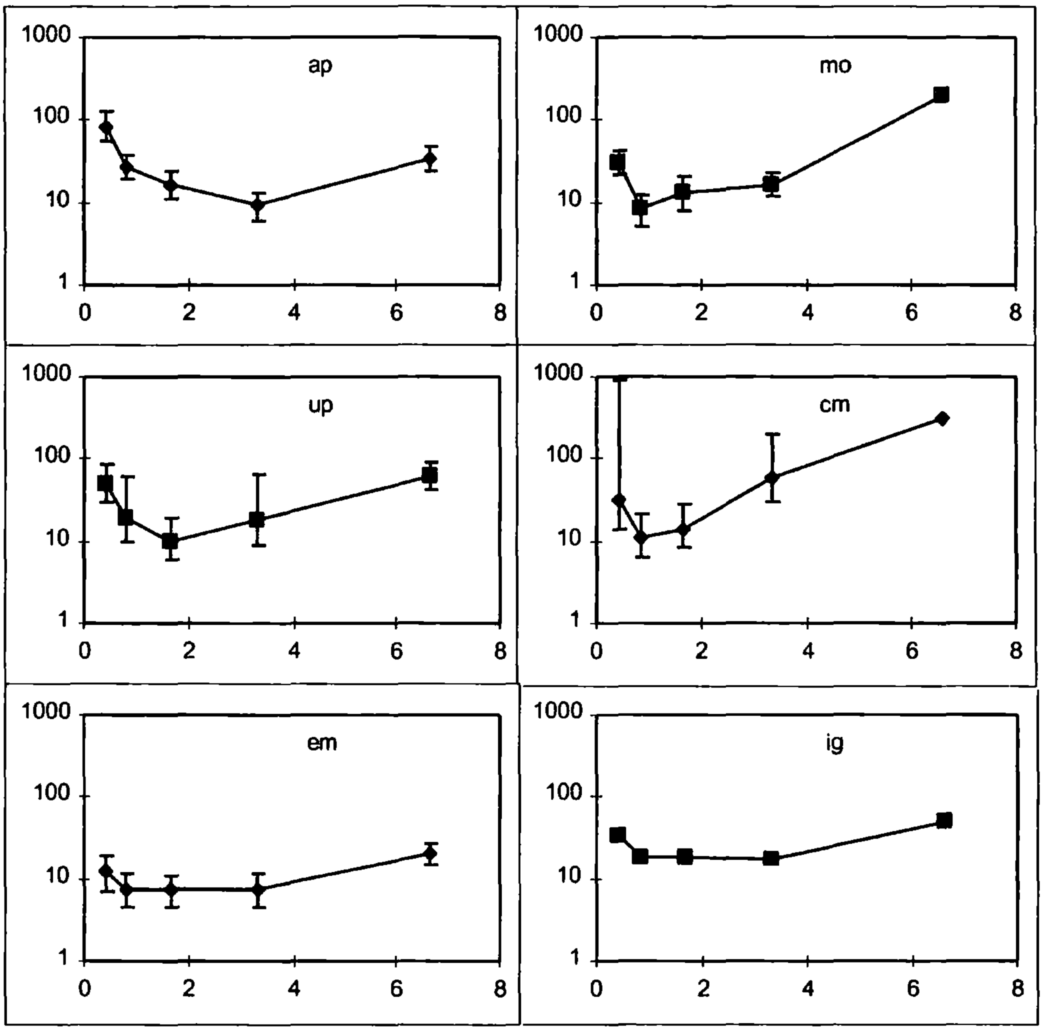


Figure 3.2. Stereoacuity thresholds (arcsec(are plotted against test disc diameter (degrees) for six observers. Error bars representing 95% confidence intervals are shown where available, in some cases error bars could not be calculated because of the heterogeneity of the data. Note the considerable individual differences in the shape of this function.

3.1.4 Discussion

Our findings for broad-band disparity stimuli (in the disparity domain) generally confirm the predictions made from published results using narrow-band stimuli, as described in the introduction. Specifically, minimum stereoacuity thresholds were obtained in the disc-diameter range from 0.9-3.4°. In our investigation, data were collected from six observers rather than the one or two common to the studies mentioned in the introduction. This may be the reason why differences between individuals were a more prominent feature of our results. To explore these differences in the context of both narrow-band and broad-band stimuli, a greater number of subjects must be studied.

The main goal of this study was to show a contrast between performance on depth perception and initial vergence as the size of a cyclopean disc is increased. This has clearly been achieved. Unlike initial vergence, which only saturated at a disc diameter wider than 5°, the best stereoacuity performance was observed when the disc was only about 2° in diameter. As discussed in chapter 2, depth perception is mediated by a finer disparity processing mechanism than is initial vergence. In addition to the differences in the acceleration of the responses at small sizes, stereoacuity shows a clear decline (i.e. stereothresholds increase) beyond an optimal size, in our broad-band disparity stimuli as in the narrow-band stimuli used in the literature. This is in contrast with initial vergence, where there was little indication of a decline after the response saturated. The rise in stereoacuity thresholds at low spatial frequencies has been taken to imply lateral interactions, which would arise from the presence of centre-surround receptive fields for disparity, as have been found for motion processing (Rogers and Graham, 1982; Tanaka et al., 1985). By implication, disparity processing for initial vergence must therefore involve classical (gaussian) receptive fields of limited size, although other possible explanations for the saturation of that response were discussed in section 2.2.

This study formed a preliminary to the investigation of the role of such coarse disparity pooling for vergence in the fine computation of depth that determines stereoacuity. From the present study, we concluded that, across subjects, stereoacuity of arcseconds was obtained for a disc diameter of 0.9-3.4°. Therefore, the intermediate size in this range (1.7°) was used for the test disc in the next experiment.

3.2 'Coarse to fine' cyclopean processing

3.2.1 Introduction

Cyclopean images are visible only by combining information from the left and right eyes, as in a random-dot stereogram (Julesz, 1971). Depth information is given by the pattern of disparities between the two eyes' views. However, only disparities separated by less than a critical limit, known as Panum's fusional area, can easily be fused simultaneously. This area is traditionally taken to be approximately ±10' from the plane of current fixation (Mitchell, 1966) although it varies with stimulus properties such as eccentricity, spatial frequency and temporal duration (for a recent review see Howard and Rogers, 1995). Additionally, there is an upper disparity limit, which is much larger (about 30 arcmin, depending on the stimulus), beyond which even diplopic stimuli do not give rise to depth. Computationally, having an upper disparity limit reduces the correspondence problem - the problem that a dot in the right eye could be matched with any one of a number of possible dots in the left. The limit also agrees with the properties of 'tuned' binocular disparity-detector cells in the primary visual cortex (Poggio and Talbot, 1981) which are maximally sensitive to disparities near fixation, and decline sharply showing almost no response beyond 1°.

In order to bring the pattern of disparities on different surfaces in the environment to within Panum's fusional area (i.e. into single vision), we make vergence eye-movements. Vergence eye-movements are disjunctive movements of the two eyes which determine the locus of their common fixation in depth, that is along the z-axis towards and away from the viewer. When the eyes move in opposite directions, the plane of fixation changes. Vergence responds directly to disparity (Rashbass and Westheimer, 1961).

Previously (Chapter 1, Popple et al., 1998) we showed that initial vergence depends on the size of a cyclopean target, such that it becomes more accurate with increase in the diameter of a random-dot disc. From that result, we predicted that stereoacuity, the ability to make fine depth discriminations on the surface of the disc, would improve similarly. This is because, with improved vergence, the surface of the disc would be brought closer to the plane of current fixation, where stereoacuity is best (Westheimer and McKee, 1978). To test this prediction, we presented a stereo target superimposed on a larger disc. The larger disc was set at a constant pedestal disparity. We varied the diameter of the disc, and measured stereoacuity thresholds for a smaller target within the disc.

3.2.2 Method

a Subjects

Eleven viewers with normal or corrected to normal vision participated. All except AP and JF were naive as to the purpose of the study.

b Apparatus and stimuli

Apparatus and stimulus generation were as in 3.1.2. The stimuli were stereograms of a 1.7° diameter test-disc in front of, or behind, a larger (2.6°-7.8°) disc set at a crossed or uncrossed pedestal disparity (±24' in experiment 3.2, ±2.4' in experiment 3.3) against a 10.3° x 10.3° square fixation plane surround.

c Design

Stereoacuity thresholds for the test disc were measured, as pedestal disc diameter varied. The pedestal disc diameter was 2.6°, 3.9°, 5.2°, 6.5°, or 7.8°. Test disc diameter was fixed at 1.7° and pedestal disparity was ±24'.

d Procedure

The procedure was similar to 3.1.2, but for the sake of completion and clarity full details are given below.

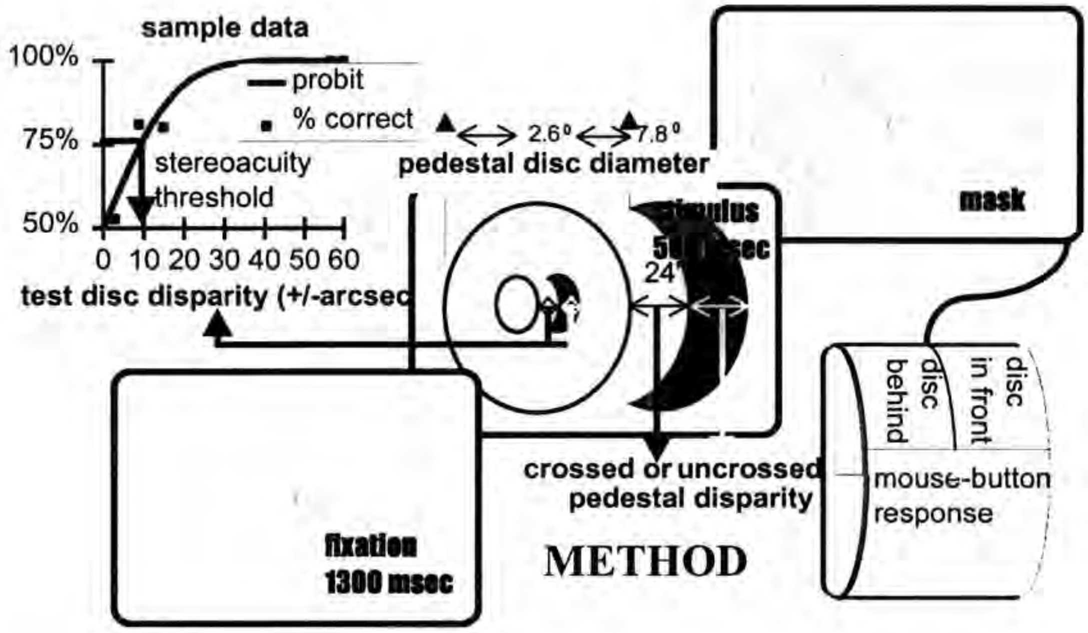


Figure 3.3. Method. Each trial consisted of a 1300 msec fixation, followed by a 500 msec stimulus which was replaced with a mask. The stimulus consisted of a large (2.6°-7.8°) disc at a pedestal disparity of ±24 in the centre of which was embeded a small (1.7°) test disc. The relative depth of the test disc was varied using a staircase procedure. Viewers pressed a mouse button to indicate whether the test disc was in front or behind the larger, pedestal disc. 75% stereoacuity thresholds for each size of the pedestal disc were determined by fitting a cumulative normal to the data pooled across crossed and uncrossed disparities (see example).

As shown in figure 3.3, each trial began with a 1.3 second fixation period during which the stereogram was calculated off-screen. This was followed by the stimulus, presented for 500 msec, which in turn was replaced with a mask. Viewers were instructed to respond with a right mouse button-press if the central disc was in front of its immediate surrounding, and left if behind. Errors were signalled by a tone to provide continuous feedback throughout.

In the experiment, there were 4 blocks of 200 trials. Throughout the experimental sessions, the depth of the test disc was judged against the larger disparity pedestal disc surrounding it. Disc size varied randomly, with equal numbers of presentations in each of the 5 sizes displayed. The disparity sign of both pedestal and test disc were crossed or uncrossed on half the trials, each chosen randomly on every trial. Test disparity was determined by a staircase procedure for each diameter of the pedestal disc. The five staircases for the different pedestal disc sizes operated simultaneously.

Results, in terms of number correct out of total number of presentations, were fitted with a cumulative normal to determine the 75% stereoacuity threshold. Data from crossed and uncrossed disparities of both test and pedestal discs were pooled. The latter is based on the assumption that fixation disparities were minimal, and crossed and uncrossed vergence follow a similar pattern as a function of cyclopean target size.

3.2.3 Results

Figure 3.4 shows the results from all 11 observers. Stereoacuity thresholds fell as the size of the pedestal disc increased. Stereoacuity thresholds are plotted on a logarithmic axis, because this produced the best linear representation of the data (see plot of means).

The log-linear relationship between stereoacuity and pedestal disc size was confirmed using repeated-measures ANOVA on the data from the 10 subjects who provided complete sets. There was a significant effect of disc size on log stereoacuity threshold ($F(4,36)=24.1$, $p<0.001$). This could be accounted for by the significant linear trend ($F(1,9)=43.6$, $p<0.001$).

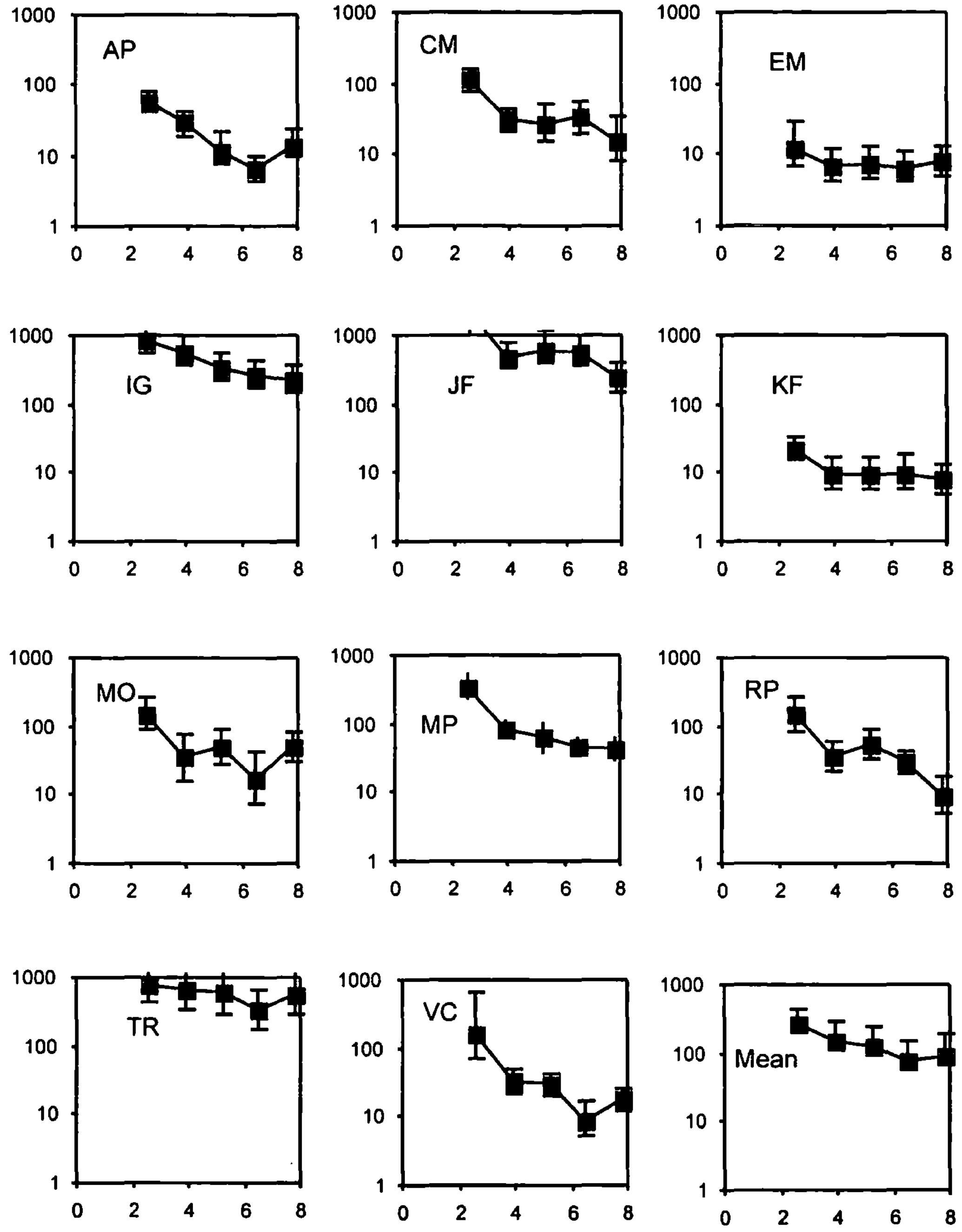


Figure 3.4. Results of experiment 3.2. Stereoacuity improved as the size of the pedestal disc was increased. All the viewers show a consistent pattern of reduced stereoacuity thresholds for larger pedestal discs. The plot of means across subjects illustrates the log-linear trend in the data. See text for details of statistical analyses.

The improvement in stereoacuity with increasing disc size was considerably greater in subjects with poor stereoacuity (IG, JF and TR), as would be expected if the function is logarithmic. Subjects with exceptionally good stereoacuity (KF and EM) showed little improvement with increasing diameter of the disparity pedestal disc beyond 3.9°. Because the plots are steeper for the interval from 2.6° to 3.9°, we wanted to exclude the possibility that this alone accounted for the significant results. A partial comparison was carried out using only the data from 3.9°-7.8°. There was still a significant effect of size ($F(3,27)=4.8$, $p=0.008$).

3.2.4 Discussion

Stereoacuity improves with increasing the spatial extent of a pedestal disparity disc. We are better at telling a bump from a dent on a large surface than on a small surface. This can be taken as evidence of coarse to fine processing.

Initial vergence pools disparities over a large (at least 6°) area (chapter 2, Popple et al., 1998), and the best stereoacuity was found for such large pedestal discs. This optimum size for the pedestal disc is considerably greater than the optimum size of the stereoacuity target (see experiment 3.1). Rogers and Graham (1982) found that stereoacuity peaks at a spatial frequency of disparity modulation from 0.2-0.5 cycles/°, equivalent to a target 1°-2.5° across.

A separation between coarse and fine mechanisms of disparity processing was proposed by Richards (1970, 1971), following his work with stereoanomalous observers. In these subjects, an impairment in coarse stereopsis for either near or far targets resulted in an inability to perceive transient depth in a briefly flashed, coarse disparity target, although the same subjects were able to perform normally on a fine stereoacuity task. It was suggested that the coarse mechanism was responsible for bringing vergence onto a disparate target, and the fine mechanism for stereoacuity tasks.

However, the results of Experiment 1 might be explained without recourse to vergence. The advantage of a larger pedestal disc could be the result of spatial interactions. Neighbouring luminance contours interact laterally, suppressing one another at close proximity and enhancing at a slightly greater distance (Toet and Levi, 1992; Polat and Sagi, 1993). Lateral interactions have been reported in the cyclopean domain (Anstis et al., 1978; Brookes and Stevens, 1989; Lunn and Morgan, 1995), as have interactions across spatial frequencies of disparity modulation (Tyler, 1983). Perhaps the improvement in stereoacuity with the size of the disparity pedestal was the result of such interactions. The 2.6° pedestal disc was close in size to the 1.7° test disc. The contour of the pedestal disc might have suppressed the test disc□s contour, rendering it less visible. Both discs may have excited similar spatial-frequency tuned mechanisms, making it hard to select which was the target of the depth discrimination task. For larger pedestal discs, these inhibitory interactions would be reduced, and possibly even replaced by facilitatory ones.

Moreover, the existence of a peak in sensitivity for ~0.3 cycles/deg disparity modulation (Rogers and Graham, 1982) might account for the advantage of a 5.1 deg disparity pedestal, which, with a 1.7 deg target disc creates a circularly-symmetric square-wave-like profile of wavelength 3.4 deg. This has energy over a range of different spatial frequencies, but will have substantial energy at ~0.3 cycles/deg (the fundamental frequency of the square-wave).

3.5). Thresholds were, generally, lower for the 2.4′ disparity than at 24′(F(1,5)=8.4, p=0.034). Pedestal disc size still had a significant effect when the standing disparity was 2.4’ (F(4,20)=3.73, p=0.02), however this effect was entirely accounted for by the difference between discs of diameter 2.6° and 3.9°, and disappeared when only discs 3.9° and larger were compared (F(3, 15)=0.23, not significant) (for similar analyses with the larger pedestal disparity, see the results section of 3.2).

For EM and KF, the functions of stereoacuity with pedestal disc diameter were almost parallel for 2.4’ and 24’ standing disparity, with the only clear improvement in stereoacuity when the pedestal disc was enlarged from 2.6° to 3.9°. Both these subjects had unusually good stereoacuity, and the lack of variation in their depth thresholds may have been due to a floor effect.

KF, who completed the 2.4’ condition (3.3) before the 24’ condition (3.2), nevertheless had stereoacuity thresholds consistently lower for the 2.4’ than the 24’ pedestal, as did the other subjects, indicating that this was not a practice effect.

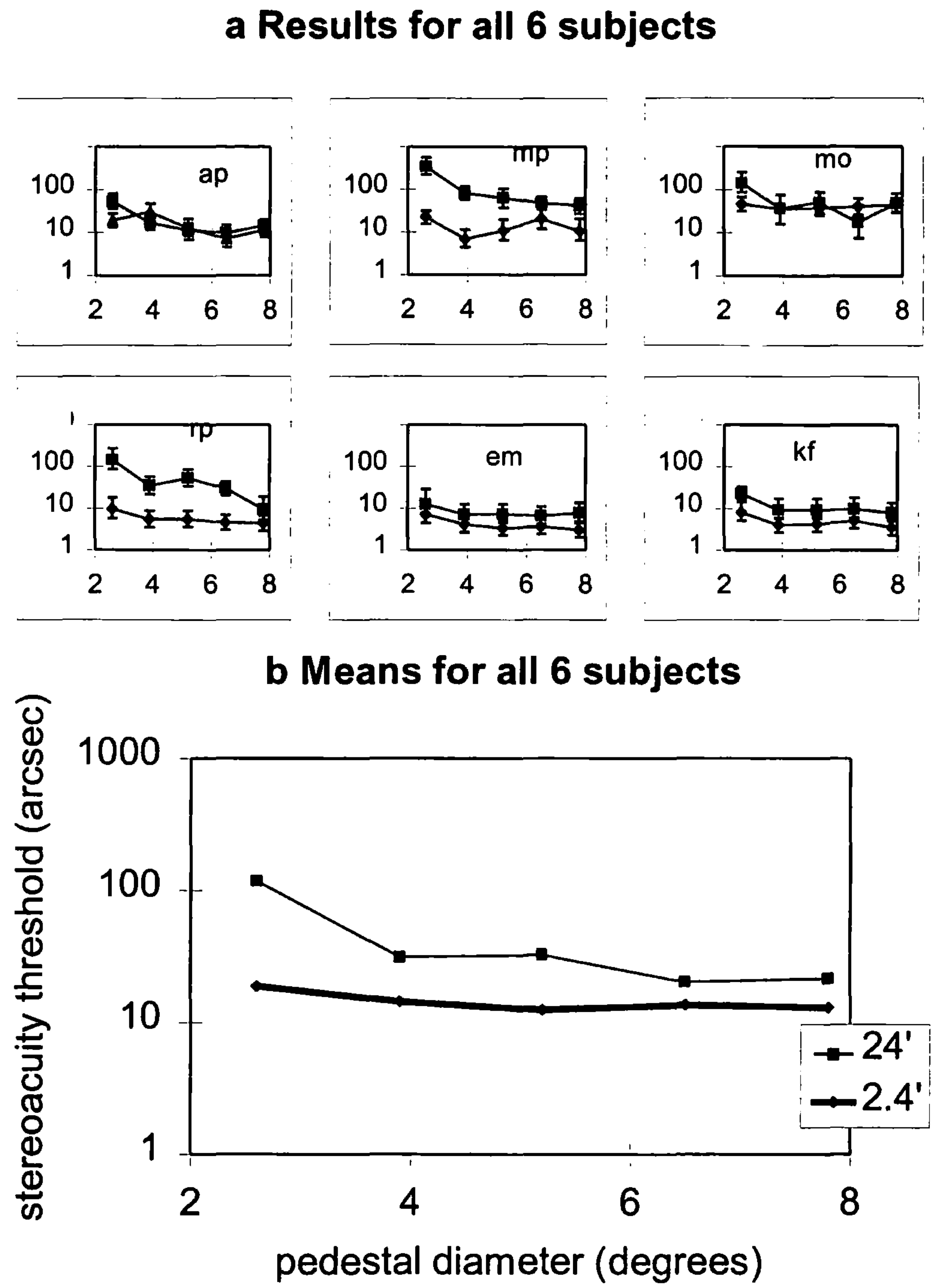


Figure 3.5. Results of experiment 3.3. The effect of pedestal disc size on stereoacuity was much reduced when the pedestal disparity was reduced to 2.4'. a) The interaction between pedestal disc size and pedestal disparity is evident in most subjects' data. b) This pattern is clearer when looking at the means. At 2.4', there is little effect of pedestal disc diameter.

3.3.4 Discussion

The effect of pedestal disc diameter on depth thresholds depends on standing disparity, and can largely be explained by impairments in stereoacuity resulting from inadequate vergence to small cyclopean discs. This finding favours the acceptance of vergence as the major constraining factor in coarse cyclopean processing. If the results of experiment 3.2 had been due to lateral interactions, a similar dependence of acuity on disc diameter would have been expected for the smaller pedestal disparity in experiment 3.3. However, cyclopean interactions cannot be ruled out as contributing to the effect of pedestal disc diameter, especially for the smallest 2.6° disc, where stereoacuity was relatively poor even at a 2.4′ pedestal disparity. Our results will be discussed in the context of existing theories of disparity processing for vergence and depth perception.

Vergence is unlikely to be the only cause of the pedestal size effect observed in Experiment 1, and, to a lesser degree (literally only for small pedestals) in Experiment 2. Cyclopean interactions cannot be ruled out as contributing to the effect of pedestal disc diameter, especially for the smallest 2.6° disc, where stereoacuity was relatively poor even at a 2.4′ pedestal disparity. It also remains possible that this effect was due to reduced acuity away from the horopter. One must simply assume that the 2.6° pedestal disc at 2.4′ disparity stimulated negligible vergence, whereas vergence was accurate for the larger discs.

Experiment 2 was carried out to determine whether the pedestal size effect observed in Experiment 1 was due to vergence, or lateral cyclopean interactions. Because of the floor effect found, this comparison is difficult to make. It could be argued that both vergence and lateral interactions might vary quantitatively (but not qualitatively) as a result of changing the magnitude of the disparity step. An alternative test of the two hypotheses is to render vergence ineffective by shortening the stimulus interval. Intervals of 100 msec and less are needed to exclude short-latency vergence (Masson et al., 1997). At such brief stimulus durations, the depth discrimination task proved too difficult for most subjects. Data from the author (AP), who had trained extensively while creating the stimuli, are presented below for 100 and 500 msec intervals.

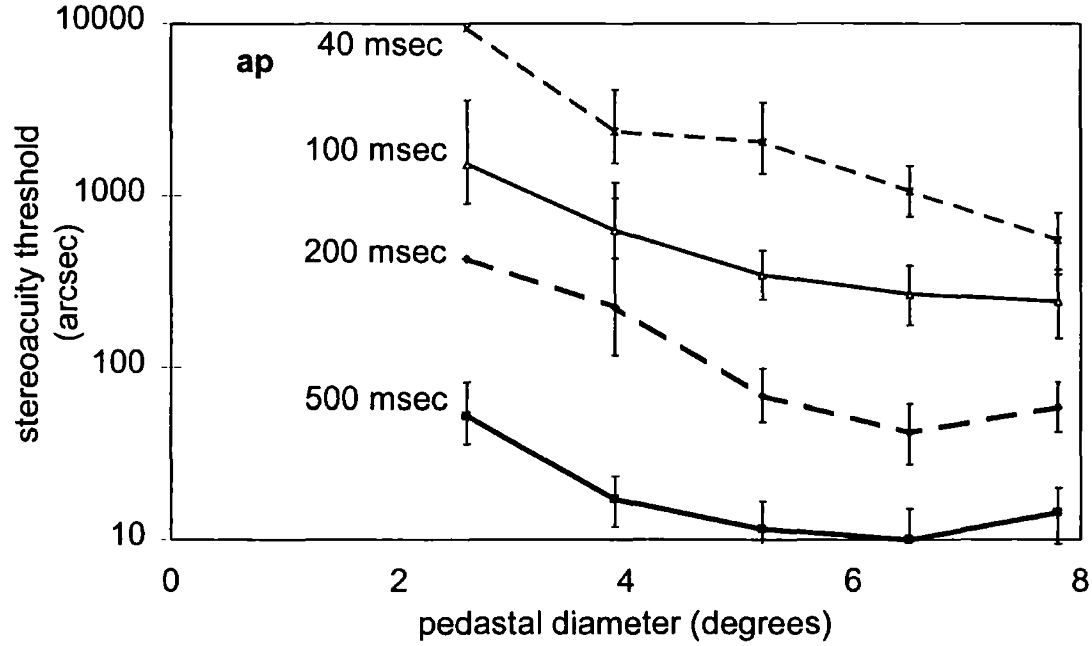


Figure 3.6. Effects of stimulus duration in AP. The effect of pedestal diameter persists even at stimulus intervals as brief as 40 msec, when vergence is not possible. However, comparison with stereoacuity thresholds in the fixation plane (figure 3.2) shows that vergence was needed for fine stereoacuity away from the horopter. These results suggest that vergence is constrained by the spatial preferences of early disparity processing, which also drives 'transient', coarse stereopsis (Richards, 1970).

AP repeated experiment 1 with briefer stimulus intervals (200, 100 and 40 msec). Stereoacuity improved considerably with increased presentation interval (figure 3.6). Depth thresholds at brief (≤100 msec) stimulus intervals were in the order of minutes rather than seconds of arc. However, these thresholds were still dependent on the spatial extent of the disparity pedestal, and indeed visual inspection shows the functions of depth threshold against pedestal disc diameter on log-linear axes were near parallel for 100 and 500 msec intervals. (As data from only one subject were available, no statistical tests were carried out). This result should be treated with caution until it is replicated by additional trained observers).

Stereoacuity at a pedestal disparity of 24′ improves considerably when the stimulus interval increases from 40 msec to 500 msec, much more so than stereoacuity in the fixation plane. The 100-fold improvement in stereoacuity with a 10-fold increase in stimulus duration is an order of magnitude larger than that obtained for lines near the horopter (Watt, 1987). Vergence, although it has a latency of 50-100 msec and a peak velocity at 150 msec (Semmlow and Carpenter, 1996) takes at least 500 msec to reach completion (Carpenter, 1988). The need for longer durations of stereoscopic stimuli at a pedestal disparity clearly implicates the role of vergence in this task.

However, the apparent continuity between the functions of stereoacuity with disc diameter at intervals ranging from 40 msec (no vergence) to 500 msec (vergence nearing completion) is evidence against the dependence of coarse-to-fine constraints in disparity processing on effecting vergence eye-movements. Instead, it supports the notion that vergence takes its input from early, coarse disparity processing, as does transient stereopsis (Richards, 1970, 1971). In their 1979 model, Marr and Poggio suggested that, as well as stimulating vergence, coarse disparities are registered in a spatial memory buffer (the 2½D sketch). Our results are in agreement with this aspect of the model.

Vergence is rapid, and under certain circumstances can have latencies as short as 50-100 msec (Masson et al., 1997). Initial vergence must therefore rely on the outcome of disparity processing in an interval perhaps as short as 50 msec. At such brief intervals, a shift in disparity spatial frequency tuning for perceived depth towards coarser modulation has been suggested (Glennerster, 1996). It is possible that initial vergence registers the outcome of the same disparity processing stream as perceived depth, but at an earlier moment in time.

In the luminance domain, low spatial frequencies are processed more rapidly (Breitmeyer, 1975; Parker and Dutch, 1987) and are thought to constrain the interpretation of higher frequencies. It is the computational efficiency of this strategy for image analysis that the Marr-Poggio model brings to disparity processing. An important feature of the Marr and Poggio model is the association between luminance and disparity spatial frequencies. It may be that because of the association between low frequencies and coarse disparity modulation, that the latter is more readily visible at brief intervals (Glennerster, 1996). Similarly, Schor et al. (1986) found that high spatial frequency stimuli were better at driving vergence to small disparities. Other evidence of this association comes from Smallman and McLeod (1997), who demonstrated that the improved stereoacuity for low-frequency stimuli at large pedestal disparities could be explained by coarse-coding of depth, as in the Marr-Poggio model (based on Richards' (1970, 1971) Three Pools hypothesis). However, the dependence of pedestal stereoacuity on spatial frequency has a direct analogy in vernier acuity, where the effect of separation is similarly dependent upon spatial frequency (Whitaker and MacVeigh, 1991). Therefore, this spatial frequency dependence may result from a problem of relative localisation rather than binocular processing per se. Both peak luminance frequency (Smallman and McLeod, 1997) and peak frequency of disparity modulation (Schumer and Julesz, 1984) for stereoacuity judgments are shifted to lower frequencies away from the horopter. In effect, this means there may be multiple disparity channels with different spatial frequency selectivity. Empirical evidence for multiple channels was found by Stevenson et al. (1992), who showed that the adaptation tuning-width to disparity is as small as 5-10 arcmin in a decorrelation detection task on broad-band stimuli.

The association between low frequencies and coarse disparities has been questioned because there is evidence that we use 'second order' information, from the pattern of high-frequency stimuli, to compute depth. Such second order information is used only for relatively coarse depth judgments (Mowforth et al., 1981; Wilcox and Hess, 1995; Statham and Georgeson, 1997) and does not necessarily signal rich monocular image recovery prior to disparity processing. In our stimuli, coarse depth judgements may have been based on first and higher order disparity. The recombination of spatial channels in the cyclopean domain to disambiguate matching noise, in metaphorical analogy with Fourier analysis in the luminance domain, and regardless of any actual association between the two, might prove to be the lasting contribution of the Marr-Poggio model.

In this experiment, vergence was not measured. Therefore, any interpretation of the results as due to vergence is entirely hypothetical, made by inference from similar experiments; there may have been no oculomotor vergence in either experiment. The evidence for vergence under similar stimulus conditions comes from Popple et al. (1998), which is chapter 2 (experiment 2.1) in this thesis, and even there vergence was measured using a 'subjective' psychophysical technique. In subsequent work, we have measured vergence objectively (Popple and Findlay, 1998b; chapter 5). A clear association between such 'subjective' vergence estimates, and concurrent 'objective' estimates obtained using binocular dual-Purkinje trackers was found.

Assuming vergence took place in Experiment 1, the similar effect of pedestal size without vergence (figure 3.6, 40 and 100 msec durations) makes it unlikely that vergence and perceived depth are processed autonomously. The literature does, however, provide considerable evidence for autonomous processing. Horizontal vergence, unlike perceived depth, responds well to absolute disparity modulation (for review see Collewijn and Erkelens, 1990). Vertical vergence has no perceptual correlate, and yet must rely the same kind of disparity processing as horizontal vergence (Stevenson et al., 1997). The non-volitional component of horizontal vergence would, according to this explanation, provide the coarse disparity code which constrains fine stereoacuity.

Previously, experimenters have failed to find coarse to fine shifting in the matching range of compound spatial frequency stimuli (Rohaly and Wilson, 1993; Smallman and MacLeod, 1997) although fine-scale stimuli can disambiguate coarse-scale information (Smallman, 1995). This may be simply because stereoacuity is not possible away from the horopter, and their experiments were designed to preclude vergence. We have shown that such shifting takes place in broad band stereograms, possibly through the motor intervention of vergence and via the scale of disparity modulation.

Chapter 4

The spatial integration of interocular correlation

4.0 Abstract

Chapters 2 and 3 explored the effects of spatial integration on the processing of stepped disparities for vergence and depth perception respectively. A cyclopean figure can be defined by interocular correlation, as well as disparity. The uncorrelated dots may be matched at a range of disparities, or undergo binocular rivalry. In this chapter, the spatial integration of decorrelation was studied.

Disparities larger than the matching limit, in random-dot stereograms, are equivalent to a region of decorrelation, because the dichoptic half-images cannot be correlated within the matching range. A vergence response to binocular decorrelation may therefore serve to bring stimuli into this range. Such a response was found in most subjects (4.1), and subsequently the spatial integration underlying this response was studied (4.1-4.4). Integration within the central area of the stimulus was smooth (4.3, 4.4), similar to disparity integration (chapters 1 and 2).

Unlike depth perception, previous studies (Cormack et al., 1994; Tyler and Julesz, 1978) have shown no spatial bounds to the integration of perceived decorrelation. However, when attempting to replicate these results (4.5) the same spatial tuning was found for decorrelation detection as for a depth detection task performed under similar conditions (3.1). Both functions were compatible with existing estimates of spatial tuning to disparity modulation in stereoacuity (e.g. Rogers and Graham, 1982).

Overall, the results in this chapter support the conclusion that decorrelation is

spatially integrated (as activity at a range of disparities) by the same neural mechanisms responsible for disparity integration per se.

4.1 The effect of surround correlation on vergence integration

4.1.1 Introduction

In the previous chapters, the data indicated that disparities for the computation of vergence are pooled over a wider area than disparities for the computation of stereoscopic depth. In the latter case, there is evidence of lateral interactions (Anstis et al., 1978; Tyler, 1983). Depth perception, unlike vergence, requires the presence of relative disparities (Erkelens and Collewijn, 1985ab). However, as pointed out in the discussion in chapter 2, optokinetic following, which like vergence can be stimulated by full-field motion, was found to be subject to lateral interactions (Miles et al., 1986). To test whether lateral interactions played a role in the computation of initial vergence, the effects of target size when the surround was either correlated and in the fixation plane, or uncorrelated, were compared. The aim was to test the hypothesis that initial vergence is enhanced by a fixation plane surround. A fixation plane surround might enhance initial vergence if the response were geared to correct for a small error in fixation, as might occur if a disjunctive saccade were aimed near the target. The role of initial, short-latency vergence in correcting small errors in fixation was suggested by Bussetini et al. (1996, see also chapter 2). In contrast, an uncorrelated surround resembles a distant (nearer or farther) environment. If the initial vergence mechanism blindly pooled disparities in the central 6° of view, the same response would be expected in both cases as the uncorrelated dots would have a mean disparity of zero, assuming that the matching range is symmetrical about the fixation depth. A target diameter of 8° was chosen because of the slight downturn in initial vergence beyond this limit, discussed in chapter 2. This was compared with a 2° target, much smaller than the integration region proposed.

4.1.2 Method

a Subjects

Eight adults aged between 23 and 30 (5 female and 3 male) participated in this experiment. All had normal or corrected to normal vision.

b Apparatus and stimuli

Apparatus and stimuli were identical to those used in experiment 2.1 (see 2.1.2). The only exception was that, in the uncorrelated condition, the dots surrounding the 12.5 arcmin crossed disparity disc were randomly positioned on the left and right monitors.

c Design

Vergence was estimated when two parameters of the stimulus were varied. As in experiment 2.1, the size of a central cyclopean disc was varied. The disc diameter was set at either 2° or 8°. In addition, the random-dot correlation in the surround was varied. The surround was either fully correlated with zero disparity, or uncorrelated.

d Procedure

For procedure, see 2.1.2. As in experiment 2.1, trials were blocked according to disc diameter. In addition, separate blocks were used for the different correlation conditions. However, unlike in experiment 2.1, a staircase procedure was used to determine the position of the dichoptic nonius lines on each trial. The centre portion of the percentage left responses was analyzed using probit as before, that is by fitting a cumulative normal function to the data. During debriefing, viewers were asked whether they saw the disc present in the stereogram (although as in the previous vergence experiment, the instructions were to attend only the dichoptic nonius display).

4.1.3 Results

The results for the eight viewers are shown in figure 4.1. Surprisingly, all except AP and IG diverged strongly when the 2° wide, 12.5' crossed disparity disc was surrounded by uncorrelated dots. AP responded no differently to correlated and uncorrelated surrounds, whereas IG converged to the uncorrelated dots beyond his already large convergent fixation disparity. As with the majority of the subjects, his response was greater when the diameter of the disc was small.

To establish whether the interaction between surround correlation and disc size was statistically significant, the data were entered into a 2 x 2 within-subjects ANOVA. Outlier IG was excluded from this analysis because of his large convergent fixation disparity, and because the interaction in his data was opposite to the other subjects'. The visually evident interaction between surround correlation and disc size was found statistically significant ($F(1,6)=17.8, p=0.006$), as were the main effects of disc size ($F(1,6)=50.5, p<0.001$) and surround correlation ($F(1,6)=11.7, p=0.014$).

In addition to the divergent response to uncorrelated dots when the disc was small, half the subjects still showed a significant difference in response to the large, 8° disc depending on its surround. Across subjects, however, this difference was not significant.

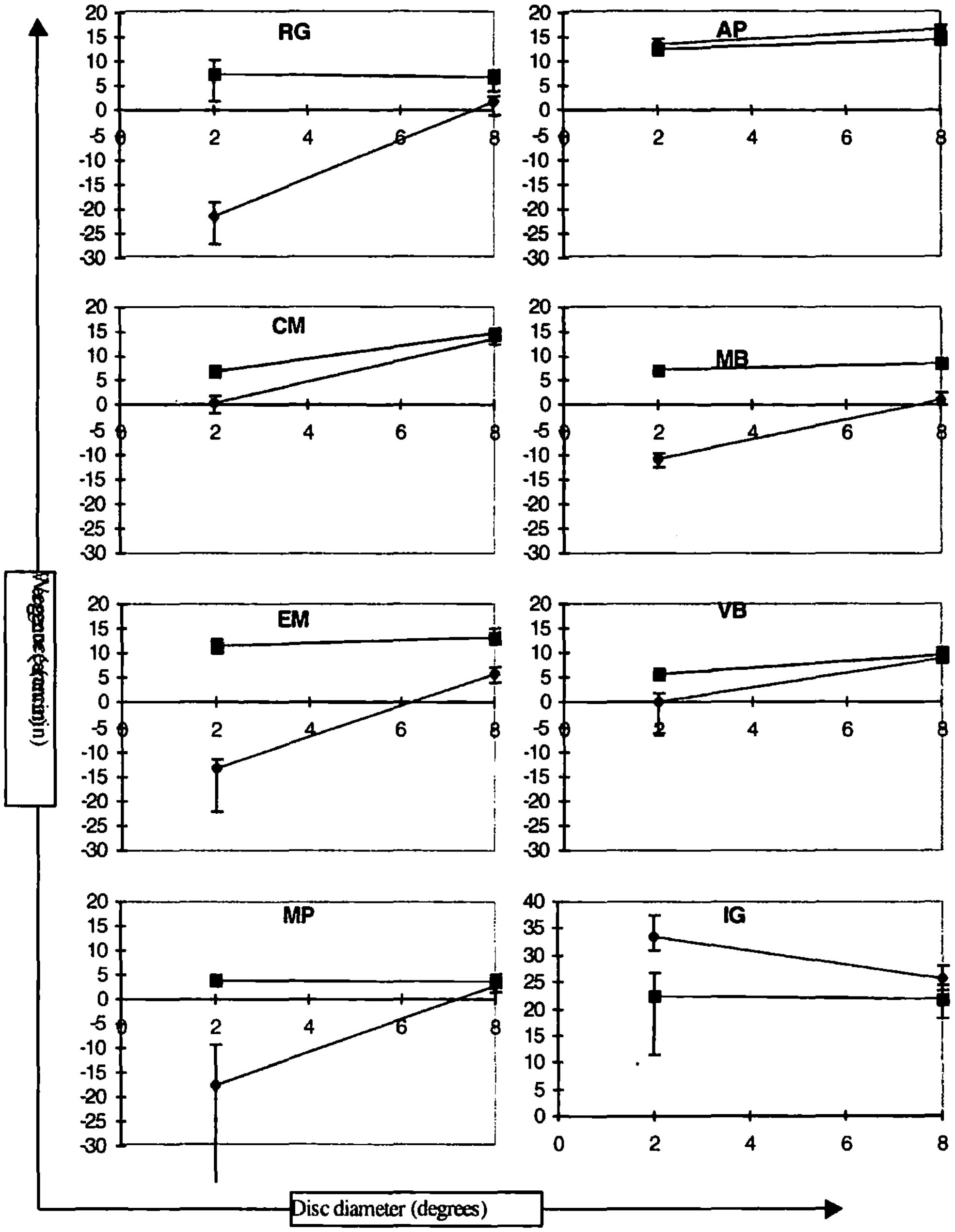


Figure 4.1. Initial vergence response (arcmin) to a 12.5 arcmin crossed disparity disc, plotted against disc diameter (degrees) for a cyclopean disc seen against a correlated surround (squares) or an uncorrelated surround (diamonds). The divergent response to the uncorrelated surround with the small disc was surprising. Note exceptions AP and IG. AP responded equally to the discs whether surrounded by correlated or uncorrelated dots. IG, whose data are plotted on a different scale from the other subjects because of his large convergent fixation disparity, converged rather than diverging in the presence of an uncorrelated surround.

Only IG reported seeing the disc in the uncorrelated surround condition, due to a lower dot density in the disc. The remaining subjects generally reported they saw the disc at a depth step when the surround was correlated. (IG is therefore an exception to the paradox whereby the same dot density is seen in uncorrelated and correlated random-dot stereograms, whereas the presence of monocular unpaired dots might be expected to increase perceived density under reduced correlation. This phenomenon was coined 'binocular frequency reduction' by Tyler in 1976).

4.1.4 Discussion

The initial purpose of this experiment was to establish whether lateral interactions involving a fixation-plane surround input (beyond an 8° target size) played a role in generating initial vergence. Across subjects, vergence to the large disc remained constant regardless of its surround. Therefore, this hypothesis can be rejected. The possibility remains that lateral interactions between, for instance, crossed and uncrossed disparities may play a role in the generation of initial vergence. However, the balance of evidence both from the present section and from chapters 1 and 2 points to a simple pooling mechanism when two disparities in the central field of view are integrated. Although the effect of disc size with a correlated surround appears smaller in figure 4.1 than it did in figure 2.2, this is because the scale of the y-axis has been enlarged to accommodate the large divergent responses made by most subjects when the surround was uncorrelated. The main effect of disc size remained significant, suggestive of spatial integration as before.

In contrast, when the surround was uncorrelated, the initial vergence response obtained was inconsistent with a centrally weighted averaging process. To reiterate the point made in the introduction, uncorrelated random dots can be matched equally across a range of disparities distributed about the fixation plane, and therefore have a mean disparity of zero. However, unlike in the case of a correlated zero-disparity surround, when the surround of a small (2°) disc was uncorrelated most naive observers diverged their eyes. This divergent response often went beyond the zero-disparity plane, despite the presence of a foveal, crossed (convergent) disparity stimulus. The present findings could be explained within the averaging model only if the matching range were biased towards divergent disparities. Then the uncorrelated dots would weigh heavier in an uncrossed direction. However, this seems unlikely as most viewers show the opposite asymmetry, with better performance both in stereoacuity and vergence towards crossed disparities.

Stevenson et al. (1994) thoroughly investigated vergence eye-movements made in response to differing disparities as the interocular correlation was varied. They found that correlation contributed to signal strength for vergence, in a similar way to luminance contrast. No initial vergence response to decorrelation was reported in their study, but such a response would be unlikely under the conditions of their experiments for a number of reasons. First, viewers were instructed to maintain fixation on a zero-disparity line abutted by dichoptic nonius to monitor their performance. All were experienced psychophysical observers, and able to follow this instruction as demonstrated by a later experiment on voluntary vergence (Stevenson et al., 1997). Secondly, even when one subject repeated the experiment with no conflicting vergence requirements, the stimulus interval was preceded by an interval containing uncorrelated dots. Any response to this stimulus, if present, would have formed the baseline for subsequent vergence calculations (these were the starting conditions throughout the experiment). Concluding the discussion of the aforementioned paper, although apparently addressing a similar subject area to the present experiment their focus was very different from ours.

The finding of a vergence response to decorrelation is not, however, without precedent. Schor and Howarth (1985) reported a tendency to verge in response to interocularly uncorrelated stimuli. O'Shea and Blake (1987) conducted a survey of a phenomenon they called 'rivaldepth', whereby a depth impression is created by the decorrelation of a region in a random-dot stereogram without introducing disparity in this region. They found that the perceived depth of the rivalrous, or uncorrelated region followed a bimodal distribution in most subjects, typically peaking at either a crossed or an uncrossed disparity beyond the fusion limit. The preferred depth was negatively associated with the subject's fixation disparity or phoria, but later found to be positively associated with their vergence response to the stimulus (phasic fixation disparity). Only the direction, and not the magnitude of phasic fixation disparity was recorded. O'Shea and Blake suggested that vergence changes were made towards the apparent depth of the stimulus. However, the results here 4.1 indicate that such vergence to decorrelation can be triggered without depth perception.

No subjects reported 'rivaldepth' in the present experiment. Several reasons for this can be put forward. O'Shea and Blake confirmed that rivaldepth is weaker with sparse stereograms, as were those here, during brief stimulus intervals, and when it is the surround rather than the centre which is uncorrelated. In addition, viewers in the present experiment were not instructed to attend to the stereogram, but to the subsequent nonius task. In conclusion, whatever triggers the initial vergence response to decorrelation, it is rapid, automatic and independent of any perceived depth.

What might be the function of a large, automatic initial vergence response to decorrelation? In a natural environment, objects or surfaces well separated in depth from the plane of fixation cannot be correlated between the two eyes' views within the disparity limits imposed by the interocular matching process. Therefore, a large vergence eye-movement that brings the plane of fixation to beyond those limits may be adaptive. If no fusion can be reached, a further 'searching' vergence movement might be triggered by the same mechanism. Such a sequence of sweeping vergence eye-movements has been suggested as a method of homing in to the appropriate vergence angle for viewing 'magic-eye' stereograms (Reimann et al., 1995). If this were the case, the data of subject IG could be explained by a preference to initiate this search with a convergent movement. Such individual differences in the search pattern could account for viewers' preferences to cross-fuse or uncross-fuse autostereograms.

One remaining question is that of how decorrelation is processed. Stevenson et al. favoured a statistical model, appropriate to their data, where the goal of early disparity processing is to compute a cross-correlation between the two eyes. The peak of this profile corresponds to the stimulus disparity, whereas the width of the distribution depends on the interocular correlation. When interocular correlation is zero, this function is completely flat. It may be the overall absence of disparity information that triggers a vergence response. If so, the response could be a default of the vergence system or driven by some aspect of the stimulus. In the next experiment, the aim was to find out which aspect of the uncorrelated stimulus might drive initial vergence.

4.2 Interocular integration of initial vergence triggered by non-fused stimuli

4.2.1 Introduction

Documented in the previous experiment is an initial, divergent response to a small (2°) crossed-disparity disc in the midst of uncorrelated dots. A similar response had previously been noted by O'Shea and Blake (1987) to 'rivaldepth' stereograms containing an uncorrelated central patch. They suggested that vergence to uncorrelated dots followed their perceived depth. However, experiment 4.1 shows such vergence in the absence of depth perceptions.

What aspect of the stimulus generated this automatic vergence response? It can be surmised that the response was stimulated by the uncorrelated dots themselves, rather than the cyclopean figure present. In a paper on vergence to anticorrelated stereograms, Masson et al. (1997) published data showing a significant non-zero saturation level of initial vergence as disparity was increased. This saturation level corresponds to vergence to uncorrelated stimuli, since, as mentioned, random-dot stereograms showing disparities beyond the matching range, that is the range where vergence is dependent disparity, are uncorrelated within this range. In this experiment, the question addressed was whether a field of uncorrelated dots alone would also trigger divergence. There were five stimulus conditions. The response to a field of uncorrelated dots was compared with fixation disparity to a field of zero-disparity, correlated dots. An additional control condition consisted of a briefly presented, blank interval where vergence was free to move from fixation disparity towards the tonic dark phoria position. Monocular fields of dots were also presented to each eye on separate trials.

A number outcomes were possible in this experiment, supporting different alternative hypotheses concerning the stimuli for vergence to uncorrelated dots. 1) The divergent response found in the previous experiment may have been due to the spread of binocular matches in the absence of a peak disparity, in which case it would require binocular stimulation and occur only with uncorrelated dots. 2) The response might be the vergence system's default in the absence of any disparity information, giving divergence in all conditions excepting a field of fixation plane dots. 3) A further possibility is that divergence is triggered by overall visual stimulation without a peak disparity, but not necessarily requiring the activity of disparity-tuned mechanisms. This latter possibility would be supported if divergence were found for monocular, as well as uncorrelated dots, but not following fixation plane dots or a blank interval.

4.2.2 Method

a Subjects

Eight viewers with normal or corrected to normal vision participated. All except the author, AP, were naive.

b Apparatus and stimuli

The apparatus was identical to that used in 4.1. Monocular stimuli consisted of a field of 2000 white dots randomly positioned in a 16° x 21° field. In the binocular stimuli these dots were either located in the same position on both left and right monitors (correlated condition) or at different locations (uncorrelated condition). In the blank condition, both screens were clear during the stimulus interval. As mentioned in 4.1.3, 2000 monocular dots will have the same apparent density as 4000 uncorrelated dots because of binocular frequency reduction (Tyler, 1976).

c Design

There were five different stimulus conditions. These were as follows: 1) binocular, uncorrelated dots; 2) binocular, correlated dots; 3) monocular dots in the left eye; 4) monocular dots in the right eye and 5) a blank interval following the fixation stimulus.

d Procedure

The procedure was similar to previous initial-vergence experiments (chs. 1,2; 4.1), i.e. each trial started with a 1 second fixation, followed by a 230 msec stimulus, and a 160 msec nonius interval. However in this case the five stimulus conditions were randomised across blocks of 200 trials. Five concurrent staircases determined the position of the nonius stimulus for the five conditions, and again only the central portion of the left response frequency was used to determine the central tendency of the response.

4.2.3 Results

Figure 4.2 shows the results for all eight subjects. Excepting AP, whose data contain no difference in vergence between the different conditions, viewers diverged somewhat following the monocular stimuli, and considerably more after the uncorrelated stimuli. Conversely, there was little or no change in fixation disparity following a blank interval.

Statistical tests carried out on the data showed a significant effect of stimulus condition (F(4,28)=10.32,p<0.001). There were no significant differences between the two monocular conditions, or between the two control conditions (correlated dots, blank interval). However, all other partial comparisons yielded significant differences (correlated vs. uncorrelated: F(1,7)=12.99, p=0.009; uncorrelated vs. monocular: F(1,7)=11.64, p=0.011; correlated vs. monocular: F(1,7)=5.79, p=0.047). The divergence to uncorrelated binocular stimuli was greater than that found for monocular stimuli, which in turn was greater than the control conditions.

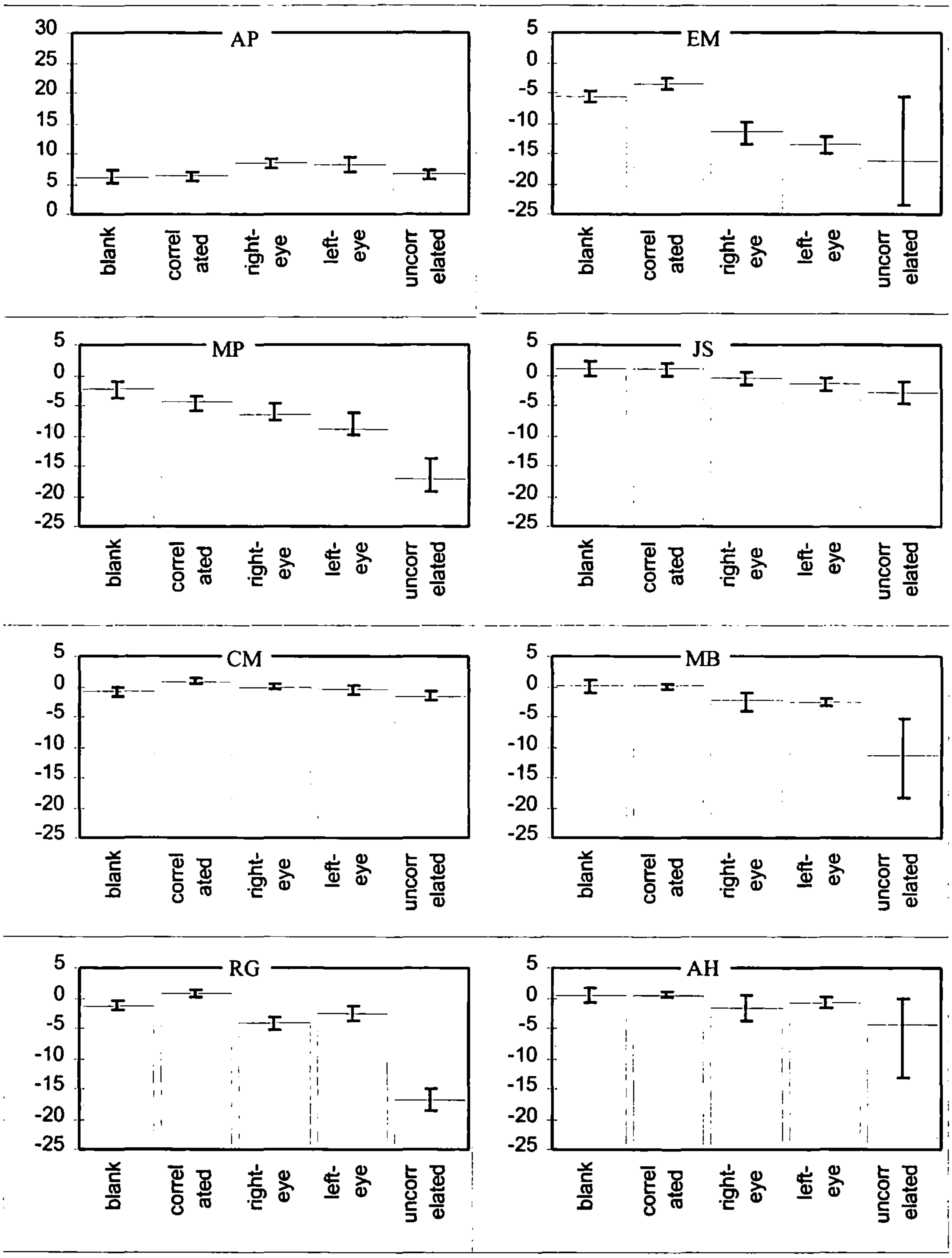


Figure 4.2. Vergence to brief, uniform random-dot stimuli. The initial vergence response (arcmin) to uncorrelated and monocular stimuli is shown. In addition, the two control conditions (correlated dots and blank interval) are illustrated.

The response to uncorrelated dots was generally divergent, in the same direction as the responses in the monocular conditions, but larger in magnitude. This is shown in the table below:

	avp	emc	mgp	js	cm	mb	rg	ah	MEAN
left-eye	8.27	-13.5	-8.83	-1.49	-0.53	-2.48	-2.53	-0.75	-2.73
right-eye	8.56	-11.4	-6.46	-0.53	-0.02	-2.42	-4.16	-1.53	-2.25
binocular	6.62	-16.3	-17.00	-2.76	-1.52	-11.40	-16.80	-4.46	-7.95

Table 4.1. The divergent response to uncorrelated dots is generally larger than in both monocular conditions.

4.2.4 Discussion

As expected, divergence was found to a field of uncorrelated random dots. This was not due simply to the lack of binocular fusion, as evidenced by the viewers' tendency to maintain vergence position at fixation disparity in the blank interval condition. Divergence was also found with monocular stimulation, however this was less than in the uncorrelated condition. This result supports a hypothesis that divergence to uncorrelated dots is stimulated by monocular activity in the absence of binocular fusion or correlation. However, it also remains possible that the divergent responses to monocular and binocularly uncorrelated stimuli were driven by entirely different mechanisms.

If this were the case, perhaps the vergence response to uncorrelated stereograms is a function of the initial vergence found in response to anticorrelated dots (Masson et al., 1997). The latter followed the disparity tuning profiles of those binocular neurones which responded to anticorrelated stereograms (Cumming and Parker, 1997). Uncorrelated dots might weigh more heavily on anti-correlated than correlated patterns, because the greater number of positive nodes (Ohzawa et al., 1997). This makes chance matches likelier with the inverted receptive field profiles, explaining the results reported here. A more parsimonious theory would, however, integrate the divergent responses found to monocular and uncorrelated stimuli.

A vergence response to brief (200 msec) monocular stimuli was recorded by Jones (1976). Little has been made of this finding in the literature, but as discussed in 4.1 such a response would be adaptive. If an object were foveated in one eye, a tendency to make a vergence eye-movement could bring that object into fusion range in the other eye. However, such eye-movements would be wasteful in the absence of any visual input.

The integration of activity across monocular neurones at an early level of visual processing, for example V1 or below, might trigger a reflex, searching vergence eye-movement. This reflex could be suppressed by the process which determines a stable disparity for fusional vergence movements. Vergence recordings taken from subjects under the influence of drugs which disinhibit reflex responses may provide support for this idea, if sudden deviations from fixation disparity at the onset of visual stimuli were observed. Indeed, many viewers habitually verge to transient stimuli even under normal sensory conditions (Edwards et al., 1998).

The aim in the next section was to explore precisely how disparity driven vergence is integrated with the initial vergence response to decorrelation.

4.3 Integrating correlated and uncorrelated dots at the same location

4.3.1 Introduction

Very little has been written about initial vergence to binocularly non-corresponding stimuli, although many viewers respond in this way (Jones, 1976; Edwards et al., 1998). Such a response can be considered adaptive, because it can potentially bring stimuli outside Panum's fusional area and outside disparity processing limits into correspondence. To serve this function, an all-or-nothing response would be ideal. There would be little point in making a small vergence eye-movement when some degree of decorrelation was present. There were quite large responses to uncorrelated stereograms in most viewers (4.1 and 4.2), but smaller changes in vergence following monocular stimulation (4.2). This suggested an integration of monocular activity in the initiation of such searching vergence eye-movements. Although less adaptive than an all-or-nothing response, such integration may operate on the same principles as the disparity integration seen in chapter 1. This kind of integration is suited to the one-stage automatic disparity pooling process which determines the goal of initial vergence, as suggested in chapter 2.

The aim was to determine how decorrelation is integrated. The stimuli chosen were 1-bit random-dot stereograms, as the percent decorrelation in these is easy to quantify. The following equation describes the cross-correlation (φ) between the two eyes' images (I_R, I_L) as a function of disparity(d):

$$\varphi(d) = \int I_R(x) I_L(x + d) dx$$

(1)

With uniform 1-bit stereograms, where 50% of dots are white and 50% black, this reduces to a function of the proportion of matching dots at each disparity (P_d):

$$\varphi(d) = 2P_d - 1$$

(2)

Clearly, when the left and right images are uncorrelated, P_d is always equal to 50%, therefore the cross-correlation is distributed about zero. Because of sampling limitations, noise from false matches will cause spurious little peaks in this function. Fifty percent is also the probability of chance matches at any disparity except the stimulus disparity, when this is defined. When the images are identical, the cross-correlation at zero disparity is equal to one, a delta function with a width equal to the size of a stereogram element surrounded by sampling noise. As the percentage correlation is reduced, this peak is submerged until it eventually becomes enveloped in the surrounding noise.

Vergence was estimated as a function of percent correlation. This is analogous to chapter 1 as a stereogram containing a reduced percentage of correlation can be interpreted (although not literally seen) as a zero-disparity fixation plane overlapping an anticorrelated 'rivaldepth' plane. An alternative description would be a plane of dots surrounded by a cloud of dots at different depths. On the cross correlation profile, the transparent stimuli in chapter 1 appeared as a bimodal distribution. With reduced correlation, a multimodal distribution is obtained.

4.3.2 Method

a Subjects

Only four could be found for this experiment. These included the author AP who, in any case, did not respond markedly to decorrelation as can be seen from 4.1 and 4.2. Additionally, data from four new paid student subjects had to be discarded as the variability in their responses suggested they were simply not following the instructions.

b Apparatus and stimuli

The apparatus, as before, was an Ambra 486 PC with a VSG stimulus generator card with two EIZO Flex-6500 monochrome monitors. This time, the monitors were positioned at a distance of 1.05 m to ensure each pixel subtended almost exactly 1 arcmin. Stimuli were 10° square binary stereograms composed of 1 arcmin elements drawn from a random pattern of numbers from 0-99. The numbers were reinterpreted on each trial to give the desired percentage of correlation between the two monitors and create a new random pattern. The fixation stimulus consisted of a 10° mid-grey square. The dichoptic nonius consisted of abutting 1 x 20 arcmin white lines on this grey square, positioned centrally on the left and right monitors. The monitors were viewed through the large 25 x 20.5 cm mirrors described in previous chapters, mounted at right angles to each other and positioned directly in front of the viewer, whose head rested on a chin rest.

c Design

The percentage dot correlation was varied. There were six levels of correlation; 0, 20%, 40%, 60%, 80% and 100%. The latter, of course, is a measure of fixation disparity. The dependent variable in all cases was initial vergence. At 0% correlation, the images in the left and right eyes are statistically uncorrelated. This means that, because of chance matches in the binary pattern, about 50% of the dots match and 50% do not match. If all the dots did not match, this would be a correlation of −100%.

d Procedure

There were 300 trials in each session, divided into four blocks of 75 trials. Fifty trials in each of the six conditions were randomised across each session. Each trial consisted of an approx. 1 second fixation during which the stimulus was calculated off-line, a 200 msec stimulus interval followed by a 160 msec nonius interval. The position of the nonius on each trial was determined by a staircase procedure that operated independently for each of the experimental conditions. A judgment of whether the top line was left or right of the lower line was indicated using a mouse-button press. The central portion of the results under each condition was fitted with a cumulative normal using Probit, to determine the estimate of subjective vergence.

4.3.3 Results

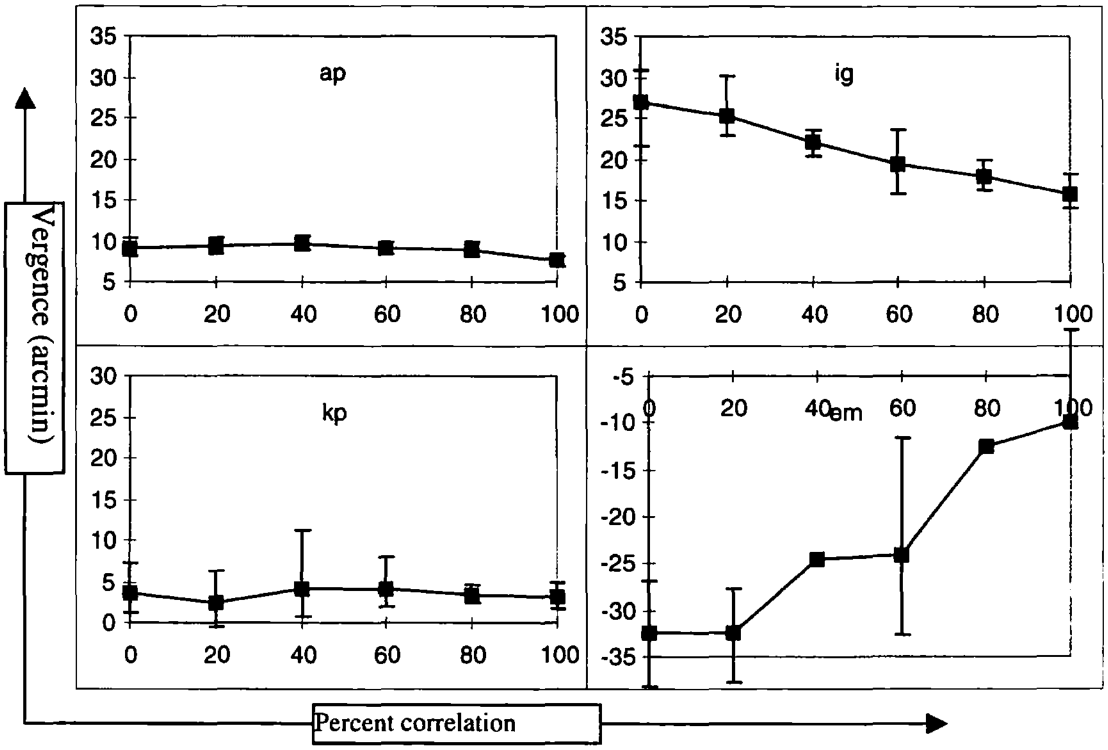


Figure 4.3 The integration of decorrelation in triggering initial vergence. Percent correlation is plotted against vergence (arcmin). AP and KP did not respond to decorrelation (i.e. their response in the uncorrelated conditions was no different from the response in the 100% correlated condition). EM diverged increasingly as decorrelation increased (going from right to left on the graph). IG converged as a smooth function of percentage decorrelation. (These results are described as a function of decorrelation rather than correlation, since the 100% correlated response, which represents the viewers' fixation disparity, is here the norm)

The results are shown in figure 4.3. Of the four subjects, AP and KP did not make a vergence response to the decorrelated stimuli. A general conclusion from these data is that the response to decorrelation, if present, is smoothly integrated with the weight of stimulation at the fixation disparity. This is the case both for a divergent response to decorrelation (EM) and a convergent response (IG).

4.3.4 Discussion

The results suggest responses to uncorrelated and correlated elements of the stimulus are integrated, like the integration of different disparities found in chapter 1. The vergence response could be put down to the prevalence of 'near' and 'far' disparity-detectors in the primary visual cortex. This would account for considerable individual variation. Richards (1970,1971) reported selective crossed and uncrossed stereoanomalies in the detection of coarse disparities in brief stimuli, which Jones (1977) later employed in his attempt to explain vergence responses to transient stimuli. Neurones classified as 'near' and 'far' show greater ocular dominance than 'tuned' cells (Poggio and Talbot, 1981), and therefore might respond more vigorously to monocular stimulation. It was suggested in chapter 2 that initial vergence might result from pooling the responses of disparity sensitive cells over a fairly large region. Does the integration of decorrelation follow similar spatial bounds? The results of 4.1 indicated that this is so. Decorrelation beyond 8° had little effect on vergence. However, the function of vergence to decorrelation within the pooling region had not been studied. Therefore, this function was explored in greater detail in the next section.

4.4 Integrating neighbouring correlated and uncorrelated regions

4.4.1 Introduction

Chapter 2 showed that the spatial integration of disparities determines the initial vergence response. In section 4.3, the local integration of decorrelation was likened to the integration of two transparent disparity planes (chapter 1). If decorrelation is processed by the same spatial pooling mechanism as disparity, and indeed represented as a bias in the disparity spectrum due to the uneven individual distribution of disparity detectors, then the vergence response to it should follow the same spatial function as disparity integration. Evidence that this is so was given in section 4.1, where decorrelation, like disparities, beyond an 8° region made no difference to the vergence response. However, as only one other data point was present in that experiment, it was necessary to repeat the main experiment of chapter 2 with an uncorrelated central figure. In this case, the same stimuli were used as in the previous section, but instead of varying the percent decorrelation over the entire stimulus, the size of a central uncorrelated region was varied.

4.4.2 Method

a Subjects

Three viewers participated. As before, these included the author AP who, in any case, did not respond markedly to decorrelation as can be seen from 4.1 and 4.2.

b Apparatus and stimuli

Apparatus and stimulus generation were as in 4.3.2b. The stimuli were 100% correlated except for a 0% correlation central square of variable size.

c Design

Initial vergence was estimated as the size of the uncorrelated square was varied. The extent of a square uncorrelated figure was 0.4°, 0.8°, 1.6°, 3.3° or 6.7°. In this experiment, there was no measure of fixation disparity.

d Procedure

There were 300 trials in each session, divided into five blocks of 75 trials. Sixty trials in each of the five conditions were randomised across each session. As in 4.3.2d, each trial consisted of an approx. 1 second fixation during which the stimulus was calculated off-line, a 200 msec stimulus interval followed by a 160 msec nonius interval. The position of the nonius on each trial was determined by a staircase procedure that operated independently for each of the experimental conditions. A judgment of whether the top line was left or right of the lower line was indicated using a mouse-button press. The central portion of the results under each condition was fitted with a cumulative normal using Probit, to determine the estimate of subjective vergence.

4.4.3 Results

The results are shown in figure 4.4. Of the three subjects, AP and MO did not make a vergence response to the decorrelated stimuli. However, MP responded smoothly to the increasing stimulus size, showing a saturation of the vergence response as the area of the decorrelated figure was increased.

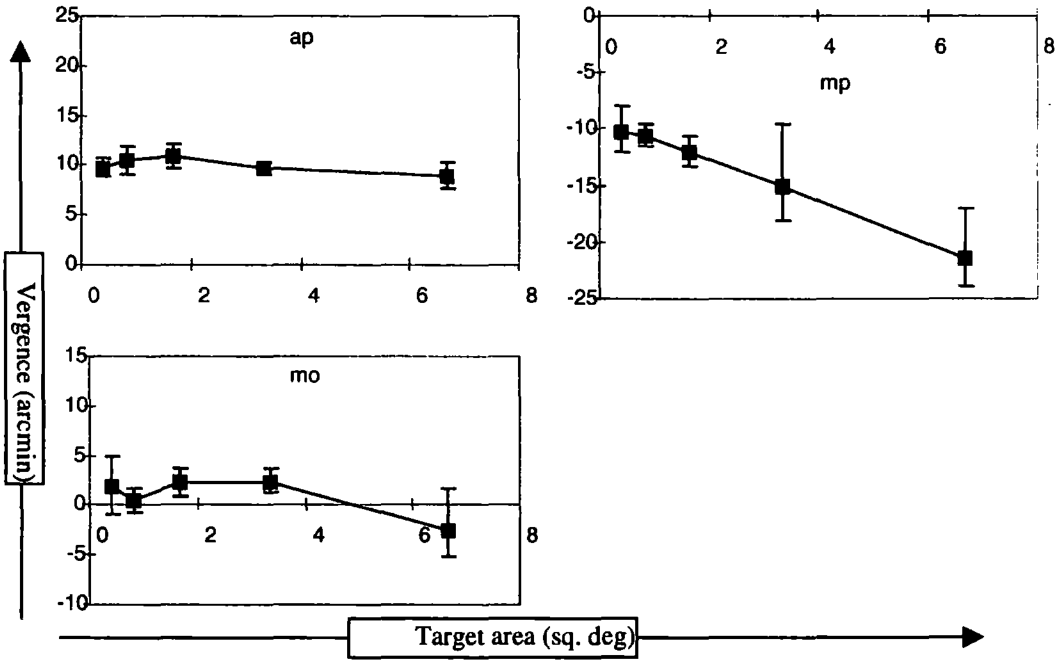


Figure 4.4 The spatial integration of decorrelation. Target area (sq. deg.) is plotted against vergence (arcmin). AP and MO did not respond to decorrelation. Note different scales on y-axis.

4.4.4 Discussion

As predicted in the discussion section 4.3.4, spatial integration of decorrelation was comparable to the spatial integration of disparities. This new theory can now be phrased more neatly. Initial vergence is directed towards a spatially integrated estimate of disparity in the centre of view (chapter 2). This is probably based on the operation of low spatial frequency and/or coarse disparity selective neurones, as are depth judgments to very brief stimuli (chapter 3). As some individuals have an imbalance between the very coarse, 'near' and 'far' types of neurones which are also more subject to ocular dominance, local monocular activity, whether from a monocular stimulus or from a region of decorrelation between the two eyes, could bias this pooled disparity estimate towards the predominant (near or far) disparity. Both simple and complex binocular cells exist, and their differing temporal preferences could account for the difference between transient and sustained responses reported by Jones (1976), assuming the imbalance of near and far pools could be reversed across these two groups of cells. Alternatively, the transient vergence documented here might be governed by the integration mechanism described, whereas phoria and the closely related measure of fixation disparity could be determined by growth (including the changing position of the eyes in the head) and the maturation of response patterns.

Although the data provide some support for this theory, a much larger number of subjects would have to be tested before it could be generally accepted. If groups of subjects were found with such convergent and divergent response tendencies, these would have to be associated with corresponding anomalies in depth judgment. Animal studies, which have purported to show imbalances between 'near' and 'far' neurones, might further determine whether these are associated with similar biases in transient vergence. Yet another explanation would have the response to uncorrelated or monocular stimulation weighted against fusional stimulation at a later stage in the vergence pathway.

In chapter 3, the suggestion was made that vergence is generated by the early stages of a process which later resulted in perceived depth. Is this also the case with the integration of decorrelation, which might be processed by the very same mechanism that pools disparity for vergence? The spatial properties of decorrelation detection were investigated in the next section so a comparison could be made with depth detection and stereoacuity (3.1). This is similar to chapter 3, except, of course, that the central figure of the stereogram consisted of decorrelated rather than disparate dots.

4.5 Spatial limitations of decorrelation hyperacuity

4.5.1 Introduction

Section 4.3 showed that decorrelation and disparity might be pooled by the same spatial integration mechanism in the generation of an initial vergence response. The notion was raised that decorrelation resulted in a skewed coarse disparity value being integrated together with other local disparities. Although such imbalances in stereoacuity between crossed and uncrossed disparities are less common (Richards, 1970;1971), the same principle could hold for the detection of decorrelation providing it is not assumed this must have a determinate depth. Thus, a region containing a small percentage of decorrelation might appear at many different disparities, and therefore less flat than its correlated surround.

Tyler and Julesz (1978; Julesz and Tyler, 1976) showed that viewers are acutely sensitive to a small amount of decorrelation (<2%), and also full (100%) decorrelation at extremely short (<5 msec) intervals. They too explored the spatial properties of correlation integration, as did Cormack et al. (1994). Neither group found any evidence of lateral interactions, as with stereoacuity. Both suggested that the detection of correlation is based on statistical sampling of individual elements, integrated both spatially and temporally until some limit is reached. However, the largest stimulus used by either group was only 5° wide. By going slightly beyond this limit, there was a chance of finding some evidence of spatial selectivity as found with stereoacuity (4.1). An additional difference between these published studies and the present experiments was that the stimuli others used were dynamic random-dot stereograms rather than the static random-dot stereograms employed here.

4.5.2 Method

a Subjects

Four observers, who except AP were naive, participated in this study. All had normal vision.

b Apparatus and stimuli

The apparatus and stimuli were essentially identical to those used in 4.3.2b. However, there was no dichoptic nonius stimulus, and instead the percentage decorrelation of the decorrelated square was set at five fixed values. These values were 0 (no square), 4%, 8%, 16% and 20% on the easy scale or 0 (no square), 2%, 4%, 8% and 10% on the hard scale. Note that these are the percentage of *de*correlation. For the percentage correlation, they must be deducted from 100% i.e. a 4% decorrelation implies 96% correlation.

c Design

As in 4.4, the size of a decorrelated square was varied (0.4°, 0.8°, 1.6°, 3.3° or 6.7°). However, this time the dependent variable was decorrelation sensitivity thresholds. This was defined as the percent decorrelation in the square required for detection.

d Procedure

As in 4.3.2d, sessions consisted of 300 trials presented in four blocks. There were ten trials in each condition, randomised across the sessions. Viewers were instructed to respond (with a mouse-button press) only if they definitely saw a square, so that the results could neatly be fitted with a cumulative normal from 0 to 100% response frequencies (rather than the 50% to 100% which would have been appropriate had the subjects been allowed to guess). In this way, the task equivalent of a 75% decorrelation detection threshold was obtained.

4.5.3 Results

All observers had the lowest thresholds for a 0.8°-3.3° square, and less well when the square was either 0.4° or 6.7° wide. The function of decorrelation thresholds plotted against stimulus size was similar to that found for stereoacuity thresholds (3.1). To make a visual comparison more convenient, the stereoacuity data are shown alongside the decorrelation data for those subjects who completed both experiments (AP and EM in figure 4.5).

4.5.4 Discussion

The results confirm the hypothesis that decorrelation is integrated by the same mechanism as disparity, in the determination of perceptual as well as oculomotor outputs. This can be concluded from the similar functions for both stereoacuity and decorrelation 'hyperacuity' as target size is varied. As in the previous section, however, these results must be treated with caution because of the small number of subjects who participated in the experiment. Nevertheless, the trough at about 1.5° in the case of stereoacuity was supported by a volume of evidence on the spatial frequency selectivity to disparity modulation in stereoacuity, described in chapter 3. The similarity between the two functions, although qualitative, is striking.

4.5.4 Discussion

The results confirm the hypothesis that decorrelation is integrated by the same mechanism as disparity, in the determination of perceptual as well as oculomotor outputs. This can be concluded from the similar functions for both stereoacuity and decorrelation ‘hyperacuity’ as target size is varied. As in the previous section, however, these results must be treated with caution because of the small number of subjects who participated in the experiment. Nevertheless, the trough at about 1.5° in the case of stereoacuity was supported by a volume of evidence on the spatial frequency selectivity to disparity modulation in stereoacuity, described in chapter 3. The similarity between the two functions, although qualitative, is striking.

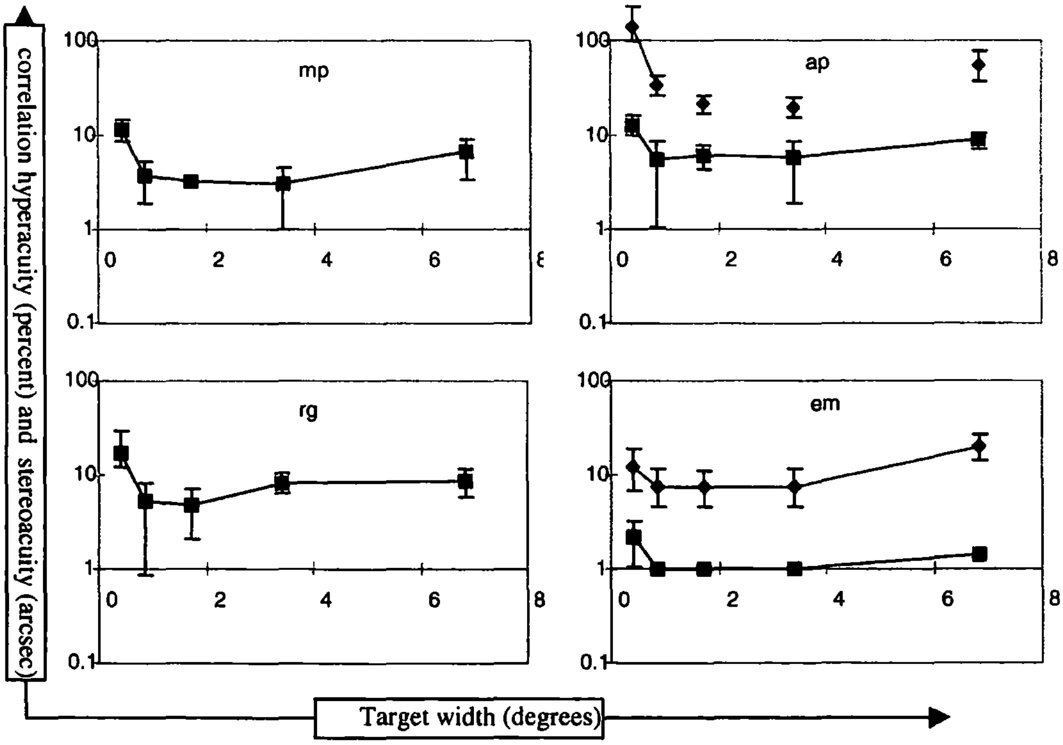


Figure 4.5. Decorrelation thresholds with increase in size of a decorrelated square (% decorrelation threshold against size in degrees). For AP and EM, stereoacuity thresholds from figure 3.4)diamonds) are plotted alongside decorrelation thresholds (squares)(stereoacuity thresholds in arcmin against disparity disc diameter in degrees).

These results stand in contradiction with previous studies (Tyler and Julesz, 1978; Cormack et al., 1994) where the detection of decorrelation following a correlated interval, and correlation following an uncorrelated one, was found to be dependent on the number of elements (integrated both spatially and temporally) up to about 10000. Based on their results, decorrelation thresholds were expected to reach a saturation point with the 1.6° square, which consisted of 10000 elements. A number of explanations for this discrepancy between the present data and others' can be suggested. First, the elements used here were smaller (1 arcmin instead of 10 and 3 x 5 arcmin). Perhaps, therefore, the edge of the stimulus could not be resolved sufficiently beyond a certain eccentricity. The same explanation would also apply to the stereoacuity data, where 0.8 arcmin elements were used. Another possibility is that decorrelation integration in dynamic random dot stereograms is subject to a different process from static integration. Further, and finally, as mentioned in the introduction the largest stimuli in the present study were bigger than any used by others, and therefore might have picked up a downturn in decorrelation hyperacuity beyond the scope of previously published studies.

Although the theory of decorrelation detection using a stereo mechanism seems appealing, there may be other processes which can also be used to detect decorrelation. The ability to detect decorrelation in an entire field (Tyler and Julesz, 1978) apparently conflicts with the inability to perceive full field disparity change (Erkelens and Collewijn, 1985ab). However, we are able to perceive transparency between two disparity planes (see chapter 1 for references). A decorrelated field could be represented as multiple planes. Nevertheless, one would expect full field decorrelation thresholds to be higher than those for a decorrelation step. Such a difference has been found between stereoscopic hyperacuity (for a depth step) and superresolution (for a thickening in depth) (Stevenson et al., 1989). No such decrement has been reported in the case of decorrelation.

To conclude this chapter, an adaptive initial vergence response to decorrelation was found (4.1). This was first described in terms of monocular activity triggering a large, generally divergent, eye-movement that brought fixation to beyond Panum's area, perhaps as part of a search for interocular correspondence (4.2). However, the spatial integration for this response could more easily be described if the monocular activity was processed by coarse-disparity binocular neurones which were then pooled using the same integration mechanism as disparity vergence (4.3). The idea that binocular visual neurones are used to process monocular stimuli is explored further in the next chapter.

Chapter 5

Visual direction and local disparity integration

5.0 Abstract

Ono (1991; Shimono et al., 1998) reported a difference in the perceived position of alignment between two dichoptic vernier targets located in the figure and background areas of a Julesz random-dot stereogram. His finding implies that monocular alignment across a disparity step is pulled in the direction of the shifted half-images. We tested this prediction in experiment 5.1. Stimuli were dense 8° cyclopean stereograms containing a single, central horizontal 11' depth step between two flat, abutting panels. A monocular vernier target consisting of two 29' vertical lines separated by 15' was in the centre of the image. The position of perceived alignment of this target was determined for 8 viewers using a forced-choice paradigm. There were two independent variables: the vernier target was presented to the left or right eye, and the top panel was behind or in front of the lower panel. The interaction between eye and depth profile was significant. The perceived alignment of the monocular vernier was shifted in the direction of the surrounding half-images by 1'-6'. Monocular alignment is dependent on the surrounding binocular context. This conclusion is in contradiction with a simple local-sign account of vernier acuity, and with Hering's Laws of Visual Direction. However, our results and others' can be modelled in terms of the binocular neurones described by Ohzawa et al. (e.g. 1996). The topographic organisation of the visual cortex makes it ideally suited for the computation of alignment and visual direction.

Experiment 5.1 suggested that the nonius procedure may give a misleading estimate of vergence, because of the role of disparity integration in the determination of

visual direction. Therefore, in experiment 5.2 we tested the accuracy of this method as used in previous chapters. Unlike the dichoptic nonius used by Ono (1991; Shimono et al., 1997) we used a brief (160 msec) dichoptic nonius which was subsequent to, and temporally distinct from, the presentation of a disparity target, to estimate vergence. A concurrent estimate based on the output of binocular dual-Purkinje trackers was also calculated. The subjective estimates correlated well (98%) with objectively measured vergence for two observers who responded appropriately to the disparity stimulus, and gave an indication of vergence error for a third subject, when the response to the stimulus was poor (72% correlation). We can conclude that the nonius method, when used carefully as in this thesis, can provide a viable estimate of initial vergence. What is more, the nonius method is not subject to the vagaries of eye-tracker operation and the tedious analysis of reams of data. These may be reserved for more complex tasks than the estimation of initial vergence.

5.1 Is monocular visual direction computed by binocular neurones?

5.1.1 Introduction

Ono (1991) reported that dichoptic nonius lines presented in the figure and background areas of a Julesz random-dot stereogram aligned at different 'vergence' angles, although vergence eye-movements were precluded by the brief (100 msec) stimulus interval. He also showed that the extent to which the lines had to be non-corresponding covaried with the disparity in the stereogram. As figure 5.1 illustrates, the difference in alignment between the dichoptic nonius lines in the figure and background areas implies that the two monocular portions of each dichoptic vernier are themselves misaligned across the disparity step, when they are perceived to be collinear.

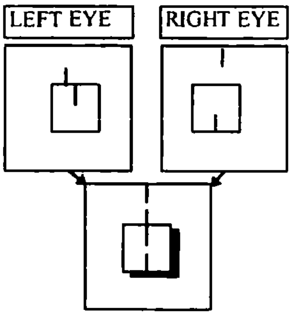


Figure 5.1 (Adapted from Ono, 1991). Schematic representation of a Julesz random-dot stereogram showing the actual positions of two pairs of dichoptic nonius lines in the two eyes' half images, and their perceived positions in the combined image.

This finding violates the first of Hering's (1879|1942) 'laws of visual direction' (0.4) explicitly, since the monocular visual direction of the dichoptic nonius lines does not transfer unaltered to the cyclopean eye, and implicitly, as the two lines in the right eye appear misaligned when they in fact have the same oculocentric direction, and conversely appear aligned when they do not. (The same is true of the two lines in the left eye). The laws of visual direction assign a unique monocular direction which should be unaffected by any binocular relationship. If the monocular misalignment is also large enough to fall outside the range of likely errors of judgment, it brings into question existing theories of vernier acuity. To accommodate the finding, the processing of monocular alignment must be subject to interaction with binocular processing, perhaps even computed in the same population of neurones. Ohzawa et al. (1996) recorded from a large number of binocular simple cells in cat V1 whose receptive field profiles in the two eyes make them suited to the first stages of calculating both binocular disparity and 2-dimensional form. Cooperative interactions between such neurones tuned locally to the stereogram disparity might explain why a monocular stimulus is shifted in the same direction as the monocular half-image of the stereogram. Shimono et al. (1998) explored in detail the effects of the spatial proximity of the dichoptic lines to the disparity stimulus. In his study, the monocular portions within each eye were widely spaced and although one was adjusted in relation to the other, the details of these adjustments were not reported.

We set out to quantify and model this monocular misalignment.

5.1.2 Method

a Subjects

Eight subjects participated in this study. All except the author were naive as to the purpose of the experiment. All had normal or corrected to normal vision. Their report of normal binocular vision was confirmed by their ability to see the depth step present in the stimuli.

b Apparatus and Stimuli

Stimuli were 7.3° square 16-shade dense stereograms composed of 0.73 arcmin random dots. There was a horizontal disparity step of 11 arcmin in the centre of the stereogram. Twenty such stereograms were stored in memory and presented in a random order, at either disparity profile. A vernier target consisting of two 29 arcmin long, 0.73 arcmin wide maximum luminance bars separated by 16 arcmin was located centrally on one half-image of the stereogram. Small (< 0.73 arcmin) displacements of the vernier lines were effected using sub-pixel shifting (see chapter 3). Stimuli were calculated on an Ambra PC with a CRS VSG card and displayed on two EIZO Flex-6500 monochrome monitors, viewed through a modified Wheatstone stereoscope. The effective screen size was 25.5 x 34 cm, with a resolution of 768 x 1024 pixels. The mirrors were 25 x 20.5 cm. The viewing distance was 155 cm (10 cm to the mirrors, and a further 145 cm to the monitors), such that each pixel subtended a visual angle of approximately 0.73 arcmin. A mid-grey 7.3° square replaced the stimulus between trials.

c Design

A separated monocular vernier was presented in the left or right eye. This was superimposed on a random-dot stereogram containing a cyclopean disparity step between the upper and lower halves of the screen. The direction of the disparity step was varied.

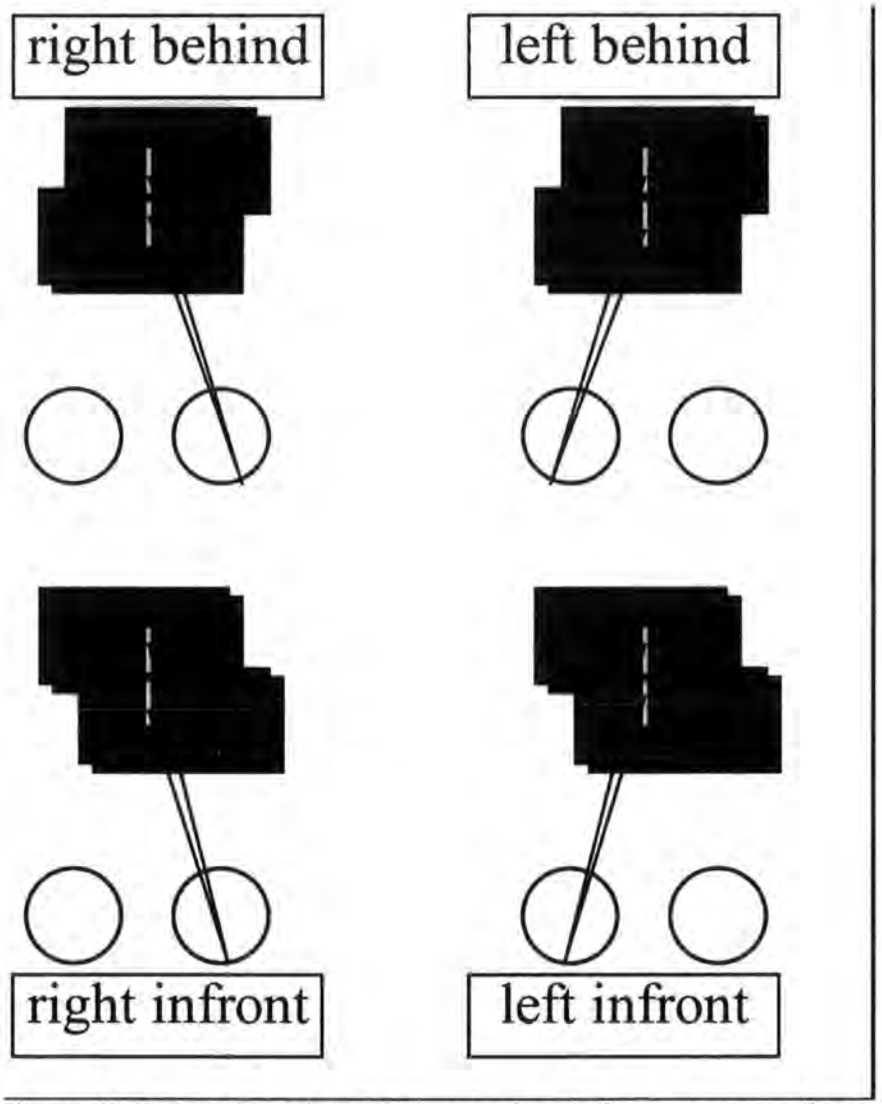


Figure 5.2. Design of experiment 5.1 - monocular misalignment across a disparity step.

d Procedure

Each subject completed one session. Trials were presented in four randomised blocks of 100. Overall, in every session 100 trials were presented in each of the four conditions. Each trial consisted of a fixation (approx. 1.3 s) while the stereogram was retrieved from memory, followed by a 100 msec stimulus interval. Concurrent staircases were used to estimate the perceived alignment of the monocular vernier in each condition. Viewers responded with a mouse-button press whether the top line was left or right of the lower line, and the position of the line was adjusted accordingly on the next trial in the same condition. Response frequencies at each angle presented were later fitted a cumulative normal function using probit, to determine the central tendency and hence the position of perceived alignment.

5.1.3 Results

As predicted, we found an interaction between the eye in which the vernier was presented, and the depth profile depicted. This interaction was statistically significant (f(1,7)=33.75,p=0.001), and visually compelling (see figure 5.3). In the pooled data from all eight subjects, this interaction crosses over the central line although most of the individual subjects show some bias in mean alignment away from the centre. These results confirm that the top line is perceptually shifted in the same direction as the surrounding stereogram half-image.

The error bars on the individual data in figure 5.3 show 75% vernier acuity thresholds. These are quite wide (1-6 arcmin), probably because of the short duration of the stimulus (100 msec), the low contrast of the vernier target and the separation between the two lines (15 arcmin). (The effects of visibility and timing on vernier acuity were studied by Waugh and Levi in 1993). Despite this limited acuity, the error bars in many cases do not contain the point of zero misalignment. The misalignment therefore not only follows a regular pattern depending on the stimulus, but can also fall outside the normal acuity margins.

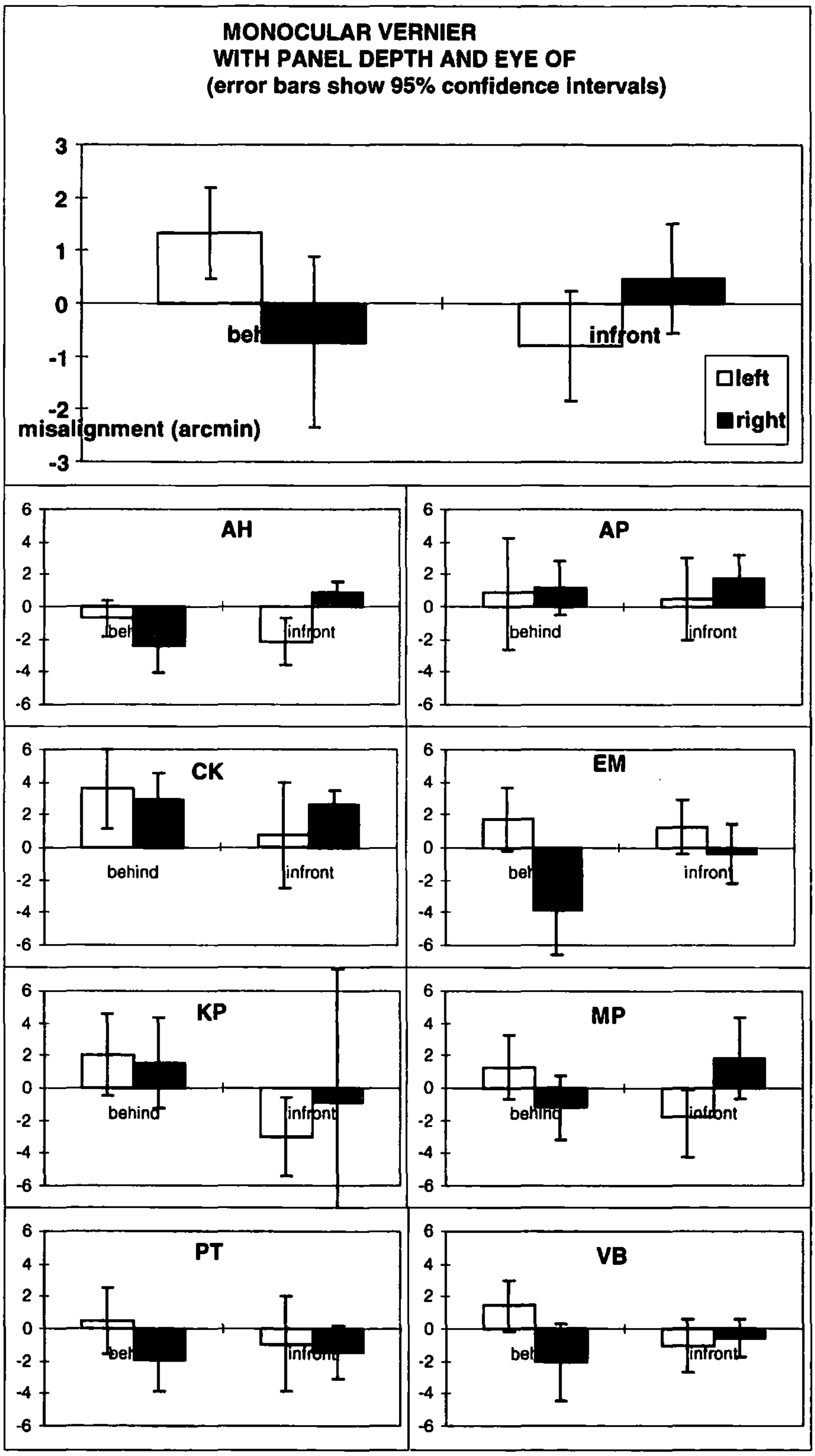


Figure 5.3 Results of experiment 5.1. Mean results for all 8 subjects are followed by the individual results. Positive values denote a shift to the left, negative a shift to the right. Error bars for individuals signify 75% vernier acuity thresholds. In overall data, error bars show standard deviation between subjects. Both across subjects, and for each subject, the interaction between eye and depth profile is clear.

In contrast, differences in alignment between the two eyes, and between the two depth profiles, although present in the individual data, showed no regular pattern across observers. The main effects of eye and depth profile were small and non-significant (see graph of pooled data).

5.1.4 Discussion

The results show that Hering's law I (the law of oculocentric direction) is invalid in the presence of binocular disparities. Lines that stimulate horizontally coincident regions of the retina can appear misaligned. This finding invalidates the notion that the visual direction of monocular objects is represented entirely by a population of monocular neurones, found only in the input layer of the primary visual cortex. Instead, we must accept that binocular interactions, either feeding back to the monocular layer or superseding its activity, determine the apparent directions of monocular objects. Similar results were recently described by Ono (Ono et al., submitted).

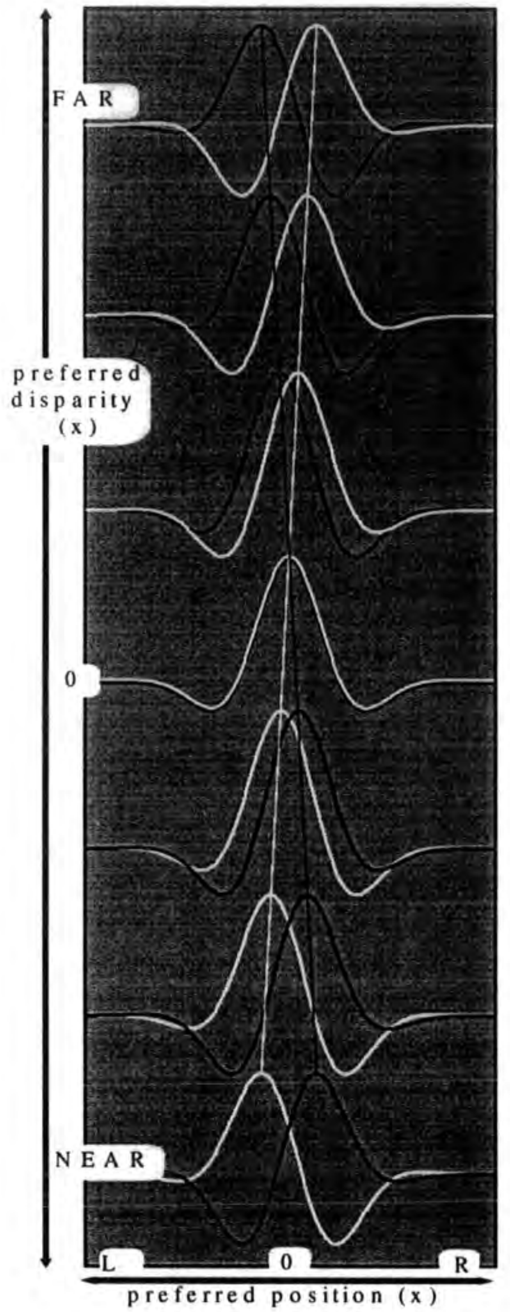


Figure 5.4. To model the results, all we need is the assumption that both position and disparity are coded by the same population of cells. Left and right response profiles of the simple binocular neurones described by Ohzawa and Freeman (1996) are shown. These are phase-shifted gabors, however the same would hold for a shift in position. Any activity at the surrounding disparity will mean a shift from zero to the black (white) line is required for alignment in the left (right) eye. A similar distribution of disparity-sensitive cells is assumed at locations throughout the visual field, providing a population code for position as well as disparity.

Many experimental findings can be explained by the assumption that monocular localisation is effected by binocular neurones (figure 5.4). The dichoptic misalignment observed by Ono (1991; Shimono et al., 1998) is one. The 'capture of binocular visual direction' described by Erkelens and Van Ee (1997ab) is also accounted for. Erkelens and Van Ee (1997ab) reported that a monocular line between two random-dot squares changing in disparity had to be given motion in the same direction as the monocular half-image of the dots to appear stationary, more than simply required to null the monocular component of the vergence eye-movements stimulated by the dots. Like Ono's findings, this result violates Hering's law IVb. The law describes the situation where a binocular object is the target of gaze, and lies on the 'common axis' which is the intersection of the 'visual axes' of the two eyes. An unfused monocular object is judged to deviate from the common axis by the angle subtended between the visual axis and the 'visual line' containing the object. This implies that a monocular object in the vicinity of a binocular one will have a visual direction equivalent to the monocular angle between the two objects. Excitatory interactions between cells tuned to the same disparity might account for violations of this 'law', and additionally for the decline in the effect of binocular disparity on monocular visual direction as the size of the gap between the two panels (Erkelens and Van Ee 1997ab) and around the dichoptic nonius lines (Shimono et al., 1998) was increased.

Why is the obtained misalignment (approx. 3 arcmin) so much smaller than the disparity difference (approx. 20 arcmin) This can be explained by considering the response of neurones such as those illustrated in figure 5.4 to a monocular stimulus. In this case, activity would be distributed between the different disparities, and different locations, peaking broadly at zero disparity and the stimulus location in monocular coordinates. With a combination of monocular and binocular stimuli, the activity would be a weighted mean between this and the response at stimulus disparity. The 'model' also accounts for the perceived location of different-contrast stimuli in the two eyes, whether they are fused (Mansfield and Legge, 1995) or unfused (Smallman and McKee, 1995). With similar-contrast stimuli, activity will peak at stimulus disparity and mean direction weighted by contrast (in idealised form, this is like the quadrature model suggested by Mansfield and Legge). When the contrast of the monocular stimuli is very different, there will be a position peak at the location of the bright monocular stimulus (centred at zero disparity, consistent with the results of Smallman and McKee) and a smaller peak at the location of the faint stimulus.

Excitatory interactions between disparity-tuned neurones bring to mind Marr and Poggio's (1976) stereoalgorithm, mentioned in the introduction. The constraint of continuity was implemented in this algorithm by positing precisely such interactions. This algorithm has many disadvantages, however these have been resolved in more recent cooperative stereoalgorithms (e.g. Marshall et al., 1996) which would nevertheless account for the dichoptic and direction-capture results, providing it is assumed that the same population of neurones represents both depth and location.

To account for our own findings, however, all that would be required is a large filter of the kind manifested by the simple cells of Ohzawa et al. (1996), which could blur the 0.7 arcmin wide vernier lines with the surrounding binocular dots. At the location of the vernier, such cells tuned to the stimulus disparity will be more active than those tuned to other disparities (with some error introduced by the unmatched vernier lines), and the additional activity of these cells resulting from the higher contrast of the vernier line, integrated along the length of the line, might account for the perceived location of the line. However, this more economical model cannot explain monocular dislocation when there is a large gap between monocular and binocular elements (Shimono 1998, Erkelens and Van Ee 1997ab).

Indeed, the interactions proposed may well underlie stereoscopic depth perception, which requires relative disparity information (Erkelens and Collewijn, 1985b). However, the effect found in Erkelens and Van Ee's (1997ab) study shows that these interactions take place even when there is no perceived depth, that is when only absolute disparities are varied. Perhaps such enhancement of disparities plays a role in their spatial integration by secondary filters, which might differ in their surround properties between the coarse filters driving vergence and the tuned ones that determine stereoacuity (chapters 2, 3, see also Summary and Discussion).

5.2 A comparison between subjective and objective estimates of initial vergence

5.2.1 Introduction

In the early chapters of this thesis (1 and 2), the Nonius method was used to estimate the angle of vergence of the two eyes. This method is termed subjective, because it relies on the subject's perceived alignment between two dichoptically presented lines. In contrast, an objective method uses a physical recording of the position of the eyes. The physical position of the eyes can be inferred from the apparent alignment of the nonius stimuli, only if Hering's Laws of visual direction (cited in 0.4) are correct. The accuracy of the subjective method has been questioned by many recent (Shimono et al., 1998; Erkelens and Van Ee, 1997ab; Fogt and Jones, 1998) as well as older publications (Remole et al., 1985,1986; Kertesz and Lee, 1987). Before that, large fixation disparities had been obtained using some objective methods, absent from the subjective measurement (e.g. Tani et al., 1956) but not with other varieties of objective eye-tracking (Rashbass and Westheimer, 1960). The general finding that fixation disparities were subjectively reduced in some way, caused speculation that there may be a cortical fusional mechanism independent of oculomotor vergence, which could shift the apparent relative directions of the two eye's images much as vergence eye-movements would do (e.g. Schor, 1983).

Many of the earlier studies of this issue are inherently flawed, either because subjective and objective measurements were not taken during the same trials (Remole et al., 1986) or because the objective position of vergence is inferred from monocular eye-tracking, and the subjective position inferred from monocular visual direction (Kertesz and Lee, 1987). However, the possibility of more than one nonius alignment position in the absence of eye-movements has clearly been demonstrated (Ono, 1991; discussed in 5.1). This makes it a potentially unreliable tool for measuring vergence. Additionally, variation in the visual direction of a monocular line does not consistently follow the objectively recorded position of that eye (Erkelens and Van Ee, 1997ab). In all these studies (except Kertesz and Lee) monocular stimuli were presented at the same time as binocular stimuli. Shimono et al. (1998), Erkelens and Van Ee (1997ab), and Remole et al. (1986) showed that monocular shift in the direction of the corresponding binocular half-images declined as the distance between monocular and binocular stimuli was increased. This finding indicates that there is a spatial integration of monocular fusional shifts. It raises concern about the nonius procedure employed in chapters 1 and 2, as although the nonius was subsequent to the binocular stimulus such wide-ranging spatial integration (up to 8°, according to Erkelens and Van Ee) may reflect some temporal integration as well.

Therefore, the aim in this experiment was to validate the procedure for subjective vergence estimation used in chapters 1 and 2, against a simultaneous objective binocular eye-movement recording. Subjective and objective estimates of vergence were taken as a function of the disparity of a prior stimulus. If temporal integration of any monocular fusional shifts were taking place, the slopes of the two functions could be expected to differ, with a steeper slope (and less fixation disparity) in the case of the subjective estimate.

5.2.2 Method

a Subjects

Three experienced psychophysical observers with normal or corrected to normal vision took part in this study. Only the author, AP, knew the purpose of the experiment.

b Apparatus and stimuli

As in previous chapters, stimuli were generated using a CRS VSG card on an Ambra 486 PC, and displayed on two monochrome Eizo FLEX-6500 monitors viewed through a modified Wheatstone stereoscope which consisted of two large (20 x 25 cm)mirrors mounted at right angles. AP and EM viewed the monitors at a distance of 80 cm, whereas VB had a viewing distance of 110 cm. Two dual-Purkinje trackers (Fourward Engineering) recorded the positions of the left and right eyes, sending digital data to an Apple-Mac Quadra computer. Stimuli were 300 x 300 pixel fields of 200 1 x 1 pixel white dots, given a disparity of -40 to 40 pixels. The nonius stimuli consisted of horizontally abutting, central 1 x 20 pixel full-contrast vertical bars. A nine-point calibration consisting of 20 x 20 pixel crosses was used. The dimensions of the display were 25.5 x 34 cm, or 768 x 1024 pixels, giving a pixel size in visual angle of 1.4 arcmin for AP and EM, and 1 arcmin for VB.

c Design

Vergence was estimated subjectively and objectively at each of a number of different stimulus disparities. The stimulus disparities were 0 and ±7, 14, 21, 28 for EM, and, additionally 42, and 56 arcmin (approx.) for AP. VB was presented disparities of 0 and ±5, 10, 15, 20 and 25 arcmin. (Stimulus disparity 0 used to measure fixation disparity).

d Procedure

Each session began with a nine-dot calibration, after the eye trackers had been positioned. This was followed by 80 trials at a given disparity from the fixation plane, half with a crossed disparity and half with an uncrossed disparity (except, of course, when the disparity was zero). The order of trials in the two conditions was random. Each trial consisted of a 1 second fixation, while the stimulus was calculated off-line, followed by a 230 msec stimulus interval and finally a 160 msec nonius display succeeded by a blank screen while the viewer made his/her response. Response was a mouse-button press to indicate whether the top line was left or right of the lower line. The position of the lines on each trial was controlled by a staircase procedure, and response frequencies were subsequently fitted a cumulative normal using Probit (see chapter 1 for further details). At the end of each session there was a second calibration. Subjects completed a number of sessions in a single sitting, taking breaks as required.

To obtain the objective vergence estimate, records of horizontal eye-position (fed to the Mac every 5 msec) were scaled according to the two calibrations. For each trial, the vergence movement between the averaged first 50 msec of stimulus onset, and the first, second and third 50 msec of stimulus offset were calculated. As with the subjective estimates, these were averaged over the forty trials with the same disparity.

5.2.3 Results

Figure 5.5 contains vergence estimates for the different temporal intervals, and the subjective estimates. The constant difference between subjective and objective estimates is due to static fixation disparity, which cannot be measured objectively without tracking the actual position of the fovea. Looking at the change in vergence during the nonius interval (the three dark lines in figure 5.5), AP and VB continue to verge in the direction of the stimulus disparity, although there is some deceleration of the eye movement (visual inspection shows the gap wider between the first two intervals than the second two intervals). For EM, however, there is no such deceleration and his eyes continue to drift in a divergent direction regardless of the stimulus disparity.

For each subject, the 'instant' of best correspondence between the two measures was used for further comparisons. This turned out to be the first 50 msec of the nonius stimulus for AP, the second 50 msec in the case of EM, and the third 50 msec for VB. The reason for subjects not basing their judgment on the first appearance of the nonius lines may be the visual suppression which occurs during vergence (Manning and Riggs, 1984). This may be why only AP, who has the least difference in vergence between the three intervals, has the best correspondence between subjective and objective estimates during the first interval. However, it seems more likely that individual differences are due to observers using different response strategies to cope with any apparent motion of the lines.

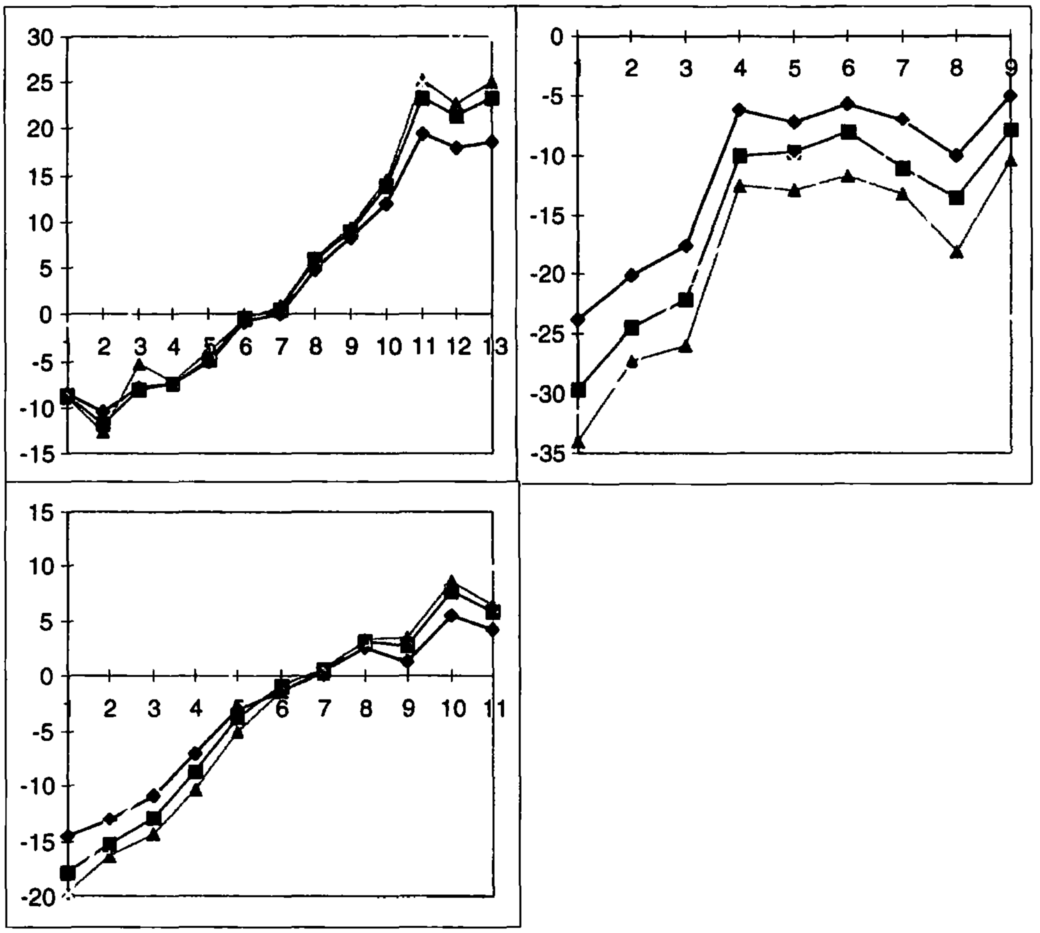


Figure 5.5. Vergence measured subjectively (white line, crosses) is shown together with objective vergence estimates averaged over the different temporal intervals (diamonds=stimulus offset to 50 msec, squares=50 -100 msec delay, triangles=100-150 msec delay).

There was a close correlation between subjective and objective estimates of vergence at all the given stimulus disparities. This correlation was over 98% for AP and VB, who responded consistently to the stimulus disparity. EM diverged effectively to the uncrossed stimuli, but failed to converge when crossed disparities were presented. Nevertheless, there is a significant (p=0.03) 72% correlation between the subjective and objective measurements in his case, and it is the objective measurement that corresponds more closely to the stimulus disparity. However, correlation is not the appropriate measure if trying to see whether the slopes of objective and subjective responses are different. For a comparison of the slopes, see figure 5.7.

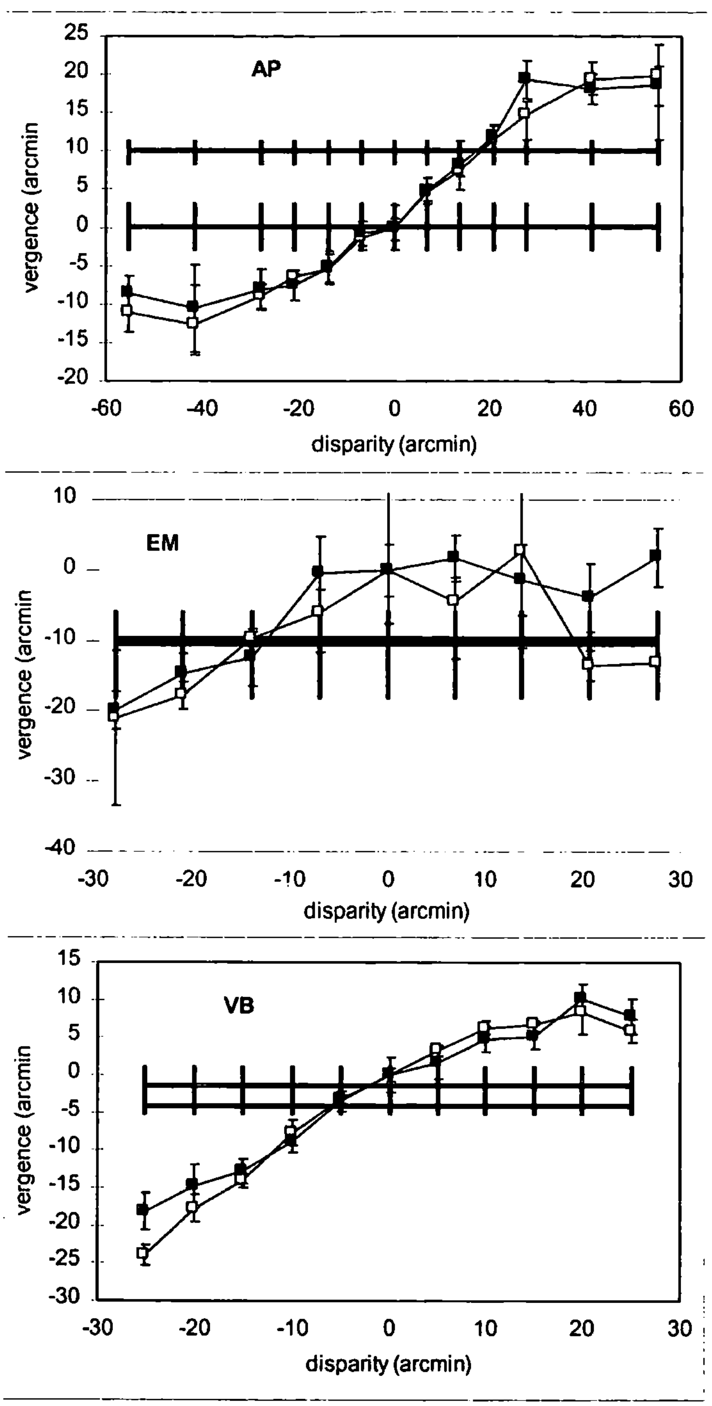


Figure 5.6. Subjective vergence estimates (open symbols) and objective vergence estimates (closed symbols) are plotted against stimulus disparity. To facilitate a visual comparison between the slopes, and the individual data points, the subjective and objective functions have been superimposed, as follows: The two measures were normalised to zero vergence at zero disparity, by subtracting a constant value. The constant subtracted from the objective measure is shown by the dark horizontal bar, the constant subtracted from the subjective measure is shown by the light horizontal bar. These constants are the respective fixation disparities, or responses to zero disparity, from the two measures. The error bars on the horizontal bars show the error associated with the vergence estimates at zero disparity.

In none of the subjects did a t-test between the two sets of data reveal any significant difference between their means (once static fixation disparity had been discarded). These results are illustrated in figure 5.6, where the data have been normalised to coincide at zero disparity. Because the objective measurement is vergence change, it cannot reflect any standing fixation disparity. Such standing fixation disparity, shown by the horizontal lines, was large in the case of AP (approx. 10 arcmin) here as in the previous chapters.

Visual inspection of figure 5.6, where error bars show 95% confidence intervals on each data-point, indicates that for AP and VB there were no statistically significant differences between the objective and subjective estimates in the range tested, the only exceptions being at approximately +30 arcmin (AP), and at -25 arcmin (VB). With EM, however, the two measures differed for crossed disparities such that the magnitude of vergence error was overestimated by the subjective procedure.

The good correlation between subjective and objective vergence estimates does not preclude a difference in magnitude between the two as a function of stimulus disparity, although inspection of figure 5.6 suggests that the slopes are similar. To test this, however, the slope of the regression of subjective on objective measurements was taken. This was close to the predicted value of 1 for AP and VB, and did not differ significantly from 1 even in the case of EM, as figure 5.7 shows. Indeed, the slope for EM was considerably less than 1, showing that, if anything, the objective measurement increased more rapidly as a function of stimulus disparity than did the subjective one.

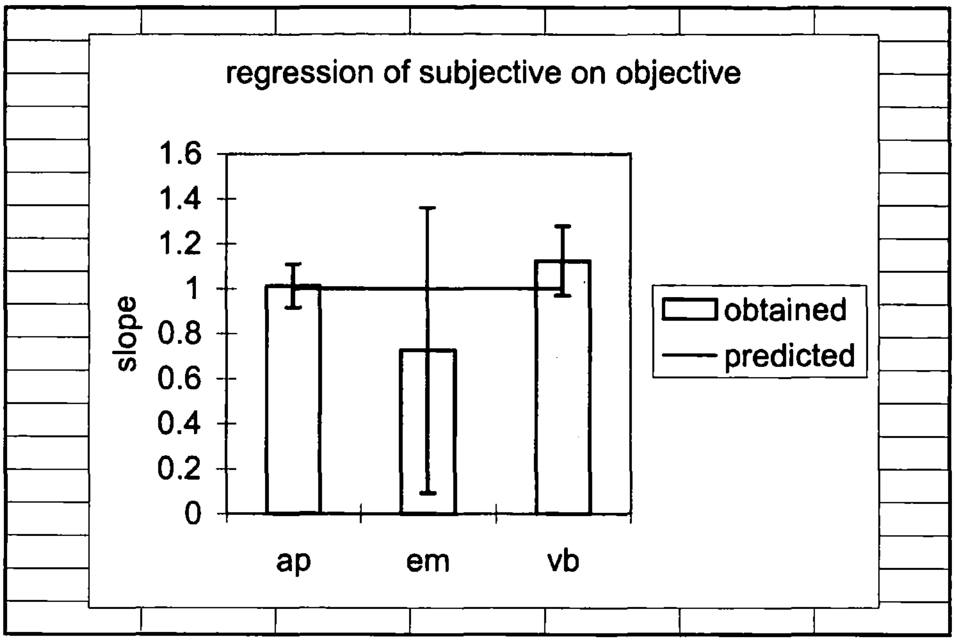


Figure 5.7. The slopes of subjective against objective fixation disparity did not differ significantly from unity, although all were significantly correlated, as shown by this histogram (error bars represent 95% confidence)

A fourth way of representing the data is as an FVFD (forced-vergence: fixation-disparity) curve (Ogle, 1950). Figure 5.8 reveals underlying trends in the pattern of fixation disparity as a function of stimulus disparity. Although not directly relevant to the comparison between objective and subjective measurements, they show that the three observers come from two different sub-populations of this function. Whereas AP belongs to type-I, EM and VB tend towards type-III (Ogle, 1950).

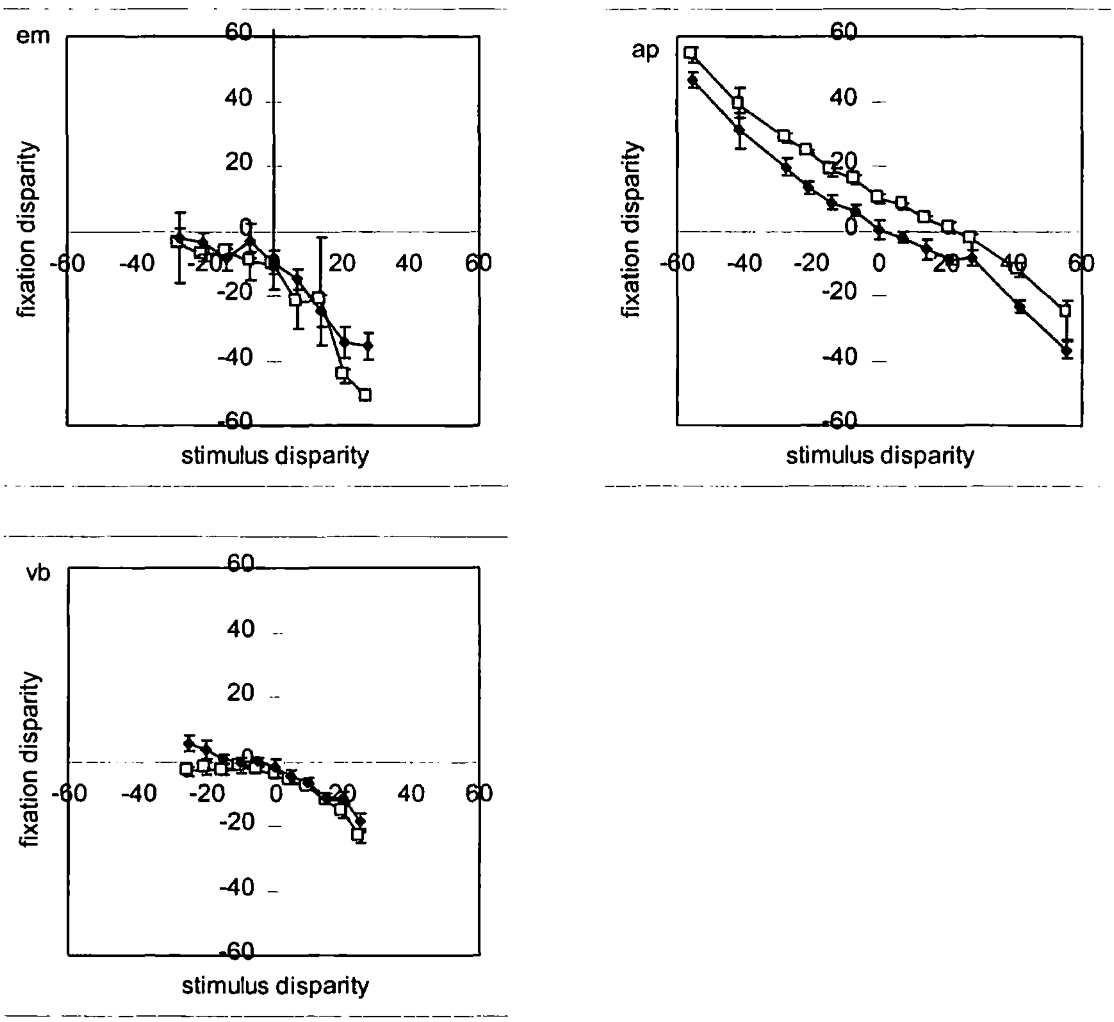


Figure 5.8. The data from figure 5.6 have been replotted as fixation disparity(i.e. the discrepancy between vergence and stimulus disparity) against stimulus disparity. Here, data were not normalized. The gap between subjective (squares) and objective (diamonds) estimates is due to static fixation error, eliminated from the objective series because of the calibration procedure. Note, however, that the series are broadly parallel, and deviations from parallelism do not always favour the subjective estimate(e.g. EM at large crossed disparities).

5.2.4 Discussion

The results show subjective estimation of vergence using the nonius procedure, as employed in chapters 1 and 2, to be unbiased in comparison with objective vergence measurement. They clearly demonstrate that the apparent direction of dichoptic lines flashed briefly in the absence of other stimulation is closely related to the physical position of the two eyes. These results are not in contradiction with extant criticisms of the nonius procedure (see introduction and 5.1), they merely specify that temporal contiguity might be required for any fusional shift in monocular visual direction. The stimuli here differed from those used in chapters 1 and 2, where two disparities were presented at the same time. However, biases in monocular visual direction have been reported both for relative disparity (Ono, 1991; Shimono et al., 1998) and absolute disparity modulation (Erkelens and Van Ee, 1997ab). Therefore, there is no reason to suppose a difference between the two situations, at least in their influence on monocular localisation.

There are a number of drawbacks to the present experiment. Only three observers participated, because of the difficulty of taking objective binocular eye-movement records from unpractised subjects they were all experienced psychophysical observers. There are four known categories of FVFD functions (Ogle, 1950) and the subjects here represented only two of these. It is possible that discrepancies between subjective and objective vergence are subject to individual differences dependent on these patterns. Indeed, Fogt and Jones (1998) presented data from six subjects, only three of whom showed the predicted discrepancy. Nevertheless, the results here pose a problem for any theory that this discrepancy is due to the influence of a monocular shift on the subjective data, as the only subject who diverged from correspondence between the two measures, actually recorded more accurate vergence in the objective than the subjective estimate (EM, see figs. 5.5-5.7).

How can this slight divergence be explained? It is possible that, because trials were blocked in randomised sequences of crossed and uncrossed disparities, EM developed an uncrossed response-bias. The staircase procedure used in the nonius method, may have converged towards an uncrossed peak in his data at crossed disparities given their possibly bimodal distribution. The objective estimate was, however, based on the average across trials and therefore would reflect the mean rather than the mode vergence across trials. Because the difference was significant at certain disparities, EM's data in other experiments were reviewed. In chapter 4, for example, unlike here, he provided considerable vergence change between conditions during a similarly blocked procedure. Therefore, perhaps the present results were due to fatigue or eyestrain resulting from the additional eye-tracking apparatus. (In the Purkinje tracker the stimulus must be viewed through a partially reflecting surface which bounces an infra-red beam onto the lens and back to a video camera and tracking device. This surface provides an inappropriate proximal vergence cue. Viewers are expected to ignore it by fixating the experimental stimulus instead, but some might conceivably get in the habit of diverging their eyes to overcome the misleading cue).

The discrepant measures of fixation disparity, normalised in figure 5.6, give rise to a serious methodological criticism of this study. The subjective measurement indicates presumed divergence from the fixation plane, which may be static or reactive to the stimulus presented. Objectively, one can only measure reactive fixation disparity as the 'static' component is eliminated by the calibration procedure. In this case, only vergence change during the stimulus interval was measured objectively, and hence clearly no underlying static fixation disparity was computed. However, even an absolute measure of vergence using the eye-trackers must rely on an assumption concerning the vergence position during the calibration procedure. Therefore, objective eye-trackers cannot be used to measure a stable static fixation disparity, only change in vergence. The only exception would be if it were possible to track the actual position of the foveas in the two eyes, and the torsional angle of the eyes. If this experiment were to be repeated, it would be preferable to take an analogous interval in the subjective measurement. This could be done by preceding each stimulus with a nonius interval, and measuring the change in nonius alignment rather than its absolute position, in analogy with the objective measurement. This procedure was not used because the experiments in previous chapters (1, 2 and 4) had only a single nonius interval per trial, and similarly what was measured was the sum of static and reactive fixation disparities.

In conclusion, and to lead back to chapter 1, nonius alignment can provide a viable estimate of oculomotor vergence. It is economical, and easy to use with untrained subjects. Providing no concurrent binocular stimuli are present, this subjective estimate of vergence is no different from an objective estimate obtained using two binocular dual-Purkinje trackers. However, when binocular stimuli are present monocular visual direction is affected (see 5.1), and this close relation between the two measures must break down. The continuum between these apparently contradictory results concerns the temporal rather than the spatial integration of disparities, and therefore lies outside the scope of this thesis.

6 Discussion

How are disparities integrated?

6.0 Summary

The aim of this thesis was to study the spatial averaging of disparities in brief random-dot stereograms, using the two possible outcomes of vergence and perceived depth. To summarise the results, disparities were averaged over a large (at least 5°) region to determine the target for automatic, initial vergence (chapter 2). This integration took place between both overlapping and adjacent disparities (chapter 1). However, in both these chapters vergence was estimated using the nonius method, which, although validated in 5.2, still has many potential drawbacks, discussed in these chapters. Particularly where trials were blocked by condition, there was the potential for subjects to develop a fixation disparity appropriate to that condition, rather than change their vergence in response to the stimulus on each trial. Disparities were integrated over a smaller (1-3°) area to determine stereoacuity (chapter 3). A region of interocular decorrelation, which can be regarded as multiple matches at different disparities, was similarly integrated over a large area for vergence, but over a smaller area when decorrelation detection thresholds were studied (chapter 4). This result must also be qualified, as many observers, including the author, did not have a vergence response to decorrelation. Indeed, the small number of subjects in the experiments of chapter 4, and the large individual differences between them, preclude any firm conclusion concerning the processing of decorrelated stimuli. Instead, the results hint at an interesting and previously unreported phenomenon which might, if studied more appropriately, bring theoretical advancement to the topic of disparity processing.

Binocular combination and disparity summation were found to precede the determination of monocular alignment (chapter 5).

In this chapter, theories of disparity integration will be examined (6.1), in their relation to the findings reported in this thesis (6.2). Following this will be some suggestions for further study (6.3), including suggested methodological improvements on the experiments of this thesis (6.3.5) and finally a conclusion relating to the issues raised in the introduction (6.4).

6.1 Theories of disparity integration

The main conclusion that can be drawn from this research is that vergence and stereoacuity result from differential spatial integration of disparities. There are three distinct ways in which this could be achieved. First, disparity detectors at increasing spatial frequencies, sensitive to decreasing disparities, might be responsible for the reduction in the pooling region from vergence to stereoacuity (Richards, 1970,1971; Marr and Poggio, 1978)(6.1.1). Put simply, large objects in one eye can be matched with large objects in the other eye, at large disparities to stimulate vergence. Smaller objects are only matched at smaller disparities, the finest giving stereoacuity. This theory is based on the linear processing of incoming luminance patterns. The second possibility is that non-linear monocular processing can be carried out over regions of different size prior to binocular combination (Wilcox and Hess, 1995)(6.1.2). In other words, the matching process takes in large windows for vergence and small windows for stereoacuity, regardless of the windows' spatial content. The third option is that the output of more or less linear primary binocular filters could be integrated by secondary disparity filters which selectively feed the resultant oculomotor and perceptual outcomes. This means matching large objects at large disparities, small objects at small disparities, and then pooling the obtained disparities from all scales in different-sized pools (6.1.3). These three theories are discussed in the sections that follow.

6.1.1 The linear model

The first theory is the simplest. This linear model would put the point of departure between vergence and stereoacuity at the initial filtering stage. Incoming visual information is known to be filtered at a number of different spatial frequencies. Coarse filters might extract the large disparities governing vergence, while fine filters extract the small disparities of stereoacuity. This model is illustrated in figure 6.1.

Marr and Poggio's (1979) model was based on the physiological presence of 'tuned' and 'untuned' disparity detectors (Poggio and Fischer, 1977), which were likened to the fine and coarse disparity pools suggested by Richards (1970,1971). Marr and Poggio suggested that these were associated with high and low spatial frequency channels. They applied the principle of spatial-frequency analysis to disparity processing, achieving a reduction in matching noise. The rapid processing of low spatial frequencies of luminance modulation is assumed to lead to the rapid formation of a coarse-disparity map, which can be used to direct vergence.

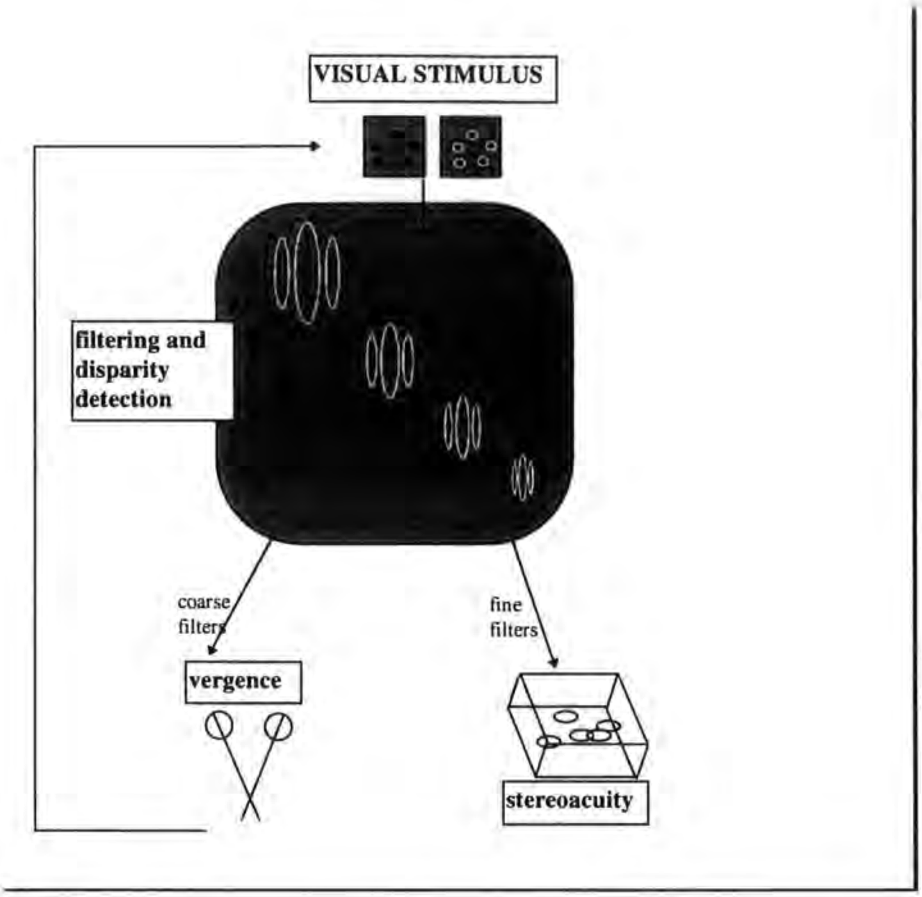


Figure 6.1. Illustrating the linear model. White lines show the right-eye's stimuli and receptive fields, black lines the left-eye's. Low frequency, coarse-disparity filtering determines vergence. Vergence position feeds back into the disparity stimulus, nulling the mean disparity to enhance activity in the finer filters.

Evidence of an association between disparity and luminance spatial tuning comes from the bandwidth of spatial frequency combination for disparity averaging (Rohaly and Wilson, 1994). Additionally, Smallman and McLeod (1997) showed that pedestal stereoacuity is limited by spatial frequency, consistent with the lack of large disparity representation in high frequency channels (see introduction and chapter 3 for further discussion of this paper).

Marr and Poggio's (1979) theory does not specify a mechanism for the extraction of relative disparity, necessary for depth perception in all but the coarsest disparity channels (Ogle, 1950; Erkelens and Collewijn, 1985ab).

6.1.2 Non-linear monocular pre-processing

According to the second theory, there is non-linear monocular processing prior to binocular combination (Hess and Wilcox, 1994; Wilcox and Hess, 1996). The non-linear channels are additional to the linear channels described in 6.1.1. One such scheme is presented in figure 6.2.

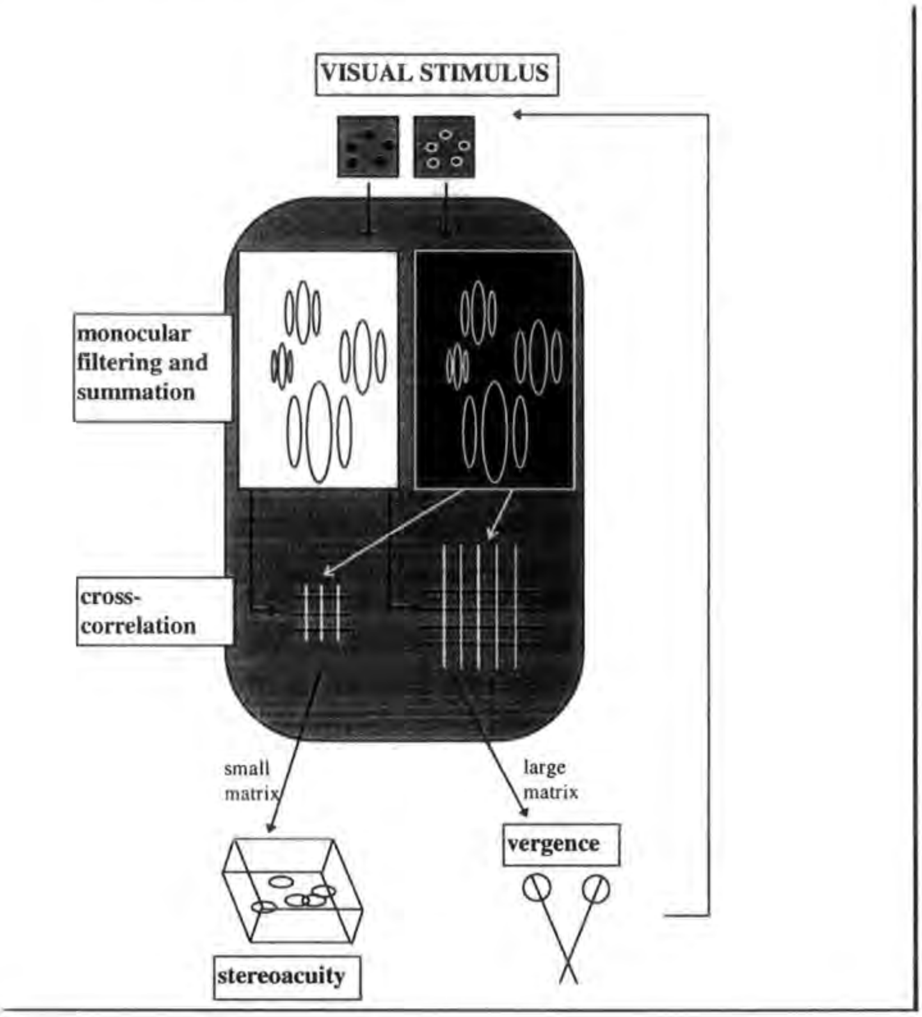


Figure 6.2. A model showing the spatial integration of disparities following monocular processing. The left and right images (indicated as in figure 6.1) are filtered independently with non-linear summation prior to cross-correlation between the two eyes. Vergence and stereoacuity differ in the size of the cross-correlation matrix.

In 1994 Hess and Wilcox showed that stereoacuity depends on the envelope size, as well as the spatial frequency, of a gabor target. They used a three-gabor task, where the disparity of the central gabor (a sine wave in a gaussian envelope) was modulated, and depth had to be detected. They argued that the results showed non-linear monocular preprocessing, similar to that evident in monocular alignment (Hess and Holiday, 1992) which must precede binocular combination. This is physiologically plausible, as non-linear channels are evident in early visual processing (e.g. complex retinal parasol cells, which project to magnocellular layers of the LGN, and complex cells in V1; for further detail see introduction). The argument was furthered in 1996, when Wilcox and Hess demonstrated stereoacuity with interocularly uncorrelated noise patches; one-dimensional band-limited noise in a gaussian envelope. Linear channels would be unable to extract a disparity between such patches. Nevertheless, stereoacuity in this task was measurable, although thresholds were 10 times higher than for correlated gabor patches.

However, non-linear monocular processing is not the only possible explanation for their results. This is easily argued in the case of stereoacuity for binocularly uncorrelated patches. Such patches will have an average disparity equal to the envelope disparity. This can be calculated by working out all the chance matches between the linear elements of the display, that is cross correlating the left and right images, and then taking a central measure of the disparity distribution (mean, median or mode). According to this argument, the findings of Wilcox and Hess (1996) are not incompatible with the theory advanced in the next section (6.1.3).

Wilcox and Hess (1996) also found that depth detection, although possible with either vertically or horizontally oriented noise carriers in the two eyes, failed when one patch contained a vertical carrier, and the other a horizontal carrier. They explained this by means of orientationally tuned envelope extraction. A much simpler explanation is that the vertical and horizontal bars simply could not be matched to extract a mean disparity from linear binocular combination. Only one result fits their account better than the above alternative. This is that stereoacuity for uncorrelated patches, as a function of viewing distance, was identical to monocular vernier acuity. This, they suggested, was because the two use a common monocular envelope extraction process. However, vernier acuity is good for vertical-horizontal comparisons but poor for ±45° comparisons (Keeble and Hess, 1998). The opposite is true of stereoacuity, which breaks down for vertical-horizontal matches (Wilcox and Hess, 1996) but not for ±45° matches (Tyler, 1995).

It could be argued that the envelope-size dependence described by Hess and Wilcox (1994) reflects the effect of frequency of disparity modulation (see chapters 2 and 3 for details). Hess and Wilcox found stereoacuity thresholds decreased down to an envelope sigma of about 10 arcmin for the highest spatial frequency carrier they used (5.24 cycles/deg.). As they were using a target-reference separation of 8 sigma, this is equivalent to a disparity modulation frequency of about 0.4 cycles/deg, similar to the value obtained by Rogers and Graham (1982). Interestingly, peak performance for lower frequency carriers was obtained with larger envelope sizes (and hence lower disparity modulation frequencies) showing an association between spatial frequency of disparity modulation and luminance spatial frequency (see previous section for a discussion of this topic). As in the previous section, the two alternative explanations cannot be distinguished using the spatially broad-band stimuli employed in this thesis. One suggestion for comparing them would be to vary the separation between the gabors used by Hess and Wilcox, another to replicate Rogers and Graham's study with spatially filtered stimuli.

Non-linear processing, although not necessary to explain the findings of Wilcox and Hess (e.g. 1996), might still take place prior to binocular combination. There is a multitude of stereo-phenomena related to Da Vinci stereopsis and occlusion (Nakayama and Shimojo, 1990; Liu et al., 1994), which are most easily explained by postulating monocular processing and interactions (Anderson and Nakayama, 1994; Anderson and Julesz, 1995). Occlusion relations are affected by left-right, temporal-nasal asymmetries, and matching occurs between junction elements distinguishable only monocularly. This provides strong evidence for the argument favoured by Wilcox and Hess (e.g. 1994), that complex (non-linear) monocular information is the substrate for binocular combination and matching.

6.1.3 Secondary filters

The third theory is that the outputs of initial filters are pooled, and perhaps recombined, by different-sized secondary filters. This approach uses the principle of noise-reduction from spatial frequency analysis, as Marr and Poggio did in 1979, but in the disparity domain and not the luminance domain (figure 6.3). The differences between disparity integration observed for vergence and depth outcomes can be attributed to differentially sized secondary filters for disparity, which are not necessarily tuned in the domain of luminance spatial-frequency. It seems probable that any such secondary filters would derive their input predominantly, but not exclusively, from similarly tuned primary filters, because they would carry less ambiguous information.

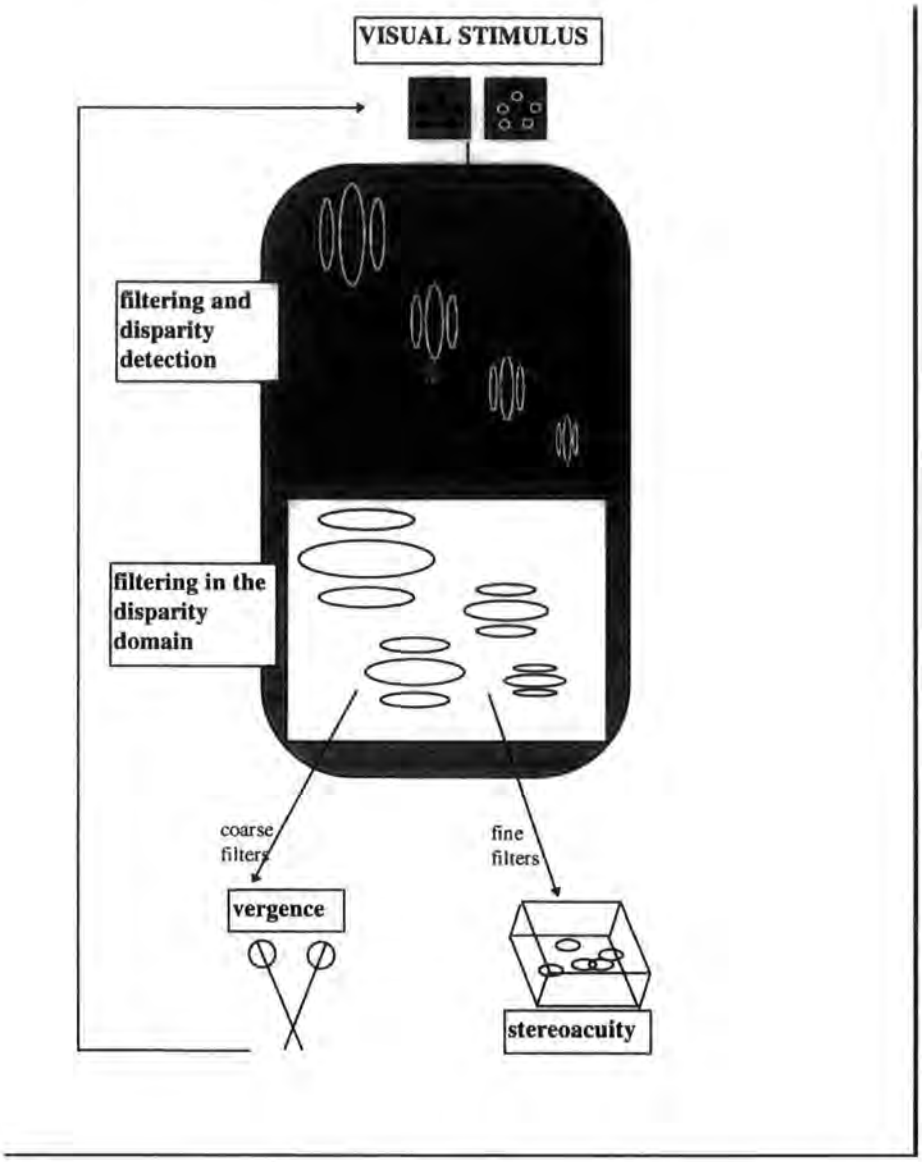


Figure 6.3. A model showing the spatial integration of disparities. Local disparity is obtained through linear, binocular filters, and subsequently filtered through different size (and depth) filters in the disparity domain to determine vergence and stereoacuity.

Before accepting the scheme shown in figure 6.3, a number of arguments must be considered. There is evidence that luminance frequencies are combined prior to the determination of depth judgments. Westheimer (1998) found a stereoacuity advantage for broad-band over narrow-band visual stimuli. The binocular neurones described by Ohzawa et al. (e.g. 1996, 1997) were, however, spatial-frequency tuned. To explain this discrepancy, the role of such neurones in depth perception must be considered. Parker and Cumming (1997) showed that neurones in an alert monkey respond ambiguously to an ambiguous local disparity, unlike the animal who favours the disparity consistent with global cues. Additionally, they point out that although the neurones continue to respond when no relative disparity is present, depth perception is unavailable in this situation (Erkelens and Collewijn, 1985a). This discrepancy clearly demonstrates that perceived depth is not represented at the level of such single neurones. It is, however, possible that the pattern of activity across a population of binocular neurones, each tuned for disparity and spatial frequency, might correspond with our perceptions. If so, they may be encoded by cooperative computation among the neurones, and additionally or alternatively by a second layer of filters such as proposed here.

Miles et al. (1997), following Parker and Cummings (1997) showed that vergence, like the cell responses, and unlike perceived depth, reversed for anti-correlated stimuli. How does single-cell activity affect vergence? One possibility is that vergence pools these cell responses summatively. The difference between vergence and stereoacuity could be in the nature of disparity pooling, with stereoacuity processed by smaller, subtractive filters. For a review of 'disparity receptive fields' for stereopsis, see Howard and Rogers (1995).

Subtractive filters form the basis of image segmentation in the luminance domain. Such centre-surround cells are often used as edge-detectors in vision algorithms (e.g. Marr, 1982). The role of stereopsis in viewing random-dot stereograms is akin to breaking camouflage (Ames leaf room, Ogle, 1950). Subtractive disparity filters would form an ideal mechanism for cyclopean image segmentation. Vergence, in contrast, requires a single (absolute or continuous) signal. As this must be reasonably close to the disparity of the attended or fixated target, an additive filter of limited extent could easily provide an adequate initial goal.

Secondary filters for disparity have been found by Ohzawa et al. (1997) in the form of complex cells in V1. Their output can be described as a function of the outputs of four or more simple cells, however there is as yet no physiological evidence of such hierarchical connections within V1. These neurones remain tuned in the spatial frequency domain, making them unlikely candidates for the formation of global perceptual judgments. According to Ohzawa et al. (1997) they could play a role in solving the correspondence problem by matching local image segments beyond a single element. Secondary filters for form from motion have been found in MT (e.g. Tanaka et al., 1986). Hierarchical connections from V1 to prestriate areas are well documented (for review see Zeki, 1993), and therefore these other areas are a more likely location for the secondary disparity filters proposed.

6.2 Theories and thesis

The way in which the three models can explain the findings presented in this thesis will be discussed in the next three sections. This discussion will follow the order in which the models were presented in the previous section.

6.2.1 Disparity in different spatial frequency channels

According to this model, vergence has access to coarse luminance channels which can provide information about large disparities. This explains both the large integration area we found for vergence (chapter 2), and the averaging of transparent disparity planes (chapter 1). The vergence response to binocularly uncorrelated stimuli can also be accounted for (chapter 4). Stereoacuity, in contrast, is processed by the finest channels as these give the cleanest information about small disparities. That is reflected by the smaller integration region found for depth discrimination (chapter 3). Similarly, decorrelation discrimination is likely to rely on fine channels, when the scale of decorrelation is fine, as in chapter 4.

Coarse processing precedes fine processing, and therefore vergence modulates subsequent analysis of disparities at higher spatial frequencies. Evidence of such modulation was found in chapter 3.

Even the results of chapter 5 can be accounted for by this theory, providing we envisage sufficiently large filters for detecting the disparity step, which are still able to pick up the signal of the monocular lines.

Without using spatially narrow-band stimuli, the Marr-Poggio model cannot be rejected. However, it is an unlikely explanation of differential disparity integration in our stimuli, which were composed of small random dots, as these have little energy at low spatial frequencies.

6.2.2 Vergence from large non-linear filters

Wilcox and Hess (1995) showed that Dmax for stereopsis, the maximum disparity that can be used to determine depth direction, depended on the envelope size of gabors, and was relatively unaffected by their spatial frequency content. A similar relationship might be predicted for vergence, which can also depend on the processing of large disparities. This is in contrast to stereoacuity, which was influenced more by the carrier spatial frequency, especially for narrow-band stimuli (Hess and Wilcox, 1994).

The differential integration regions found in this thesis for vergence and stereoacuity might, therefore, reflect a difference not only in the underlying spatial frequency channel, but a difference between linear channels for stereoacuity and non-linear channels for vergence. Vergence might match the pattern of dots, extracted over a large area, between the left and right images. This explanation overcomes the problem of the limited availability of low frequency information from random dots, which was a drawback for the model described in the previous section (6.1.2). It accounts for the effect of envelope disparity on the perceived depth of gabors (Statham and Georgeson, 1998), and similarly the effect of pedestal disparity on stereoacuity (chapter 3). Alternatively, the difference between the disparity integration regions of stereoacuity and vergence might be one of window size, rather than a difference between linear and non-linear channels. Both may be the result of monocular pre-processing, but over a small scale for stereoacuity and a large scale for vergence.

There is some evidence against monocular pre-processing in this thesis. If considerable monocular processing takes place prior to binocular combination, it is hard to imagine why monocular alignment cannot be assessed veridically in the presence of a binocular disparity (5.1). The possibility remains, however, that binocular combination takes place inevitably after whatever monocular processing (including nonlinear envelope extraction), which stops short of relative localisation.

6.2.3 Different sized secondary filters

The first two theories described (the simple linear model, 6.1.1; non-linear monocular preprocessing, 6.1.2) cannot be rejected outright on the basis of the evidence presented in this thesis (see 6.2.1 and 6.2.2). However, the former seems unlikely because of the spatial frequency spectrum of the random-dot stimuli used here, and the latter seems unlikely because of the effect of binocular disparity on monocular alignment. This leaves the disparity tuning scheme (6.1.3) as the best candidate.

The empirical evidence presented in the different chapters supports the integration of disparities by secondary filters, possibly subject to non-linear interactions or feedback at the early stages of visual processing (chapter 5). These filters will differ in their spatial properties (at least in the cyclopean domain) for the two measured outcomes of vergence (chapters 1 and 2) and stereoacuity (chapter 3). However, vergence modulation serves to bring the finer filters into register, providing a reduction in matching noise (chapter 3).

The complex binocular neurones described by Ohzawa et al. (1997) can be modelled as secondary disparity filters, that is they might result from a combination of the outputs of simple binocular neurones. It is hard to speculate, based on the properties of these neurones' receptive fields in cats, whether they could match the specifications for the secondary filters proposed here. The cells described are narrowly tuned in their disparity profile, but broadly tuned across spatial location. Therefore, although they might underlie the pooling of disparities for vergence, they could not account for the effect of disparity on apparent visual direction.

As discussed in chapter 5, the large scale interaction of disparity and monocular visual direction (up to 4°, Erkelens and Van Ee, 1997ab) can only be explained by assuming excitatory interactions among neighbouring disparity-tuned neurones, or feedback from a secondary filtering process. The neural-network model that was referred to (Marshall et al., 1996) was similar to Marr and Poggio□s (1976) model (see introduction, 0.3), however it contained a third layer of surface-patch neurones. The model contained feedback from this layer to the disparity-detector layer, which could explain the far-reaching effects of binocular objects on monocular ones described in chapter 5. Surface patch neurones are secondary filters operating on disparities. For further discussion of the relation between this model and the theory outlined here see 6.3.1.

6.3 Suggestions for further studies

How might these different possible mechanisms of disparity integration be distinguished? One simple modification to the experiments reported here would be the use of spatially filtered stimuli, as has already been suggested. If disparity integration depended on a single linear filter, the removal of low spatial frequencies would have detrimental consequences on any response to disparity. This would not, however, be sufficient to distinguish between the second two theories (6.1.2 and 6.1.3). The large-scale spatial integration of high spatial frequency stimuli might be the result either of non-linear monocular processing or of secondary disparity integration. These two theories are harder to discriminate empirically. One possibility would be to create a stimulus where two different (linear) depths were presented in the same (non-linear) envelope, and examine whether transparency could be perceived. Using similar stimuli, one could determine whether vergence depended on the envelope disparity, or the contrast weighted mean of the two planes when their relative contrast was varied. This is the case for unbounded stimuli (Mallot et al., 1996).

There are several research areas beyond the scope of this thesis which are relevant to the topic of disparity integration. Four will be discussed in further detail in the next few sections. These are as follows: 1)Disparity gradients, slant and inclination perception. 2) Vertical disparities and differential perspective. 3)The dynamics of disparity processing. 4) The neurophysiological study of disparity processing. Finally, in 6.3.5, methodological improvements on the experiments reported in the thesis are suggested.

6.3.1 Disparity gradients, slant and inclination perception

Horizontal disparity gradients define slant. Vertical disparity gradients define inclination. Howard and Rogers (1995) provide a comprehensive review of both. A model of 3D tilt perception was referred to in chapter 5 (5.1; Marshall et al., 1997). 3D tilt is a combination of slant and inclination. Disparity integration for depth perception serves two complementary purposes, the first is identify and locate differences in disparity, and the second is to discriminate patterns of continuity from the abrupt changes caused by surface boundaries. In this thesis, the emphasis has been on the former; however, the latter must also be considered in a complete account of disparity integration. The mechanism suggested in 6.1.3 could serve as a differencing operator to obtain a disparity gradient over the visual field, with subsequent grouping to give a representation of local tilt. Alternatively, primary linear disparities might be pooled directly by different tilted disparity receptive fields. To test this, tilt thresholds might be compared with appropriately oriented disparity grating stereoacuity thresholds (from horizontal for inclination to vertical for slant). The use of a differencing operator based on the modulation thresholds would limit tilt perception.

6.3.2 Vertical disparities and differential perspective

Vertical disparities are pooled over a large (20°) region to control vertical vergence, and unlike horizontal disparities the absolute signal in a small region, even without surrounding visual stimuli, is insufficient to drive a full response (Fang and Howard, 1998). This difference may have developed because vertical vergence movements are required to correct any misalignment of the eyes, and the best estimate of this misalignment can be obtained from sampling the entire field. Horizontal vergence movements, on the other hand, are needed to selectively bring into correspondence the object that is the target of attention, leaving other parts of the field disparate.

Objects not equidistant between the two eyes will have a larger projection on the retina of the eye to which they are closest. This is termed differential perspective. Like vergence or absolute disparity, the pattern of vertical disparities can be used to estimate the viewing distance of an object. This estimate is used by the visual system to scale both size and depth (Rogers and Bradshaw, 1993). Such vertical disparity gradients are integrated over a region of up to 20° (Kaneko and Howard, 1995).

Most points in the visual field have a disparity vector which is intermediate between horizontal and vertical (see example in introduction, 0.1). There is no evidence, however, that the angle of disparity is used by the visual system. Many models of stereo-matching use an epipolar constraint, that is they match horizontally only elements up to a certain limit of vertical disparity. This is equivalent to a vertical estimate of Panum□s fusional area. It appears that the visual system computes depth using orthogonal axes of disparity, with the horizontal axis used for depth and horizontal vergence, and the vertical axis used for vertical vergence, differential perspective and cyclovergence (evidence for the latter is reported by Rogers and Bradshaw, submitted).

The differential integration and processing of vertical and horizontal disparities might be hard-wired, or may have developed as a result of early visual experience. These two options could be distinguished by rearing animals with optically rotated visual fields, although the ethics of an experiment of this kind are questionable. Such a study would shed light on the plasticity of the pooling and implementation of horizontal and vertical disparities in vergence and depth perception, and also determine whether the use of orthogonal axes of disparity can be modified.

6.3.3 The dynamics of disparity processing

All the experiments in this thesis have used brief, static stimuli to exclude temporal integration. Disparities are, however, integrated temporally as well as spatially both to improve static performance (Glennerster, 1996) and to obtain a dynamic experience of motion in depth (Harris and Watamaniuk, 1995) or a dynamic vergence response (Erkelens and Collewijn, 1985a). Such temporal integration may reflect the temporal properties of the spatial integration mechanisms already discuss, or might involve completely different mechanisms.

Fang et al. (1998) replicated the spatial limitations described in chapter 2, using dynamic stimuli and measuring the gain in vergence velocity. They suggest that, at least for vergence, spatial integration is consistent for static and dynamic stimuli. However, recently Howard et al. (1998) showed that motion in depth could be obtained from dynamic stereograms which contained no static disparity. Interocularly uncorrelated, but temporally correlated patches (seen through an aperture) moved in depth in the direction of relative monocular motion. This could be achieved either by a mechanism which cross-correlates the monocular velocity signals (similar to 6.1.2), or by one that pools the change in disparity of randomly paired dots. The latter is compatible with the spatial pooling proposed in 6.1.3. This could be tested by investigating the spatial properties of the motion in depth phenomenon described by Howard et al., and also the spatial limitations of any vergence resulting from such stimuli.

The effect of disparity on monocular visual direction (5.1) was found to be short lived (5.2). The decay of such effects must be studied to specify the temporal properties of disparity integration. In a visual environment where both observer and observed are commonly moving, the spatiotemporal integration of disparities is necessary for the use of disparity information to compute depth, or to perform image segmentation.

6.3.4 Neurophysiological studies

Functional magnetic resonance imaging (fMRI) has provided useful information about the prestriate visual areas used in a variety of visual modalities, for example colour, motion and form (for review see Zeki, 1993). However, no area has been highlighted for processing disparity. Disparity sensitive neurones can be found in most visual areas, however so can neurones sensitive to the other modalities. One possibility is that this lack of positive findings reflects not a lack of imagination in the direction of performing image subtractions involving stereo, but a lack of interesting findings when disparity exclusive areas are extracted. This might suggest that disparity processing takes place at the earliest possible stages in the hierarchy (perhaps even V1 or V2) before the other modalities are segregated. A review and meta-analysis of any such null findings could be useful.

Additional support for this idea comes from the dearth of neuropsychological patients with specific deficits in disparity processing. Although such disturbances have been reported in the literature (Poppelreuter, 1917; Holmes, reprinted 1979) defects in stereoacuity and form from disparity often accompany other deficits, for example akinetopsia (Rizzo et al., 1995). However, the study of disparity processing would benefit if a patient were found with selective cortical stereo-blindness. Therefore, it may be useful to test potential patients□ stereovision selectively in different areas of the visual field, using measures of both acuity and cyclopean form perception.

Our knowledge of disparity processing in the brain has grown through a number of studies using single-cell recording techniques (Poggio, e.g. 1991; Ohzawa, e.g. 1997).

So far, such studies have mostly addressed the disparity (and other) tuning properties of neurones in V1 and V2. These neurones are invariably sensitive to absolute disparity (Cumming and Parker, 1997). Subsequent visual areas such as V4 and V5 also contain disparity-sensitive neurones, and their tuning properties have yet to be investigated. Perhaps some of the latter might be selective for relative rather than absolute disparity, reflecting psychophysical data.

6.3.5 Suggestions for methodological improvements

The chapters on disparity averaging for vergence (chs. 1,2 and 4) use the nonius method of vergence estimation. This method is subsequently validated in chapter 5. However, the method, as used, has many drawbacks. Since only a single measurement is taken, nothing can be determined about the dynamics of vergence, so the results obtained may be due to adapatation of fixation disparities, rather than vergence change. There is no conclusive evidence that vergence change actually occurred. Additionally, the results may represent differences in vergence velocity, rather than the final state of vergence. To address these issues, it would be worthwhile to attempt a replication of the major findings using an objective eye-movement measuring technique, such as with the binocular dual-Purkinje trackers employed in chapter 5.

In many cases (experiments 1.1, 1.2, 2.1, 2.2 and 4.1) trials were blocked by condition. This may introduce artefacts in the results, such as the adaptation of the fixation disparity, as already mentioned. These experiments in particular should be replicated using a randomised trial sequence. Then a comparison could be made to determine the effect (if any) of the order of the trials.

As mentioned in the acknowledgements at the front of the thesis, most of the participants in the various experiments were the author's colleagues, friends and family, who volunteered their time for little or no monetary reward. Often, only a few subjects could be found. Many of the experiments where there were fewer than five subjects, also revealed considerable individual differences. As already discussed in the relevant sections, these experiments should be explored further using a larger sample of observers. They include experiments 1.2, 2.2, 2.3, 4.3, 4.4, and 5.2.

6.4 Conclusion

The problem of seeing the world with two eyes was discussed in the introduction. Hering□s laws of visual direction were summarised. These geometrical rules, like the geometry of disparity processing, can be attributed to the way in which visual information from the two eyes is processed by the brain (5.1). The disparity between the images in the two eyes is not absolute, such that it could be corrected by appropriate vergence, but varies over the visual field because of the layout of surfaces in the environment. In this chapter, theories have been outlined concerning the way the brain deals with this variation - how disparities are integrated, to stimulate vergence and perceived depth. Presumably, the ultimate goal of such processing is to recover efficiently something of the environmental layout which caused the pattern of disparities, at least sufficient to respond to salient stimuli. To progress towards understanding how this goal is achieved, however, vergence and depth perception must be investigated using richer stimuli than in this thesis, such as those described under suggestions for further study. The relationship between neurophysiology and even the low-level mechanisms proposed here is still unclear. It is hoped that the discussion in the various chapters has shown how simple neural mechanisms might account for apparently elaborate computations, without the need to postulate symbolic abstracts such as the visual angles Berkeley objected to in Descartes' account of binocular depth perception (Introduction).

This thesis has addressed empirical constructs of visual processing. A theoretically unsurprising difference in pooling regions for vergence and stereoacuity has been found (Chs. 1-3) and made to account also for responses to interocular decorrelation (Ch. 4). Chapter 5 shows, indirectly, that this integration of disparities takes place surprisingly early in visual processing.

REFERENCES

Adams, W., Frisby, J. P., Buckley, D., Gårding, J., Hippisley-Cox, S.D., and Porril, J. (1996). Pooling of vertical disparities by the human visual system. *Perception, 25,* 165-176.

Akerstrom, R. A., and Todd, J. T. (1988). The perception of stereoscopic transparency. *Perception and Psychophysics, 44,* 421-432.

Alpern, M. (1957). The position of the eyes during prism vergence. *American Medical Association: Archives of Ophthalmology, 57,* 345-353.

Alpern, M. (1962). Types of Movement. In: Davson, H. (Ed.) *The Eye. Volume 3: Muscular Mechanisms, 5,* 63-151. London: Academic Press.

Alpern, M., and Ellen, P. (1956). A quantitative analysis of the horizontal movements of the eyes in the experiment of Johannes Mueller. I. Method and results. *American Journal of Ophthalmology, 42,* 289-296.

Anderson, B. L. (1992). Hysteresis, cooperativity, and depth averaging in dynamic random-dot stereograms. *Perception and Psychophysics, 51,* 511-528.

Anderson, B. L., and Julesz, B. (1995). A theoretical analysis of illusory contour formation in stereopsis. *Psychological Review, 102,* 705-743.

Anderson, B. L. and Nakayama, K. (1994). Toward a general theory of stereopsis: binocular matching, occluding contours, and fusion. *Psychological Review, 101,* 414-445.

Anstis, S., Howard, I. P., and Rogers, B. (1978). A Craik-O'Brien-Cornsweet illusion for visual depth. *Vision Research,18,* 213-217.

Arndt, P. A., Mallot, H. A., B‚lthoff, H. H. (1995). Human stereovision without localized image features. *Biological Cybernetics, 72,* 279-293.

Atherton, M. (1990). *Berkeley's Revolution in Vision.* Ithaca: Cornell University Press.

Baizer, J. S., Ungerleider, L. G. and Desimone, R. (1991). Organisation of visual inputs to the inferior temporal and posterior parietal cortex in macaques. *Journal of Neuroscience, 11,* 168-190.

Banks, M. S., VanEe, R., and Backus, B. T. (1997). The computation of binocular visual direction: a re-examination of Mansfield and Legge (1996). *Vision Research, 37,* 1605-1610.

Berkeley, G. (1709). An Essay towards a New Theory of Vision. In: *Berkeley - Philosophical Works* (1975). London: Dent.

Berry, R. N. (1948). Quantitative relations among vernier, real depth and stereoscopic depth acuities. *Journal of Experimental Psychology, 38,* 708-721.

Blakemore, C. (1970). The range and scope of binocular depth discrimination in man. *Journal of Physiology, 211,* 599-622.

Bradshaw, M. F., Glennerster, A. and Rogers, B. J. (1996). The effect of display size on disparity scaling from differential perspective and vergence cues. *Vision Research, 36,* 1255-1264.

Breitmeyer B. G. (1975). Simple reaction time as a measure of the temporal response properties of transient and sustained channels. *Vision Research, 15,* 1411-12.

Brookes, A., and Stevens, K. A. (1989). The analogy between stereo depth and brightness. *Perception, 18,* 601-614.

Brown, R. (1958). *Words and things.* New York: Free Press.

Bruce, V., and Green, P. (1990). *Visual Perception: Physiology, Psychology and Ecology.* 2nd Edition. London: Lawrence Erlbaum Associates.

Bülthoff, H. H., and Mallot, H. A. (1988). Integration of depth modules: stereo and shading. *Journal of the Optical Society of America A,* 5, 1749-1758.

Busettini, C., Miles, F. A. and Krauzlis, R. J., (1996). Short-latency disparity vergence responses and their dependence on a prior saccadic eye movement. *Journal of Neurophysiology, 75,* 1392-1410.

Campbell, F. W. C., and Robson, J. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology, 197,* 551-566.

Carpenter R. H. S. (1988). *Movements of the eyes.* London: Pion.

Ciuffreda, K. J. (1991). Accommodation. In W. N. Charman, (Ed.), Visual optics and Instrumentation, in J. R. Cronly-Dillon (Ed.), *Vision and Visual Dysfunction.* London: MacMillan.

Ciuffreda, K. J., and Kenyon, R. V. (1983). Accommodative vergence and accommodation in normals, amblyopes and strabismics. In C. M. Schor and K. J. Ciuffreda (Eds.), *Vergence Eye Movements: Basic and Clinical Aspects* (pp. 101-195). Boston: Butterworth.

Collewijn, H. and Erkelens, C. J., (1990). Binocular eye movements and the perception of depth. in E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes* (pp. 213-261). Elsevier Science.

Collewijn, H., Erkelens, C. J. and Steinman, R. M., (1995). Voluntary binocualr gaze-shifts in the plane of regard - dynamics of version and vergence. *Vision Research, 35,* 3335-3358.

Collewijn, H., Van der Steen, J., and Van Rijn, L. J. (1991). in A. Gorea (Ed.), *Representations of Vision* (pp. 165-183). Cambridge: Cambridge University Press.

Cormack, L. K., Stevenson, S. B., and Schor, C. M. (1994). An upper limit to the binocular combination of stimuli. *Vision research, 34,* 2599-2608.

Cumming, B. G., and Parker, A. J. (1997). Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature, 389,* 280-283.

DeAngelis, G. C., Ohzawa, I., and Freeman, R. D. (1991). Depth is encoded in the visual cortex by specialized receptive field structure. *Nature, 352,* 156-159.

DeAngelis, G. C., Ohzawa, I., and Freeman, R. D. (1993). Spatiotemporal organisation of simple-cell receptive fields in the cat□s striate cortex. II. Linearity of temporal and spatial summation. *Journal of Neurophysiology, 69,* 1118-1135.

DeValois, R. L., Albrecht, D. G., and Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research, 22,* 545-559.

Duwaer, A. L. and van den Brink, G. (1981). Diplopia thresholds and the initiation of vergence eye-movements. *Vision Research, 21,* 1727-37.

Edwards, M., Pope, D. R., and Schor, C. M. (1998). Luminance contrast and spatial-frequency tuning of the transient-vergence system. *Vision Research, 38,* 705-717.

Enright, J. T. (1984). Changes in vergence mediated by saccades. *Journal of Physiology, 350,* 9-31.

Enright, J. T. (1986). Facilitation of vergence changes by saccades: influences of microfocused images and of disparity stimuli in man. *Journal of Physiology, 371,* 69-87.

Enright, J. T. (1996a). Sequential stereopsis: a simple demonstration. *Vision Research, 36,* 307-312.

Enright, J. T. (1996b). Slow-velocity asymmetrical convergence: a decisive failure of Hering's Law. *Vision Research, 36,* 3667-3684.

Erkelens, C. J. and Collewijn, H. (1985a). Motion perception during dichoptic viewing of moving random dot stereograms. *Vision Research, 25,* 583-588.

Erkelens, C. J. and Collewijn, H. (1985b). Eye movements and stereopsis during dichoptic viewing of moving random dot stereograms. *Vision Research, 25,* 1689-1700.

Erkelens, C. J., and Collewijn, H. (1989). Ocular vergence under natural conditions. II. Gaze shifts between targets differing in distance and direction. *Proceedings of the. Royal Society, B, 236,* 441-465.

Erkelens, C. J. and Collewijn, H. (1991). Control of vergence: gating among disparity inputs by voluntary target selection. *Experimental Brain Research, 87,* 671-678.

Erkelens, C. J., and Van Ee, R. (1997a). Capture of visual direction: an unexpected phenomenon in binocular vision. *Vision Research, 37,* 1193-1196.

Erkelens, C. J., and Van Ee, R. (1997b). Capture of the visual direction of monocular objects by adjacent binocular objects. *Vision Research, 37,* 1735-1745.

Erkelens, C. J., and Van de Grind, W. A. (1994). Binocular visual direction. *Vision Research, 34,* 2963-2969.

Fahle, M., Henke-Fahle, S., and Harris, J. (1994). Definition of thresholds for stereoscopic depth. *British Journal of Opthalmology, 78,* 572-576.

Falmagne, J. C. (1985). *Elements of psychophysical theory.*

Fang, X., Howard, I. P., Allison, R. S., and Zacher, J. E. (1998). Effects of stimulus size and eccentricity on horizontal vergence. *Investigative Opthalmology and Visual Science, 39,* 4, s4829.

Findlay, J. M. and Kapoula, Z. (1991). Eye-movments and vision. In A. Gorea (Ed.) *Representations of Vision,* (pp. 153-154). Cambridge: Cambridge University Press.

Finney, D. J. (1947). *Probit Analysis.* Cambridge: Cambridge University Press.

Flax, N. (1983). Strabismus diagnosis and prognosis. In C. M. Schor and K. J. Ciuffreda (Eds.), *Vergence Eye Movements: Basic and Clinical aspects,* (ch. 18, pp. 579-604). Boston: Butterworth.

Fogt, N., and Jones, R. (1998). Comparison of fixation disparities obtained by objective and subjective methods. *Vision Research, 38,* 411-421.

Foley, J. M. (1978). Primary distance perception. In: R. Held, H. Leibowitz, and H. L. Teuber (Eds.). *Handbook of Sensory Physiology, Vol. VIII; Perception,* (pp. 181-213). Berlin: Springer.

Foley, J. M. (1980). Binocular distance perception. *Psychology Review, 87,* 411-434.

Freeman, R. D., and Ohzawa, I. (1990). On the neurophysiological organisation of binocular vision. *Vision Research, 30,* 1661-1676.

Frisby, J. P., Catherall, C., Porrill, J., and Buckley, D. (1997). Sequential stereopsis using high-pass spatial frequency filtered textures. *Vision Research, 37,* 3109-3116.

Gilbert, C. D. and Wiesel, T. N. (1989). Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex, *Journal of Neuroscience, 9,* 2432-2442.

Glennerster A, (1996). The time-course of 2-D shape-discrimination in random-dot stereograms. *Vision Research, 36,* 1955-68.

Gregory, R. L. (1958). Eye movements and the stability of the visual world. *Nature, 182,* 1214-1216.

Gregory, R. L. (1966). *Eye and Brain.* London: World University Library.

Hampton, D. R. and Kertesz, A. E. (1983). Fusional vergence response to local peripheral stimulation. *Journal of the Optical Society of America, 73,* 7-10.

Harris, J. M., McKee, S. P., and Smallman, H. S. (1997). Fine-scale processing of human binocular stereopsis. *Journal of the Optical Society of America A, 14,* 1673-1683.

Harris, J. M., and Watamaniuk, S. N. J. (1995). Speed discrimination of motion-in-depth using binocular cues. *Vision Research, 35,* 885-896.

Harwerth, R. S., Smith, E. L. III, Crawford, M. L. J., and von Noorden, G. K. (1997). Stereopsis and disparity vergence in monkeys with subnormal binocular vision. *Vision Research, 37,* 483-493.

Heinemann, E. G., Tulving, E., and Nachmias, J. (1959). The effect of oculomotor adjustments on apparent size. *American Journal of Psychology, 72,* 32-45.

Hering, E. (1879). *Spatial sense and movements of the eye.* Trans. C. A. Radde, (1942). Baltimore: American Academy of Optometry.

Hess, R. F., and Holliday, I. E. (1992). The coding of spatial position by the human visual system: effects of spatial scale and contrast. *Vision Research, 32,* 1085-1097.

Hess, R. F., and Wilcox, L. M. (1994). Linear and non-linear filtering in stereopsis. *Vision Research, 34,* 2431-2438.

Heywood, C. A., Cowey, A. and Newcombe, F. (1991). Chromatic discrimination in a cortically blind observer. *European Journal of Neuroscience, 3,* 802-812.

Hirsch, J. A., and Gilbert, C. D. (1991). Synaptic physiology of horizontal connections in the cat□s visual cortex. *Journal of Neuroscience, 11,* 1800-1809.

Holmes, G. (reprinted 1979). VI.8 Disturbances of spatial orientation and visual attention with loss of stereoscopic vision. In: C. G. Phillips (Ed.), *Selected Papers of Gordon Holmes.* Oxford: Oxford University Press.

Howard, I. P., Allison, R. S., and Howard, A. (1998). Depth from moving uncorrelated random-dot displays. *Investigative Opthalmology and Visual Science, 39, 4,* s3084.

Howard, I. P., and Rogers, B. J. (1995). Binocular Vision and Stereopsis. Oxford: Oxford University Press.

Hung, G. K., Semmlow, J. L., and Ciuffreda, K. J. (1986). A dual-mode dynamic model of the vergence eye-movement system. *IEEE Transactions on Biomedical engineering, 33,* 1021-1036.

Ittleson, W. H. and Ames, A. (1950). Accommodation, convergence and their relation to apparent distance. *Journal of Psychology, 30,* 43-62.

Jones, R. (1977). Anomalies of disparity detection in the human visual system. *Journal of Physiology, 264,* 621-640.

Julesz, B. (1960). Binocular depth perception of computer generated patterns. *Bell System Technical Journal, 39,* 1125-1162.

Julesz, B. (1971). *Foundations of Cyclopean Perception.* Chicago:University of Chicago Press.

Julesz, B., and Tyler, C. W. (1976). Neurontropy, an entropy-like measure of neural correlation, in binocular fusion and rivalry. *Biological Cybernetics, 23,* 25-32.

Kaneko H. and Howard I. P. (1995). Spatial organisation of vertical disparity pooling. *Investigative Opthalmology and Visual Science, 36,4,* S230.

Kaufman, L. (1974). Sight and mind. An introduction to visual perception. Oxford: Oxford University Press.

Kawano, K., and Miles, F. A. (1986). Short-latency ocular following responses of the monkey. II. Dependence on a prior saccadic eye-movement. *Journal of Neurophysiology (suppl.), 56,* 1355-1380.

Keeble, R. T., and Hess, R. F. (1998). Orientation masks 3-gabor alignment performance. *Vision Research, 38,* 827-840.

Kertesz, A. E., and Lee, H. J. (1987). Comparison of simultaneously obtained objective and subjective measurements of fixation disparity. *American Journal of Optometry and Physiological Optics, 64,* 734-738.

Koenderink, J. J., and Van Doorn, A. J. (1976). Geometry of binocular vision and a model for stereopsis. *Biological Cybernetics, 21,* 29-35.

Lehky, S. R., and Sjenowski, T. J. (1990). Neural model of stereoacuity and depth interpolation based on a distributed representation of stereo disparity. *Journal of Neuroscience, 10,* 2281-2299.

Liu, L., Stevenson, S. B., and Schor, C. M. (1994). Quantitative stereoscopic depth without binocular correspondence. *Nature, 367,* 66-69.

Logothetis, N. K., Leopold, D. A., and Sheinberg, D. L. (1996). What is rivaling during binocular rivalry. *Nature, 380,* 621-624.

Lunn, P. D. and Morgan, M. J. (1995). 'The analogy between stereo depth and brightness': a reexamination. *Perception, 24,* 901-904.

Mallot, H. A., Roll, A., and Arndt, P. A. (1995). Disparity-evoked vergence is directed towards average depth. *Max-Planck-Institut Technical Report No. 19, June.*

Mallot, H. A., Roll, A., and Arndt, P. A. (1996). Disparity-evoked vergence is driven by interocular correlation. *Vision Research, 36,* 2925-2937.

Manning, K A., and Riggs, L. A. (1984). Vergence eye movements and visual suppression. *Vision Research, 24,* 521-526.

Mansfield, J. S., and Legge, G. E. (1996). The bioncular computation of visual direction. *Vision Research, 33,* 1645-1658.

Mapp, A. P., and Ono, H. (1998). The wandering cyclopean eye. *Vision Research,* accepted.

Marr, D. (1982). *Vision.* San Francisco: Freeman.

Marr, D., and Poggio, T. (1976). Cooperative computation of stereo disparity. *Science, 194,* 283-287.

Marr D., and Poggio T. (1979). A Computational theory of human stereo vision. *Proceedings of the Royal Society of London B, 204,* 301-328.

Marshall J. A., Kalarickal, G. J., and Graves, E. B. (1996). Neural model of visual stereomatching: slant, transparency and clouds. *Netowrk: Computation in Neural Systems, 7,* 635-669.

Masson, G. S., Bussetini, C., and Miles, F. A. (1997). Vergence eye movements in response to binocular disparity without depth perception. *Nature, 389,* 283-286.

Mayhew, J. E. W., and Frisby, J. P. (1980). The computation of binocular edges.

Perception, 9, 69-86.

Mayhew, J. E. W. and Longuet-Higgins, H. C. (1982). A computational model of binocular depth perception. *Nature, 297,* 376-379.

Mays, L. E., and Gamlin, P. D. R. (1995). A neural mechanism subserving saccade-vergence interactions. In J. M. Findlay, R. Walker and R. W. Kentridge (Eds.), *Eye Movement Research.* Elsevier Science.

McKee, S. P., and Harrad, R. A. (1993). Fusional suppression in normal and stereoanomalous observers. *Vision Research, 33,* 1645-1658.

McKee, S. P., and Levi, D. M. (1987). Dichoptic hyperacuity: the precision of nonius alignment. *Journal of the Optical Society of America A, 4,* 1104-1108.

McKee, S. P., Welch, L., Taylor, D. G. and Bowne, S. F. (1990) Finding the common bond: stereoacuity and the other hyperacuities. *Vision Research, 30,* 879-891.

McLin, L. N. Jr., Schor, C. M., and Kruger, P. B. (1988). Changing size (looming) as a stimulus to accommodation and vergence. *Vision Research, 28,* 883-898.

Miles F. A., and Kawano, K. (1987). Visual stabilization of the eyes. *Transactions in Neuroscience, 10,* 153-158.

Miles, F. A., Kawano, K. and Optican, L. M. (1986). Short-latency ocular following responses of monkey. I. Dependence on temporospatial properties of the visual input. *Journal of Neurophysiology, 56,* 1321-1354.

Mitchell, D. E. (1966). Retinal disparity and diplopia. *Vision Research, 6,* 441-451.

Morgan, M. J., and Aiba, T. S. (1986). Vernier acuity predicted from changes in the light distribution of the retinal image. *Spatial Vision, 1,* 151-161.

Morgan, M. J., and Fahle, J. (1992). Effects of pattern element density upon displacement limits for motion detection in random binary luminance patterns. *Proceedings of the Royal Society of London B, 248,* 189-198.

Mowforth, P., Mayhew, J. E. W., and Frisby, J. P. (1981). Vergence eye movements made in response to spatial frequency filtered random-dot stereograms. *Perception, 10,* 299-304.

Müller, J. (1826). Zur vergleichenden Physiologie des Gesichtssinnes des Menschen und der Thiere. Cnobloch: Leipzig. Trans. W. Baly, (1843). *Elements of Physiology,* (pp. 1136-1150). Taylor and Walton: London.

Nakayama, K., and Shimojo, S. (1990). Da Vinci stereopsis: depth and subjective occluding contours from unpaired image points. *Vision Research, 30,* 1811-1825.

Ogle, K. N. (1932). Analytical treatment of the longitudinal horopter. *Journal of the Optical Society of America, 22,* 665-728.

Ogle, K. N. (1950). *Researches in Binocular Vision.* Philadelphia: Sanders.

Ogle, K. N. (1952). On the limits of stereoscopic vision. *Journal of Experimental Psychology, 44,* 253-259.

Ogle, K. N. (1953). Precision and validity of stereoscopic depth perception from double images. *Journal of the Optical Society of America, 43,* 906-913.

Ohzawa, I., DeAngelis, G. C., and Freeman, R. D. (1996). Encoding of binocular disparity by simple cells in the cat□s visual cortex. *Journal of Neurophysiology, 75,* 1779-1805.

Ohzawa, I., DeAngelis, G. C., and Freeman, R. D. (1997a). Encoding of binocular disparity by complex cells in the cat□s visual cortex. *Journal of Neurophysiology, 77,* 2879-2909.

Ohzawa, I., DeAngelis, G. C., and Freeman, R. D. (1997b). The neural coding of stereoscopic depth. *Neuroreport, 8,* R3-R12.

Olzak, D. A., and Thomas, J. P. (1986). Chapter 7: Seeing spatial patterns. In K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of Perception and Human Performance VI: Sensory Processes and Perception.* New York: Wiley.

Ono, H. (1983). The combination of version and vergence. In: C. M. Schor, and K. J. Ciuffreda (Eds.) *Vergence Eye Movements: Basic and Clinical Aspects* (pp. 373-400). Boston: Butterworth..

Ono, H. (1991). Binocular visual directions of an object when seen as a single or double. In: D. Regan, (Ed.) *Binocular Vision* (ch. 1, pp. 1-18). In: J. Cronly-Dillon, (Ed.) *Vision and Visual Dysfunction, Volume 9.* London: Macmillan.

Ono, H., and Nakamizo, S. (1978). Changing fixation in the transverse plane at eye level and Hering□s law of equal innervation. *Vision Research, 18,* 511-519.

Ono, H., Shimono, K., and Saida, S. (submitted). Transformation of the visual line value in binocular vision: stimuli on corresponding points can be seen in two different visual directions.

O'Shea, R. P., and Blake, R. (1987). Depth without disparity in random-dot stereograms. *Perception and Psychophysics, 42,* 205-213.

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

Panum, P. L. (1858). *Physiologie Untersuchungen ‚ber das Sehen mit zwei Augen.* Schwers Buchandlung: Kiel.

Parker, D. M., and Dutch, S. (1987). Perceptual latency and spatial-frequency. *Vision Research, 27,* 1279-1283.

Poggio, G. F. (1991). Physiological basis of stereoscopic vision. In D. Regan, (Ed.) *Binocular Vision,* (Ch. 11, pp. 234-238). In Cronly-Dillon (Ed.) *Vision and Visual dysfunction, Vol. 9.*

Poggio, G. F., and Fischer, B. (1977). Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. Journal of Neurophysiology, 40, 1392-1405.

Poggio, G. F., and Poggio, T. (1984). The analysis of stereopsis. *Annual Review of Neuroscience, 7,* 379-412.

Poggio, G. F., and Talbot, W. H. (1981). Mechanisms of static and dynamic stereopsis in foveal cortex of the rhesus monkey. *Journal of Physiology, 315,* 469-492.

Polat U., and Sagi D. (1993). Lateral interactions between spatial channels - suppression and facilitation reveald by lateral masking experiments. *Vision Research, 33,* 993-99.

Pollard, S. B., Mayhew, J. E. W., and Frisby, J. P. (1985). PMF: a stereo correspondence algorithm using a disparity gradient limit. *Perception, 14,* 449-470.

Poppelreuter W. (1917). Translated Zihl (1990). *Disturbances of lower and higher visual capacities caused by occipital damage,* Section 9: Disturbances of depth perception, 69-80. Oxford: Oxford Science Publications.

Popple, A. V., Smallman, H. S., and Findlay, J. M. (1998). The area of spatial integration for initial horizontal disparity vergence. *Vision Research, 38,* 319-326.

Rashbass, C., and Westheimer, G. (1960). Recording rotational eye movements independently of lateral displacement. *Journal of the Optical Society of America, 50,* 642-644.

Rashbass, C., and Westheimer, G. (1961). Disjunctive eye movements. *Journal of Physiology, 159,* 361-364.

Reimann, D., Ditzinger, T., Fischer, E., Haken, H. (1995). Vergence eye-movement control and multivalent perception of autostereograms. *Biological Cybernetics, 73,* 123-128.

Remole, A. (1985). Fixation disparity vs. binocular fixation misalignment. *American Journal of Physiology and Physiological Optics, 62,* 25-34.

Remole, A., Code, S. M., Matyas, C. E., McLeod, M. A., and White, D. J. (1986). Objective measurement of binocular fixation misalignment. *American Journal of Optometry and Physiological Optics, 63,* 63-68.

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

Richards, W. (1971). Anomalous stereoscopic depth perception. *Journal of the Optical Society of America, 61,* 410-414.

Rizzo, M., Nawrot, M., and Zihl, J. (1995). Motion and shape perception in cerebral akinetopsia. *Brain, 118,* 1105-1127.

Rogers, B. J., and Bradshaw, M. F. (1993). Vertical disparities, differential perspective and binocular stereopsis. *Nature, 361,* 253-255.

Rogers, B. J., and Bradshaw, M. F. (submitted). Disparity minimisation, cyclovergence and the validity of nonius lines as a technique for measuring torsional alignment. *Perception, special issue following AVA Depth Perception meeting.*

Rogers, B. J., and Graham, M. E. (1982). Similarities between motion parallax and stereopsis in human depth perception. *Vision Research, 22,* 261-270.

Rogers, B. J., and Graham, M. E. (1985). Motion parallax and the perception of three-dimensional surfaces. In D. Ingle, M. Jeannerod, and D. Lee (Eds.) *Brain mechanisms and spatial vision.* (pp. 95-111). The Hague: Martinus Nijhoff.

Rohaly, A. M., and Wilson, H. R. (1993). The nature of coarse-to-fine constraints on binocular fusion. *Journal of the Optical Society of America A, 10,* 2433-441.

Rohaly, A. M., and Wilson, H. R. (1994). Disparity averaging across spatial scales. *Vision Research, 34,* 1315-1325.

Schmidt, K. E., Goebel, R., Lowel, S., and Singer, W. (1997). The perceptual grouping criterion of colinearity is reflected by anisotropies of connections in the primary visual cortex. *European Journal of Neuroscience, 9,* 1083-1089.

Schor, C. M. (1983). Fixation disparity and vergence adaptation. In C. M. Schor, and K. J. Ciuffreda (Eds.) *Vergence Eye Movements: Basic and Clinical Aspects.* (Ch. 14, pp. 465-516). Boston: Butterworth.

Schor, C. M., and Howarth, P. A. (1985). Suprathreshold stereo-depth matches as a function of contrast and spatial frequency. *Perception, 15,* 249-258.

Schor, C., Heckmann, T., and Tyler, C. W. (1989). Binocular fusion limits are independent of contrast, luminance gradient and component phases. *Vision Research, 29,* 821-835.

Schor, C., Robertson, K. M., and Wesson, M. (1986). Disparity vergence dynamics and fixation disparity. *American Journal of Optometry and Physiological Optics, 63,* 611-618.

Schor, C. M., Wood, I. C., and Ogawa, J. (1984). Spatial tuning of static and dynamic local stereopsis. *Vision Research, 24,* 661-665.

Schumer, R. A., and Julesz, B. (1984). Binocular disparity modulation sensitivity to disparities offset from the plane of fixation. *Vision Research, 24,* 533-542.

Semmlow, J. L., and Carpenter, T. (1996). Interaction of initial and late components in disparity vergence. *Investigative Opthalmology and Visual Science, 37(3),* s780.

Semmlow, J. L., and Hung, G. K. (1983). The near response: theories of control. In C. M. Schor and K. J. Ciuffreda (Eds.), *Vergence Eye Movements: Basic and Clinical aspects*, (ch. 6, pp. 175-195). Boston: Butterworth.

Semmlow, J. L., Hung, G., Horng, J-L, and Ciuffreda, K. J. (1994). Disparity vergence eye-movements exhibit preprogrammed motor control. *Vision Research, 34,* 335-343.

Sheedy, J. E., Bailey, I. L., Buri, M., and Bass, E. (1986). Binocular vs. monocular performance. *Americal Journal of Optometry and Physiological Optics, 63,* 839-846.

Shimono, K., Ono, H., Saida, S., and Mapp, A. P. (1998). Methodological caveats for monitoring binocular eye position with Nonius stimuli. *Vision Research, 38,* 591-600.

Smallman, H. S. (1995). Fine-to-coarse scale disambiguation in stereopsis. *Vision Research, 35,* 1047-1060.

Smallman, H. S., and MacLeod, D. I. A. (1997). Spatial scale interactions in stereo sensitivity and the neural representation of binocular disparity. *Perception, 26,* 977-994.

Smallman, H. S., and McKee, S. P. (1995). A contrast ratio constraint on stereo matching. *Proceedings of the Royal Society of London B, 260,* 265-271.

Sperling, G. (1970). Binocular vision: A physical and a neural theory. *American Journal of Psychology, 83,* 461-534.

Stark, L. (1983). Normal and Abnormal Vergence. In: C. M. Schor, and K. J. Ciuffreda (Eds.) Vergence Eye Movements: Basic and Clinical Aspects.(ch. 1, pp. 3-13). Boston: Butterworth.

Statham, A. K., Georgeson, M. A. (1997). Stereo depth perception from linear combination of first and second-order disparities. *Perception, 26,* 10, 29.

Stevenson, S. B., Cormack, L. K., and Schor, C. M. (1989). Hyperacuity, superresolution and gap resolution in human stereopsis. *Vision Research, 29,* 1597-1605.

Stevenson, S. B., Cormack, L. K., and Schor, C. M. (1994). The effects of stimulus contrast and interocular correlation on disparity vergence. *Vision Research, 34,* 383-396.

Stevenson, S. B., Cormack, L. K., Schor, C. M., and Tyler, C. W. (1992). Disparity tuning mechanisms of human stereopsis. *Vision Research, 32,* 1685-1694.

Stevenson, S. B., Lott, L. A., and Yang, J. (1997). The influence of subjective instruction on horizontal and vertical vergence tracking. *Vision Research, 37,* 2891-2898.

Tanaka, K., Hikosaka, K., Saito, H., Yukie, M., Fukada, Y., and Iwai, E. (1986). Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. *Journal of Neuroscience, 6,* 134-144.

Tanaka, S. (1995). Topological analysis of point singularities in stimulus preference maps of the primar visual-cortex. *Proceedings of the Royal Society of London B, 261,* 81-88.

Tani, G. T., Ogle, K. N., Weaver, R. W., and Martens, T. G. (1956). On the precise objective determination of eye movements. *American Medical Association Archives in Opthalmology, 55,* 174-185.

Thompson, A. (1997). An evolved circuit, intrinsic in silicon, entwined with physics. In T. Higuchi, and M. Iwata (Eds.), *Proceedings of the First International Conference on Evolvable Systems: from Biology to Hardware*, (pp 390-405). Springer Verlag LNCS 1259.

Toet, A., and Levi, D. M. (1992). The two-dimensional shape of spatial interaciton zones in the parafovea. *Vision Research, 32,* 1349-1357.

Tyler, C. W. (1975). Spatial organisation of binocular disparity sensitivity. *Vision Research, 15,* 583-590.

Tyler, C. W. (1976). Observations on binocular frequency reduction in random noise. *Perception, 4,* 305-309.

Tyler, C. W. (1983). Sensory processing of Binocular Disparity. In C. M. Schor and K. J. Ciuffreda (Eds.), *Vergence Eye Movements: Basic and Clinical Aspects,* (pp. 199-295). Boston: Butterworth.

Tyler, C. W. (1991). The horopter and binocular fusion. In: D. Regan, (Ed.). *Binocular Vision*, (ch. 2, pp. 19-37). In: J. Cronly-Dillon, (Ed). Vision and Visual Dysfunction, Vol. 9. London: Macmillan.

Tyler, C. W. (1995). Chapter in: T. V. Papathomas, C. Chubb, A. Gorea and E. Kowler (Eds.). *Early Vision and beyond.* Cambridge, MA: MIT Press.

Tyler, C. W., and Clarke, M. B. (1990). The autostereogram. *Proceedings of the International Society for Optical Engineering, 1256,* 182-197.

Tyler, C. W., and Julesz, B. (1978). Binocular cross-correlation in time and space. *Vision Research, 18*, 101-105.

Ungerleider, L. G., and Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, and R. J. W. Mansfield (Eds.) *Analysis of Visual Behavior*, Cambridge MA: MIT Press.

Van Ee, R. (1995). *Stability of binocular depth perception.* Doctoral thesis. The Netherlands: Helmholtz Institute.

Van Ee, R. and Erkelens, C. J. (1996). Stability of binocular depth perception with moving head and eyes. *Vision Research, 36,* 3827-3842.

Watt, R. J. (1987). Scanning from coarse to fine spatial scales in the human visual system after the onlset of a stimulus. *Journal of the Optical Society of America A, 4,* 2006-2021.

Waugh, S. J., and Levi, D. M. (1993). Visibility, timing and vernier acuity. *Vision Research, 33,* 505-526.

Westheimer, G. (1979). Cooperative neural processes involved in stereoscopic acuity. *Experimental Brain Research, 36,* 585-597.

Westheimer, G., and McKee, S. P. (1978). Stereoscopic acuity for moving retinal images. *Journal of the Optical Society of America, 68,* 450-455.

Westheimer, G. (1998). Lines and gabor functions compared as spatial visual stimuli. *Vision Research, 38,* 487-491.

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

Whitaker, D. and MacVeigh, D. (1991). Interaction of spatial frequency and separation in vernier acuity. *Vision Research, 31,* 1205-1212.

Wilcox, L. M., and Hess, R. F. (1995). D-max for stereopsis depends on size, not spatial frequency content. *Vision Research, 35,* 1061-1069.

Wilcox, L. M., and Hess, R. F. (1996). Is the site of nonlinear filtering in stereopsis before or after binocular combination. *Vision Research, 36,* 391-399.

Wilcox, L. M., and Hess, R. F. (1997). Scale selection for second-order (non-linear) stereopsis. *Vision Research, 37,* 2981-2992.

Wilson, H. R., Blake, R., and Halpern, D. L. (1991). Coarse spatial scales constrain the range of binocular fusion on fine scales. *Journal of the Optical Society of America A, 8,* 229-236.

Wilson, H.R., Levi, D., Maffei, L., Rovamo, J. and DeValois, R. (1990). The perception of form: retina to striate cortex. In: L. Spillmann, and J. S. Werner (Eds.), *Visual Perception: The Neurophysiological Foundations* (pp. 231-272). San Diego: Academic Press.

Young, M. J., Landy, M. S., and Malony, L. T. (1993). A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision Research, 33,* 2685-2696.

Zee, D. S., Fitzgibbon, E. J., and Optican, L. M. (1992). Saccade-vergence interactions in humans. *Journal of Neurophysiology,* 1624-1641.

Zeki, S. (1993). *A vision of the brain.* Oxford: Blackwell Scientific Publications.

Zuber, B. L., and Stark, L. (1968). Dynamical characteristics of the fusional vergence eye-movement system. *IEEE Transactions on Systems, Science and Cybernetics, 4*, 72-79.

Zurek, W. H. (1989). Thermodynamic cost of computation, algorithmic complexity and the information metric. *Nature, 341*, 119-124.