FIRST AND SECOND ORDER STEREOSCOPIC PROCESSING OF FUSED AND DIPLOPIC TARGETS

DEBI STRANSKY

A DISSERTATION SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN PSYCHOLOGY
YORK UNIVERSITY
TORONTO, ONTARIO

September 2014

Abstract

Depth from stereopsis is due to the positional difference between the two eyes, which results in each eye receiving a different view of the world. Although progress has been made in understanding how the visual system processes stereoscopic stimuli, a number of questions remain. The goal of this work was to assess the relationship between the perceptual, the temporal and the 1st- /2nd- order dichotomies of stereopsis and in doing so, determine an appropriate method for measuring depth from large disparities. To this end, stereosensitivity and perceived depth were assessed using 1st- and 2nd- order stimuli over a range of test disparities and conditions. The main contributions of this research are as follows: 1) The sustained/transient dichotomy proposed by Edwards, Pope and Schor (2000) is best considered in terms of the spatial dichotomy proposed by Hess and Wilcox (1994). At large disparities it is not possible to categorize performance based on exposure duration alone; 2) There is not a simple correspondence between Ogle's (1952) patent / qualitative perceptual categories and the 1st- /2nd- order dichotomy proposed by Hess and Wilcox (1994); 3) Quantitative depth is provided by both 1st- and 2nd- order mechanisms in the fused range, but only the 2nd- order signal is used when stimuli are diplopic; 3) The quantitative depth provided by a 2nd- order stimulus scales with envelope size; and 4) The monoptic depth phenomenon may be related to depth from diplopic stimuli, but for conditions tested here when both monoptic depth and 2nd- order stereopsis are available, the latter is used to encode depth percepts. The results reported here expand on earlier work on 1st- and 2nd-

order stereopsis and address the issues in the methodologies used to study depth from large disparities. These results are consistent with the widely accepted filter-rectify-filter model of 2nd- order processing, and 1st- and 2nd- order stimuli are likely encoded by disparity-sensitive neurons via a two-stream model (see Wilson, Ferrera, and Yo (1992); Zhou and Baker (1993)).

Acknowledgements

I am especially grateful to my supervisor, Laurie Wilcox, whose continued support, understanding, expertise and encouragement made it possible to complete this work. I will be forever grateful to Laurie to helping me get to the finish line.

Thank you to my supervisory members, Hugh Wilson and Robert Allison for providing me with their guidance and advice. I would also like to acknowledge Robert Hess who was a collaborator in the first experiment described in this dissertation. Also, a big thank-you to Bob Hou for all of his help, and to the observers who contributed many hours of their time participating in my experiments.

I would like to thank my family for supporting me through this long process. I am very grateful to my husband, Hanan, and to my parents who have always supported me. Lastly, thank you to my daughters, Katherine and Jessica who always make me smile.

Table of Contents

List of Figures

Chapter 1

Introduction

1.1. Geometry of Stereopsis

The lateral separation of the eyes results in each retina receiving a different view of the world. As a result, the image of a particular object will fall on different or *disparate* retinal locations. This difference in location is referred to as retinal disparity, and is the primary information used by the stereoscopic system to determine relative depth. The geometry of binocular disparity is outlined in Figure 1.1, assuming fixation on a point F. Since images of the fixation point, F, fall on the fovea in each eye, the binocular disparity at this location is zero. The angular disparity, η, between the images of point P and fixation is the binocular subtense of the fixation point (F) minus the binocular subtense of the point P, $\eta=\phi_l-\phi_r=\theta_F-\theta_P$.

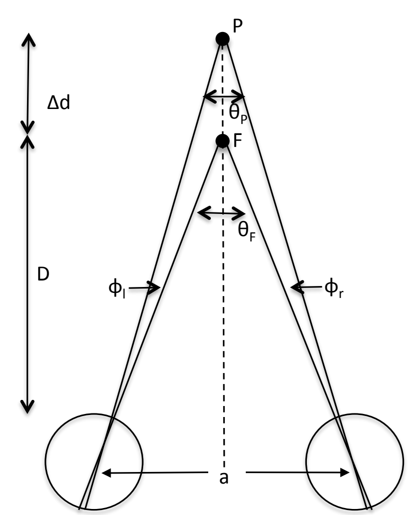


Figure 1.1. Illustration of the geometry defining the disparity of a point, P, relative to fixation point, F, at a distance, D. θ_p is the binocular subtense of the eyes at P, θ_F is their binocular subtense at F, a is the interocular distance, Δd is the distance between points P and F, and the difference between ϕ_l and ϕ_r is the angular disparity (η).

The visual system uses retinal disparities to determine a stereoscopic percept of depth. Absolute disparity is the difference in the two retinal locations corresponding to a single point (P) and provides a depth estimate relative to where the observer is fixating (Howard & Rogers, 2002, p. 217). The object at point P appears fused at small disparities, but appear diplopic, or double, when the retinal disparity is large. Relative disparity, on the other hand, is independent of fixation and is represented by the difference between the absolute disparities of two objects. It has been shown (among others, Westheimer (1979)) that compared to absolute disparity, relative disparity provides more precise depth judgments.

The Vieth-Müller defines the locations in space where objects stimulate corresponding points on the retinas and will have zero relative disparity. As illustrated in Figure 1.2, a circle passing through the nodal points of the right eye (ϕ_l), left eye ($\phi_{r)}$ and the fixation point (F) represents the theoretical horopter. With Wheatstone's (1838) invention of the stereoscope, it became possible to measure the empirical horopter. One method to determine the empirical horopter is by asking observers to judge whether objects in varying directions in depth and lateral distances from fixation appear fused or diplopic (Howard & Rogers, 1995, p. 53). When tested empirically, the points that yield single vision are not the same as those predicted by the theoretical horopter. Instead, the empirical horopter is not perfectly circular (Howarth, 2011) and is less concave at close viewing distances (Howard & Rogers, 1995, p. 57).

The empirical horopter defines the locations on the retina that yield singleness of vision. Images appear fused when they fall on corresponding points or within a range of points near the point of correspondence (Howard & Rogers, 1995). Panum (1858) demonstrated that there is a range of disparate points that appear single, or fused. He showed that an image on a retinal point of one eye fuses with a similar image located anywhere within a small area on the other retina (Ogle, 1964b). This idea was contrary to the discoveries described in the early work by Ptolemy and Alhazen who posited that only objects stimulating corresponding points on the retina were seen as single, while disparate points were perceived as diplopic (Howard, 1996; Howard & Wade, 1996). Panum's region of single vision is known as 'Panum's fusional area'. Within this zone, as retinal disparity increases, so does relative depth although the image still appears single. Disparate points beyond this region will be seen as two points, or diplopic (Figure 1.2). While there is a dramatic difference in the appearance of stimuli that are fused vs. diplopic, it is well established that depth percepts can be obtained from both types of stimuli (Blakemore, 1970; Foley, Applebaum, & Richards, 1975; Helmholtz, 1910; Hering, 1942; Hess & Wilcox, 1994; Ogle, 1952b, 1953; Richards, 1971; Richards & Kaye, 1974; Westheimer & Tanzman, 1956; Wilcox & Hess, 1995). Properties of these percepts will be discussed in Chapters 3 and 4.

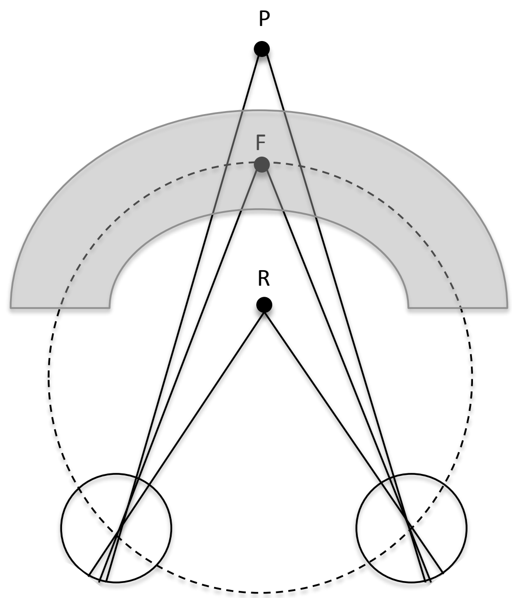


Figure 1.2. Illustration of Panum's fusional area (not to scale). The fixation point, F, falls on the fovea and is perceived as single. Points P and R produce images that correspond to more disparate retinal points which fall outside of Panum's fusion area, and are perceived as diplopic.

1.2. Physiological Basis of Stereopsis

1.2.1. Receptive fields and binocular disparity

Hubel and Wiesel (1959, 1962) were the first to show that cells that are sensitive to binocular stimulation have similar receptive fields (RF) for each eye. That is, aside from ocular dominance, the sub-regions of a simple cell's RFs in each eye were activated by the same property, for example the same orientation of a line stimulus. Disparity-selective neurons were first discovered in the cat's primary visual cortex by Barlow, Blakemore, and Pettigrew (1967). They found that cells in area 17 were responsive to binocular disparity and that there was an optimal distance between RF centres for maximal activation. Similar findings were reported by Pettigrew, Nikara, and Bishop (1968).

Disparity-selective cells were further investigated in subsequent experiments (DeAngelis, Ohzawa, & Freeman, 1991; Fischer & Krueger, 1979; Freeman & Ohzawa, 1990; Hubel & Wiesel, 1973; Joshua & Bishop, 1970; LeVay & Voigt, 1988; Maske, Yamane, & Bishop, 1984). Maske et al. (1984) confirmed Barlow and colleagues' (1967) findings and showed that disparity-selective neurons have similar RFs in each eye and that disparity selectivity results from slight differences in the position of the RFs. Freeman and colleagues (1986, 1991, 1995) demonstrated that disparity selectivity could also result from RFs that have different shapes, or phases, in each retina. Position and phase coding is discussed in more detail in Section 1.2.3.

1.2.2. Types of disparity-sensitive neurons

Poggio and colleagues (1995; 1977; 1988) recorded activity from individual cells in the cat and the monkey visual cortex and found that cells in the primary visual cortex (V1) could be divided into distinct classes. Tuned zero (TO) and tuned inhibitory (TI) cells responded to zero or near-zero disparities, and both types of neurons exhibited a symmetric response profile about zero disparity. They found that when an object was close to fixation (within ± 6 arc min), TO cells were maximally excited while TI cells were suppressed. In V1, tuned near (TN) and tuned far (TF) neurons had broad asymmetric profiles (Poggio, 1995) and had peak responses in the crossed (TN) or uncrossed (TF) direction for disparities larger than

6 arc min but less than 12 arc min (Poggio & Fischer, 1977). In a later study, TN and TF neurons in more peripheral V1 areas, V2 and V3-V3A were found to respond to disparities up to 35-40 arc min (Poggio et al., 1988).

The final category of neurons identified by Poggio and his colleagues was reciprocal neurons, labeled near (NE) and far (FA) cells. NE cells exhibited maximal excitation for a wide range of large crossed disparities and maximum suppression for all uncrossed disparities, and FA cells had the opposite response profile. These six classes of cells (T0, TI, TN, TF and NE/FA) were also found in area V2 (Poggio & Fischer, 1977)and V3-V3A (Poggio et al., 1988). Furthermore, the proportion of disparity-selective neurons appears to increase in cortical areas beyond the early processing areas (Poggio, 1995). For example, Maunsell and Van Essen (1983) found that two thirds of neurons in MT are disparity sensitive, while 90% of neurons are disparity-sensitive in MST (Roy, Komatsu, & Wurtz, 1992).

Prince, Cumming, and Parker (2002) recorded responses to binocular disparity (0 to 1°) in random-dot stereograms from 787 V1 neurons in the macaque. Their data is based on the responses from a selection of 180 neurons that were well fit by a Gabor function. They concluded that the distinct classes of neurons described by Poggio are not found in V1, rather a continuum of neurons exists. However, because they eliminated neurons with asymmetric profiles, they likely excluded near and far cells. The cells that were selected were most likely *only* tuned cells that were selective for disparities less than 1°. Therefore, these results do not provide definitive evidence for a continuum because an entire population of neurons that likely account for depth percepts at large disparities was excluded.

1.2.3. Phase coding and the energy model

As outlined above, early investigators argued that disparity sensitivity results from the relative displacement (positions) of RFs in the two eyes (Barlow et al., 1967; Hubel & Wiesel, 1962; Joshua & Bishop, 1970; Maske et al., 1984), which are identical in shape (Hubel & Wiesel, 1962; Maske et al., 1984). Thus, this model assumes that the preferred disparity for a disparity-selective simple cell is determined by the displacement of the binocular RFs in the right and left eyes.

An alternative to position coding in simple cells posits that RFs in the left and right eyes are centered at corresponding retinal locations but are different in shape, or phase (Freeman & Ohzawa, 1990). According to this model, disparity is determined by the difference in phase between two RFs in each eye (DeAngelis, Ohzawa, & Freeman, 1995) The disparity energy model, proposed by Ohzawa, DeAngelis, and Freeman (1990), postulates that at least four simple cells converge onto one complex cell. The input consists of pairs of RFs from simple cells, each in quadrature phase, which means that each simple cell's spatial frequency phase is shifted (within the pair) by 90°. Therefore, when viewing a stimulus at a specific disparity, a given complex cell will be activated if at least one of the simple cells is responsive. The disparity energy model assumes that all of the RFs have the same centre spatial period, orientation, and are in quadrature with each other (Ohzawa, DeAngelis, & Freeman, 1997). Importantly, a given phase-based disparity detector will only be able to reliably encode disparities less than one cycle of the spatial frequency to which it responds. However, one problem with this area of research where anesthetized animals are used is the lack of a satisfactory method for determining zero disparity.

The disparity energy model assumes that the input from the left and right retinas are summed linearly by a subsequent binocular neuron, therefore this model cannot account for depth from anti-correlated stimuli. Cumming and Parker (1997) and Read, Parker, and Cumming (2002) showed that there is a reduction in amplitude of the disparity tuning curve as well as phase shifts other than π (although they were mostly close to π) when viewing binocularly anti-correlated stimuli. Therefore they suggested a modification to Ohzawa and colleagues' (1990) model to account for the nonlinearity required to encode the disparity of anti-correlated stimuli. Read et al. (2002) suggested that input from each RF initially projects to a monocular simple cell, where it undergoes a half-wave rectification before binocular combination. Subsequently, information from the two monocular simple cells feed into a binocular complex cell. Read and colleagues' (2002) modified energy model can account for responses of disparity-tuned neurons to both correlated and anti-correlated stimuli.

1.3. Evidence of a Continuum and a Dichotomous System in Stereopsis

As described above, Poggio and colleagues identified several classes of disparity-tuned cells with distinct response properties. Their work and that of Ferster (1981) and others showed that within a relatively small range of disparities, tuned inhibitory and excitatory neurons have symmetric, and graded response profiles. However, the NE/FA neurons respond only to relatively large disparities;

they do not have a graded response to disparity and their response profiles are asymmetric. Given these fundamental differences it seems plausible that these neurons form the substrate for depth percepts from large or coarse disparities, including those beyond the fusional range. The substantial differences between the tuned cells and the NE/FA cells have supported the notion that there may be two streams of disparity processing, depending on the disparity of the stimulus: a fine high-resolution system subserved by the Tuned neurons, and a coarse low-resolution but broad range system that relies primarily on the NE/FA neurons.

Phase and position coding provide for different predictions for the mechanisms underlying stereopsis. The energy model necessitates that binocular disparity be processed using a continuum of spatial frequency tuned neurons whose range of disparity to which they respond scales with the size of the RF. Thus, disparity selective cells that are tuned to low spatial frequencies encode large disparities while cells tuned to high spatial frequencies encode small disparities (DeAngelis et al., 1995; Maffei & Fiorentini, 1977; Marr & Poggio, 1979). A position coding model does not require a size disparity correlation. There does not have to be a correlation between the size of matching receptive fields and the cells' preferred disparity (Howard & Rogers, 1995, p. 140). That is, small receptive fields can encode large disparities if the RFs are sufficiently offset. Similarly, neurons with large RFs can encode small disparities (Pettigrew et al., 1968).

Modern computational models of stereopsis have relied on a combination of phase and position coding to accurately describe the physiological literature and human performance. In particular, they rely on position coding for disparities that appear diplopic, as the displacements tend to be large relative to the scale of the stimulus. Fleet, Wagner, and Heeger (1996) provided a mathematical analysis of a binocular energy model and measure the relative contributions of phase and position coding. There is a limited range of unambiguous phase disparities that can be detected by phase shifted neurons (-π to π). This makes phase coding particularly restrictive for stimuli with high spatial frequencies, as a very limited range of disparities can be detected. Therefore, they propose a hybrid model where disparity is represented by both a phase shift and a position shift. Position shifts must be present because a phase-model, especially for stimuli with high spatial frequencies, cannot account for depth perception beyond the fusion limit. Anzai, Ohzawa, and Freeman (1999) recorded neural activity from simple cells of the cat's striate cortex to estimate phase and receptive field position disparities. They too found evidence of both phase and position shifts. Phase shifts were prevalent at smaller disparities but could not account for the encoding of large binocular disparities. Furthermore, they argued that stimuli with high spatial frequencies are restricted to a very narrow range of phase shifts. Therefore position shifts are essential for encoding depth from high frequency stimuli. Neurophysiological and computational studies have shown that the way disparity is encoded might be dependent on the amount of retinal disparity. Thus, phase coding could account for depth at small disparities, while position encoding may be necessary for processing depth from large disparities.

1.3.1. Psychophysical evidence for a continuum subserving stereopsis

Prior to the work conducted by Poggio and colleagues, Richards and Kaye (1974) looked for psychophysical evidence of separate disparity mechanisms. They defined the mechanisms as 'local' stereopsis, which operates at small disparities (less than 1/2°) when stimuli are fused, and 'global' stereopsis, which operates at large disparities (greater than 1/2°) when stimuli are diplopic and/or unmatched. In their first experiment, they used a depth estimation task and presented observers with a range of stimulus bar widths and disparities (0.06, 0.1, 0.25, 0.5, 1, 2, 4, 8, and 16 degrees). They predicted that if two mechanisms exist, they should find two peaks (a bimodal distribution) in perceived depth as a function of disparity. Their data did not seem to exhibit two peaks so they concluded that a single mechanism underlies stereopsis across the entire range of disparities. Unfortunately, they mainly tested very large disparities, which were likely just at and above the fusion limit; thus it was not surprising that their depth from disparity function was not bimodal.

In their second study, Richards and Kaye (1974) tested a stereoanomalous observer using the same depth estimation task. They reasoned that if two mechanisms exist then this observer's stereoanomaly should be present for only large but not small disparities. They concluded that the anomaly was evident across the various stimulus sizes and disparities tested, therefore they supported a continuum rather than a local/global dichotomy. However, it should be noted that their conclusions are based partly on the results of *one* stereoanomolous subject (IB), who did not perceive depth from disparities in only the uncrossed range. This is not surprising since they reported in a previous experiment (Richards, 1971) that this observer was deficient in uncrossed depth. Richards and Kaye's (1974) study is one of the most cited psychophysical works in support of a continuum of disparity processing from fine to coarse scales, however it is clear that this study's flaws make it difficult to interpret.

1.3.2. Psychophysical evidence for a perceptual dichotomy

While it is possible that the categories of neurons identified by Poggio are not strict categories but rather lie on a continuum of disparity processing, there is a large body of psychophysical evidence that favours a two-mechanism[1] description underlying stereopsis (Badcock & Schor, 1985; Blakemore, 1970; Edwards, Pope, & Schor, 2000; Hess & Wilcox, 1994; Kovács & Fehér, 1997; Langley, Fleet, & Hibbard, 1999; Lin & Wilson, 1995; McKee, Verghese, & Farell, 2004, 2005; Ogle, 1952b; Sato, 1983; Schor & Wood, 1983; Shimono, 1984; Wilcox & Hess, 1995; Wilcox & Hess, 1996, 1997, 1998). In the early 1950s, Ogle conducted a number of depth magnitude estimation experiments in which he documented changes in depth percepts as a function of disparity. In one of Ogle's (1952b) classic experiments, observers fixated on an object while the disparity of a neighbouring object was gradually increased. Initially, observers reported that the distance between the two objects appeared to increase with increasing disparity. This proportional increase in perceived depth continued even as the stimulus became diplopic. Eventually, a point was reached,

[1] Many researchers describe a system based on two completely separate mechanisms as a dichotomous system. For clarification, they do not necessarily function in opposition to each other, but differ in their perceptual, spatial, temporal or physiological properties.

which Ogle termed the ‘patent limit’, where the perceived depth magnitude plateaued. Importantly, at these large test disparities observers were still able to indicate the direction or sign of the displacement in depth (ie. nearer or further).

From this and subsequent experiments, Ogle argued that there are two types of stereoscopic depth percepts that arise from binocular disparity: patent and qualitative. He stated that patent depth percepts are quantitative, and result from disparities within and slightly beyond the fusional zone. At larger disparities, qualitative depth percepts dominate, and observers lose the ability to accurately report the amount of depth (figure 1.3).

Later experiments have confirmed Ogle's results (Badcock & Schor, 1985; Blakemore, 1970; Westheimer & Tanzman, 1956). For example, Westheimer and Tanzman (1956) investigated whether qualitative depth is present at large disparities using spots of light presented at crossed and uncrossed disparities ranging from 1° to 10°. Observers judged the direction of the depth offset. Although individual differences were observed, on average, observers reliably indicated depth sign for disparities up to 7°, and one observer performed well up to 10°. Blakemore (1970) used a depth matching task and like Ogle, found high precision at small disparities and a transition to qualitative depth marked by a plateau at large disparities. The following section provides evidence to link these psychophysical results with the physiological data discussed in Section 1.2.

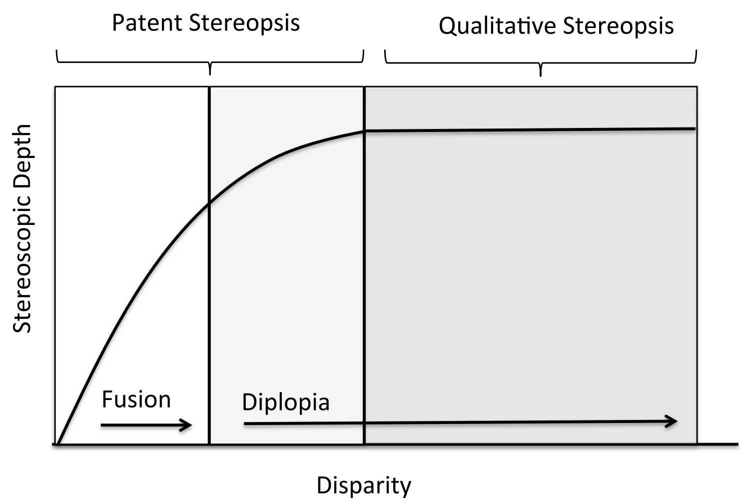


Figure 1.3. Summary of Ogle's Patent/Qualitative Distinction. Patent depth percepts are obtained at small disparities when stimuli appear fused as well as within a limited range of large disparities when targets appear diplopic. Beyond the patent limit, observers perceive qualitative depth only.

1.3.3. Linear and non-linear systems for visual processing

Before describing linear and non-linear disparity processing, it is important to first understand the linear systems approach to visual processing. The following section provides an overview of linear processing in the visual system. In 1822, Baron Jean Fourier demonstrated that any periodic waveform could be broken down into a linear sum of sine waves of a specific phase, frequency, and amplitude that are harmonically related. This is referred to as a Fourier series. Conversely, an inverse Fourier series allows for a complex waveform to be synthesized by adding together multiple waveforms of a specific phase, frequency and amplitude. When a system behaves approximately linearly, Fourier analysis is useful for deconstructing images into their component sinusoidal waveforms.

Campbell and Robson (1968) conducted the seminal work investigating the linear nature of visual processing. They measured contrast detection thresholds for sine-wave gratings with different spatial frequencies in isolation and in combination. They then analyzed the thresholds in terms of the harmonic components of the waveforms. For instance, in one case observers viewed a waveform that alternated between a sine-wave and a square-wave grating. The sine/square contrast ratio was initially set to $4/\pi$ and observers were asked to adjust the contrast of each grating until the two gratings were distinguishable. Campbell and Robson (1968) predicted that a square-wave grating would be distinguishable from the sine-wave grating when the third harmonic of the square grating reached its own threshold. This hypothesis was confirmed and their results demonstrated that perception depends critically on spatial frequency. They argued that the spatial frequency is a fundamental stimulus attribute, and the visual system consists of different spatially tuned channels that respond maximally to a particular spatial frequency. Their results (along with others') led to a working model that approached the early stages of visual processing as a form of Fourier analysis. That is, that neural RFs act as spatial frequency filters, which decompose complex natural scenes into their component frequency, phase and orientations. Thus all aspects of early vision (such as texture perception, contrast sensitivity, motion perception and even stereopsis) could be described in terms of a linear combination of Fourier components.

The early work of Campbell and Robson (1968) spawned a vast literature on the Fourier-like properties of the human visual system. Many subsequent publications reported evidence of independent spatial frequency tuned channels in human vision (Blakemore, Nachmias, & Sutton, 1970; Carter & Henning, 1971;

Graham, 1972; Henning, Hertz, & Hinton, 1981; Stromeyer & Julesz, 1972; Thompson & Movshon, 1978). While this approach proved fruitful for understanding some properties of visual processing, attempts to apply the linear systems approach showed that the approximation was incomplete (Blakemore & Campbell, 1969; Derrington & Henning, 1993; Henning, Hertz, & Broadbent, 1975; J. Nachmias, 1989; J. Nachmias & Rogowitz, 1983). One of the principal implications of the “vision as a Fourier analysis” approach is that the visual system will only respond to changes in luminance that can be specified in the Fourier spectrum of the stimulus. Changes in luminance can be easily analyzed by simple cells in V1 via linear filtering, but non-Fourier (nonlinear) processing is necessary for higher level processing, such as the perception of curvature, texture boundaries, illusory contours or Glass patterns (Wilson, 1999). Similarly, as outlined below, experiments by Chubb and Sperling (1989) and by Cavanagh and Mather (1989) showed that a linear filtering model could not account for a range of phenomena in motion perception. This led them to make a distinction between linear (1^{st}-order) and non-linear (2^{nd}- order) motion processing. Examples of RF processing of a 1^{st}-order luminance defined stimulus and a 2^{nd}-order contrast defined stimulus are shown in figure 1.4.

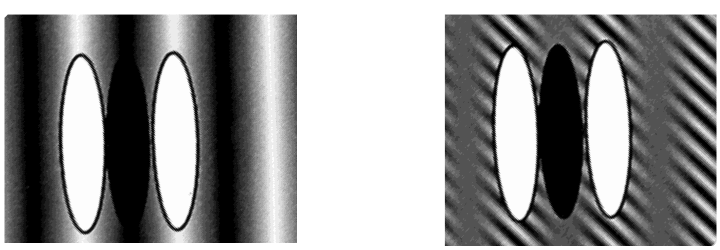


Figure 1.4. Examples of RF stimulation in 1st- and 2nd-order stimuli. In the example on the left, a representation of a neuron's receptive field is superimposed on a 1st-order stimulus consisting of a sine-wave grating, which should elicit a strong neural response. On the right, the sine-wave (carrier) is modulated by a sinusoid which oriented off the preferred axis of the neuron. There is no Fourier component at the spatial frequency of the contrast modulation and therefore, the 2nd-order information is invisible to linear filters.

1.3.4. 1st/2nd order processing in human vision

Chubb and Sperling (1989) demonstrated that while motion perception could be driven by luminance-based spatial frequency information, reliable percepts of motion could also be obtained in so called non-Fourier stimuli. These patterns were carefully constructed so that a linear filter would not be able to extract the motion signal. Instead, a non-linear or 2nd- order operation was needed to correctly identify the direction of the motion. They examined the role of 1st- and 2nd- order mechanisms in the detection of moving stimuli. In one condition, observers were presented with a moving, contrast-reversing bar, which provided a strong signal for the 1st- order motion system. In another condition, a contrast reversing grating was presented, which provided ambiguous motion via 1st- order (linear) processing, but could provide coherent motion via 2nd- order system that implements full-wave rectification. Observers could reliably detect motion in both conditions; therefore Chubb and Sperling concluded that there are two types of motion processing that involve linear processing of 1st- order luminance variation as well as nonlinear full wave rectification of 2nd- order contrast modulation. Subsequently, many motion perception studies have shown that the visual system relies on two sources of motion signals, which are variously described as Fourier/Non Fourier, 1st-order/2nd- order, and linear/nonlinear (Chubb & Sperling, 1988; Chubb & Sperling, 1989; Derrington & Henning, 1993; Henning et al., 1975).

In 1992, Wilson, Ferrera and Yo proposed a two-stage model to describe how the motion energy of a moving two-dimensional grating ("plaid") is processed. Their model proposed that simple motion energy is computed by a Fourier pathway and is detected by orientation-selective filters followed by a motion energy computation. The motion of texture boundaries is processed by a Non-Fourier pathway and initially undergoes orientation filtering, followed by rectification and finally a later stage of filtering tuned to a different orientation and lower spatial frequency. The filter-rectify-filter model is also described by Wilson (1999) and is illustrated in Figure 1.5. Wilson, Ferrera, and Yo (1992) argue from their modeling results that in the human visual system the Fourier pathway projects from V1 directly to MT, while the Non-Fourier pathway projects from V1 - V2 - MT. Zhou and Baker (1993) provided electrophysiological support for Wilson and colleagues' (1992) filter-rectify-filter model by recording neural activity from cells in areas 17 and 18 in the cat. They found that neurons are sensitive to the spatial frequency of both the carrier and envelope but for a different range of spatial frequencies for each component. In summary, Zhou and Baker concluded that their electrophysiological data are well modeled by a filter-rectify-filter model and cells are independently sensitive to the envelope and carrier spatial frequency. However, it should be noted that there is evidence for an alternative to the two-stream model. Rosenberg and colleagues (2010, 2011) demonstrated that neural responses of LGN Y-cells and cortical neurons (Area 18) to non-Fourier features are similar, suggesting 1st- and 2nd- order processing arise from a common neural pathway that begins at the retina.

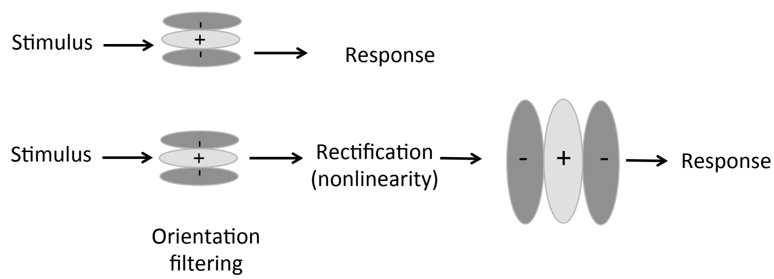


Figure 1.5. Linear (Fourier) and Non-Linear (Fourier) processing.

In a series of publications, Schofield and Georgeson have conducted detailed psychophysical studies of the properties of neural mechanisms that respond to luminance and contrast modulation using a variety of methods and stimuli. For example, they demonstrated that detection thresholds for 1st- and 2nd- order stimuli are affected differently by noise. In a 2AFC task (Schofield & Georgeson, 1999), stimuli were presented at a range of spatial frequencies and viewing distances and observers distinguished vertical modulations from either a blank or noise-only field. When the luminance was modulated with noise, the function was flatter compared to the no-noise condition. Sensitivity was lower in the contrast-modulated condition and decreased more steeply compared to the luminance conditions. Using an interval discrimination task, Schofield and Georgeson (2003) evaluated the effect of the carrier type on sensitivity to 2nd order contrast-modulated stimuli. They measured sensitivity for contrast-modulated stimuli using different noise carriers (ie. white (binary),1/f, and high pass noise). Observers indicated the interval that contained the contrast-modulated stimulus in addition to the noise (the other interval only contained noise). Similar modulation sensitivity functions were obtained for all carrier types presented with the 2nd order contrast-modulated stimuli. They concluded that when controlling for 1st- order artifacts, the 2nd- order system is insensitive to the carrier. While it has been argued that 2nd order processing is optimal with high frequency carriers (Dakin & Mareschal, 2000), Schofield and Georgeson (2003) work indicates that 2nd- order sensitivity does not depend on the carrier's spectral content, nor does the carrier mask the detection of the 2nd- order signal. However, it should be noted that these studies do not rule out the possibility that individual 2nd- order channels might be highly carrier-dependent (Zhou & Baker, 1993; Mareschel & Baker, 1998).

1.3.5. 1st- /2nd- order properties in stereoscopic vision

As outlined in the previous section, 1st- and 2nd- order signals are processed by different filtering operations (Chubb & Sperling, 1988; Chubb & Sperling, 1989; Schofield, 2000; Schofield & Georgeson, 1999, 2003; Tanaka & Ohzawa, 2006; Wilson et al., 1992; Zhou & Baker, 1993). Sato and Nishida (1994) and Hess and Wilcox (1994) investigated the possibility that linear and nonlinear filtering operations also contribute to stereoscopic processing. Using Gabor stimuli, Hess and Wilcox (1994) showed that depth discrimination depended on the contrast envelope when the frequency of the carrier grating exceeded 2.5 cycles per envelope. Thus, there was a critical bandwidth that determined whether the 1st- order carrier or 2nd- order envelope signal was used to make depth judgments. They posited that this was due to the fact that as the number of cycles per envelope increased, the possibility of false matches also increased, making the signal less reliable. In a series of publications, Wilcox and Hess (2006, 2008; 1995; 1996) demonstrated that stereoscopic mechanisms can be distinguished based on their spatial properties, similar to the distinction reported in the motion (Chubb & Sperling, 1989; Lu & Sperling, 1995) and spatial vision (Schofield, 2000; Schofield & Georgeson, 1999, 2003; Wilson et al., 1992; Zhou & Baker, 1993) literature.

Tanaka and Ohzawa (2006) investigated neural responses of the cat to 1st- and 2nd- order stereoscopic stimuli. Nearly all of the neurons responded to the 1st-order luminance stimuli (148 out of 151), and 70 of these neurons also responded to the 2nd-order contrast-envelope stimuli. In one experiment, tuning measurements were made in area 18 when the interocular phase of the carrier was fixed while the interocular phase of the envelope varied in a 2nd-order contrast-envelope stimulus. Their results indicated that binocular responses were modulated by the interocular phase, and the optimal phases were different for each neuron tested. This suggests that these neurons are tuned to disparity in contrast-envelope stimuli. In another experiment using the contrast-envelope stimuli, the interocular phase of the carrier was varied while the interocular phase of the envelope was fixed at the optimal value. Here, neurons (n=3) showed no change in response contingent on the phase of the carrier, suggesting that the 2nd-order disparity tuning is independent of the 1st order response properties. Their work forms an important link between that of Baker and colleagues on physiology of 2nd-order motion processing, and the psychophysics of 2nd-order stereopsis.

As in 2D spatial vision, the main difference between the two systems depends on the information used to compute retinal disparity: the 1st- order system uses luminance-based features (bars, edges) to extract a disparity signal. However, the 2nd- order system is able to provide disparity information based on contrast modulation or, in the case of Gabor patches, the contrast envelope of the stimulus (Schofield & Georgeson, 1999). In the case of 2nd- order stereopsis, the underlying luminance signal, or carrier, does not have to be correlated in the two eyes, making this mechanism resilient to differences in shape and polarity that would make it difficult or impossible for the 1st- order system to provide a meaningful disparity estimate (Mitchell & O'Hagan, 1972; Wilcox & Hess, 1995; Wilcox & Hess, 1996).

One feature of the 2nd- order system is that envelope extraction determines depth percepts at the upper limit for stereopsis (D_{max}); the point at which the disparity is so large that depth percepts are lost (Wilcox & Hess, 1995). Wilcox and Hess (1995) used the method of adjustment and a staircase method to determine which properties have the most significant effect on D_{max}. They manipulated contrast, carrier spatial frequency, and envelope size in separate conditions and found that the envelope size had the strongest influence on D_{max}. This result differs from earlier findings by Schor and Wood (1983) who reported that D_{max} increased with decreasing spatial frequency. However, it should be noted that they used difference of Gaussian (DoG) stimuli. In these stimuli, size and spatial frequency co-vary as changes in spatial frequency also result in changes in overall size. Given the results of Wilcox and Hess (1995), it is likely that it was size, and not spatial frequency, that determined D_{max} in Schor & Wood's study.

Wilcox and Hess (1996) investigated whether the non-linear filtering operation used to extract the stimulus envelope occurs before or after binocular combination. Their stimuli consisted of vertically-oriented 1-D spatial noise multiplied by a 2-D Gaussian envelope, that were presented as either matched (correlated) or unmatched (uncorrelated) stereopairs. Both types of stimuli contain a 2nd- order envelope disparity signal, but in the unmatched case different random samples of noise form the carrier, and so this stimulus can not provide a reliable disparity signal via 1st- order processing. They reason that if binocular combination occurs before envelope extraction, it should be impossible to extract depth from the uncorrelated stimulus. Since observers are able to make reliable depth judgments using the uncorrelated stimuli, Wilcox and Hess argue that the envelope information is derived monocularly, before binocular correlation, and involves a nonlinear operation, such as rectification. This is consistent with both the modified energy model proposed by Read et al. (2002) and the filter-rectify-filter model (Wilson et al., 1992; Zhou & Baker, 1993).

In sum, there is considerable evidence that, as in motion and texture perception, disparity relies on more than one type of processing. The existing body of work suggests that when the stimuli have well-defined luminance edges and are matched in the two eyes, the 1st- order mechanism is used. However, if the stimuli are not well matched in the two eyes, are presented at the upper disparity limit, or are defined by contrast modulation alone, the 2nd- order system is used.

1.3.6. Sustained/Transient properties in stereoscopic vision

Jones (1980) recorded eye movements to compare disparity-induced vergence when viewing fusable or non-fusable stimuli. The fusable stimulus consisted of vertical lines presented to each eye, while the non-fusable (unmatched) stimulus consisted of vertical line presented to one eye and a horizontal line presented to the other eye. When the fusable stimulus was presented for an extended time, the vergence response began at 200 msec and took approximately 1 second to complete. For the unmatched stimulus, the response was identical to the fusable stimulus when the exposure duration was set to 200 msec, but the vergence response was degraded if the stimulus was presented for an extended time. Based on this and other work, it has been suggested that fusional vergence occurs in two stages: a transient fusion-initiating phase, followed by a fusion-sustaining phase (Jones, 1977, 1980; Westheimer & Mitchell, 1969).

Edwards and colleagues (1999; 2000; 2001) suggested that stereoscopic mechanisms are best differentiated on the basis of their sustained or transient temporal properties. In a series of publications, they argued that a sustained stereoscopic system is available to process stimuli that are presented at relatively small disparities for extended durations (>200 msec). In contrast, they argued that stimuli presented briefly, at large binocular disparities, are processed via a transient mechanism. They argued that instead of 1st- /2nd- order processing, the sustained/transient dichotomy underlies stereopsis because they found that depth from both 1st- and 2nd- order stimuli are processed by the transient system (Edwards et al., 2000).

Unfortunately, the conclusions drawn by Edwards et al. (2000) are not fully justified by their own experiments. That is, they did not directly assess if there is a transition between the sustained and transient systems in stereoscopic tasks or where it occurs. Others have shown that vergence and discrimination performance in correlated (1^{st}- order) and uncorrelated (2^{nd}- order) conditions are similar when the exposure duration is 200 msec (Hess & Wilcox, 2006, 2008; Jones, 1980). In this study, Edwards and colleagues tested only one duration and one disparity offset, which is not sufficient to establish if a sustained/transient can describe stereopsis comprehensively, or if they only represent one aspect of 1^{st}- and 2^{nd}- order stereoscopic processing. Therefore, more work must be conducted to carefully test if the sustained/transient mechanisms fully describe stereopsis at all disparities and over a range of exposure durations.

1.4. Objectives

It is possible that stereoscopic vision is supported by multiple dichotomous systems, however it seems more parsimonious that the categories outlined above reflect either a continuum or a single dichotomy with different spatial and temporal characteristics. Despite the considerable literature on these different types of classification, the use of different stimuli, methods, and disparity ranges make direct and useful comparisons virtually impossible. Thus, the goal of my dissertation research is to use specifically designed stimuli and test conditions to assess the relationship between the proposed categories of stereopsis. A major focus of this work was to determine an appropriate method for measuring depth from large disparities. Ultimately, this work will lead to a more complete and coherent understanding of stereoscopic vision across the full range of binocular disparities.

In Chapter 2, the relationship between 1^{st}- /2^{nd}- order stereopsis and sustained/transient properties was evaluated. Hess and Wilcox (2008) presented relatively small disparities to assess stereoacuity for 1^{st}- and 2^{nd}- order patterns for a range of exposure durations. However, Edwards et al. (2000) used larger disparities (near 1 deg) in their studies of sustained/transient stereopsis. It is possible that their temporal dichotomy is specific to relatively large disparities and/or to the duration tested. In the study presented here, 1^{st}- and 2^{nd}- order stereopsis was assessed for a range of large disparities and exposure durations (90 to 1,400 ms).

The purpose of Chapter 3 was to consolidate the work of Ogle (1952b, 1953), Blakemore (1970), Badcock and Schor (1985), and McKee, Levi, and Bowne (1990) using both a discrimination and a depth estimation paradigm. This chapter formed the foundation for subsequent experiments, and permits direct comparison between the two-part functions reported by Blakemore (1970) and by Badcock and Schor (1985), as well as the transition from patent to qualitative depth percepts identified by Ogle. In addition, a major focus of this chapter was the assessment and evaluation of the separation information in diplopic stimuli.

The relationship between Ogle's patent/qualitative distinction and 1^{st}- /2^{nd}- order stereopsis was evaluated in Chapter 4. Previous studies have found that observers perform significantly better when presented with correlated stimuli compared to uncorrelated stimuli (Wilcox & Hess, 1996) and that 2^{nd} order processing is coarse (Hess & Wilcox, 1994; McKee et al., 2004, 2005; McKee, Verghese, Ma-Wyatt, & Petrov, 2007; Wilcox & Hess, 1995; Wilcox & Hess, 1997). Discrimination experiments using 2nd-order uncorrelated stimuli illustrated that depth sign is perceived quite reliably (Wilcox & Hess, 1996), however, it is unclear if metric depth information can be extracted. In Chapter 4, a depth estimation task was used initially to evaluate the nature of the depth percept provided by 1st- and 2nd- order stimuli. In the second experiment, a two interval forced choice task was used to compare the discrimination of metric depth from 1st- order and 2nd- order stimuli.

Ohzawa and colleagues (1999; 1995; 1990) have argued that depth is encoded by spatial frequency tuned cells, where low spatial frequency tuned neurons encode large disparities and high spatial frequency tuned neurons encode small disparities. This property implies a continuum of depth processing which relies on linear processing of the luminance information of the carrier. Wilcox and Hess (1995) found that the size of the envelope influences the upper limit of stereopsis (D_{max}), and that D_{max} is determined solely by the 2nd order contrast envelope of the stimulus. In Chapter 5, the role of stimulus size on metric depth obtained from 2nd order uncorrelated stimuli was evaluated. This was achieved by varying the width of the envelope and measuring stereosensitivity using a 2IFC task.

Both 2nd order stereopsis and monoptic depth can be used to process depth at large disparities. Studies by Hering (1861), Kaye (1978), Wilcox, Harris, and McKee (2007), and Fukuda, Wilcox, Allison, and Howard (2009) provided empirical evidence that depth can be extracted from monocular targets. In Chapter 6, these two possibilities (depth from 2nd order stereopsis vs. monoptic depth) were evaluated using conditions designed to isolate a binocular disparity cue, a monoptic depth cue as well as a condition that included both cues to depth.

Chapter 2

Temporal Dynamics of 1st- and 2nd- order Stereoscopic Processing

2.1.1. Introduction

There has been debate concerning whether stereopsis is subserved by one or more distinct mechanisms and the basis for the categorization into different systems (Hess & Wilcox, 1994; Kovács & Fehér, 1997; Langley et al., 1999; Lin & Wilson, 1995; McKee et al., 2004, 2005; Sato, 1983; Wilcox & Hess, 1995; Wilcox & Hess, 1996, 1997, 1998)). As outlined in Section 1.3.2 of the Introduction, Ogle (1952a, 1952b, 1953) provided a description of the categorical nature of depth percepts that result from different amounts of disparity. He argued that the patent and qualitative categories reflect different neural mechanisms but did not test this. Subsequent electrophysiological work seems consistent with Ogle's proposals. For example, Poggio and colleagues (Poggio, 1995; Poggio et al., 1988) showed that there are disparity-selective neurons that differ in terms of their excitability to specific ranges of disparities. As outlined in the Introduction (Section 1.2.2.), tuned neurons are sensitive to small disparities close to zero, while near/far cells are sensitive to large disparities. These classes of neurons have been shown to exhibit different response profiles (Poggio, 1995).

More recently, other ways of categorizing stereoscopic processing have been forwarded; one proposal is based on mechanisms with different temporal properties (Edwards, Pope, & Schor, 1998) and another is based on their spatial properties (Hess & Wilcox, 1994). The temporal dichotomy postulated by Edwards, Pope and Schor (1998) is based on the sustained/transient distinction in the vergence system (Jones, 1980). In a series of experiments Edwards and colleagues (Edwards et al., 1998, 1999, 2000; Schor, Edwards, & Pope, 1998; Schor et al., 2001) have argued that a sustained stereoscopic system is available to process stimuli that are presented at relatively small disparities for extended durations. In contrast, stimuli presented briefly, at large binocular disparities, are processed via a transient mechanism. The link between timing and disparity is fundamental to their proposal.

Edwards and colleagues have pointed out that the disparity dependence of the sustained and transient mechanisms may map onto the patent/qualitative distinction proposed by Ogle (1952b). Indeed, although Ogle did not vary exposure duration experimentally, he noted that brief viewing times were necessary to see depth from diplopic stimuli, and that depth percepts became more ambiguous with increased exposure duration. This observation was confirmed in later studies (Mitchell & O'Hagan, 1972; Ogle, 1964a; Reading, 1970; Westheimer & Tanzman, 1956). This property stands in sharp contrast to the temporal properties of stereopsis at relatively small disparities. For such stimuli, it is well established that performance improves as exposure duration is increased, at least up to 1 sec (Harwerth, Fredenburg, & Smith, 2003; Harwerth & Rawlings, 1977; Hess & Wilcox, 2008; Ogle & Weil, 1958). Other distinguishing properties of the putative sustained and transient systems include sensitivity to interocular polarity differences and spatial frequency in the sustained but not in the transient system (Edwards et al., 1999; Schor et al., 1998).

As outlined in the Introduction, Hess and Wilcox (1994) argued that stereopsis is processed by two independent mechanisms: a 1st- order mechanism, which detects disparity based on luminance information and a 2nd- order mechanism, which relies on contrast modulation. Note that in their original formulation, they did not consider the relative temporal properties of these mechanisms. While the evidence for the existence of these two mechanisms is not disputed ((Edwards et al., 2000; Hess & Wilcox, 1994; Kovács & Fehér, 1997; Langley et al., 1999; Lin & Wilson, 1995; McKee et al., 2004, 2005; Sato, 1983; Wilcox & Hess, 1995; Wilcox & Hess, 1996, 1997, 1998) their contribution to depth perception as a function of disparity is not fully understood. That is, while Wilcox and Hess (1995) have shown that at the upper disparity limit only 2nd-order stereopsis is used, they did not assess its role in signaling depth from disparities between the upper fusion limit and D_{max}.

Hess and Wilcox (2008) and Edwards et al. (2000) have conducted experiments designed to explicitly evaluate the proposed temporal vs. spatial dichotomies. For example, Edwards et al. (2000) assessed depth identification (percent correct) for 1st- and 2nd- order stimuli. They found that observers could make depth sign judgments reliably for 1st- or the 2nd- order stimuli, and concluded that the transient system can process both types of patterns. However, they restricted their measurement to a single exposure duration, 200 ms. This duration is long in terms of temporal summation (Harwerth et al., 2003; Harwerth & Rawlings, 1977; Hess & Wilcox, 2008; Ogle & Weil, 1958) and has been shown to elicit similar vergence responses when viewing fusible and nonfusible targets (Jones, 1980;

Mitchell, 1970). Therefore, this study does not adequately distinguish between the two proposed dichotomies.

In a more recent set of experiments, Hess and Wilcox (2006, 2008) showed that within Panum's fusional area, 1^{st}- and 2^{nd}- order stereopsis have distinct temporal characteristics. Hess and Wilcox (2008) used a set of stimuli designed to access 1^{st}- and 2^{nd}- order mechanisms and assessed stereoacuity for a wide range of exposure durations. By measuring the full temporal response function for each stimulus, they demonstrated that when the 1^{st}- order system is used, performance improves with increasing exposure duration. However, when the 2^{nd}- order system is activated, thresholds remain relatively constant over the full range of viewing times (80 msec to 1.3 sec).[2] There was no improvement with increasing duration, and in some cases sensitivity was poorer at the longest durations. Their results suggest that the temporal properties are but one dimension along which 1^{st}- and 2^{nd}- order processing differ. That is, over the range of durations and disparities tested, 1^{st}- order stereopsis exhibited sustained properties, while 2^{nd}- order stereopsis appeared to exhibit more transient characteristics. However, Hess and Wilcox (2008) pointed out that at any given exposure duration it is possible to obtain a depth percept from *either* 1^{st}- or 2^{nd}- order stereopsis so temporal properties are not sufficient to characterize the two types of processing.

[2] The range of test durations used here and by Hess and Wilcox (2006, 2008) is larger than typically used to assess 'transient' processing. However, effectively the viewing time was much shorter because they (unlike Edwards and colleagues, 1998, 1999, 2000) used a post-stimulus mask to limit processing.

2.1.2. Rationale

An important difference between the experiments of Hess and Wilcox (2006, 2008) and those of Edwards and colleagues is the range of test disparities employed. That is, Hess and Wilcox (2008) measured stereoscopic *thresholds* for both 1st- and 2nd- order stimuli using the method of constant stimuli; test disparities were never larger than 10 arc min and were well within the fusable range. Edwards and colleagues measured observers' ability to make near/far depth judgments for targets presented at 1 deg of disparity, an offset that was large relative to the extent of the stimuli they used. This difference in disparity range could be the basis for the discrepancy in their results. Therefore in this study we measured the temporal response of 1st- and 2nd- order stereopsis for a range of large disparities, comparable to those used by Edwards, Pope & Schor (1998, 1999, 2000). These disparities were chosen to be well above threshold, but not diplopic. If Edwards and colleagues are correct, and the exposure duration can be used to characterize the fine and coarse disparity mechanisms, then in this disparity range, there should be a clear improvement in performance with increasing exposure duration, for *both* 1st- and 2nd- order conditions. However, if spatial properties are the critical attribute, then we should replicate the results of Hess and Wilcox (2008) in showing no change, or a reduction in performance as exposure duration increases for 2nd- order stimuli. The hypotheses are affected by the disparity range used in this experiment. Because large disparities were used, stimuli were not completely fused and it was unclear if the disparity was too large relative to the scale of the 1st- order information. Therefore, if observers could use the 1st- order information, I would predict that results from this experiment would be identical to the results found by Hess & Wilcox (2008). Alternatively, if observers could not use the 1st- order information at this disparity range, then performance on both conditions should be the same and should not improve as a function of duration.

Note that in the noise patterns used here there is a 2nd- order contrast envelope available in both stimuli. The stereoscopic system has been shown to use the stimulus envelope when the carrier component does not provide a reliable disparity signal (as in the uncorrelated noise stimuli). Therefore it is possible that given the large disparity range used in this study, the disparity signal provided by the 2nd- order component will be used in the correlated noise condition as well. If so, we would expect little or no improvement in performance as a function of exposure duration in either condition.

2.2. Experiment 2.1: Methods

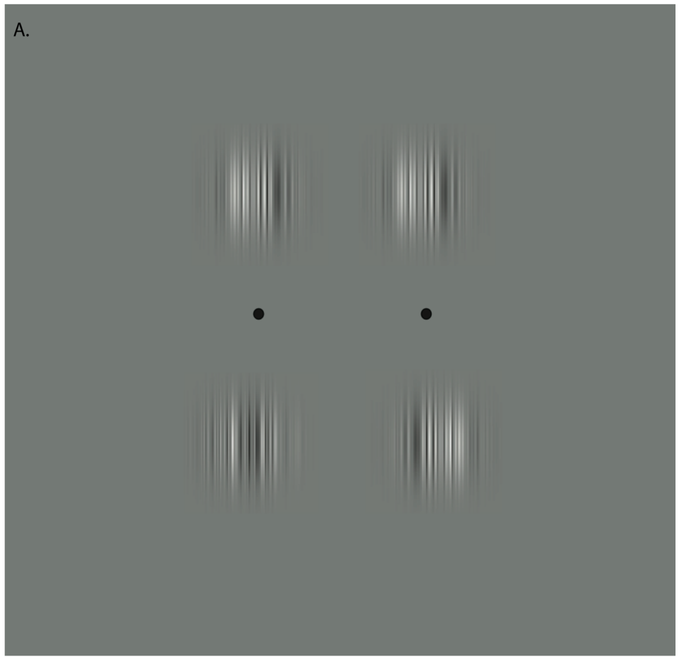
Apparatus and Observers

Four observers with normal or corrected to normal vision were tested extensively. Before participating in the experiment, all subjects were assessed using the Adult Randot™ Stereo test to ensure that they could appreciate depth from binocular disparities of at least 40 arc sec. All observers had prior experience as participants in stereoscopic experiments but were unaware of the hypothesis tested here. The study was approved by the York Research Ethics board and followed the guidelines of the Declaration of Helsinki.

Stimuli were generated using a Cambridge Research Systems VSG2/3 graphics card which implements a resistor network to sum DAC outputs to allow a pseudo 12 bit grey-level representation after gamma correction. Stimuli were presented using on a single, fast phosphor, Clinton monitor. The display (1024 x 768 pixels) was viewed from 1.08 m and subtended 10.7 °of visual angle with mean luminance of 69 cd/m^2. The frame rate was 120 Hz, so the effective frame rate for each eye was 60 Hz. Stereo-pairs were displayed on alternate frames and restricted to one eye using Display Tech™ light valves mounted in optical trial frames. The use of a single phosphor and the fast extinction rate of the light valves, combined with the use of low-contrast stimuli to reduce the amount of interocular cross-talk between the images presented to the two eyes.

Stimuli

The stimuli consisted of 1D white noise windowed with a 2D Gaussian envelope with a Gaussian sigma of 1.69 deg. Each noise stimulus subtended approximately 1.2 x 1.2 deg, and they were separated (edge to edge) by 1.2 deg. To ensure that we could use large test disparities, but still measure stereoscopic thresholds, we assessed discrimination about a pedestal with a crossed disparity of 34 min. Correlated and uncorrelated stereopairs were generated with a range of disparities that bracketed this pedestal (Figure 2.1).



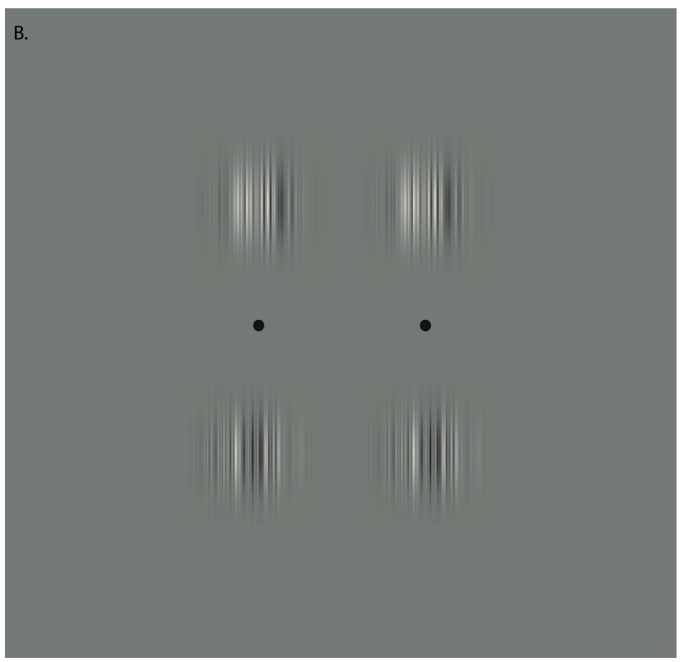


Figure 2.1. Stimulus configuration used in Experiment 2.1. Observers fixated at the centre, a reference patch was presented above, and a target below, fixation. The target in A is uncorrelated and presented at a crossed disparity. The target in B is correlated, and also at a crossed disparity.

A 50 msec mask was presented immediately following the stimulus to ensure that stereoscopic processing was limited to the duration that the stimuli were visible (Breitmeyer & Ogmen, 2000; Liss, 1968; Sperling, 1963). The mask stimulus was a horizontal 3.6 x 7.6 deg strip of vertically oriented 1D luminance noise . It was generated by assigning each row of pixels a random luminance value, and then selecting a different sample to present to each eye on each trial. The random nature of the structure in each eye produced the percept of a 3D volume of noise. The mask was the same contrast as the test stimuli, and, when present, completely covered the stimulus area.

Procedure

Contrast Thresholds

Contrast thresholds were first measured for each observer, for both the correlated and uncorrelated conditions, and at each test duration, prior to assessing thresholds. We used the method of adjustment with a randomized starting point to obtain 7 binocular detection thresholds, which were then averaged. Subsequently, the contrast of the stereoscopic test stimulus was set at three times the Michelson contrast of that particular stimulus/duration combination at detection threshold. This ensured that all stimuli were equally detectable, regardless of exposure duration.

Depth Discrimination

The method of constant stimuli was used to assess discrimination thresholds with a set of 11 test disparities, centered at the pedestal disparity (34 arc min, crossed). The disparity range was chosen individually for each subject to ensure adequate sampling of the psychometric function. The stimuli were presented within a temporal Gaussian envelope the size of which (sigma) was varied across sessions to manipulate exposure duration. The observer's task was to identify if the target was in front or behind the reference stimulus that was at a distance of 34 arc min in front of fixation. Both the reference and test stimuli were presented within a range of disparities where the two half images did not appear diplopic. Within a single session, each disparity was presented 40 times in random order. One exposure duration was tested per session, for either the correlated or uncorrelated condition, and observers completed four sessions for durations ranging from 233 to 1340 msec. Sensitivity was computed from the resulting psychometric function by fitting the error function (cumulative normal), ERF (*x*), of the form:

$$P(x) = A(0.5 + 0.5ERF((x - B)/(\sqrt{2}C))) \qquad 2.1$$

Where A is the number of presentations per stimulus condition, B is the offset of the function relative to zero, and C is the standard deviation of the underlying normally distributed error function. The reciprocal of the standard deviation parameter was used as the index of sensitivity, and confidence intervals were computed by bootstrapping each observers' data (Wichmann & Hill, 2001a, 2001b) with 1000 replications.

2.3. Experiment 2.1: Results

Figure 2.2 depicts results for four observers. Stereosensitivity is plotted as a function of duration for both the correlated and uncorrelated test conditions. For all observers it appears that the slope of these functions are generally higher in the correlated condition, and there is little difference between the correlated and uncorrelated results in terms of their dependence on duration.

The results were evaluated using an analysis of variance (ANOVA) and a significant main effect of condition was obtained ($F(1)=6.675$, $p<0.05$, partial $\eta^2=0.218$, observed power=0.698). There was no significant effect of duration nor was the interaction of duration by stimulus type (correlated vs. uncorrelated) significant. Paired-sample t-tests were used to compare the correlated vs. uncorrelated condition at each duration. A significant difference was found only at the longest duration, 1340 msec ($t(3)=3.518$, $p<0.05$) where there was a slight improvement in sensitivity in the correlated noise condition.

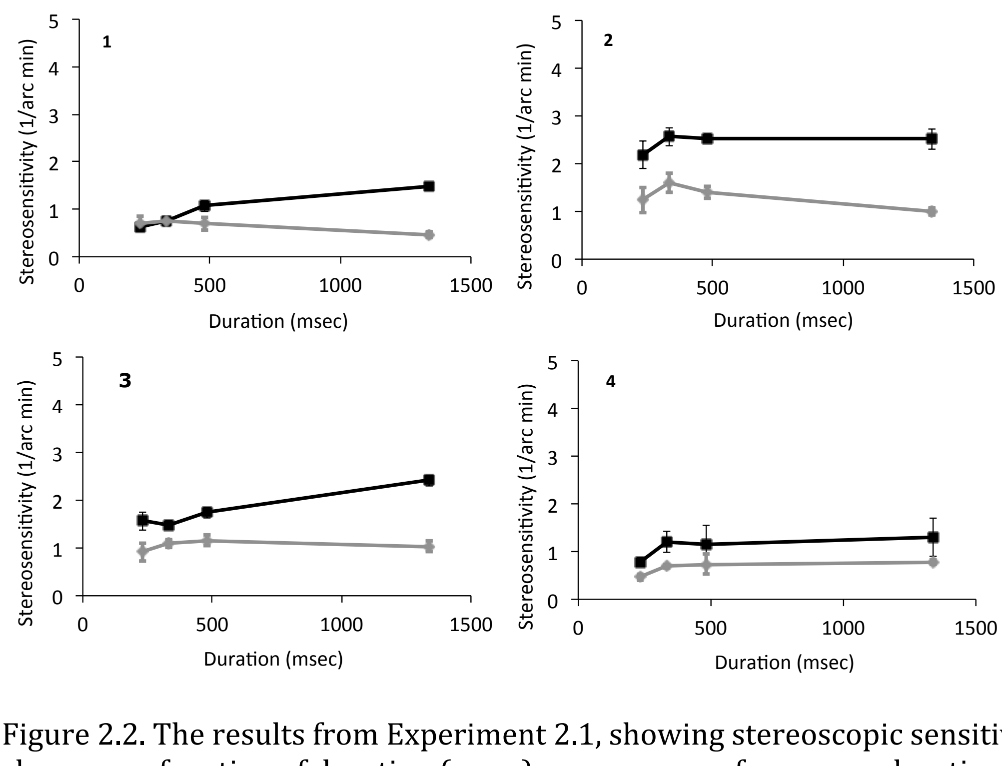


Figure 2.2. The results from Experiment 2.1, showing stereoscopic sensitivity is shown as a function of duration (msec) over a range of exposure durations (233 to 1340 msec). The correlated and uncorrelated conditions are represented by black and grey lines respectively. Error bars represent confidence intervals estimated using a bootstrapping procedure.

To contextualize the difference between the correlated and uncorrelated test conditions in this study, it is useful to compare the results with those reported by Hess and Wilcox (2008). Figure 2.3 depicts the ratio of the correlated vs uncorrelated thresholds obtained in our study and in that of Hess and Wilcox (2008) at two exposure durations: 340 and 1340ms. The difference is striking. For small threshold disparities at the longest duration (1340 msec), sensitivity is 20-35 times higher in the correlated condition than it is in the uncorrelated condition (Hess & Wilcox, 2008), while in the large disparity range tested here, at the same duration, performance in the correlated condition is only 1.7 to 3 times higher than performance in the uncorrelated condition (figure 2.2). Likewise, at a short duration (334 msec), observers were 8 to 12 times more precise when viewing correlated stimuli in the small disparity range in Hess and Wilcox (2008) experiment, and only 1-1.8 times more precise in the correlated condition in the large range tested here.

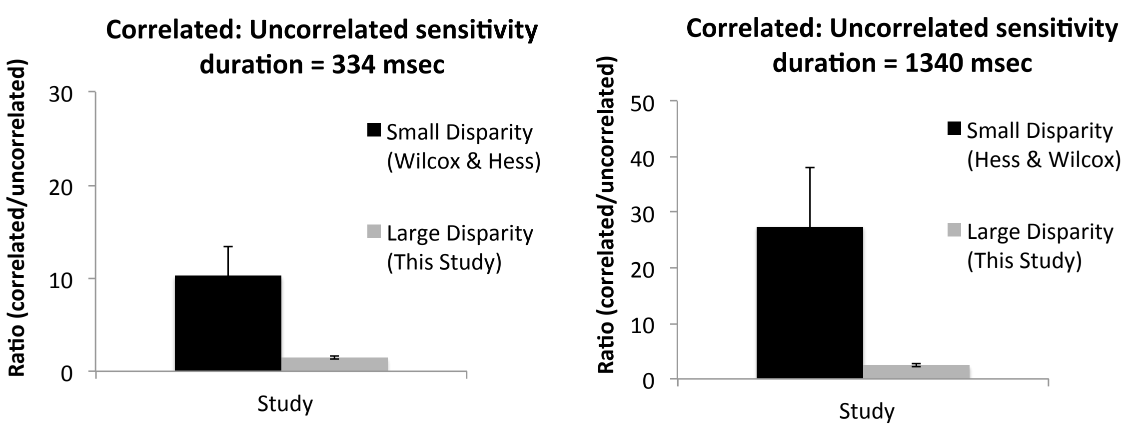


Figure 2.3. Comparison of the ratio of stereosensitivity when stimuli were correlated or uncorrelated, in the fine disparity range (black bar, extracted from Hess and Wilcox (2008) and in the large disparity range (grey bar, data from the present study). The graph on the left represents performance at a moderate

duration (334 msec) and the graph on the right represents performance at a relatively long duration (1340 msec). Data from Wilcox and Hess represent an average of results for 2 observers, and the data from this experiment represents an average across 4 observers. The range of results for each experiment is reported in the Results section.

2.4. Discussion

In the small disparity range evaluated by Hess and Wilcox (2008) it is clear that 1st- and 2nd- order signals are processed independently as reflected by a difference in their dependence on exposure duration. In Experiment 2.1, this relationship was re-examined at larger disparities, and shows that a small difference in performance in the correlated and uncorrelated conditions is present though remarkably diminished at these large disparities (Figure 2.3). Importantly, again no evidence for a stereoscopic dichotomy based on the temporal properties of the stimuli was found here. According to Edwards et al. (1998, 1999, 2000) performance should vary significantly with exposure duration regardless of the stimulus type, with high sensitivity at longer durations and vice versa. Instead the results from Experiment 2.1 show that sensitivity is relatively consistent, across exposure duration (with slight increases in performance for the correlated stimuli at long durations and slight decreases in sensitivity for uncorrelated stimuli at long durations).

Edwards et al. (2000) report similar performance under 1st- and 2nd- order test conditions at a disparity of 1 degree, and at a 200 msec exposure duration. They concluded that the transient system processes both 1st- and 2nd- order signals equally. The data in Experiment 2.1 support a different interpretation; that the effect of exposure duration is specific to the nature of the spatial information used to do the task. It is very clear that the expected improvement in sensitivity is not seen for 1st- order stimuli, even at our longest exposure duration. While it is possible to create conditions under which similar thresholds are obtained using 1st- and 2nd- order test patterns, like Hess and Wilcox (2008), no mechanistic conclusions should be drawn without measuring the dependence on duration over a wide range.

These findings extend the results from the existing literature on the temporal dynamics of stereoscopic processing, which suggests that at small disparities observers use the 1st- order carrier information in the correlated condition which results in a sustained dependence on exposure duration, while performance in the uncorrelated condition is dependent on the 2nd- order contrast envelope and elicits a transient dependence on exposure duration. When fusion is compromised it appears that the matched carrier information is not as useful as it is when stimuli are perfectly fused and observers rely on the contrast envelope instead. There is some elevation in stereosensitivity in the correlated condition compared to the uncorrelated condition. It is unclear why this difference exists, however it could be due to rivalry of the 2nd-order pattern or perhaps 1st- order contribution. Nonetheless, these results suggest that at large disparities, stereosensitivity is influenced by the 2nd- order information, does not vary much with exposure duration and as a result, the pattern of performance in both conditions is similar as a function of duration. This explains why the stereosensitivity in the 1st- order condition did not improve to the extent that is seen in the experiment conducted by

Hess and Wilcox (2008). The dependence on the 2nd- order stereoscopic system at large disparities will be explored in more detail in future chapters.

The data shown here help consolidate the stereoscopic literature, by suggesting that the transient/sustained dichotomy proposed by Edwards and colleagues (Edwards et al., 1998, 1999, 2000; Schor et al., 1998; Schor et al., 2001) reflects the processing of 1st- and 2nd- order mechanisms. While there are differences in the effect of exposure duration on stereoscopic sensitivity assessed using these two classes of stimuli, it is possible to make use of either system at any exposure duration. Thus, this work confirms and extends that of Hess and Wilcox (2008). It is notable that the substantially lower sensitivity reported by Hess and Wilcox (by a factor of up to 35), for 1st– order patterns, is absent in this experiment. This difference is likely due to the disparity range used here, which puts the stimuli near the upper limit for fusion. In previous experiments, Wilcox and Hess (1995) showed that only the 2nd- order system contributes to depth percepts at the upper disparity limit (D_{max}) for stereopsis. It is possible (but as of yet not tested) that near the fusion limit the visual system begins to rely more heavily on 2nd- order processes. The correlated and uncorrelated noise stimuli used here were windowed with identical 1D Gaussians resulting in the same stimulus envelope. It is likely that this 2nd- order information was used in both conditions to perform the task, thus producing very similar sensitivity functions. The relationship between 2nd- order processing and fusion/diplopia is explored in more detail in Chapters 3 and 4.

Chapter 3

Evaluation of Patent/Qualitative percepts Using Depth Discrimination and Depth Estimation tasks

3.1.1. Introduction

In Chapter 2 I confirmed that the spatial dichotomy proposed by Hess and Wilcox (1994) encompasses the temporal dichotomy proposed by Edwards, Pope and colleagues. The next step in consolidating the literature is to re-evaluate Ogle's classification of stereoscopic percepts and finally, to determine the relationship between the spatial dichotomy and Ogle's perceptual dichotomy (which will be discussed in Chapter 4). As outlined in the Introduction (Section 1.4.2), Ogle (1952b) argued that depth percepts are either quantitative (patent) or qualitative, depending on the disparity range. Ogle (1953) mentioned an important caveat: "care must be taken to be sure that equal subjective depths are being equated, and not equal separations of the double images of the reference and test lines". However, it is difficult for an experimenter to ensure that observers make relative depth judgments when the separation information is fully confounded with the binocular disparity signal. This issue is illustrated in Figure 3.1.

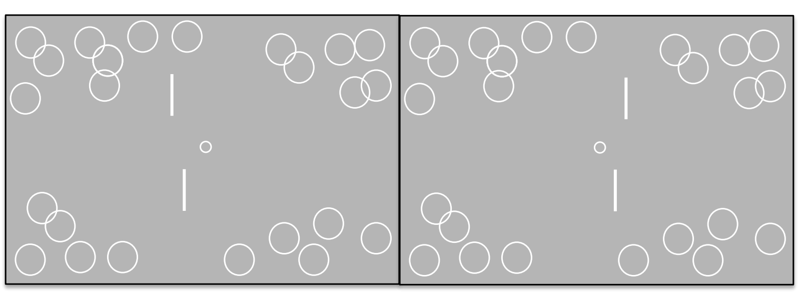


Figure 3.1. The separation confound. At large disparities, when the stereopairs are cross fused, and the central dot is fixated, observers can use the separation information instead of stereopsis to make relative depth judgments.

This same issue was faced by researchers (Badcock & Schor, 1985; Blakemore, 1970) who followed up on Ogle's experiments and assessed stereothresholds as a function of disparity. Blakemore (1970) measured relative depth discrimination between two trans-illuminated slits and the depth sign varied across trials, and the stimuli were presented for 100 msec. Like Ogle (1953), Blakemore found lower stereoacuity thresholds at smaller pedestal disparities and higher stereoacuity thresholds at larger pedestal disparities. He argued that stereoacuity over a full range of test disparities is represented by an exponential function, which becomes linear when plotted on logarithmic axes. However, it is possible that a plateau in the diplopic range was not observed because he did not sample a sufficient number of pedestals or a large enough disparity range. Close examination of both Ogle (1953) and Blakemore's (1970) data shows that their test disparities were primarily in the diplopic range. Also, it is possible that separation information is present in the diplopic range. Thus it is not clear if the reported linear increase in thresholds is due to an increase in perceived depth or an increase in the separation between the half images.

Blakemore (1970) raised the concern regarding the possible use of the lateral separation of the diplopic stimuli in place of depth from disparity and explained that he attempted to make this cue less salient by varying the horizontal position of the stimulus configuration across trials. While this manipulation did not eliminate the presence of the separation information, Blakemore believed that his observers did not rely on the lateral separation because they were naïve, were not given feedback, and seemed to be unaware of the double images in the diplopic conditions. Finally, he argued that the error rates were similar across the full range of test disparities; therefore it was unlikely that observers used different strategies for fused and diplopic stimuli. While this may be true, it was not tested and therefore is not verifiable.

In a later study that used a wider range of disparities (both fused and diplopic), Badcock and Schor (1985) evaluated the relationship between stimulus size of DoG stimuli and discrimination thresholds to infer receptive field properties of disparity sensitive neurons. Stereoacuity thresholds were measured using a 2AFC staircase task with an exposure duration of 750 msec. The direction and magnitude of the disparity pedestal varied within a session, but not within a staircase. Two conditions were tested: in the first a bright bar was used as the comparison stimulus and a DoG as the test stimulus. In the second condition a DoG stimulus was used as both the comparison and test stimulus. When both stimuli were DoGs, they found that depth increment thresholds increased with increasing pedestal disparity in the fused range, and then plateaued in the diplopic range. When there was no disparity pedestal, stereoacuity was similar at high spatial frequencies (>2.4 c/deg), but it deteriorated at lower spatial frequencies. At pedestal disparities above 20 arc min, stereosensitivity was not influenced by increasing pedestal disparity. Furthermore, the authors report that stereopsis was absent for high spatial frequency (>2.4 c/deg) stimuli at large pedestal disparities.

Unlike Gabors, the size of a DoG stimulus varies with changes in spatial frequency. Thus, the effects of spatial frequency reported by Badcock and Schor (1985) may be related to the fact that for these stimuli, the fusion limits varied along with the changing spatial frequency (scale). The low (wide) and high (narrow) spatial frequency stimuli most likely did not become diplopic at the same pedestal disparity, and as a result higher spatial frequency stimuli would elicit better stereoacuity over a larger range of disparities. Badcock and Schor did not measure diplopia thresholds for each stimulus frequency and did not scale the pedestal disparity to coincide with the desired percept for each stimulus width so it is not possible to know exactly how these data relate to Panum's fusional area or if the transition point from patent to qualitative depth occurred at the fusion limit.

Unlike Blakemore (1970), Badcock and Schor report that an exponential function fit their data within the fused range, but another function was needed to fit the results in the diplopic range. Further, they replicated these results with the same exposure duration as Blakemore (100 msec), so the difference was not due to viewing time or associated vergence eye movements. Importantly, Badcock and Schor comment that this difference may be due to the fact that observers did not perceive depth within the diplopic range for high spatial frequencies (narrow stimuli) and were forced to use a "diplopia criterion". However, this fact cannot account for the observed difference because according to the authors, depth *was* perceived when lower frequencies (wide stimuli) were presented, and an exponential function still did not fit the results.

The experiments conducted by Ogle (1952b, 1953), Blakemore (1970) and Badcock and Schor (1985) have measured perceived depth and/or depth discrimination with the assumption that reliable relative depth judgments can be made in the diplopic range. In 1990, McKee, Levi and Browne showed conclusively that beyond 60 arc min, observers can and do make relative separation judgments instead of relative depth judgments. They measured the precision of stereopsis by calculating Weber fractions for a range of disparities. Weber fractions were calculated by dividing stereoacuity thresholds by the target's distance from fixation, or in this experiment by the standing disparity. Unlike other experiments that used a pedestal paradigm to assess large disparities, McKee and colleagues (1990) used standing disparities: one line was presented in the fixation plane while another line was placed at a standing disparity (D). They also included a monocular condition which was spatially equivalent to viewing the two half images of the stereopair in one eye. Thus they were able to compare stereoscopic performance in the diplopic range to lateral (2D) separation judgments. Like Badcock and Schor (1985), McKee and colleagues obtained a 2-segment function where thresholds obey Weber's law in the small-disparity range, but thresholds remain constant at large disparities up to 60 arc min. Beyond 60 arc min, thresholds from the binocular and monocular conditions were identical, suggesting a reliance on the separation cue at these disparities. These data are the first to demonstrate a clear transition from using stereoscopic depth to using lateral image separation at very large disparities.

3.1.2. Rationale

The experiments presented in this chapter were exploratory in nature, therefore there was no directional hypothesis. Here, I re-evaluated the literature concerning depth percepts from fused and diplopic targets. Certain flaws exist in previous studies that make it impossible to draw mechanistic conclusions. For example, studies by Ogle (1952b) and by Blakemore (1970) did not sample a sufficient range of fused and diplopic disparities, and other studies (Badcock & Schor, 1985; McKee et al., 1990; Ogle, 1952b, 1953) did not identify the diplopia threshold. Also, in most cases the separation cue was present in the diplopic range but its use was not evaluated. To address these issues I used a broadband bar stimulus, measured diplopia thresholds and presented disparities within and outside Panum's fusional area. Performance was be assessed with both threshold and suprathreshold tasks and results were compared to those reported by Ogle (1952b). Furthermore, the visibility and use of the separation information in each task was evaluated.

In Experiment 3.1 discrimination thresholds were measured for a range of fused and diplopic broadband bar targets using a conventional method of constant stimuli paradigm with a disparity pedestal. A suprathreshold magnitude estimation task was used in Experiment 3.2, where observers reported perceived depth for stimuli presented over a large range of disparities. In addition, each experimental design was assessed for the presence and usability of the separation information in diplopic test conditions. The data and methodologies used here will form a foundation for subsequent experiments. Importantly, they will allow careful mapping of the 1^{st}- /2^{nd}- order dichotomy onto Ogle's patent/qualitative distinction.

3.2. Experiment 3.1: Depth discrimination thresholds

3.2.1. Introduction

Depth discrimination thresholds were measured to determine if a two-segment function describes depth discrimination of fused and diplopic targets. As described below, separate blocks of trials were presented at one of five pedestal disparities. Two blocks of trials consisted of disparity ranges where stimuli appeared fused, and three consisted of disparity ranges where stimuli appeared diplopic.

Siderov and Harwerth (1993) argued that by interleaving crossed and uncrossed trials within a single session, it is possible to eliminate the separation information in diplopic targets. They reasoned that without eye of origin information, the binocular image is the same for a given disparity in the diplopia range. Therefore, the separation of the two half images of each pair would not signal the position of the test stimulus relative to the reference stimulus. Therefore, an interleaved 2AFC task was used in Experiment 3.1 in an effort to remove this confound.

3.2.2. Experiment 3.1: Methods

Subjects and Apparatus

Four experienced observers with normal or corrected-to-normal vision were tested. All observers had excellent stereopsis as assessed by the Adult Randot™ stereotest and previous participation in other stereoscopic experiments. Stimuli were generated using a Cambridge Research Systems VSG2/3 graphics card which implements a resistor network to sum DAC outputs to allow a pseudo 12 bit grey-level representation after gamma correction. Observers sat 1.1 m away from the monitor and a chin rest was used to maintain a stable head position. The screen resolution was 1024 x 768 pixels and subtended 10.7° of visual angle. The mean luminance was 69 cd/m^2. Stereoscopic stimuli were displayed on alternate frames and displayed to one eye using a Display Tech™ light valves that alternated at 120 Hz, so the effective frame rate to each eye was 60 Hz. The use of a single phosphor and the fast extinction rate of the light valves, combined with the use of low-contrast stimuli ensured that there was no visible cross-talk between the images presented to the two eyes.

Stimuli

Within a trial, observers fixated on a zero disparity fixation spot positioned at the centre of the screen. An array of circles was positioned at zero disparity along the top and bottom of the screen similar to the configuration shown in figure 3.2. The stimuli were presented in the blank central horizontal region (1.2 x 7.6 deg). On each trial, a reference stimulus, which consisted of a line (1.69 x 33.8 arc min) multiplied by a Gaussian with a standard deviation of 1 arc min was presented above the fixation spot at a fixed pedestal disparity. The test stimulus was presented below fixation and eleven test disparities (at equal step sizes) were individually chosen for each subject to bracket the pedestal disparity and to ensure that stimuli were perceived as fused or diplopic.

Procedure

A 2AFC increment detection paradigm was used with a disparity pedestal to measure depth discrimination across a large range of disparities. The pedestal disparities were 0, 17, 34, 51 and 68 arc min. Stimuli presented at disparity pedestals of 0 and 17 arc min were perceived as fused, and stimuli presented at pedestals 34, 51 and 68 appeared diplopic. Each test disparity was tested 40 times. The direction of the disparity offset (crossed or uncrossed) was randomly interleaved within a session, and each pedestal disparity was presented in separate blocks. On each trial, observers indicated if the target (on the bottom) was in front of or behind the reference stimulus (always on top).

Results were analyzed using conventional curve fitting techniques to obtain the slope of the psychometric function. Each psychometric function was fit to the error function (cumulative normal) as described in equation 2.1. Like in Chapter 2, Variance was estimated using a bootstrapping procedure (Wichmann & Hill, 2001a, 2001b), and iterated 100 times for each dataset, which determined 95% confidence intervals for each sensitivity measurement.

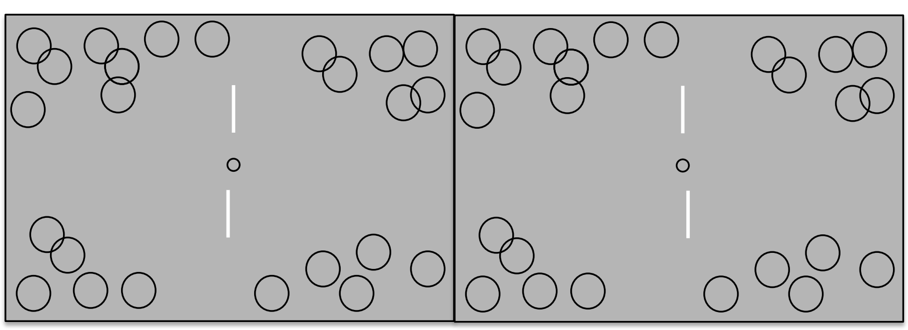


Figure 3.2. Representation of the stimulus configuration for the depth discrimination task. An array of circles that served as a fixation lock was presented at the top and bottom of the screen. A reference stimulus was positioned at a fixed

pedestal disparity (here, the pedestal is 0 arc min) above the fixation marker, and a test stimulus was presented below the fixation marker.

3.2.3. Experiment 3.1: Results and Discussion

In agreement with previous experiments (Badcock & Schor, 1985; McKee et al., 1990; Ogle, 1952b, 1953), these results suggest that a two-part function best represents stereosensitivity across a range of fused and diplopic disparities (Figure 3.3). Slopes of the psychometric functions were steep in the fused range and shallow in the diplopic range. Individual observers are numbered in all the experiments described in this dissertation so that they can be identified for their participation across different experiments. Here, subjects 1, 2, 3, 5 were tested here. For 2 out of the 4 observers (1 and 3) in the crossed direction, performance in the fused (0,17 arc min) and diplopic (34, 51, 68 arc min) conditions was significantly different, as indicated by the lack of overlap of the bootstrapped confidence intervals. For observers 1 and 3 in the crossed direction, performance in the fused range (pedestals 0 and 17 arc min) was similar, and better than performance in the diplopic range (pedestals 34, 51 and 68 arc min). For the remaining observers (2, 4), the transition in performance occurred at the smallest pedestal, where performance was significantly better when there was no disparity pedestal compared to the other 4 conditions. The direction of the depth offset was interleaved across trials in an attempt to eliminate the separation confound. However, as discussed in the next section, it is not clear from this experiment if this manipulation was effective. There is an asymmetry in performance in the crossed and uncrossed directions (at large disparities), which may be due to the difference in theoretical depth predicted by the geometry.

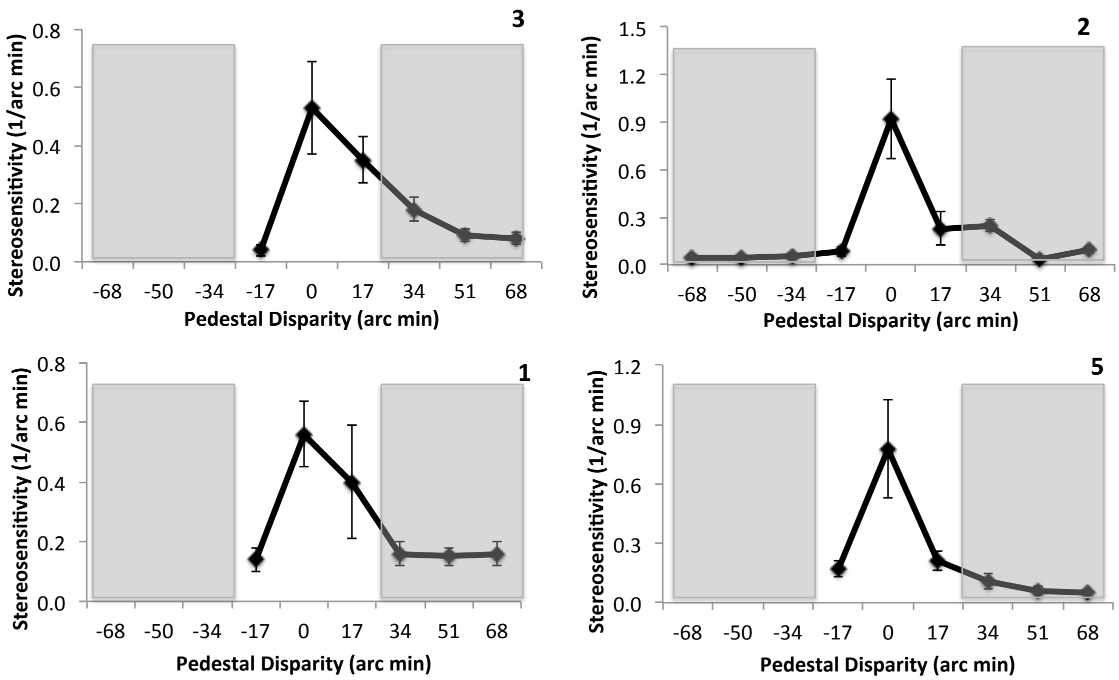


Figure 3.3. The results from Experiment 3.1, showing stereosensitivity (arc min) is plotted as a function of pedestal disparity for 4 observers (subjects 3, 1, 2 and 5). The error bars represent the bootstrapped 95% confidence intervals. The shaded region represents the diplopia range. Note that the scale on the ordinate varies because the range of test values depended on the discrimination thresholds for each observer.

3.3. Experiment 3.1.1: Follow-up to Experiment 3.1 to evaluate if observers accurately discern depth sign at large disparities

The results from Experiment 3.1 suggest that a 2-part stereosensitivity function best describes perceived depth for disparities encompassing fused and diplopic percepts, thus supporting previous findings (Badcock & Schor, 1985; McKee et al., 1990; Ogle, 1952b, 1953). However, this conclusion assumes the separation confound is eliminated by interleaving crossed and uncrossed pedestal disparities, as posited by Siderov and Harwerth (1993). It is true that in the diplopic conditions the line separation between the stereopairs will be the same for crossed and uncrossed disparities. If observers relied on this information alone, they would not be able to judge the depth of the target relative to the reference stimulus (at the pedestal disparity). However, if they were able to access the sign of the depth offset, then observers could use the separation of the half images to guide a judgment of 'relative depth'. For instance, if the target pair has a larger separation than the reference pair, and the sign is crossed, then the target is closer than the reference. This relationship is reversed if the disparity sign is inverted. To use this strategy, observers would have to have access to the direction of the pedestal disparity offset on all trials in Experiment 3.1. To evaluate whether this strategy would be possible, a 2AFC task was designed where observers simply reported the sign of the depth offset for the entire arrangement of the reference and target.

3.3.1. Experiment 3.1.1: Methods

The experimental design for the first follow up experiment was identical to Experiment 3.1, where observers performed a depth discrimination task in which crossed and uncrossed trials were interleaved. In this study, observers were asked if the whole arrangement of both the reference and test stimuli was in front or behind the fixation plane. The rationale was that if observers were able to discern the overall depth offset, they could subsequently use the separation information to make their judgment. While this very deliberate strategy may not be obvious to naïve observers, experienced participants may be able to exploit it.

3.3.2. Experiment 3.1.1: Results and Discussion

Observers (subjects 1 and 2) performed the depth discrimination task at three different pedestals in the diplopic range and percent correct was recorded for each test disparity. Two observers were tested and their average results are shown in Figure 3.5. These graphs represent percent correct as a function of disparity at a pedestal of 34, 51 and 68 arc min respectively. The blue curve represents trials in which the stimuli are in front of the fixation plane, and the red curve represents trials when the stimuli are beyond the fixation plane. It is clear that observers judged the direction of the depth offset with nearly 100% accuracy for the full range of disparities tested in Experiment 3.1. Given this, it is possible that experienced observers could determine the direction of the offset and use the separation information to correctly identify the depth of the target relative to the reference. While one aspect of the judgment involves extraction of depth sign, the other being the separation judgment, does not rely on stereopsis and could be used to make inferences regarding quantitative depth in diplopic stimuli.

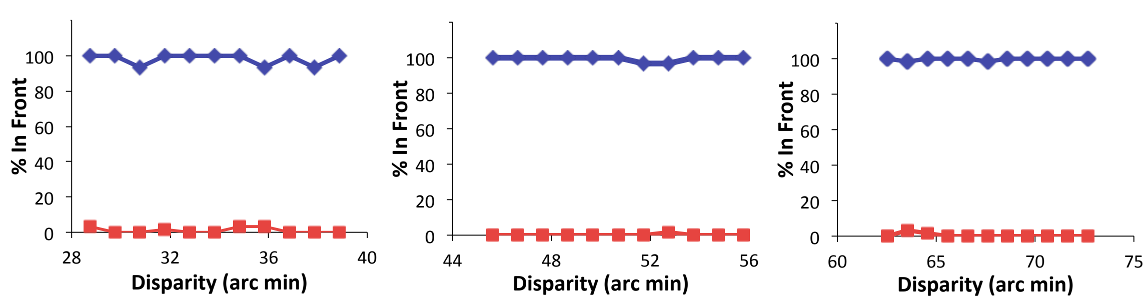


Figure 3.4. The results from Experiment 3.1.1; Percent In Front as a function of disparity for two observers for pedestal disparities of 34, 51 and 68 arc min respectively. The blue curve represents stimuli that were in front of the fixation plane, and the red curve represents stimuli that were positioned beyond the fixation plane.

3.3.3. Experiment 3.1.2

Interpretation of Experiment 3.1 relies on the assumption that it is possible to make the separation information less reliable by interleaving crossed and uncrossed trials, as suggested by Siderov and Harwerth (1993). The second follow up experiment (3.1.2) will evaluate if the use of the interleaved disparity protocol in Experiment 3.1 is critical to the pattern of results.

Methods

To evaluate the effectiveness of interleaving crossed and uncrossed trials, I re-tested three observers (subjects 1, 2 and 3) using the same range of disparities, stimuli and task, but in the crossed direction only. As in Experiment 3.1, observers judged whether the test stimulus (on the bottom) was in front or behind the reference stimulus that was positioned at a fixed pedestal disparity above the fixation marker.

3.3.4. Experiment 3.1.2: Results and Discussion

Stereosensitivity as a function of pedestal disparity (crossed direction only) is presented in Figure 3.5 (grey curve). Similar to the results of Experiment 3.1, for most observers, stereosensitivity is higher in the fused range (0 and 17 arc min pedestals) than it is in the diplopic range (34, 51 and 68 arc min pedestals). Results from Experiment 3.1 are also plotted on figure 3.5 to allow for a direct comparison between interleaved and crossed-only presentations. As results in Experiment 3.1 3.1.2 are so similar, it is likely that observers used similar strategies in the interleaved and non-interleaved paradigms. It remains possible that separation-based judgments are useful whether direction is interleaved or not.

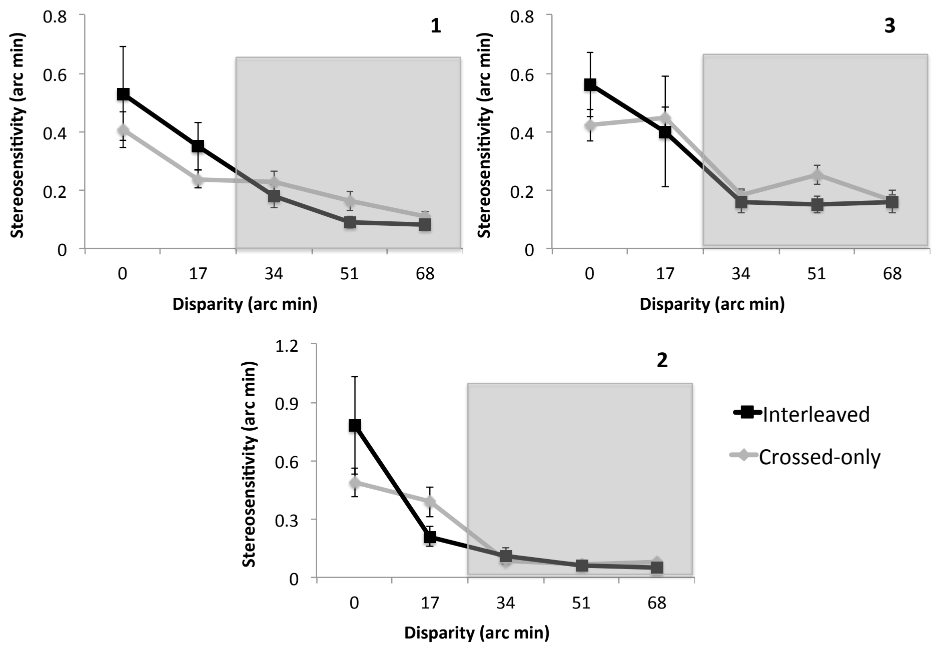


Figure 3.5. Results from Experiment 3.1.2; depth discrimination results for n=3 subjects, when both directions were interleaved in Experiment 3.1. (black curve) or when presented the crossed direction only (grey curve). Error bars represent bootstrapped 95% confidence intervals. The grey shaded region represents disparities that were perceived as diplopic.

3.4. Discussion of Experiment 3.1, 3.1.1 and 3.1.2

In Experiment 3.1, depth discrimination across a large range of disparities was measured using an interleaved 2AFC task as recommended by Siderov and Harwerth (1993). The results were consistent with previous studies (Badcock &

Schor, 1985; McKee et al., 1990; Ogle, 1952b). Results from Experiment 3.1.1 indicated that observers reliably discern the direction of the entire arrangement of the reference and test stimulus. However, Experiments 3.1.2 showed that performance is identical whether trials are interleaved or not. Therefore since observers are aware of the direction of the depth offset on each trial, they could, and likely do, employ a strategy to subsequently use the separation information to perform the task. This suggests that results from Experiment 3.1, as well as results from earlier studies, may possibly be confounded by the separation information.

While observers might not always rely on the separation information, it is present, and at the very least could add noise. Therefore a traditional depth discrimination method, whether the sign of the depth offset is interleaved or not, may not be appropriate for assessing depth from diplopic targets. In the following section, a depth estimation method will be used to investigate how perceived depth varies over a large range of disparities and as in Experiment 3.1, the reliability of this method will be evaluated in a follow-up experiment.

3.5. Experiment 3.2: Depth Estimation

Ogle (1952b) used a depth estimation task where observers reported depth percepts as a needle moved further away from fixation. However, he did not use a broad range of disparities, rather his test range consisted mainly of large disparities that were near or beyond the diplopia threshold. As a result it was possible that the separation information influenced observers' responses. Therefore, the goal of Experiment 3.2 was to evaluate Ogle's (1952b) findings using a broader range of disparities and a paradigm that makes the separation information less salient. A simple broadband bar stimulus was presented for a short duration (160 msec), at a range of disparities, which encompassed fused and diplopic percepts. A strong reference plane was provided to ensure that observers fixated on the zero disparity reference plane. Unlike Ogle's procedure where observers verbally reported their depth percept, observers in this experiment reported the amount of perceived depth by adjusting a horizontal line as shown in Figure 3.6. Since the same stimulus parameters and disparity range were used in Experiments 3.1 and 3.2, a comparison can be made between depth discrimination results and perceived depth estimates.

3.5.1. Experiment 3.2: Methods

Subjects and Apparatus

Four observers were tested using the depth estimation paradigm, all of whom had normal or corrected to normal vision and excellent stereopsis. As in Experiment 3.1, a bar stimulus (4.0 x 100.0 arc min) whose edges were blurred by a horizontal Gaussian envelope (sigma=2.0 arc min) was used. Stimuli were presented on a pair of LCD monitors arranged in a Wheatstone mirror stereoscope at a viewing distance of 57.0 cm, and a chin rest was used to maintain a stable head position. The screen resolution was 1920 x 1200 pixels and the screen size was 51.7 x 32.4 cm so each pixel subtended 1.6 arc min. The displays were refreshed at a 120 Hz and their mean luminance was approximately 46 cd/m^2. The room was dark except for light from the stereoscopic display. Nonlinearities in the luminance output from the monitors were gamma-corrected using software look-up tables. The appropriate correction was determined from the relationship between pixel value and screen luminance obtained at a range of contrasts using a Konica Minolta LS-110 (78423013) digital luminance meter.

Procedure

Since the shape of Panum's fusional area varies among individuals, diplopia thresholds were measured for each observer to correctly classify stimuli within a session as either "fused" or "diplopic". Observers fixated a zero disparity cross, and a test line appeared for 160 msec. A diplopia threshold was measured using a forced choice task where observers judged whether the stimulus appeared fused or diplopic. For the main experiment, a nonius cross was presented where one eye viewed the top and the other eye viewed the bottom half of the cross. A fusion guide consisting of circles displayed on the top and bottom of the screen were positioned at zero disparity. The pattern of circles was updated from trial to trial so they provided no consistent position cue. A horizontal gap (1.2 x 7.6 deg) in the middle of the screen where the stimulus appeared did not contain the pattern of circles. Once the observer indicated that the fixation cross was aligned, a target stimulus appeared for 160 msec. Observers then reported the amount of perceived depth between the target and the fixation cross by adjusting a horizontal bar along a vertical ruler (figure 3.6). It is possible that using a 2D vertical ruler to indicate 3D depth biases estimations, however Hartle and Wilcox (in preparation) showed that depth estimations using a ruler, such as the one used here, are identical to estimations made when using a 3D probe or a tactile sensor strip.

The direction of the depth offset was interleaved across trials and a large range of disparities was tested from -40 to +40 arc min. Each disparity was presented 10 times, and a total of 19 disparities were presented (9 crossed, 9 uncrossed and zero disparity).

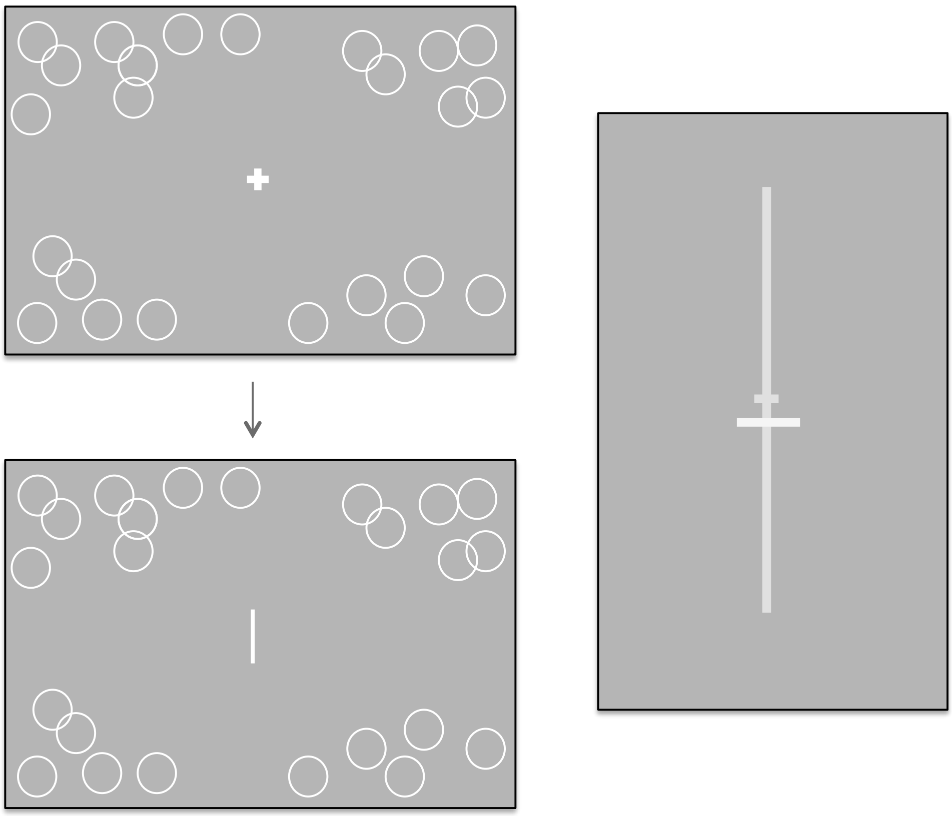


Figure 3.6. The stimulus configuration for Experiment 3.2 is illustrated here. Observers fixated a Nonius cross, and a line was shown at a given depth offset. Observers reported perceived depth with an adjustable ruler.

3.5.2. Experiment 3.2: Results and Discussion

Perceived depth estimates as a function of actual (theoretical) depth was recorded for each observer and average data are shown in Figure 3.7 (top graph). The non-shaded region indicates disparities that appeared fused and the shaded region indicates disparities that appeared diplopic. The grey curve represents predicted depth from disparity calculated using equation 3.2, described in section 3.2.5. For all observers, perceived depth was near veridical in the fused range (shown in the bottom graph in figure 3.7), but did not follow geometric predictions beyond the fusion limit.

Results from Experiment 3.2 are generally consistent with those of Ogle (1952b, 1953), and coincide well with results of the discrimination task (fused range) in Experiment 3.1. All observers perceived patent depth in the fused range and qualitative depth in the diplopic range. To test the observation that the data are best described by different functions in the fused and diplopic regions, two models were compared using the Extra sum-of-squares F-test. The null hypothesis is that a straight line fits the data from Experiment 3.2. The alternative hypothesis is that a non-linear model (Naka-Rushton function) is necessary to fit the data. Results from this analysis indicate that the Naka-Rushton function fits the data (F91, 17)=291, R^2=0.992, p<0.001), thus the null hypothesis is rejected. Therefore, a non-linear function is needed to model the depth estimation data.

The reported depth sign was correct for all test disparities and the diplopia threshold for each observer fell between 10-18 arc min (or actual depth of 1.5-2.5 cm). An increase in perceived depth was observed in the fused range (disparities corresponding approximately to depth equal to – 2 to +2 cm), and remained constant in the diplopic range (disparities corresponding to depth larger than ±2 cm). Data from one observer continued to increase slightly beyond the fusion limit which suggests that she may have been influenced by the separation information at large disparities (individual graphs are shown in Appendix A). This possibility will be tested further in a follow-up experiment.

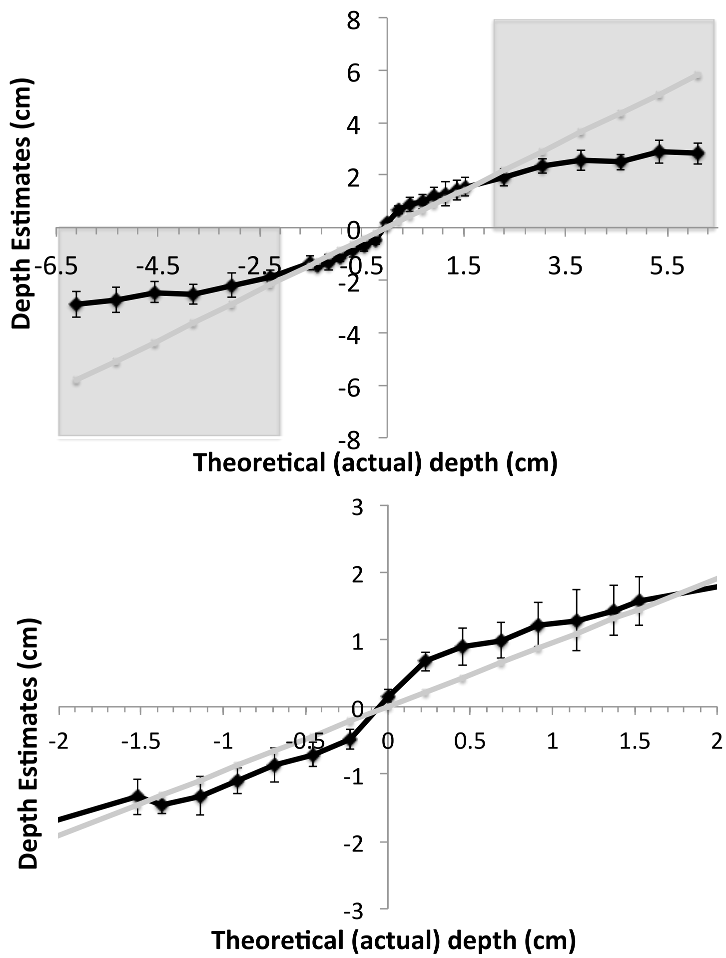


Figure 3.7. Results from Experiment 3.2 showing perceived depth estimates as a function of disparity were averaged across n=4 observers (subjects 1, 2, 6, 7) and plotted in the top graph. Targets appear diplopic when disparities are within the grey shaded regions. The grey curve represents the predicted depth. The bottom graph is an enlarged version of the fused range.

3.6. Experiment 3.2.1: Depth estimation Follow-up

In Experiment 3.2, a depth estimation paradigm was used to evaluate perceived depth over a range of fused and diplopic disparities. When assessing perceived depth in the diplopic range, it is possible that observers rely on the separation between the half images instead of stereoscopic disparity. One observer's responses continued to increase slightly in the diplopic range, suggesting that she may have relied on the separation information in the depth estimation. On the other hand, it is also possible that she did not relying on the separation of the stereo images, but continued to perceive quantitative depth over the full range of test disparities. To evaluate the extent to which her responses were influenced by separation, a follow-up experiment was conducted where observers reported the amount of lateral separation for these same 'offsets' when both sets of half images were presented to both eyes (so they appeared at zero disparity). Performance on this 2D separation task was compared to the depth judgments made when the two stimuli were presented stereoscopically. If an observer was using the separation information to indicate the amount of depth in the diplopic targets their results should be identical in the two conditions.

3.6.1. Experiment 3.2.1: Methods

The same four observers (subjects 1, 2, 6, 7) participated in this follow-up experiment. In one session, stereoscopic stimuli (Figure 3.8a) were presented at large disparities. In the second session, observers viewed 2D versions of the same stimuli (Figure 3.8b). The degree of separation between identical "half images" was varied and was equal to the disparities tested in part one. As in Experiment 3.2, observers first fixated on a central Nonius fixation cross. The top half of the cross was presented to one eye and the bottom half of the cross was presented to the other eye. Once the observer indicated that the cross appeared aligned, the test stimulus appeared for 160 msec. Following stimulus presentation, a vertical line appeared. Observers adjusted a small horizontal line along the length of the vertical line to reflect the amount of depth or separation (depending on the session) that they perceived between the target (line) and fixation (Figure 3.8c). Five disparities ranging from 4.6-10.7 cm (equal to 30-70 arc min of actual depth) were presented 10 times apiece in random order.

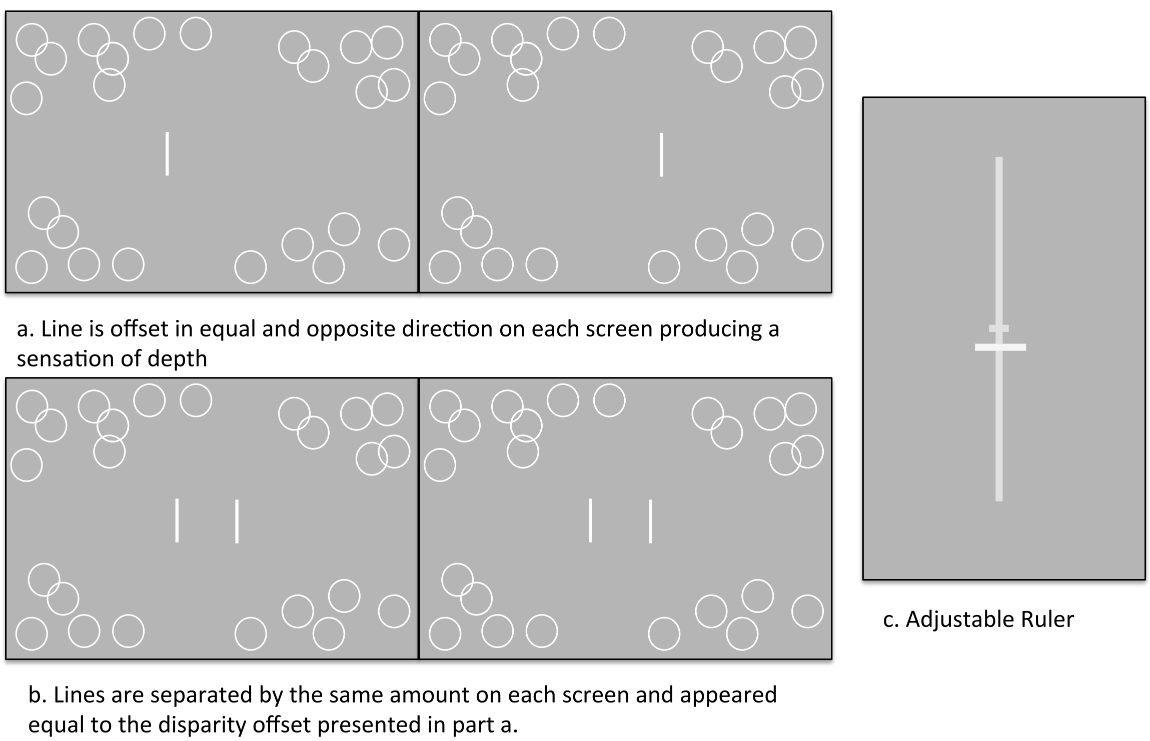


Figure 3.8. The conditions tested in Experiment 3.2.1 are shown here. Observers judge depth magnitude in (a), and the lateral separation in (b) and report the amount of depth or separation using an adjustable ruler (c).

3.6.2. Experiment 3.2.1: Results and Discussion

As shown in Figure 3.9, observers responded differently depending on whether observers were asked to judge the depth (3D condition) or the horizontal

2D separation of the stimuli ($F(1, 398)=78.642$, , $p<0.01$, partial $\eta^2=0.312$, Power= 1.0). One-way ANOVAs were run for each individual and results indicated a significant main effect for each observer. A Bonferroni post hoc analysis showed that every comparison tested was statistically significant except for one comparison at 7.6 cm for Observer 2 ($p>0.01$). The results of two observers (1 and 2) were linear in the 2D condition, which reflects an increase in reported estimates with increasing separation, while responses were somewhat flat in the 3D condition. Observer 4's performance was quite different in the two conditions ($p<0.01$) and as seen in Figure 3.9, estimates in the 3D condition were larger and more variable than responses in the 2D condition. For Observer 3, the pattern of performance in the two conditions appeared somewhat similar, but when tested statistically, performance was significantly different ($p<0.01$). It is clear from these data that observers are not simply and consistently relying on the 2D separation information to make metric depth judgments at the large depth values. However it is possible that on some trials and for some observers, this separation cue influences their estimates.

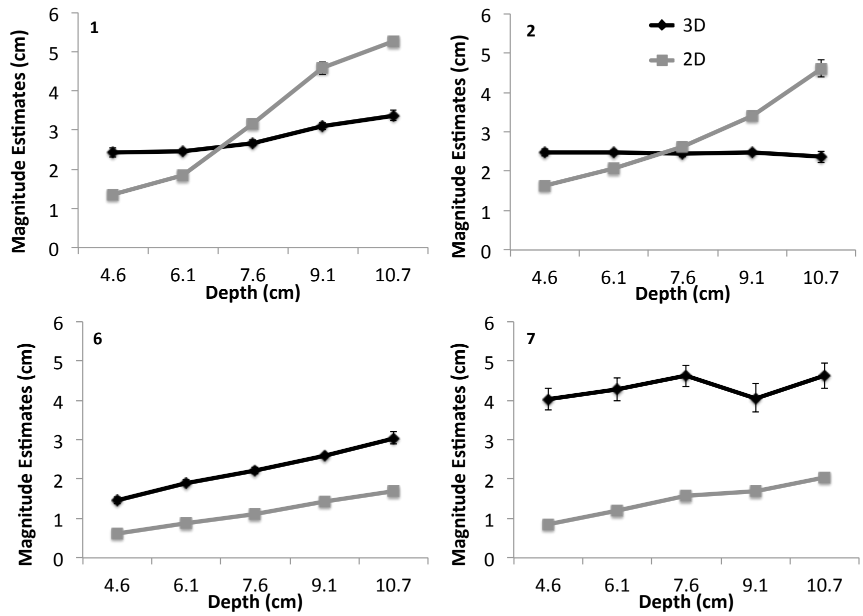


Figure 3.9. The results from Experiment 3.2.1 are shown here. Graphs for n=4 subjects, comparing magnitude estimates for 3D (black line) and 2D (grey line) conditions. Error bars represent standard error of the mean.

3.7. Discussion

Investigators have argued that a two-segment function best describes depth discrimination (Badcock & Schor, 1985; McKee et al., 1990; Ogle, 1952b, 1953) and perceived depth (Ogle, 1952b) as a function of disparity. Using a range of different stimuli and experimental designs, these authors found an increase in discrimination thresholds/perceived depth at small disparities when targets were fused, and a plateau at large disparities (diplopic targets). However, previous research consisted of subjective measures (Ogle, 1952b), inadequate sampling of disparities (Blakemore, 1970; Ogle, 1952b, 1953), no determination of a diplopia threshold and lack of assessment of the reliance on separation information for diplopic disparities (Ogle, 1952b). The purpose of this chapter was to address these omissions, and develop a method that could be used to assess depth percepts from diplopic stimuli.

Previous investigators have suggested that a simple depth discrimination paradigm may not be the ideal method for studying diplopic targets. McKee et al. (1990) found similar functions for a disparity condition and a monocular condition, suggesting a reliance on the separation between the half images at large disparities. Later, Siderov and Harwerth (1993) suggested that the reliance on the separation information could be removed if the direction of the depth offset was interleaved across trials. While this may be true for some observers, interleaving crossed/uncrossed trials did not remove the separation information in the discrimination task in Experiment 3.1. Once observers discern the overall depth sign of the entire arrangement of the reference and test stimuli, depth discrimination is potentially confounded by the separation information.

In Experiment 3.2, a depth estimation paradigm was used to determine if a two-segment function represents perceived depth across a large range of test disparities. Compared to a depth increment discrimination paradigm, this method evaluates the amount of metric depth reported by observers, and results can also be compared to theoretical predictions of depth from disparity. As reported by Ogle (1952b), an increase in perceived depth estimates as a function of disparity was observed in the fused range, and perceived depth estimates were relatively constant within the diplopic range. These results provide further evidence for the patent/qualitative dichotomy to describe depth percepts as a function of disparity. However, these results do not reflect the same transition point between patent and qualitative percepts as described by Ogle (1952b). Similar to the results of Badcock and Schor (1985), data from Experiment 3.2 show that the transition between patent and qualitative depth percepts occurs near the diplopia threshold.

To summarize, simple depth discrimination tasks may not be appropriate for assessing stereothresholds at large disparities. Conversely, the magnitude estimation paradigm used in Experiment 3.2 seemed to make the separation information less salient. Perceived depth increased as a function of disparity in the fused range, while observers transitioned to qualitative depth at the fusion limit and reports of metric depth flattened. Therefore a two-part function is necessary to model the data across the entire range of disparities, which suggests that the visual system uses two separate mechanisms to encode depth from fused and diplopic stimuli. The relationship between patent/qualitative percepts and the spatial dichotomy (Hess & Wilcox, 1994) will be evaluated in the following chapter.

Chapter 4

Evaluating the relationship between 1st-/2nd- order stereopsis and Patent/ Qualitative percepts

4.1. Introduction

As outlined in the Introduction (section 1.3.5) and Chapter 2 (section 2.1.1), researchers have demonstrated that the stereoscopic system is able to compute retinal disparity to extract a disparity signal by using either luminance variation (1st- order) or contrast modulation (2nd order). In the case of 2nd- order stereopsis, the carrier (or the underlying luminance signal) does not have to be correlated in the two eyes, making this mechanism resilient to differences in shape and polarity that would make it impossible to encode disparity via strictly 1st order processing (Mitchell & O'Hagan, 1972; Wilcox & Hess, 1995; Wilcox & Hess, 1996). Depth can be extracted from both 1st- and 2nd- order stimuli at small disparities (Wilcox & Hess, 1997, 2006, 2008) while it appears that a 2nd- order mechanism solely determines depth percepts at the upper limit for stereopsis (Wilcox & Hess, 1995). The purpose of this chapter is to investigate the relationship between 1st- /2nd- order stereopsis and Ogle's description of patent/qualitative percepts, described in the Introduction and in Chapter 3.

4.2. Rationale

The first step in relating Ogle's distinction to 1st - /2nd- order stereopsis is to evaluate the nature of the depth percept provided by stimuli that are processed by 1st- order or 2nd- order mechanisms. Wilcox and Allison (2009) proposed that 1st-order stereopsis encodes patent depth and operates in the fused range, while a 2nd-order mechanism likely encodes depth across the entire range of disparities (Figure 4.1). Discrimination experiments using uncorrelated stimuli, which activate 2nd-order processing, indicate that depth sign is perceived quite reliably (Wilcox & Hess, 1996), however, it is unclear if depth magnitude information can be extracted. The accuracy of depth estimates from uncorrelated stimuli as a function of disparity will be evaluated and compared to results from a 1st- order condition, which contain both 1st- and 2nd- order signals.

In Experiment 4.1, a depth estimation experiment was conducted where observers estimated the amount of perceived depth for correlated and uncorrelated stimuli. In Experiment 4.2, a 2-interval forced choice (2IFC) paradigm was implemented and observers made relative depth magnitude judgments using a depth pedestal discrimination task. Observers were presented with two conditions, one that activated both 1st- and 2nd- order processes (correlated condition), and a second condition that isolated 2nd- order stereopsis (uncorrelated condition).

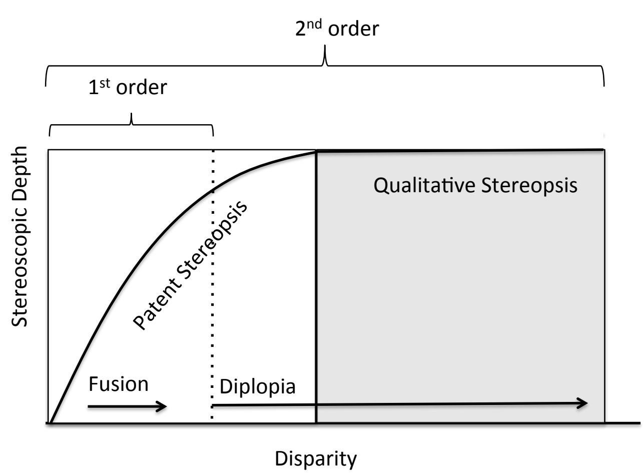


Figure 4.1. Diagram illustrating Wilcox and Allison's (2009) proposal.

4.3. Experiment 4.1: Depth Estimation Experiment

4.3.1. Introduction

In Experiment 4.1, the relationship between Ogle's perceptual dichotomy and the 1^{st}-/2^{nd}- order dichotomy was evaluated. This was accomplished by determining the magnitude of depth that could be extracted from correlated and uncorrelated stimuli. Using a depth magnitude paradigm, observers estimated the amount of perceived depth across a large range of disparities. Observers reported depth estimates in two conditions: one that activated both 1^{st}- and 2^{nd}- order processes (correlated stimuli), and a second condition that isolated 2^{nd}- order stereopsis (uncorrelated stimuli). Given Ogle's results, my hypothesis was that an increase in perceived depth would be observed in the fused range for the correlated condition. Two outcomes are possible at the transition from fused to diplopic percepts. If Ogle's description is correct, the transition from patent to qualitative depth should occur in the diplopic range. However based on the results from Experiment 3.2, the transition point should occur at the fusion limit. Finally, there is some evidence that 2^{nd}- order stereopsis provides a low resolution disparity signal (coarse), (Langley et al., 1999; McKee et al., 2007; Schor et al., 1998; Wilcox & Hess, 1996, 1997), therefore it was expected that only qualitative (signed) depth percepts would be obtained in the uncorrelated condition.

4.3.2. Experiment 4.1: Methods

Subjects and Apparatus

Stereoacuity was assessed with the Adult Randot™ stereotest to ensure that observers could detect depth from binocular disparities of at least 40 arc sec. The four observers who participated in this experiment had normal or corrected-to-normal vision. One observer (9) was inexperienced in stereoscopic tasks. Stimuli were presented on a Wheatstone mirror stereoscope with a viewing distance of 57 cm and a chin rest was used to maintain a stable head position. The screen resolution was 1280 x 960 pixels and the screen size was 40 x 30 cm. The displays were refreshed at 120 Hz. The mean luminance of the display was approximately 69 cd/m^2. The linearity of the monitors was assessed using a luminance meter (Konica Minolta, LS-110, 78423013), and then calibrated by gamma correcting nonlinearities in the luminance output response.

Stimuli

Stimuli consisted of one-dimensional luminance noise windowed horizontally with a raised-cosine (Figure 1). The width of the envelope was 60 arc min, making the dimensions of the noise samples 60 x 60 arc min. Noise samples were re-generated on every trial so that no two noise patches in the same trial (applicable to uncorrelated stimuli) or consecutive trials consisted of the same noise samples. Noise patterns were created by randomly choosing one of 256 grey levels for each band within the image. Each band was 2 arc min wide. Noise samples were balanced about the background grey level so that they averaged to mean luminance.

Prior to presentation, each stimulus was assessed to ensure that the bands at the edges did not equal mean luminance. Therefore, the edges did not blend into the background, which ensured that the width of every noise sample appeared the same.

Two conditions were tested. In the correlated condition, given the presence of a reliable carrier (1st- order) and raised-cosine envelope (2nd- order) both sources were available to perform the task. To clarify, correlated stimuli could stimulate both mechanisms because the 1st- order system is responsive to the matched luminance noise information within the carrier while the 2nd- order system is responsive to the contrast modulation of the raised cosine envelope. In the uncorrelated condition, different noise samples were presented to each eye to selectively activate 2nd- order processing. In addition, to avoid any possibility of edge-based matching, noise samples in the uncorrelated condition had different orientations as one eye received vertically oriented noise and the other received horizontally oriented noise. Previous research has shown (Hess & Wilcox, 1994; Wilcox & Hess, 1995; Wilcox & Hess, 1996) that when different randomly generated noise patches are presented to each eye as a stereopair, the disparity provided by the 2nd- order contrast envelope is used to judge depth. To verify that rivalry did not influence the results, pilot studies were conducted using vertical-only noise, which yielded similar results to those presented here (see appendix B).

An array of circles positioned at the screen plane was displayed above and below fixation to help lock vergence. The pattern of circles was updated from trial to trial so they provided no consistent localization cue. A horizontal gap (1.2 x 7.6 deg) along the middle of the screen where the stimulus appeared did not contain the pattern of circles. Once the stimulus disappeared, observers reported the amount of perceived depth between the target and the fixation cross by adjusting a horizontal bar along a vertical ruler, as described in the previous chapter (section 3.5.1).

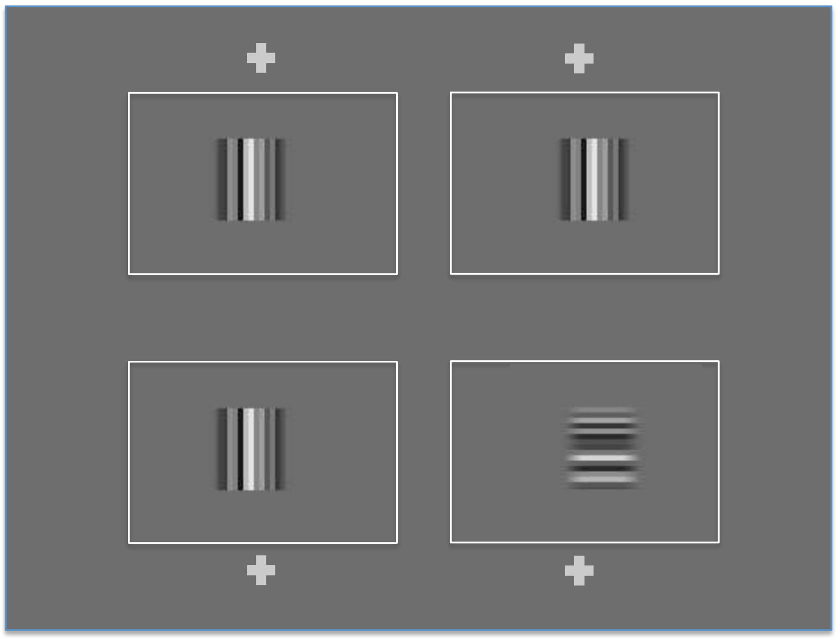


Figure 4.2. Stimuli presented in Experiment 4.1. The upper stereogram is a correlated stimulus and the lower stereogram depicts an uncorrelated stimulus. Both types of stimuli are windowed with a raised cosine envelope.

Procedure

Initially a pre-test was conducted in which observers' diplopia threshold was measured. Because Ogle's classification of patent and qualitative percepts was linked to Panum's fusional area, it was very important to carefully define diplopia. For each observer, three different disparity ranges were identified: completely fused, overlapping and completely diplopic. The overlapping region is a diplopia transitional zone, where the stimuli did not appear completely fused but were not perceived as distinct entities.

The main experimental session consisted of a depth magnitude task. Observers fixated a nonius cross which was presented such that one eye viewed the top half of the cross and the other eye viewed the bottom half of the cross. Once the observer indicated that the two half images were aligned, a trial was initiated. After a short pause (50 msec) the stimulus replaced the cross for 160 msec. This exposure duration was chosen to ensure that observers did not complete vergence eye movements during the trial (Rashbass & Westheimer, 1961). After the stimulus disappeared, a vertical ruler was presented, and observers adjusted a small horizontal line along the length of the ruler to indicate the amount of depth perceived between the stimulus and the zero-disparity fixation cross.

The disparity ranges determined during the pre-test were used to ensure that test disparities from all three perceptual categories (fused, overlapping, diplopic) were assessed. Crossed and uncrossed disparities were presented randomly across trials within a session. The test disparities were presented in random order, and observers viewed 20 repetitions and a total of 29 disparities (14 crossed and 14 uncrossed bracketing zero disparity).

4.3.3. Experiment 4.1 Results

In Figure 4.3, perceived depth as a function of disparity is plotted on a log-linear scale for four observers. Observers 1, 2, 8, 9 were tested here (note that subjects 1 and 2 also participated in the experiments described in Chapter 3). The shaded grey regions in Figure 4.3 represents the range of disparities classified as overlapping, therefore targets presented at disparities smaller than the lower limit of this region appeared fused while targets presented at disparities beyond the upper limit of this region were perceived as completely diplopic.

In general, perceived depth was similar for both the correlated and uncorrelated conditions, with slightly higher variability in the uncorrelated condition. It is clear that in the fused range, patent depth is perceived in both the correlated *and* uncorrelated conditions. Paired-sample t-tests were conducted to compare estimates at each disparity value, and for each observer. In the fused range, the majority of comparisons were statistically significant for three observers. No statistically significant difference was found between the two conditions in the overlapping or diplopic ranges ($p<0.01$). The similarity in performance at large disparities likely reflects a single mechanism, likely 2^{nd} order, encoding depth across this range. The performance of observer 4, the only completely naïve observer, was more variable than the other observers which may have resulted in fewer significantly different comparisons in the fused range compared to the other three observers.

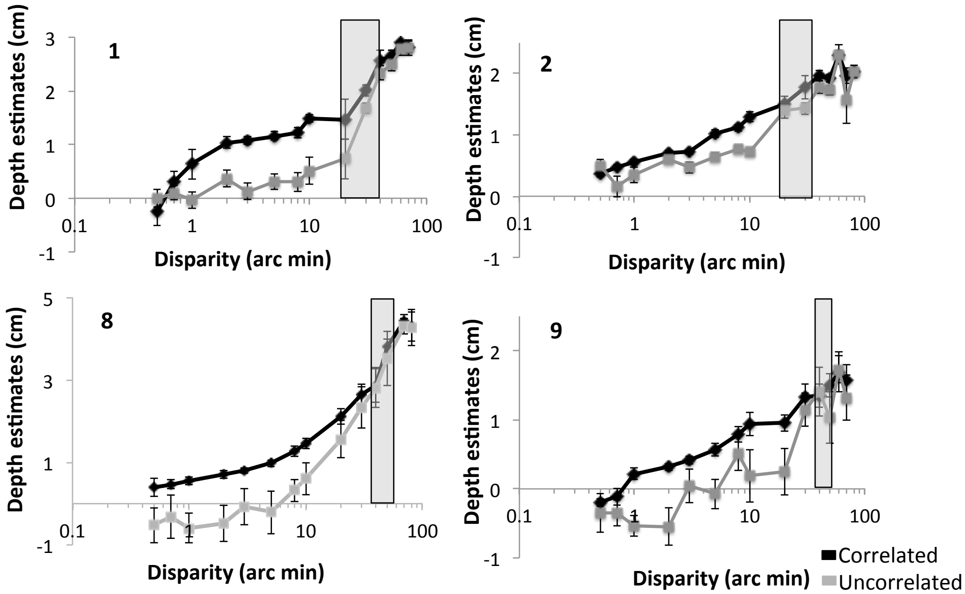


Figure 4.3. Results from Experiment 4.1 showing magnitude estimates for n=4 observers (subjects 1, 2, 8, and 9). The black curve represents data from the correlated condition and the grey curve represents data from the uncorrelated condition. The shaded grey region represents the overlapping range, therefore targets at smaller disparities appeared fused and targets beyond this region appeared completely diplopic. Error bars represent 95% confidence intervals.

4.3.4. Experiment 4.1: Discussion

The primary goal of this experiment was to determine the quantitative nature of 1st- and 2nd- order processing. The original prediction was that the 1st-order system can extract quantitative depth from correlated stimuli but only sign information could be encoded by a 2nd- order signal. Therefore it was expected that observers could reliably discern the direction of depth in the uncorrelated condition, but will not report the correct amount of perceived depth. Clearly, within the fused range, patent depth percepts are supported by both 1st- and 2nd- order stereopsis. Since only 2nd- order information can be extracted in the uncorrelated condition, this study provides the first evidence that quantitative depth is a property of both 1st- and 2nd- order stereopsis. Previous studies have found that observers perform significantly worse, or have significantly higher thresholds, when presented with uncorrelated stimuli compared to correlated stimuli (Wilcox & Hess, 1996). It has also been suggested that 2nd- order processing is coarse (Hess & Wilcox, 1994; McKee et al., 2004, 2005; McKee et al., 2007; Wilcox & Hess, 1995; Wilcox & Hess, 1996). This observation is consistent with the data depicted in Figure 4.3. Within the fused range, depth estimates are generally lower in the uncorrelated (2nd- order) test condition compared to the correlated condition.

One limitation of the depth magnitude paradigm is that observers viewed the stimuli in depth but reported the amount of depth using a ruler in the y-axis. Thus a spatial transformation was required to perform the task, which may have influenced depth estimation. However, since the same estimation method was used throughout the study, any effect of the method should apply equally to all test conditions. Less experienced observers reported that this task was very difficult and as a result, responses were variable (ie. Observer 4).

Furthermore, like the depth discrimination task conducted in Chapter 3 (Experiment 3.1), the magnitude estimation method was potentially confounded by the separation of the half images at large disparities. In Chapter 3, a comparison of a 2D and 3D condition using a line stimulus suggested that separation information did not influence metric depth judgments in a depth estimation task. It is unlikely that this information was used at large disparities because responses plateaued. However, to be certain, a control study similar to Experiment 3.2.1 was conducted in the following section (4.4).

4.4. Experiment 4.1.1: Follow-up experiment to Experiment 4.1.

As outlined previously, when stereoscopic stimuli are diplopic the visibility of the half-images becomes a potential confound. In this follow up experiment, the separation cue was assessed (as in Chapter 3) by recording estimates using a depth estimation task for 2D and 3D conditions. It is possible that observers were able to report the amount of lateral separation between the half images instead of the amount of depth as a function of disparity. It is important to re-assess observers' use of this strategy at large disparities for each paradigm and stimulus tested. In Chapter 3, results confirmed that observers do in fact rely on stereopsis in the diplopic range, not lateral separation. The goal of the following experiment is to determine if this is also true when stimuli consist of 1D luminance noise.

4.4.1. Experiment 4.1.1: Methods

The stimuli and apparatus were identical to those described in section 4.3 and the procedure was the same as that described in Experiment 3.2.1. To summarize, three observers (subjects 1, 2, 8) participated in two conditions and these observers also participated in Experiment 4.1. In the 3D condition, observers viewed correlated stimuli that were presented in depth at a range of binocular disparities (50-120 arc min, depending on their diplopia point). Each disparity was presented 10 times. In the 2D condition, the "half-images" were separated horizontally and presented simultaneously on both screens, and observers were asked to judge the lateral separation. The degree of separation between identical "half images" was varied and was equal to the disparities tested in the 3D condition.

4.4.2. Experiment 4.1.1: Results and Discussion

An analysis of variance (ANOVA) showed a main effect for condition ($F(1, 398)=93.47$, $p<0.01$, partial $\eta^2=0.193$, Power= 1.0). There was no significant interaction between condition and disparity ($F(4,390)=2.115$, $p>0.05$, partial $\eta^2=0.022$, Power= 0.62). Due to the large degree of individual variation, each observer's data was assessed separately. Paired-sample t-tests were conducted to compare performance at each disparity tested. As seen in Figure 4.4, performance on the 3D condition was significantly different from the 2D condition at some, but not all, disparities. For example, Observer 2's performance in the two conditions was significantly different when the disparity was between 50-70 arc min ($t(9)=22.457$, $t(9)=7.635$ and $t(9)=5.939$, $p<0.01$) but was similar at the two larger disparities tested. Observer 3 viewed stimuli at larger disparities than the other observers because her fusion limit was almost double the fusion limit compared to other observers. Her performance in the two conditions was significantly different when test disparities were between 80-100 arc min ($t(9)=4.149$ and $t(9)=2.425$, $p<0.01$), but performance was the same when test disparities were 110 and 120 arc min. The performance of Observer 1 in the two conditions was significantly different ($p<0.01$) for all disparities other than 60 and 70 arc min. These results suggest that depth estimation tasks are reliable overall, but reliance on the separation information is dependent on the individual, the range of disparities tested and the type of stimulus.

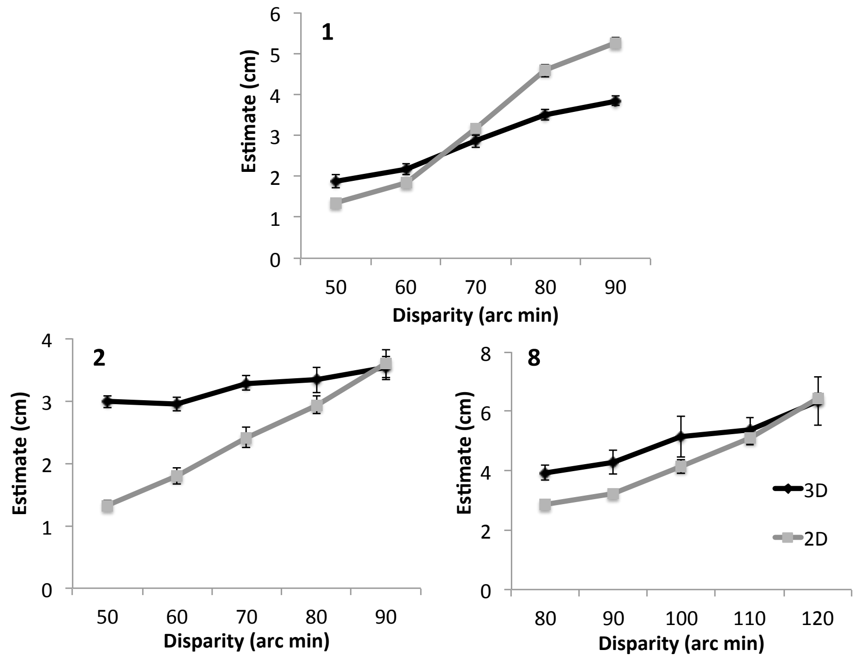


Figure 4.4. The results from Experiment 4.1.1 are shown here. The estimates in the 2D (grey curves) and 3D (black curves) conditions are plotted as a function of disparity (arc min). The error bars represent 95% confidence intervals.

4.5. Experiment 4.2: Two Interval Forced Choice Experiment

4.5.1. Introduction Experiment 4.2

Experiment 4.1 demonstrated that in the fused range, quantitative depth is obtained from stimuli that contain both 1st- and 2nd- order information (correlated) as well as from stimuli that isolate 2nd- order processing (uncorrelated). In the diplopic range, similar patterns of depth estimates are found in both the correlated and uncorrelated conditions, however the follow-up experiment suggests that there is individual variation in performance and the estimation task in Experiment 4.1 may be unreliable for certain observers at large disparities. The purpose of

Experiment 4.2 was to use a discrimination task to quantify perceived depth using a task that makes the separation information less salient. A two-interval depth discrimination (2IFC) paradigm was used where observers selected the interval that contained the larger depth offset.

4.5.2. Methods

Apparatus and Stimuli

Four experienced observers with normal or corrected-to-normal vision were tested. All observers had excellent stereopsis as assessed by the Adult Randot™ stereotest. Stimuli were presented on a Wheatstone mirror stereoscope with the same parameters as Experiment 4.1.

Observers were tested using both correlated and uncorrelated noise stimuli. In the correlated condition, identical vertically-oriented 1-D noise patches were presented to each eye. In the uncorrelated condition, different vertically oriented noise patches were presented to each eye. Stimuli in both conditions were windowed with a raised cosine envelope.

Procedure

Four observers participated in Experiment 4.2. Three of the four observers also participated in Experiment 4.1 (Subjects 1, 2, 8) while one observer did not (observer 10). Initially, each observer's diplopia threshold was measured using the task described in Experiment 4.1, in which observers indicated whether they perceived the stimulus as fused, overlapping or diplopic. This task defined the ranges of disparities necessary to carefully and completely isolate the three distinct perceptual categories. The purpose of the second pre-test was to measure the JND at each of the disparity pedestals tested. A method of constant stimuli was used and observers indicated whether two stimuli appeared in the same or different depth planes. Each function was fit with a Weibull function and the 75% JND was extracted. The target disparity used in the main experiment was set at three times the JND.

For the main experiment, observers were presented with two intervals and were asked to select the interval with the largest amount of depth between two stimuli (Figure 4.5.). Three pedestals were presented, each having 19 test disparities that were positioned in equal steps centered on the target disparity. The reference stimulus was set at a specified pedestal and was positioned either above or below a fixation cross. The reference + target, or the reference + test, was randomly presented in either the first or second interval. A fusion guide consisting of circles set at zero disparity was positioned above and below the test area.

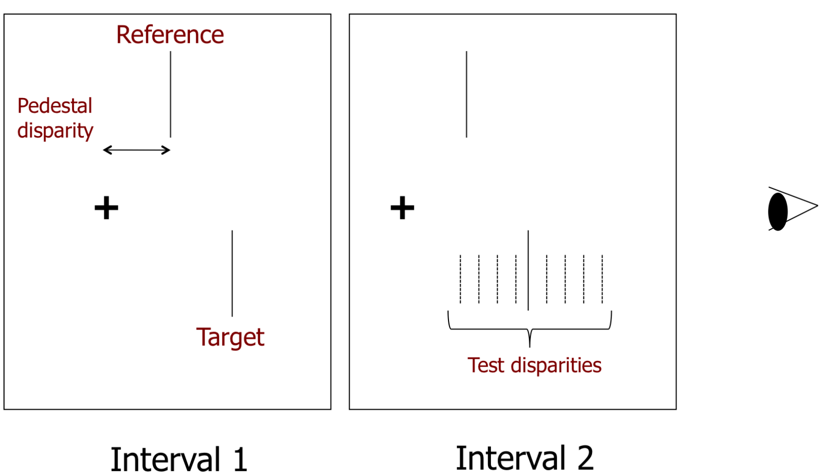


Figure 4.5. A summary of the 2IFC task, in which observers selected the interval containing the largest amount of depth between the reference and a test disparity.

4.5.3. Experiment 4.2: Results and Discussion

Psychometric functions were fit with a Weibull function and were analyzed using traditional curve fitting techniques. Sensitivity was computed from the resulting psychometric function by fitting the function:

$$F(x; k, \lambda) = 1 - \exp(-(x/\lambda)^{k}) \qquad (4.1)$$

Where x is the offset of the function relative to zero, k is the shape parameter and λ is the scale parameter. The slope of each function was used as an index of sensitivity. The slope was computed by taking the derivative of the cumulative Weibull function at 75 percent correct, and 95% confidence intervals were computed by bootstrapping each observers' data (Wichmann & Hill, 2001a, 2001b).

Stereosensitivity is plotted as a function of pedestal disparity for four observers (Figure 4.6). Each data point corresponds to a different disparity range and condition. In the fused range (pedestal = 0 arc min), observers' performance in the correlated condition was significantly different from performance in the uncorrelated condition (full psychometric functions are shown in Appendix C). There was a significant interaction between condition and disparity range ($F(5,18)=6.132$, $p<0.01$, partial $\eta^2=0.63$, Power=0.973) and a main effect for condition ($F(1,18)=6.963$, $p<0.05$, partial $\eta^2=0.28$, Power=0.704) and disparity range ($F(2,18)=8.149$, $p<0.01$, partial $\eta^2=0.48$, Power=0.923). In the fused range, steep slopes were obtained in the correlated condition while much shallower slopes were obtained when observers viewed uncorrelated stimuli. The slopes obtained from the correlated and uncorrelated conditions were similar once fusion was compromised (F(1,14)=3.017, p>0.01). Sensitivity dropped precipitously between pedestals of 0 and 40 arc min in the correlated condition (at the diplopia point) and slopes were somewhat steady across all disparities in the uncorrelated condition.

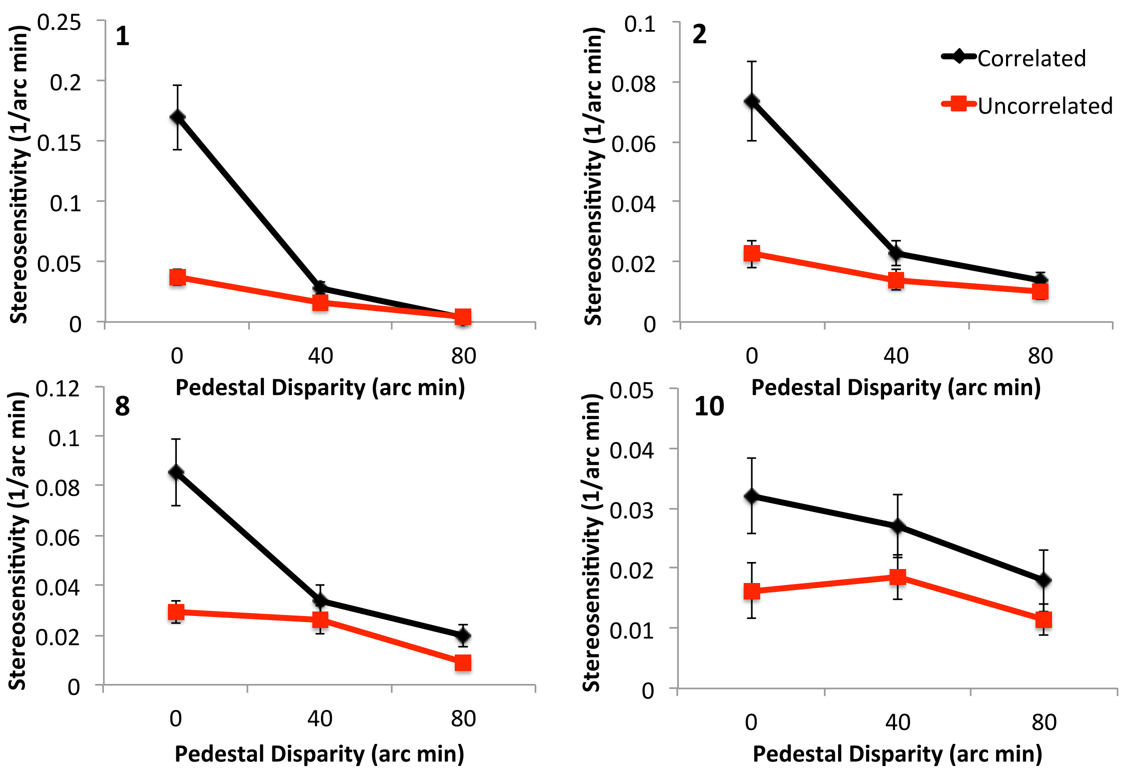


Figure 4.6. Results from the 2IFC task for n=4 subjects (1, 2, 8, 10), plotting slope as a function of the pedestal disparity. The black curves represent the correlated condition and the red curves represent the uncorrelated condition and error bars represent 95% bootstrapped confidence intervals.

Experiment 4.2 demonstrates that performance in the correlated condition was significantly different from performance in the uncorrelated conditions in the fused range only. This finding supports the results found in Experiment 4.1. Taken together, results from Experiments 4.1 and 4.2 are consistent with previous experiments that have suggested that 2nd- order stereopsis provides much coarser depth percepts than 1st- order processing (Wilcox & Hess, 1995, 1997; McKee et al., 2004, 2005, 2007). In addition to the lower stereosensitivity reported previously, these experiments suggest that observers perceive less depth from 2nd- order stimuli (uncorrelated condition) compared to 1st- order stimuli (correlated condition). When fusion is compromised this difference in performance no longer exists. These findings suggest that when viewing correlated stimuli, observers likely use a 1st- order signal to extract depth in the fused range, but transition to 2nd- order processing beyond the fusion limit. This result is consistent with Wilcox and Hess (1995) finding that stereopsis is processed by non-linear filtering at large disparities. Therefore, it is likely that performance at large disparities in Experiment 4.2 is achieved by envelope extraction, and is not determined by the carrier's luminance information.

One issue with previous literature on diplopia is that authors often fail to define it. It seems obvious that diplopia means 'double vision', and this is a relatively straightforward percept when stimuli are very thin (as in the case of Ogle's needles). However, when stimuli have some horizontal extent, there is a range of disparities in which the percept appears to overlap. Given that the stimuli used here were relatively wide, we asked observers to indicate how they perceive the stimuli across a range of disparities, which provided direct estimates of the overlapping range of disparities where the percept is not fused and not fully diplopic. This experiment provides evidence of a change in depth perception occurring at the lower bound of the fusion limit. Despite striking differences in appearance in the overlap and diplopic conditions, performance was very similar. Observers only perceived qualitative depth from stimuli in the overlapping and diplopic range. Since observers performance in this task was relatively poor in the overlapping and diplopic range, it is unlikely that results were confounded by separation information. However, because large disparities are presented here, the impact of the separation confound will be evaluated.

4.6. Experiment 4.2.1: Evaluation of Separation Confound in Experiment 4.2

Results from Experiment 4.1 indicate that some observers can and do use the separation information (at some but not all disparities) to guide depth magnitude estimates. In Experiment 4.2, observers would have had to deliberately go through a series of cognitive steps and deliberate steps. Observers could initially discern the depth sign in two separate intervals, then use the separation information in each interval to determine the amount of separation between the two stimuli, and lastly compare the amount of separation difference in each interval to determine which interval has a greater degree of separation. It seems that observers were not influenced by the separation cue in Experiment 4.2, however to be certain, we compared performance in the 3D version of the 2IFC task outlined above, with a 2D version. The logic was the same as that outlined in Chapter 3 (Experiment 3.2.1): if individuals use the separation information then performance should be the same in the 2D and 3D test conditions at large (diplopic) disparities.

4.6.1. Methods

The paradigm used here was the same as described in Chapter 3 (Experiment 3.2.1), where observers were presented with stimuli that elicited a percept of depth in one session (3D Condition), and a non-stereoscopic version of the stimuli (2D condition). The procedure used was identical to Experiment 4.2. Observers performed the task at the largest pedestal disparity (3D condition) and at equivalent separations (2D condition).

4.6.2. Experiment 4.2.1: Results and Discussion

Slopes (75%) were extracted by fitting individual data sets with a Weibull function, as described in section 4.5.3. For 3 out of 4 observers, performance in the 3D condition was significantly different from performance in the 2D condition, as shown by the lack of overlap of the 95% confidence intervals (Figure 4.7). Therefore although there are some individual differences in observer's reliance on the separation cue, most observers perform quite differently in the 2D and 3D conditions in the 2IFC task. Observer IT performed similarly in the 2D and 3D conditions, and it should be noted that this observer was also influenced by the separation information in the depth estimation task. These results suggest that although it is impossible to *completely* remove the separation cue for every observer, the 2IFC paradigm described here makes it difficult for naive observers to rely on information other than disparity, and is a promising method for the study of depth percepts from diplopic stimuli.

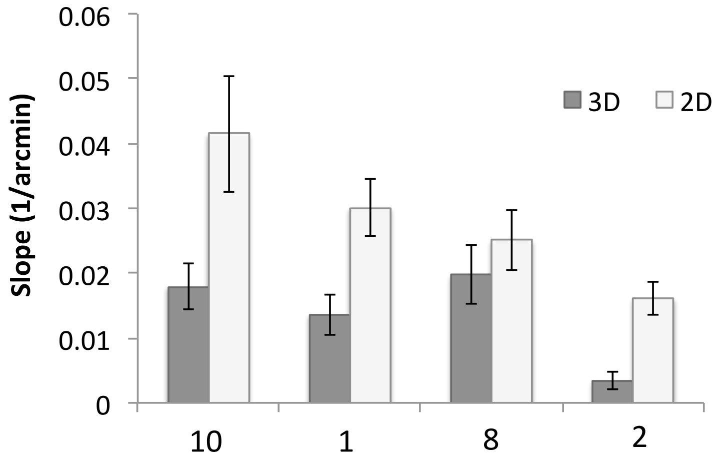


Figure 4.7. Results from Experiment 4.2.1. This graph compares slopes in the diplopic range using the 2IFC task. Dark grey bars represent the 3D condition and light grey bars represent the 2D conditions. Error bars represent bootstrapped 95% confidence intervals.

4.7. Experiment 4.2.2: Can the scale of the information account for the difference in performance in Experiment 4.1 and 4.2?

Both Experiment 4.1 (section 4.3.) and 4.2 (section 4.5.) show that quantitative depth percepts are obtained from both 1st- and 2nd- order signals in the fused range. In Experiment 4.1, perceived depth was identical for 1st- and 2nd- order diplopic targets and less depth was reported in the uncorrelated condition than the correlated condition in the fused range. Similarly, a difference in stereosensitivity in the correlated and uncorrelated conditions was found using the 2IFC task in Experiment 4.2. In the discussion of Experiment 4.2, this difference was attributed to the relative sensitivities of the 1st- and 2nd- order mechanisms, however there is a difference in the scale of the 1st- and 2nd- order information in these stimuli. That is, the 1st- order mechanism relies on the disparity of the luminance noise which consisted of 30 bands that were 2 arc min wide. 2nd- order stereopsis relies on the envelope of the noise patch, which is 30 times larger (60 arc min). The purpose of the following experiment was to determine if the difference in performance between the correlated and uncorrelated conditions in Experiments 4.1 and 4.2 was due to a difference in scale. To evaluate this, Experiment 4.2 was replicated and a stimulus with 1st order luminance information at the same scale as the 2nd- order envelope was presented.

4.7.1. Experiment 4.2.2: Methods

The 2IFC task described in Experiment 4.2 was used to assess stereoscopic sensitivity. Stimuli consisted of uniform white patches windowed with a raised cosine envelope (Figure 4.8). All testing was conducted at a pedestal of 0 arc min (fused range). If the difference in precision between the correlated and uncorrelated conditions found in Experiment 4.2 was due to the scale of the signal, then stereoscopic sensitivity assessed using this 1st- order stimulus, should be equivalent to that measured in Experiment 4.2 using the uncorrelated noise patches.

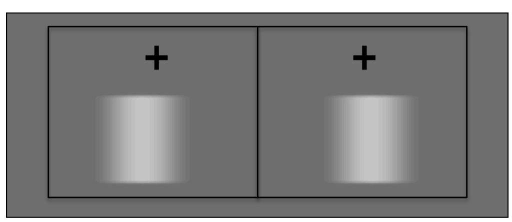


Figure 4.8. Large-scale 'White Patch' correlated noise patch. The scale of the correlated luminance information equals the size of the envelope.

4.7.2. Experiment 4.2: Results and Discussion

Slopes extracted from the original correlated and uncorrelated conditions and the 'White Patch' (large scale) condition are presented in Figure 4.9. The original correlated and uncorrelated noise results were obtained in Experiment 4.2 (section 4.4). The discrimination of metric depth using the Large-Scale Correlated stimulus was most similar to the original correlated condition for all four observers, which is illustrated by the overlapping 95% confidence intervals for 3 out of 4 observers. The similarity of results for the two correlated stimuli despite large differences in the scale of the information, suggests that the shallow slopes in the uncorrelated condition is not due to stimulus scale. Instead the difference in precision in the fused range is most likely due to the different mechanisms that encode the luminance (1^{st}- order) or the envelope (2^{nd}- order) information.

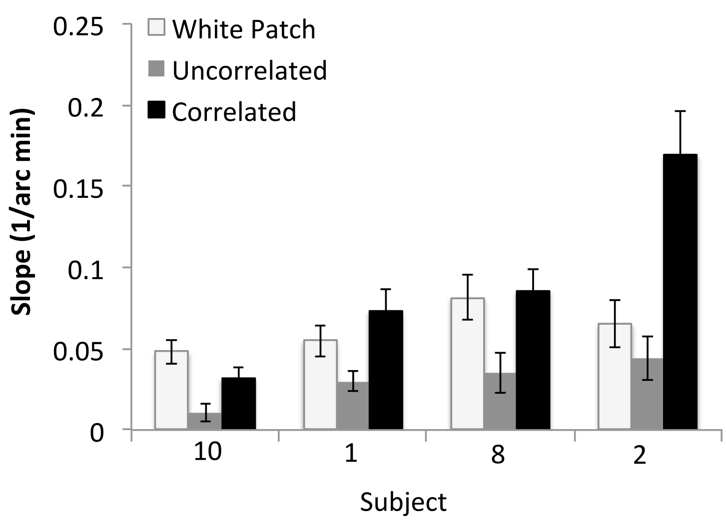


Figure 4.9. Sensitivity is shown here for four observers who participated in Experiment 2. Individual bars show slopes obtained in Experiment 4.2 (Figure 4.6) for the correlated (black) and uncorrelated patches (grey) and for the white patch (white bars) tested here. Error bars represent 95% bootstrapped confidence intervals.

4.8. General Discussion

The purpose of this Chapter was to evaluate the relationship between Ogle's perceptual dichotomy and 1^{st}- /2^{nd}- order stereopsis. To assess this relationship, a depth estimation task and a 2IFC task were used to evaluate depth percepts from 1^{st}- and 2^{nd}- order stimuli. In a series of experiments, Hess and Wilcox (2006, 2008; 1995; 1996) showed that 1^{st}- and 2^{nd}- order stereopsis differ in terms of their spatial and temporal properties. They also showed a difference in stereosensitivity: 1^{st} order stereopsis provides very fine depth percepts while 2^{nd}- order stereopsis provides coarser depth percepts. Furthermore, it is known that quantitative percepts are obtained from correlated (1^{st}- order) stimuli (Blakemore, 1970; McKee & Taylor, 2010; Ogle, 1952b), but to date the metric nature of 2^{nd}- order stereopsis has not been explicitly assessed. Results from the depth estimation experiment (Experiment 4.1) showed that 2^{nd}- order stereopsis does in fact provide quantitative depth in the fused range, however the amount of depth was reduced relative to the depth provided by a 1^{st}- order signal. For most observers, similar amounts of perceived depth were reported at larger disparities (in the overlapping and diplopic range) in the correlated and uncorrelated conditions.

There were two limitations to the estimation task used in Experiment 4.1. One was the presence and use of the separation information, and the other was that only a magnitude estimate is found for a given disparity (not a range of disparities). It would be preferable to use a psychophysical method with more precision that provides a single measure of sensitivity for each disparity range. Therefore, a 2IFC discrimination task was employed in Experiment 4.2 which addressed the concerns from Experiment 4.1 and in addition made it difficult to access and use the separation information at large disparities, as discussed in section 4.6.

Similar to the depth estimation results reported in Experiment 4.1, observers were able to perceive not only the sign, but were also able to discriminate the amount of depth using both the 1st- and 2nd- order signals within the fused range in Experiment 4.2. However, it is clear that patent depth can be extracted in the fused range, estimates/discrimination of depth was more accurate in the correlated condition. For all four observers, discrimination of metric depth at large disparities was similar in both the correlated and uncorrelated conditions. This result suggests that observers rely on the same 2nd- order signal to process depth magnitude beyond the fusion limit, lending further support to the notion that 2nd- order processing mediates stereopsis in the diplopic range. This coincides with results from Wilcox and Hess (1995) where depth percepts close to the upper disparity limit were mediated by 2nd- order stereopsis only.

In addition to exploring the nature of depth magnitude percepts resulting from 1st- and 2nd- order stereopsis, the purpose of this work was to determine how 1st- /2nd- order stereopsis relates to Ogle's perceptual dichotomy. According to Ogle (1952b), patent (quantitative) percepts occur in the fused range and for some range of disparities beyond the diplopia threshold. In both experiments reported here, results in the fused range coincided with Ogle's findings: patent depth was extracted from both 1st- and 2nd- order stimuli within the fused range. The relationship between the spatial dichotomy and Ogle's perceptual categories is less clear in the "slightly diplopic" range. Since Ogle's stimuli consisted of thin lines (needles), observer’s percepts likely transitioned quickly from a single to a double percept. The stimuli used in the experiments reported here were approximately 1 degree wide, so there was a large range of disparities for which the stimuli appeared neither fused nor diplopic. The presence of this intermediate, overlapping zone permitted a closer look at the nature of depth percepts in the transition from fusion to diplopia. As reported by Ogle, quantitative depth percepts were obtained when stimuli appeared fused. The results of Experiments 4.1 and 4.2 show that depth percepts became qualitative at disparities which produced the intermediate percept and likewise, there was no quantitative depth perceived in the diplopic range. Ogle's use of thin stimuli meant that he did not observe this transition. Lugtigheid, Wilcox, Allison, and Howard (2014) presented observers with stereoscopic afterimages (thin lines) and measured depth estimates across a large range of disparities. Depth estimates followed geometric predictions in the fused range and at the first two disparities presented in the diplopic range (up to 1 degree), however beyond 1 degree, depth estimates declined. Therefore, their results coincide with those of Ogle (1952a) but not with the data presented in Experiment 4.1.

The results from the 2IFC task (Experiment 4.2) in the fused range showed a large difference in sensitivity between the correlated and uncorrelated conditions. Since the uncorrelated noise information is unmatched, the visual system is forced to rely on the larger envelope (and therefore larger sized receptive fields), which could account for poorer performance. A follow up experiment (Experiment 4.2.2) was conducted using a large-scale correlated stimulus that did not have the fine luminance information like the correlated noise stimuli that was used in

Experiments 4.1 and 4.2. Therefore, the size of the information was equal to the envelope and would stimulate large receptive fields. Precision of depth estimates in this condition was similar to that seen when correlated noise was used, and significantly different from the uncorrelated condition. This result suggests that the difference in sensitivity seen in Experiment 4.2 is not simply a function of the system processing a different scale of information. Rather, both 1^{st}- and 2^{nd}- order processing signal patent depth in the fused range, but differ in the amount of depth that is perceived, which is reflected in observers' sensitivity.

The data presented here show that patent depth can be extracted by both the linear 1^{st}- order system and the non-linear 2^{nd}- order system in the fused range. Results from these experiments suggest that 2^{nd}- order stereopsis independently mediates depth perception beyond the fusion limit, thus extending the work of Wilcox and Hess (1995). In agreement with Ogle, quantitative depth was obtained in the fused range and qualitative depth was obtained in the diplopic range, however these experiments do not provide evidence of quantitative depth in the diplopic range, as asserted by Ogle (1952b). A remaining issue, which will be discussed in Chapter 6, is the fact that 2^{nd}- order stimuli provide different percepts depending on the disparity range. Another possibility is that the visual system encodes depth using both 1^{st}- and 2^{nd}- order processing in the fused range, but an additional mechanism is present that either mediates depth perception or interacts with 2^{nd}- order processing in the diplopic range.

Chapter 5

The role of stimulus size on the discrimination of metric depth in 2nd-order stimuli

5.1. Introduction

Stereopsis is influenced by the spatial frequency content of a stimulus (among many others, Maffei & Fiorentini, (1977); Marr & Poggio, (1979); Schor & Wood, (1983); Schor, Wood & Ogawa, (1984); DeAngeles, Freedman & Ozhawa, (1995); Wilcox & Hess, (1995). Physiological evidence suggests that there is an inverse relationship between spatial frequency, or receptive field size, and disparity selectivity, often referred to as the size-disparity correlation (DeAngelis et al., 1995; Freeman & Ohzawa, 1990; Marr & Poggio, 1979). For example, a simple cell tuned to low spatial frequencies encodes coarse, or large disparities, while a cell tuned to high spatial frequencies can only reliably localize relatively small disparities. DeAngeles and colleagues (1995) argued that there is a trade-off between the range of disparities that a population of cells can encode and the precision of depth discrimination. Their findings suggest that disparity is encoded by one mechanism along a continuum, where sensitivity varies with spatial frequency. This is in contrast to the discrete tuning types of neurons proposed by Poggio and colleagues (1988, 1995) who argued that different populations of disparity sensitive cells encode specific ranges of disparities.

Psychophysical studies have shown that stimulus spatial frequency affects observer's fusion limits (Felton, Richards, & Smith, 1972; Kulikowski, 1978; Richards & Kaye, 1974; Schor et al., 1984) and perceived depth (Richards & Kaye, 1974; Schor & Wood, 1983). For example, Richards and Kaye (1974) measured perceived depth of bars that were presented at different widths and over a range of disparities. Observers reported an increase in perceived depth as a function of disparity until they reached the upper disparity limit (D_{max}), a maximum point when perceived depth reverted back to zero. They reported that D_{max} scaled gradually with stimulus width: narrower bars elicited less metric depth than wider bars. Later, Schor and Wood (1983) used the method of adjustment to evaluate the effect of spatial frequency on D_{max} by varying the size (width) of DoG stimuli. They found lower D_{max} values when the stimulus was narrow, and D_{max} was high when the stimulus was broad. Both Richards and Kaye (1974) and Schor and Wood (1983) reported that D_{max} increased with stimulus size (width) according to a square root relationship. As discussed in Chapter 1, there were a number of issues with the experimental design in these studies, which makes it difficult to interpret their conclusions.

Wilcox and Hess (1995) assessed the relative contributions of spatial frequency and stimulus size to D_{max}. Using the method of adjustment, Gabor stimuli were initially presented at a near-zero disparity and increased incrementally until D_{max} was reached. In their first experiment, stimulus size was varied using different viewing distances; therefore both the carrier spatial frequency and the size of the Gaussian envelope changed. They found that as viewing distance was increased (which resulted in an increase in spatial frequency and a decrease in envelope size), D_{max} decreased. In their next experiment, the size of the envelope varied while the spatial frequency of the carrier was held constant. Here, they found that D_{max}

increased with increasing envelope size. Conversely, in another experiment reported in this same publication, D_{max} was not affected when the spatial frequency of the carrier varied while the envelope size was held constant. From this work, Wilcox and Hess (1995) suggested that D_{max} is determined solely by the contrast envelope of the stimulus and consequently by 2^{nd} order non-linear filtering.

In a later study, Hess and Wilcox (2008) used a 2AFC task and compared stereosensitivity when observers viewed 1^{st}- and 2^{nd}- order stimuli at a range of exposure durations at smaller (fused) disparities. Their broadband stimuli consisted of correlated (1^{st}- order) and uncorrelated (2^{nd}- order) Gaussian windowed 1D spatial noise patterns, presented at three different spatial scales. For both observers, performance in the correlated condition remained constant as a function of stimulus width, confirming their previous findings (Hess & Wilcox, 2006). However, stereosensitivity was influenced by the size of the envelope in the uncorrelated condition: sensitivity was best when stimuli had narrow Gaussian envelopes, compared to broader envelopes.

Wilcox and Hess have investigated the role of scale on the upper limit for stereopsis (Wilcox & Hess, 1995) and on the smallest resolvable disparity difference (Hess & Wilcox, 2008). However, in many of the experiments reported here, where depth pedestals are used to assess perceived magnitude, disparities fall within the upper and lower bounds of disparity processing. It is possible that stimulus dimensions have a different impact on performance for these moderate suprathreshold disparities. To investigate the effect of varying stimulus size on depth magnitude, a two-interval forced choice discrimination (2IFC) task was used.

Observers viewed uncorrelated noise patches and were asked to discriminate the interval, which contained more relative depth between the reference and test stimuli. The experiment consisted of three blocks of trials, each containing uncorrelated stimuli with one of three different envelope sizes; the carrier information remained unchanged across sessions. In a control experiment (Experiment 5.1.1), observers viewed uncorrelated stimuli with a differently scaled carrier, but the same envelope size.

The hypothesis for Experiment 5.1 was that if a size-disparity correlation exists for 2nd- order processing, then the ability for observers to discriminate metric depth should vary with stimulus width. That is, performance should be best when uncorrelated stimuli are narrow, and performance should be poor when stimuli are wide.

5.2. Experiment 5.1: Methods

Subjects and Apparatus

Four observers participated in this experiment; all had normal or corrected to normal vision. Observers were first assessed using the Randot™ stereo test and were able to appreciate depth from binocular disparities of 40 arc sec or better. Two observers had experience with stereoscopic experiments while the other two observers were naïve. Stimuli were presented on a pair of LCD monitors in a modified Wheatstone mirror stereoscope arrangement with a viewing distance of 57.0 cm. A chin rest was used to maintain a stable head position. The screen resolution was 1920 x 1200 pixels and the screen size was 51.7 x 32.4 cm so each pixel subtended 1.6 arc min. The displays were refreshed at 120.0 Hz and the mean luminance was approximately 46.0 cd/m^2. Linearity of the monitors was assessed using a luminance meter (Konica Minolta, LS-110, 78423013), and was calibrated by gamma correcting nonlinearities in the luminance output response. The room was dark except for light from the stereoscopic display.

Stimuli

Stimuli consisted of one-dimensional luminance noise windowed horizontally with a raised-cosine. Only uncorrelated (unmatched) stimuli were used in this experiment. Noise samples were re-generated on every trial so that no two noise patches in the same trial or consecutive trials consisted of the same noise samples. Noise patterns were created by randomly choosing one of 256 grey levels for each band of the image. Noise samples were balanced about the background grey level so that they averaged to mean luminance. Prior to presentation, each stimulus was assessed to ensure that the luminance of the bands at the edges did not equal mean luminance. This ensured that the edges would not blend into the background so that the width was held constant.

Trials were blocked in three sessions, each consisting of a different stimulus size. The width (size) of the stimulus was constant within a single session. Three envelope widths were tested: 20, 60 and 180 arc min (Figure 5.1). The luminance information within the carrier was chosen randomly. Therefore, the only difference between the three sessions was the size of the raised-cosine envelope, therefore any differences in performance can be attributed to stimulus size, not carrier spatial frequency.

Figure 5.1. Three different envelope sizes were presented in different sessions: (a) 20 arc min, (b) 60 arc min, (c) 180 arc min.

Procedure

Initially a pre-test was conducted to measure each observer's diplopia threshold. As in the experiments described in Chapters 4, a forced choice task was used to identify three distinct perceptual regions: fused, overlapping and diplopic. This was repeated for each stimulus width. A second pre-test was conducted to determine the target JND for each condition. Using the method of constant stimuli, observers indicated whether two stimuli appeared in the same or different depth planes. Psychometric functions were fit with a Weibull function and the 75% JND was extracted. The target disparity used in the main experiment was set at three times the target JND determined in the second pre-test.

For trials within the pre-tests and the main experiment, a nonius reference cross was presented at the plane of the screen; one eye viewed the top and the other eye viewed the bottom half of the cross. Observers indicated when the two halves of the fixation cross were aligned, and the stimuli were displayed. Stimulus half images viewed by the two eyes were offset in equal and opposite directions by the amount required by each test condition. A fusion guide consisting of circles displayed on the top and bottom of the screen were positioned at the plane of the screen. The pattern of circles was updated from trial to trial so they provided no consistent location cue. The stimuli were presented within a horizontal gap (1.2 x 7.6 deg) along the middle of the screen that did not contain the pattern of circles.

Experiment 5.1 consisted of the 2IFC task described in Chapter 4. To summarize, observers were presented with two intervals and were asked to select the interval with more relative depth between a reference and target stimulus. Here, the reference was always fixed at zero disparity and was presented in both intervals either above or below fixation. In one interval a target was positioned at a fixed disparity determined in the second pre-test. In the other interval, the test stimulus was presented at one of 19 different disparities. Each test disparity was randomly presented ten times in either the first or second interval. Performance was assessed by fitting each observer's psychometric data with a Weibull function and the variance was estimated by a bootstrapping procedure (Wichmann & Hill, 2001a, 2001b) run 100 times for each dataset.

5.3. Experiment 5.1: Results

Figure 5.2 shows stereosensitivity results for four subjects (1, 2, 6, 7). These data show that perceived depth was influenced by the envelope size of the uncorrelated stimulus. An ANOVA showed that there was a significant main effect for condition, ($F(2,9)=8.196$, $p<0.01$, partial $\eta^2=0.645$, Power=0.868). A Bonferroni corrected post-hoc analysis indicated a significant difference only between the smallest (20 arc min) and largest (180 arc min) condition ($p<0.01$). On average, observers performed two times better when stimuli were narrow (20 arc min) compared to the widest stimulus (180 arc min). The slope obtained when the envelope size was intermediate (60 arc min) was similar to the smallest size (20 arc min) condition for 2 observers (1 and 7), and it was similar to the largest size (180 arc min) condition for the other 2 observers (2 and 6).

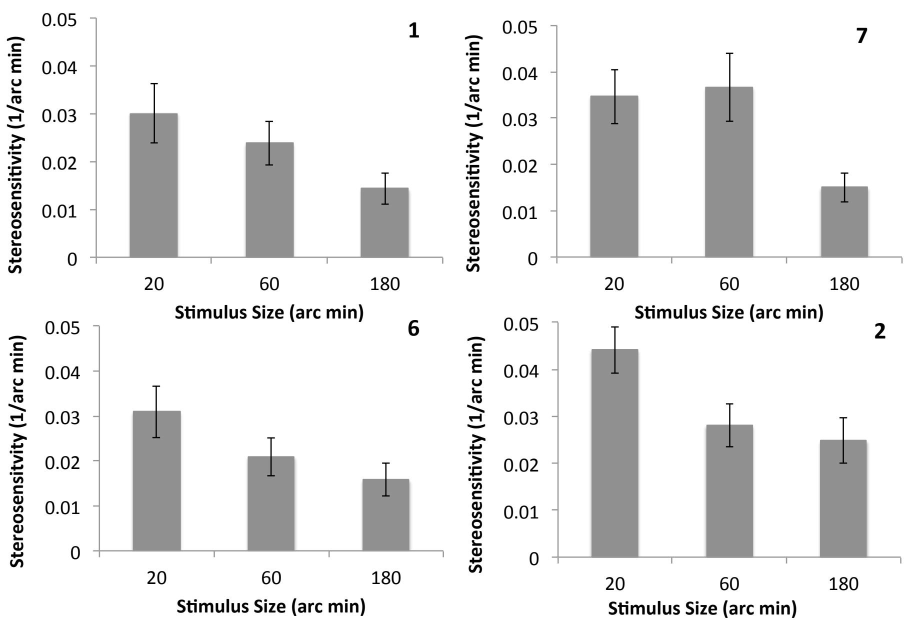


Figure 5.2. The results of Experiment 5.1 showing stereosensitivity as a function of stimulus size for four observers. Error bars represent bootstrapped 95% confidence intervals.

5.4. Experiment 5.1.1: Varying the scale of the carrier

In Experiment 5.1, the results indicated that envelope size influenced performance on the 2IFC task. The purpose of this control experiment was to ensure that *only* the envelope size influences the discrimination of metric depth. In Experiment 5.1.1, the scale of the carrier information was varied while the envelope size was held constant. The hypothesis for Experiment 5.1.1 was that performance should not change when the carrier information is varied.

5.4.1. Experiment 5.1.1: Methods

In a control experiment, a 2IFC task was used and as in the main experiment, the stimuli consisted of uncorrelated noise patches. Here, the scale of the carrier information was varied while the envelope size was fixed. As in the main experiment, the test disparity was randomly presented in either the first or second interval and each test disparity was presented 10 times. Performance was assessed by fitting each observer's psychometric data with a Weibull function and the variance was estimated by a bootstrapping procedure (Wichmann & Hill, 2001a, 2001b) run 100 times for each dataset.

5.4.2. Experiment 5.1.1: Results

Stereosensitivity (slope) as a function of band thickness (arc min) for four observers is plotted in Figure 5.3. The same four observers (1, 2, 6 and 7) who participated in Experiment 5.1 were re-tested here. An ANOVA indicated that performance was identical across the different conditions for all four observers, $F(4,15)=1.29$, $p>0.05$). One observer performed poorly on the thickest band condition compared to the other conditions ($p>0.01$). In general, these data show that when viewing uncorrelated stimuli in the fused range, the perception of metric depth is dependent on the scale of the envelope, and is not influenced by scale of the information within the carrier.

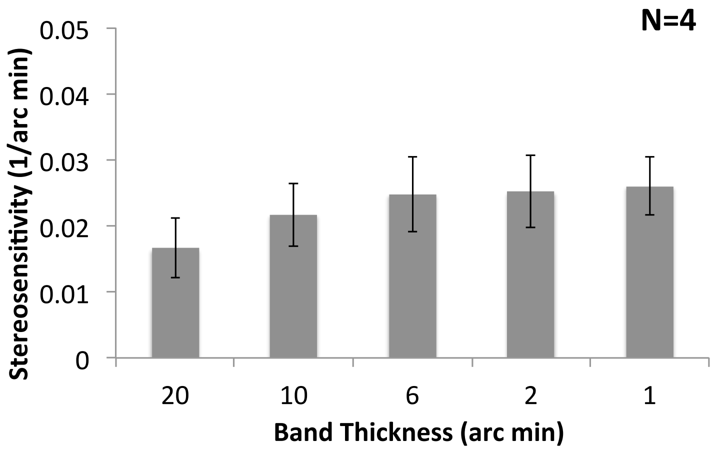


Figure 5.3. The results of Experiment 5.1.1 showing stereosensitivity as a function of the carrier's scale of information for 4 observers. Error bars represent bootstrapped 95% confidence intervals.

5.5. Discussion

When presented with fused uncorrelated stimuli, observers relied on the size of the envelope, not the information within the carrier, to successfully discern the interval with more relative depth. When the size of the envelope was varied, observers performed best when uncorrelated stimuli were narrow (20 arc min). To confirm that the depth from uncorrelated 2nd order stimuli is dependent only on the contrast envelope, a second set of conditions was tested in which the envelope size was fixed, and the noise carrier band thickness was varied. In general, results showed that varying the carrier information did not influence the discrimination of metric depth.

Previous researchers have argued that D_{max} is dependent on the spatial frequency of the stimulus (Schor & Wood, 1983) and stereoscopic processing has been shown to involve an early filtering stage that involves linear spatial frequency detectors (Marr & Poggio, 1979). Similar to the findings reported by Hess and Wilcox (1994; 1995), the results from these experiments support the existence of a non-linear mechanism that extracts depth information from the envelope instead of the spatial frequency of the carrier. When the carrier information is unavailable or unreliable, such as in the case of binocularly uncorrelated stimuli, the visual system is forced to use the 2nd order envelope signal to extract depth information. These results are consistent with the filter-rectify-filter model (as presented by Wilson, Ferrera & Yo (1992) as well as Zhou & Baker (1993), which can be used to describe how the 2nd order signal is extracted and then used by the visual system.

Physiological studies have suggested that populations of spatial frequency tuned neurons encode depth across a wide range of disparities. Cells that are tuned to low spatial frequencies encode coarse disparities, while cells that are tuned to high spatial frequencies process fine disparities (DeAngelis et al., 1995; Marr & Poggio, 1979) and it has been shown that the size of RFs is inversely related to spatial frequency (Maffei & Fiorentini, 1977). Many studies have found evidence for this relationship, for example, Schor and Wood (1983); Schor et al. (1984) and Smallman and Macleod (1994) have shown that disparity limits for stereopsis decrease with spatial frequency. Schor et al. (1984) showed that the fusion limit is influenced directly by the spatial frequency of the stimulus, a result previously reported by Kulikowski (1978).. The size-disparity correlation also applies to broadband stimuli, such as bars, which contain many spatial frequencies. It is assumed that depending on the disparity presented, a spatial frequency tuned detector is stimulated by the relevant spatial frequency within the stimulus.

Richards and Kaye (1974) showed that perceived depth was affected by the size of a bar stimulus. However, there are behavioral experiments that have reported conflicting evidence for this size-disparity correlation, such as results reported by Mayhew and Frisby (1979a, 1979b). Mayhew and Frisby (1979a) presented textured stimuli consisting of random dot patterns or 2D spatial frequency filtered random dot patterns. Stimuli were presented at a range of disparities (6 to 20.8 arc min) and perceived depth was measured by asking observers to adjust two lines, one to the depth of the surround and the other to the depth of the centre. They found that the spectral content of the stereograms did not affect perceived depth as a function of disparity. Mayhew and Frisby (1979a) concluded that spatial frequency tuned channels do not differentially encode disparities on a continuum. The results found in the experiment reported here provide evidence for a size-disparity correlation existing on a coarser scale for 2nd order processing. That is, discrimination of metric depth is best when 2nd order uncorrelated stimuli are narrow and performance is unaffected by the type of luminance information within the carrier.

The experiments described in Chapter 4 confirmed that quantitative depth percepts are obtained from 2nd order stimuli. The depth estimation task (Experiment 4.1) showed that less metric depth was reported in the 2nd order condition compared to the 1st- order condition. The 2IFC task (Experiment 4.2) showed that lower stereosensitivity is obtained in the 2nd- order condition than the 1st- order condition. Previous researchers have shown that stereopsis and rivalry can co-exist (Treisman, 1962, Mayhew & Frisby, 1976; Blake, Yang & Wilson, 1991;

Buckthought & Wilson, 2007; Su, He, & Ooi, 2009). By varying the thickness of the bands in the control experiment described in the current chapter, the amount of potential of rivalry is varied as well. For example, when the bands were wide (width=20 arc min), only 3 bands are generated in each half image of the noise patch. As mentioned in the Results section, when the band thickness was 20 arc min, LW performed significantly worse compared to the other conditions, and it is possible that rivalry disrupted stereopsis in this session.

The purpose of this experiment was to investigate the role of stimulus size on the discrimination of metric depth in 2nd- order stimuli. The scale of the envelope was shown to affect D_{max} (Wilcox and Hess, 1995) and stereo thresholds (Wilcox & Hess, 2008) and this experiment shows that the scale of the envelope also influences the discrimination of depth magnitude.

Chapter 6

Binocular Disparity vs. Monoptic depth at Large Disparities

6.1. Introduction.

As a result of the publications by Wheatstone (1838) and Panum (1858), it is generally accepted that single vision results from objects stimulating corresponding as well as non-corresponding points on the retina. Within the region of single vision, known as 'Panum's fusional area', as retinal disparity increases so does relative perceived depth. Disparate points beyond this region will result in a double, or diplopic, percept. Although there is a substantial difference in the appearance of stimuli that are fused vs. diplopic, it is well established that depth percepts are obtained from both fused and diplopic targets (Blakemore, 1970; Foley et al., 1975; Helmholtz, 1910; Hering, 1942; Mitchell, 1966, 1969; Ogle, 1952b, 1953; Reading, 1970; Richards & Foley, 1974; Westheimer & Tanzman, 1956; Wilcox & Hess, 1995; Ziegler & Hess, 1997).

As described in previous chapters, Ogle (1952b) examined the *quality* of depth percepts across a range of fused and diplopic disparities. Patent, or quantitative, depth resulted from disparities where the target appeared fused or slightly diplopic. Beyond the 'patent limit', observers did not perceive metric depth, rather only qualitative (sign) information was perceived. Ogle's investigations (1952b, 1953) represent the first attempt to categorize depth percepts based on the disparities presented.

The dichotomy proposed by Hess and Wilcox (1994) is based on the spatial characteristics of the stimuli with 1st- order processing using the luminance- based features (bars, edges) to extract depth. As shown in the Experiments reported in this dissertation and by Wilcox and Hess (1994, 2006, 2008; 1995; 1996, 1997), 2^{nd}-order stereopsis is used to provide depth information for a wide range of fused and diplopic disparities when the stimuli are not well matched in the two eyes.

Depth can also be encoded by a mechanism that relies on 'monoptic depth', a term established by Wilcox et al. (2007). Hering (1861) was the first to describe the relationship between the position of a monocular image on the retina and it's perceived depth. He proposed that when the image of a monocular target falls on the temporal side of the retina, it would be perceived in front of the fixation plane. When the image falls on the nasal side of the retina, it would appear behind the fixation plane. This is illustrated in Figure 6.1. Hering also predicted that the amount of perceived depth of a target relative to the fixation plane would increase with distance from the fovea. Helmholtz opposed Hering's findings, but later studies by Kaye (1978) and by Wilcox and colleagues (2009; 2007) showed empirically that depth is extracted from monocular targets and confirmed the relationship between retinal position and depth sign as predicted by Hering.

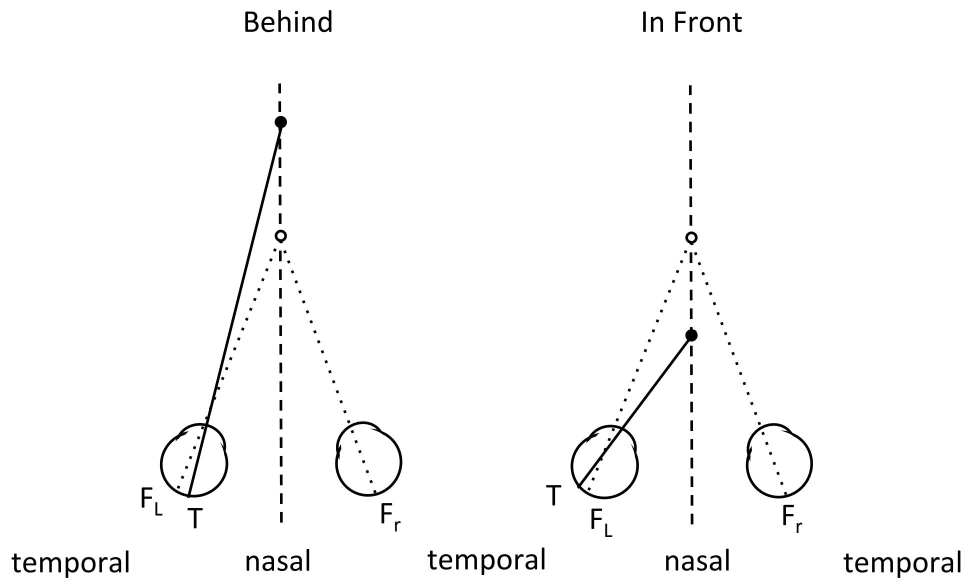


Figure 6.1. The geometry of monoptic depth. F_L and F_r represent the retinal locations corresponding to the fixation point (white circle), and T corresponds to the retinal location of the monoptic target (black circle). A monoptic image on the temporal side of the retina (left) creates the percept of an object behind the fixation plane. When the monoptic image is on the nasal side of the retina (right), the object appears in front of the fixation plane.

Kaye (1978) evaluated the apparent depth reported by observers in binocular and monocular conditions. Stimuli consisted of patterns of o (circle) or x (lines at 45° and 135°) and were flashed for 100 msec. Depth magnitude was measured using a percentage scale based on the depth of the target relative to the fixation plane. Kaye's results were consistent with the predictions made by Hering (1861). Only three eccentricities were presented to each eye, and there was not an increase in metric depth as a function of distance from the midline over the range of eccentricities tested.

Wilcox and colleagues (2007) measured percent consistent when observers viewed stereoscopic and monoptic targets at a range of offsets (0 to 60 arc min, and up to 120 arc min for one observer). Responses were scored as 'consistent' if they followed the pattern originally predicted by Hering (1861): targets were labeled in front when they were presented in the temporal retina, and were considered behind when they stimulated the nasal retina. They measured percent consistent by asking observers to indicate if the stimulus (thin white bar) appeared in front or behind the fixation plane. They found an increase in percent consistent as a function of eccentricity up to a certain point, followed by a plateau at around 80-100% in the stereoscopic condition. In the monoptic condition, observers performed at chance for small disparities and at 70% at larger disparities (>7 arc min). In a follow-up experiment, Wilcox and colleagues (2007) patched the observer's unstimulated eye to determine if monoptic depth was purely a monocular phenomenon. In this condition, performance fell to chance, confirming that monoptic depth requires binocular stimulation. They argued that monoptic depth involves a coarse match of the line of sight to the fovea in the unstimulated eye. Therefore, Wilcox and colleagues (2007) confirmed Hering's original proposal and in addition demonstrated that monoptic depth is a binocular sensation that is restricted to large offsets.

Both stereopsis and monoptic depth encode depth at large disparities. When the fixation is at the midline and the half images of a target are symmetrically positioned on either side of the fixation point, it is possible for an observer to use one of the half images to perceive depth relative to fixation. Both the geometry of stereopsis and the geometry for monoptic depth (Figure 6.1) would predict that the stimulus would appear 'in front'. The conclusions drawn from Chapters 2 and 4 where depth from diplopic targets is assessed, rests on the assumption that the depth percepts rely on binocular stereopsis (2nd order). Therefore it is important that we verify that percepts in these experiments do not reflect monoptic depth. To evaluate these two possibilities, observers viewed stimuli that contain 1) disparity-only, 2) both disparity and monoptic information and 3) monoptic-only. If observers rely on monoptic depth only, percent consistent in conditions 2 and 3 should be identical and performance should be better than the disparity-only condition (1).

6.2. Experiment 6.1: Methods

Subjects and Apparatus

Stereoacuity was assessed with the Adult Randot™ stereotest to ensure that observers could detect depth from binocular disparities of at least 40 arc sec. Four observers who participated in this experiment had normal or corrected-to-normal vision. Two observers had experience with monoptic tasks (LW and IT), one was experienced in stereoscopic tasks (DS) and the fourth (BH) was naïve. The stimuli were presented on a Wheatstone mirror stereoscope with a viewing distance of 57 cm and a chin rest was used to maintain a stable head position. The screen resolution was 1280 x 960 pixels and the screen size was 40 x 30 cm. The displays were refreshed at 120 Hz and their mean luminance was approximately 69 cd/m^2. Linearity of the monitors was assessed using a luminance meter (Konica Minolta, LS-110, 78423013), and then calibrated by gamma correcting nonlinearities in the luminance output response.

Stimuli

Similar to experiments reported in Chapters 4 and 5, stimuli in Experiment 6.1 consisted of one-dimensional luminance noise windowed horizontally with a raised-cosine. Noise samples (dimensions were 20 x 60 arc min) were uncorrelated and were re-generated on every trial so that no two noise patches were presented in the same trial or consecutive trials. Noise patterns were created by randomly choosing one of 256 grey levels for each vertical column of pixels within the image. The uncorrelated noise samples were balanced about the background grey level so that they averaged to mean luminance. Prior to presentation, each stimulus was assessed to ensure that the bands at the edges were greater than 1 standard deviation above or below mean luminance. Therefore, the edges did not blend into the background, which ensured that the width of every noise sample appeared the same.

A fusion guide was displayed above and below fixation and consisted of an array of circles positioned at the screen plane. The pattern of circles was updated from trial to trial so they provided no consistent location cue. A horizontal gap (1.2 x 7.6 deg) along the middle of the screen where the stimulus appeared did not contain the pattern of circles. Directly above and below a central stimulus, two small lines (dimensions were 4.0 x 20.0 arc min) were positioned at zero disparity, which served as a zero disparity reference. Each line was positioned 30.0 arc min above and below the location that the central stimulus would appear. These lines were not present once the stimulus appeared.

Four conditions were tested, as outlined in Figure 6.2. In the disparity only condition (1) the stereo-pair was diplopic, and the image presented to one of the two eyes was (randomly) positioned at the centre of the display. Given that monoptic depth makes no contribution to perceived depth at fixation, this arrangement isolates stereopsis. On each trial observers were asked to judge the position of the central patch relative to zero disparity, which was the location defined by the fusion guide. There was no reference patch adjacent to the stimulus, making this a difficult task, but all observers could reliably make the required judgment (see also Wilcox et al., 2007).

In condition 2 (disparity+monoptic), the same stimulus arrangement was used, and observers maintained central fixation, but were asked to report the depth of the peripheral patch. Note that in this case both stereopsis and monoptic depth are available to help perform the task. The brief exposure duration (100 msec) ensured that observers were unable to converge on the peripheral target.

In condition 3 (monoptic-only) both noise patches were presented to one eye. Observers indicated the position of the peripheral target. In condition 4 (Monoptic-centre), the configuration was the same as that used in condition 3 but observers indicated the depth sign of a central target. This condition was intended to replicate Wilcox and colleagues' (2007) result and confirm that depth is not extracted from monocular elements presented at the fovea, an assumption relied upon in the design of this experiment.

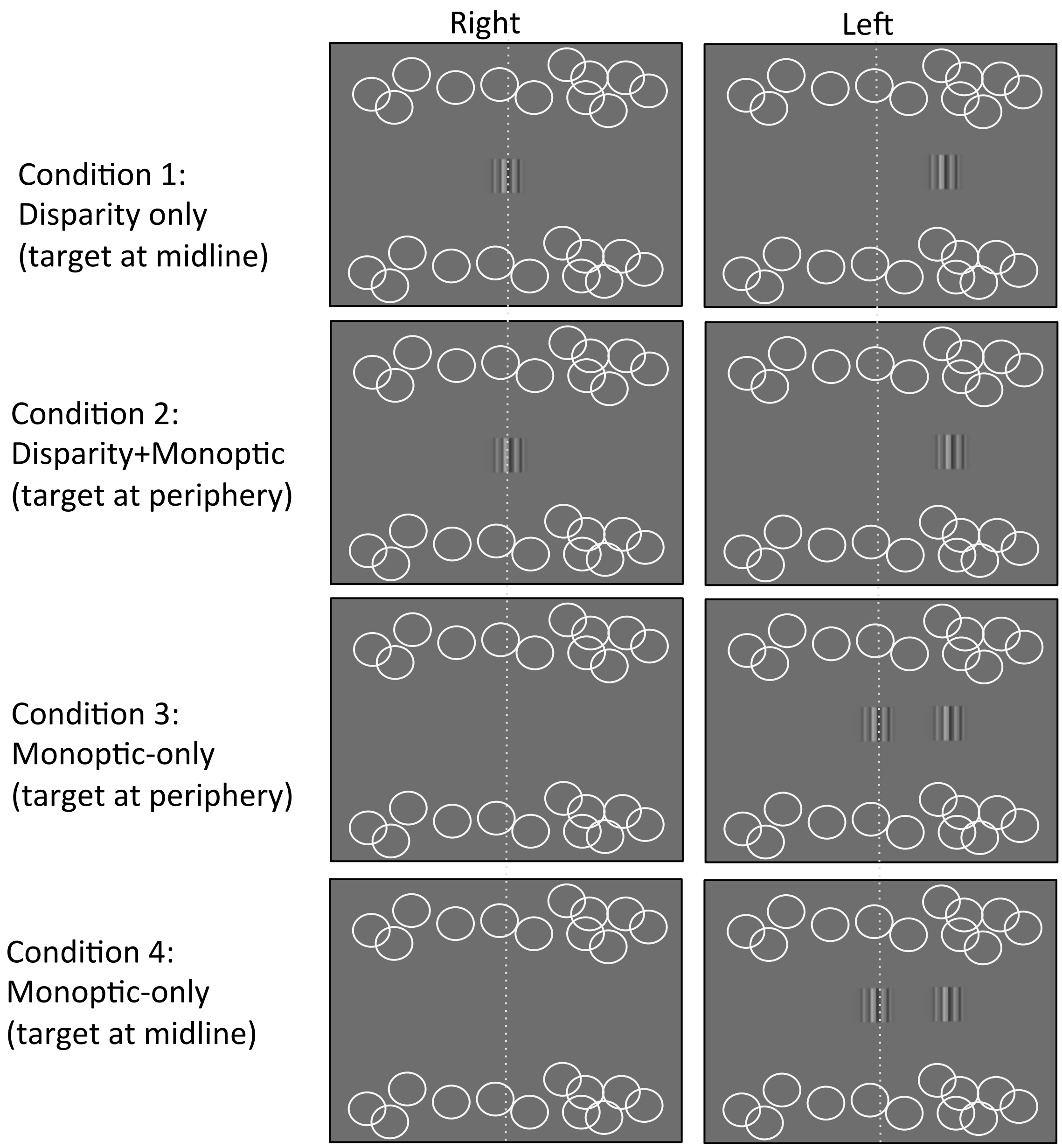


Figure 6.2. Diagram illustrating the four test conditions. Dotted vertical lines are presented here to assist with fusion, but did not appear in the experiment.

Procedure

Observers initially viewed two thin binocular lines that served as a central fixation guide. Observers were asked to click a button on the game pad to initiate a trial. Following a 50 msec delay, the lines disappeared and the stimuli were presented for 100 msec. A 2AFC task was used and observers indicated if the target was located in front or behind the fixation plane. Each condition was tested in a separate block of trials.

Crossed and uncrossed disparities, or equivalent monoptic offsets, were randomly interleaved across trials within a session. The test disparities were presented in random order, and observers viewed 20 repetitions and a total of 12 disparities (6 crossed and 6 uncrossed offsets). Six offsets (80, 100, 120, 140, 160, 180 arc min) were presented in each block.

Data Analysis

The two binocular (stereoscopic) conditions were scored in the conventional way. For example, a response was labeled 'consistent' (correct) when the target had a crossed disparity and the observer reported that it was 'in front', or when the target was uncrossed and the observer indicated that it was 'behind'. The analysis of the monocular conditions depended on whether the target stimulated the nasal or temporal retina. A response was consistent if nasal stimulation produced a far response or if temporal stimulation produced a near response (see Kaye (1978) and Wilcox et al. (2007)). This relationship is illustrated in Figure 6.1.

6.3. Experiment 6.1: Results and Discussion

Figure 6.3 shows the results from the four conditions tested in Experiment 6.1. Percent consistent was recorded for each condition. For the monocular conditions, percent consistent was recorded for each eye separately and then evaluated in terms of the location (nasal/temporal) on the retina. The results for each condition are averaged across the four observers and data was collapsed across both eyes and both directions of depth (Figure 6.3). An analysis of variance (ANOVA) was conducted to evaluate the effect of condition and disparity. A significant main effect for condition was found ($F(3)= 224.539$, $p<0.01$, partial $\eta^2 =0.14$, Power=1.0), but not for disparity ($F(5)=1.015$, $p>0.01$) or the interaction of condition x disparity ($F(15)=0.845$, $p>0.01$). A Bonferroni post-hoc analysis indicated that for all observers, performance in the two binocular conditions (1 and 2) were identical ($p>0.01$). Therefore, it is clear that when both disparity and monoptic elements are present, observers rely on the stereoscopic signal.

The post-hoc analysis indicated that all other comparisons (conditions 1 vs. 3, 1 vs. 4, 3 vs. 4) were significantly different from one another ($p<0.01$), as shown in Figure 6.3. These data were also analyzed on an individual basis because there was individual variability in observer's ability to discern depth from monoptic elements. For the monoptic-only condition (3), two observers performed at about 70% across all the offsets tested in the monoptic-only (3) condition, thus confirming previous reports of monoptic performance (Wilcox et al., 2007). Performance on this condition was significantly different from performance on the other three conditions ($p<0.01$). These results suggest that monocular targets can provide a reliable depth signal, although it is not as reliable as depth from disparity. The two remaining observers performed at about 50% in the monoptic only condition, which was not significantly different from condition 4 ($p>0.01$). These results show that there is variability among individuals in the ability to extract depth from the monoptic signal. Nonetheless, the fact that these two observers were not able to accurately perceive depth from the monoptic target but were still able to achieve nearly 100% correct in the stereoscopic conditions shows that different mechanisms underlie depth from disparity and depth from monoptic elements.

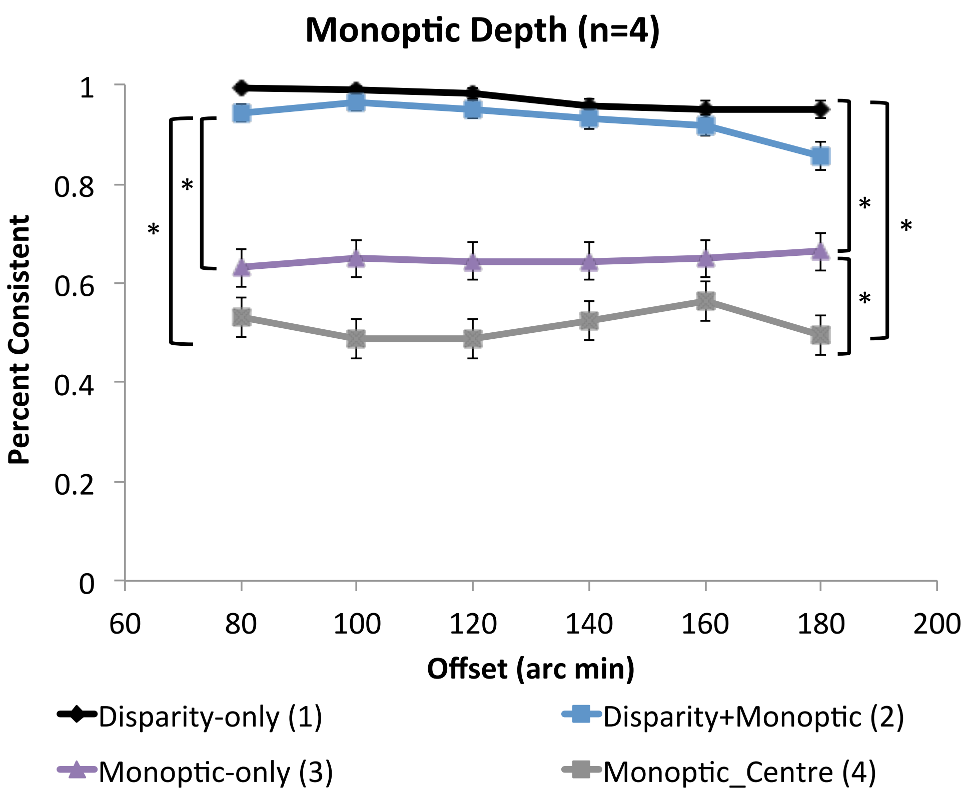


Figure 6.3. Average results for four observers. The black, blue, purple and grey curves represent the disparity only (1), Disparity+monoptic (2), Monoptic-only (3), Monoptic-centre (4) conditions respectively. The error bars represent standard error of the mean. Significantly different comparisons are marked by an asterisk.

The purpose for including Condition 4 was to confirm Wilcox and colleagues' (2007) results when the target is at, or close, to the midline. They found that monoptic depth is not available at the midline, and is only available beyond 7 arc min. Here, all four observers performed at chance when the monocular target was at the midline. This confirms Wilcox and colleagues results and provides support for the underlying assumption necessary in the interpretation of conditions 1 and 2. Monoptic depth is not available in condition 1 when the target is at the midline, and if observers used monoptic depth in this condition, percent consistent should be at chance. Since performance was nearly 100% consistent, observers were using a stereoscopic mechanism. In condition 2, both stereopsis *and* monoptic depth are available because the stimulus configuration was stereoscopic and observers judged the relative position of the peripheral target which provides a monoptic depth signal. As both types of information are available, the visual system extracts a depth signal using the more accurate and reliable cue, stereoscopic depth.

Wilcox and colleagues (2007) demonstrated that monoptic depth is a binocular phenomenon and cannot be explained by a monocular local sign hypothesis. They argued that the visual system makes a coarse match between the line of sight for the target and the line of sight for the fovea in the unstimulated eye. Matching to the fovea would require steady vergence, which is affected by fatigue, eyestrain and experience. This may explain the large variability in observer's ability to perceive monoptic depth as well as the variability across conditions (for observers who do perceive monoptic depth). The variability was higher in the monoptic condition than the binocular condition, which is another indication that a different mechanism underlies depth percepts in conditions 1 and 2 vs. condition 3.

Over the range of disparities in the diplopic range presented in this and in previous chapters, both stereopsis and monoptic depth could be available. Experiment 6.1 shows that when both sources of information are available, the disparity signal is used, and as shown in Chapters 2 and 4, this is the 2nd-order mechanism. It is clear that all observers perform better when stereopsis is present, whereas performance is poor and variable when only monoptic depth is present. The data reported in this chapter indicates that depth percepts from stereoscopic and monoptic stimuli are distinct, suggesting that they are encoded by different mechanisms. Further research is needed to understand the relationship between 2nd-order stereopsis and monoptic depth for diplopic targets.

Chapter 7

Discussion

7.1. The relationship between the proposed dichotomies for stereopsis

There is extensive evidence that stereopsis is subserved by two distinct mechanisms (Hess & Wilcox, 1994; Kovács & Fehér, 1997; Langley et al., 1999; Lin & Wilson, 1995; McKee et al., 2004, 2005; Poggio & Talbot, 1981; Sato, 1983; Wilcox & Hess, 1995; Wilcox & Hess, 1996, 1997, 1998). In the past investigators have classified stereopsis based on temporal (Edwards et al., 1998, 1999, 2000), spatial (Hess & Wilcox, 1994; Wilcox & Hess, 1995; Wilcox & Hess, 1996, 1997) and perceptual (Ogle, 1952b) properties. However, because of the use of different types of stimuli, experimental paradigms and disparity ranges, direct comparisons are impossible. The goal of this dissertation was to better understand the mechanisms that underlie stereopsis, and in doing so consolidate the existing literature. The relationship between the spatial dichotomy, the temporal dichotomy (proposed by Edwards and colleagues) and the perceptual dichotomy (proposed by Ogle) was examined. To achieve this, depth percepts across a large range of disparities were examined using depth discrimination and depth estimation tasks.

In Chapter 2, the spatial and temporal dichotomies were compared at large disparities. Previously, Edwards and colleagues (2000) assessed depth discrimination at one large depth offset (1 deg) and at one exposure duration (200 msec). Since similar results were found in both a 1st- order and a 2nd- order condition, they argued that sustained/transient stereopsis underlies both types of stereoscopic depth perception. A more recent study by Wilcox and Hess (2008) assessed stereosensitivity at near-threshold disparities and exposure durations ranging from 80 msec to 1.3 sec. They found different temporal response profiles in a 1st- order and 2nd- order condition. The purpose of Chapter 2 was to determine the temporal response of 1st- and 2nd- order stereopsis at large disparities, comparable to those used by Edwards and colleagues. Stereosensitivity as a function of exposure duration was similar in the both conditions and was quite poor. These results are similar to those obtained by Edwards and colleagues (2000), however taken together with the results from later Chapters (discussed below) and from Wilcox and Hess (1995), the results from Chapter 2 should be interpreted differently. Edwards and colleagues (2000) argued that sustained/transient mechanisms underlie stereopsis of both 1st- and 2nd- order stimuli at small and large disparities respectively. However, since stimuli were presented at large disparities, they did not appear completely, fused and depth in both conditions was most likely encoded by a 2nd- order mechanism. Therefore, Experiment 2.1 demonstrated that 2nd- order stereopsis facilitates depth perception at large disparities and depth percepts cannot be distinguished by their temporal properties alone.

The purpose of Chapter 3 was to consolidate previous research (Badcock & Schor, 1985; Blakemore, 1970; Ogle, 1952b) and evaluate Ogle's perceptual dichotomy using an appropriate method for evaluating depth at large disparities (methodology evaluation discussed in detail in section 7.2.1). Studies that have evaluated this perceptual dichotomy did not sample a sufficient range of disparities; rather, mostly large disparities were presented (Blakemore, 1970; Ogle, 1952b, 1953). Also, these studies (Badcock & Schor, 1985; McKee et al., 1990; Ogle, 1952b, 1953) did not measure the diplopia threshold. In Chapter 3, a broadband line stimulus was presented to observers using both a depth discrimination and depth magnitude task, and an increase in perceived depth was found within Panum's fusional area. At large disparities, stereosensitivity and perceived depth were constant, suggesting that only qualitative percepts occur within this range. However assessing stereoacuity or perceived depth beyond the fusion limit was problematic because responses were confounded by the visibility of the lateral separation of the half images (see section 7.2.1).

The relationship between 1st- and 2nd- order stereopsis and Ogle's perceptual classification were examined in Chapter 4. The nature of the depth percept provided by stimuli that are processed by 1st- or 2nd- order mechanisms was evaluated. Depth magnitude percepts were measured using a depth estimation task (Experiment 4.1) and a two-interval forced choice discrimination (2IFC) task (Experiment 4.2). In the depth estimation experiment, quantitative depth estimates were obtained in the fused range in both the correlated (1st- and 2nd- order information) and the uncorrelated (2nd- order only) conditions. A slightly greater amount of depth was reported in the correlated condition compared to the uncorrelated condition. Perceived depth at large disparities was relatively constant, therefore the transition from patent to qualitative found in Chapter 4 did not correspond well with Ogle's results. In Experiment 4.2, a discrimination task was used to quantify perceived depth reported in a correlated and uncorrelated condition. A two-interval depth discrimination (2IFC) paradigm was used where observers selected the interval that contained a larger amount of depth between the reference and test stimuli. In the fused range, performance in the correlated condition was significantly better than performance in the uncorrelated condition. At large disparities, performance in the correlated and uncorrelated conditions was similar. Therefore, Chapter 4 demonstrated that both 1st- and 2nd- order stereopsis encode quantitative depth in the fused range, while only 2nd- order stereopsis mediates depth beyond the fusion limit and only sign information is extracted.

The difference in sensitivity in the correlated and uncorrelated conditions in Chapter 4 most likely reflects the stimulation of two different mechanisms, however another possible explanation for this difference is the scale of the information in 1st- and 2nd- order stimuli. To assess this possibility, a follow-up experiment (Experiment 4.2.2) was conducted where observers were presented with a large-scale correlated stimulus in which the scale of the luminance information was equal to the size of the envelope. Since performance on the large-scale condition was more similar to performance in the original correlated condition, the difference in sensitivity seen in Chapter 4 is not simply a function of the system processing a different scale of information. Rather 1st- and 2nd- order stereopsis signal different degrees of depth magnitude, which can be explained in the context of the size disparity correlation. It is possible that neurons that respond to 2nd- order stimuli have larger receptive fields, and as a result, depth is encoded less precisely and observers need a larger relative depth difference to discriminate changes in metric depth.

Rivalry is another possible explanation for the difference in sensitivity found in Chapter 4. Although the interaction between rivalry and stereopsis is a controversial topic, many investigators have found evidence that observers perceive depth from stereopsis in the presence of rivalry (Mayhew & Frisby, 1976; Ogle & Wakefield, 1967; Treisman, 1962) and they can coexist in different spatial frequency and orientation bands, but not within the same band (Buckthought & Wilson, 2007). The uncorrelated stimuli used in the Experiment 4.1 (depth estimation) consisted of one vertically oriented noise patch and one horizontally oriented noise patch. Since these unmatched images were so dissimilar it is possible that they were rivalrous. A follow-up study (appendix B) was conducted in which observers were presented with uncorrelated stimuli that were vertically oriented in both eyes. Observers did not report rivalry but instead they likely matched similar neighbouring luminance bands resulting in a percept of a volume of noise. If rivalry influenced depth percepts in the horizontal/vertical stimulus, then estimates would likely be lower compared to the vertical/vertical stimulus condition. However, observers performed identically. The vertical/horizontal version of the uncorrelated stimulus was used in the experiment to avoid the presence of any potential luminance artefacts that could be used by the 1^{st}- order system. In the control experiment reported in Chapter 5 (Experiment 5.1.1), the width of the bands within the carrier of an uncorrelated stimulus was varied, which inevitably varied the degree of rivalry. For all four observers, there was no difference in performance for four out of the five band thicknesses tested. When the band thickness was largest (only three bands were visible in each eye), stereopsis was unaffected for three out of four observers. Although testing the effects of rivalry was not the purpose of that experiment, the fact that there was no effect of varying the carrier information demonstrates the resilience of 2^{nd}- order processing in spite of potentially rivalrous luminance information. Therefore the 2^{nd}- order contrast envelope provides reliable depth information and, in general, is not affected by the information within the carrier.

The purpose of Chapter 5 was to investigate how varying the scale of the envelope in a 2^{nd}- order stimulus affects perceived depth. Observers performed well when they were presented with a narrow stimulus and poor when presented with a wide stimulus. This result coincides well with physiological studies (DeAngelis et al., 1995; Freeman & Ohzawa, 1990; Ohzawa et al., 1990) that have shown that the size of a cell's RF is inversely related to spatial frequency. Many investigators have argued that neurons tuned to high spatial frequencies encode small disparities, and cells tuned to low spatial frequencies encode large disparities (DeAngelis et al., 1995; Marr & Poggio, 1979; Mayhew & Frisby, 1981; Schor et al., 1984; Smallman & Macleod, 1994). This can also be applied to broadband stimuli such as bars, and can explain the data reported in Chapter 5. Cells tuned to high spatial frequencies likely responded to the narrow 2^{nd}- order stimuli, which resulted in better performance compared to wider stimuli that stimulated neurons with a lower spatial frequency preference. Therefore, the size-disparity correlation holds for 2^{nd}- order processing of metric depth: performance on the 2IFC task is influenced by the size of the envelope, not the thickness of the bars within the carrier.

Monoptic depth is another mechanism that can account for depth perception at large disparities. In Chapter 4, both quantitative and qualitative percepts were obtained from uncorrelated stimuli. Therefore it is possible that a completely different mechanism underlies qualitative depth at large disparities. This possibility was assessed in Chapter 6 by measuring percent consistent achieved by observers in four different conditions. For the observers that were able to perceive monoptic depth, performance was around 70%, similar to that reported by Wilcox et al. (2007). For the binocular conditions, similar results were obtained in the disparity-only condition and the disparity+monoptic condition. For all four observers tested, performance on the two binocular conditions was significantly different from performance on the monoptic conditions. This suggests that different mechanisms underlie depth from disparity and depth from monoptic elements. Therefore, monoptic depth cannot account for the results found in the diplopic range in Chapters 3 and 4. This provides further evidence that 2nd- order processing mediates depth from disparity at large disparities.

7.2. Depth from Diplopic Stimuli

7.2.1. The separation confound at large disparities

In a discrimination task, observers are presented with a target and a reference stimulus. When the stimuli are diplopic, four half-images are visible and it is possible to indicate the position of the target relative to the reference stimulus by comparing the separation of the two pairs of images. It is difficult to ensure that observers make relative depth judgments when the separation information is fully confounded with the binocular disparity signal. A primary focus of Chapters 3 and 4 was evaluating the methods used to study depth from diplopia and determining an appropriate method for assessing depth across the entire range of disparities.

Siderov and Harwerth (1993) argued that when using a depth discrimination task, interleaving crossed and uncrossed trials within a single session would eliminate the separation information in diplopic targets. Their rationale for this was that the binocular image is the same for any disparity if the observer does not have access to eye of origin information. However, as shown in Experiment 3.1.1, observers reliably (near 100% correct) discern the direction of depth for the set of stimuli (reference + target). Therefore, observers could determine the direction of the offset and use the separation information to correctly identify the direction of the depth offset of the target relative to the reference.

The visibility of the lateral separation of the half images is also an issue in depth estimation tasks at large disparities. Similar to the discrimination task, observers could initially discern the direction of the depth offset (relative to the screen plane) and use the separation of the half images to guide their estimates of perceived depth. If this strategy were used, one would predict a linear increase in perceived depth in the diplopic range to coincide with the increase in lateral separation. To determine if observers relied on the lateral separation in the depth estimation task, follow up experiments were conducted in Chapters 3 (Experiment 3.2.1) and 4 (Experiment 4.1.1) where observers were presented with a non-stereoscopic (2D condition) version of the bar (Chapter 3) and noise (Chapter 4) stimuli. Most observers performed differently in the 2D vs. 3D conditions for both stimulus types. Therefore, different strategies were used in the 2D and 3D conditions, which suggests that the visibility of the lateral separation of the half images did not influence estimates of perceived depth.

Certain observers can (and do) use the lateral separation to guide their estimates of perceived depth in depth estimation tasks. Although this method is reliable for most observers, it was preferable to use a more precise psychophysical method that provides a single measure of sensitivity for each disparity range. Therefore, a 2IFC task was implemented, which meets these criteria and in addition was meant to make the separation information less salient when measuring stereosensitivity at large disparities. While the lateral separation was still available in the 2IFC task, to make use of it observers would intentionally have had to go through a number of cognitive steps. Observers would initially discern the depth sign in two intervals, then use the separation information in each interval to determine the amount of separation between the two stimuli, and lastly compare the size of the difference in the separation of the two sets of half images in each interval to determine which interval has a greater degree of separation. Most observers performed poorly in the overlapping and diplopic range, so it was not likely that they were using this strategy. However a follow-up task (Experiment 4.2.1) was conducted to evaluate this possibility. Similar to the Experiment 4.1.1, observers performed the 2IFC task for both a 2D and 3D condition. Results indicated that three out of four observers used different strategies in the 2D and 3D conditions. It should be noted that the same observer was the outlier in both follow-up experiments. Therefore, certain observers are more heavily influenced by the visibility of the separation of the half images and it is most likely impossible to remove the separation information for *all* observers at large disparities. Thus it is important to directly assess each observer's strategy using control experiments like those described here. Nonetheless, for most observers, the depth estimation and the 2IFC task provide reliable results of perceived depth and stereosenstivity measurements respectively.

7.2.2. Monoptic depth

It is clear that quantitative depth is encoded by 2nd- order stereopsis within the fused range (Chapter 4), but only sign information is obtained beyond the fusion limit (Chapters 2 and 4, Wilcox & Hess, 1995). However, if depth is processed by the same mechanism (2nd- order), performance should presumably be the same across the range of disparities tested. Because different patterns were found, it is possible that instead of 2nd- order stereopsis encoding depth from disparity in the diplopic range, an entirely different mechanism can operate at that range. Monoptic depth has been shown to occur reliably at large disparities (Wilcox et al., 2007), therefore it is possible that a separate mechanism operates in the diplopic range that processes depth of the half images, instead of the stereoscopic pair. This hypothesis was tested in Chapter 6, and results indicated that a statistically different pattern of performance is obtained in the disparity and monocular conditions. This suggests that both monoptic depth and 2nd- order stereopsis can be used to perceive depth at large disparities, but when binocular disparity is present, depth is encoded by a 2nd- order mechanism.

7.3. Summary of Main Findings

With the goal of relating the spatial dichotomy to the temporal dichotomy and to Ogle's perceptual dichotomy, the properties of 1st- and 2nd- order stereopsis were evaluated over a series of experiments reported in this dissertation. Compared to the temporal dichotomy, the spatial dichotomy remains the most likely mechanistic dichotomy for stereopsis. Sustained and transient properties can describe 1st- and 2nd- order processing in the fused range (Hess & Wilcox, 2008), but different temporal response profiles are not found at large disparities because only 2nd- order processing mediates stereopsis at the range tested here. Overall, results indicated that both 1st- and 2nd- order mechanisms independently encode quantitative depth in the fused range, although the amount of metric depth differs.

These experiments demonstrated that performance on the correlated (1st- order) and uncorrelated (2nd- order) conditions are similar at large disparities, thus providing further evidence for Wilcox and Hess' (1995) hypothesis that 2nd- order alone mediates depth from disparity near the upper disparity limit (D_{max}). These experiments suggest that 2nd- order processing not only mediates depth near D_{max}, but also over a range of large disparities starting at the fusion limit. Furthermore, in accordance with Ogle's findings (1952b), quantitative depth was perceived in the fused range, however only qualitative depth was perceived in the diplopic range. The transition from quantitative to qualitative depth occurred at the lower limit of fusion, which seems inconsistent with Ogle's observations. Figure 7.1 summarizes the relationship between 1st- /2nd- order stereopsis, Ogle's perceptual dichotomy and monoptic depth.

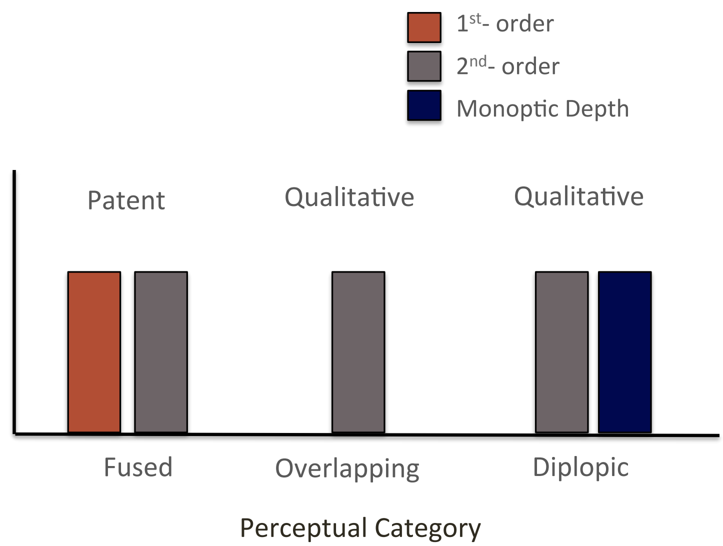


Figure 7.1. Summary of the mechanisms and categorization of depth percept within each disparity range. Both 1st and 2nd order stereopsis operate in the fused range and patent depth can be extracted from these signals. Beyond the lower fusion limit, the visual system defaults to using 2nd- order processing and only qualitative depth is extracted. When stimuli are diplopic, both 2nd-order processing and monoptic depth can be used to encode depth.

7.4. Limitations

Despite the considerably careful design of methodology, stimuli and evaluation, there are a number of limitations associated with this research project. The separation information remains a hurdle in studying depth from diplopic targets. The use of this information must be evaluated for every paradigm, stimulus and for every individual tested. Both the depth estimation and 2IFC tasks made the separation information less salient than a simple depth discrimination task, however one observer continued to rely on this information no matter what methodology was used.

Another limitation in evaluating the properties of 1^{st}- and 2^{nd}- order stereopsis is that it is impossible to isolate the 1^{st}- order signal because the 2^{nd}- order information is always available because of the external edges of the stimulus. Therefore, it is possible to isolate the 2^{nd}- order information by using uncorrelated noise patches, but the 1^{st}- order correlated condition contains both types of information. Does the visual system ignore 2^{nd}- order contrast information when stimuli are matched, or does the 1^{st}- and 2^{nd}- order information sum and contribute (even unequally) to the overall percept of depth?

7.5. Neural correlates of stereopsis

The relationship between 1^{st}- and 2^{nd}- order stereopsis and the other proposed dichotomies were examined in this dissertation. How are these mechanisms represented in the brain? The simplest comparison can be made with the neural correlates described in the electrophysiological work on alert monkeys by Poggio and colleagues (1995; 1977; 1988; 1981). They suggested that tuned neurons encode patent depth while near/far neurons encode qualitative depth (Poggio & Talbot, 1981). How does the 1^{st}-/2^{nd}- order dichotomy correspond to these distinct neural categories? There is physiological evidence that the same neuron is responsive to both 1^{st}- and 2^{nd}- order stimuli, but with a different preferred spatial frequency and these neurons encode the spatial information using distinct neural pathways (Tanaka & Ohzawa, 2006; Zhou & Baker, 1993). It is possible that tuned neurons are the neural correlate for both 1st- and 2nd- order processing and the same neuron encodes patent depth using different neural pathways. According to the findings reported here, only 2nd- order processing encodes qualitative depth in the diplopic range. If 1st-/2nd- order stereopsis were directly related to Poggio's neural classification, 2nd- order stereopsis should also be represented by the near/far neurons in the diplopic range and provide only qualitative depth (which it does) in this range.

Another possibility is that near/far cells are the neural correlate for monoptic depth. Kaye (1978) measured depth estimates for monocular stimuli when the disparities were 0, 1, 2, and 4 degrees, and did not find an increase in perceived depth as a function of eccentricity. Fukuda, Wilcox, Allison, and Howard (2008) reported a slight increase in metric depth with eccentricity but this pattern was apparent for only two out of four observers, in only the crossed direction. Therefore, although the current literature on monoptic depth is limited, thus far it seems that percepts are qualitative (Fukuda et al., 2009; Kaye, 1978), and are present at diplopic disparities (Wilcox et al., 2007, and Chapter 6). These properties coincide with Poggio and colleagues description of near/far cells.

The results of Chapter 6 indicated that depth from disparity results from a 2nd- order signal, not from a monoptic depth mechanism. If 2nd- order stereopsis is represented only by tuned neurons, why is depth from disparity in the diplopic range qualitative only? It is possible that tuned cells (which likely represent both 1st- and 2nd- order processing) actually encode depth across the entire range of disparities but patent depth simply cannot be extracted past a certain point (fusion limit). Because depth as a function of disparity increases logarithmically, the theoretical depth beyond a certain point is exceptionally large and as a result percepts are very coarse. Another possibility is that a 2nd- order target stimulates extremely large receptive fields, which produce a coarse percept that appears like constant perceived depth because the differences in perceived depth across the disparities tested are so small.

The categorization of neurons described by Poggio and colleagues (1995; 1977; 1988; 1981) may not be entirely accurate, and other experiments (LeVay & Voigt, 1988) have shown that there are many intermediate response patterns in addition to the original categories proposed by Poggio and colleagues. Their result coincides with Ohzawa and colleagues' (1995; 1990) argument that there is a continuum of spatial frequency tuned cells. It is clear from what has been presented here, and the existing literature that there is a distinct 2nd- order processing stream, and as proposed by Zhou and Baker (1993), that 1st- and 2nd- order processing is supported by the same neurons but are organized into two streams. In one stream, 1st- order stimuli are filtered linearly and in the other stream, 2nd- order stimuli are processed by a filter-rectify-filter pathway.

Therefore, 1st- and 2nd- order stimuli are likely encoded by disparity-sensitive neurons via a two-stream model (proposed by Wilson, Ferrera & Yo, 1992 and Zhou & Baker, 1993) which consists of an early luminance processing filter and a late stage spatial frequency filter which corresponds to the envelope's spatial frequency, separated by a pointwise nonlinearity. According to Zhou and Baker (1993), a separate linear stream is still necessary to account for the cell's luminance response properties. From the results thus far, monoptic depth seems to represent a separate phenomena, and is processed by a distinct class of cells, perhaps by near/far cells, however not enough is known about depth percepts from monoptic elements to determine their neural correlate.

7.6. Conclusions

The results in Chapter 2 confirmed that the sustained/transient distinction is a re-description of the 1st-/2nd- order dichotomy and cannot be used to explain the underlying mechanisms of stereopsis. Hess and Wilcox (2008) showed that in the fused range, 1st- and 2nd- order processing have distinct temporal properties. At large disparities, correlated and uncorrelated stimuli exhibited the same temporal dynamics because 2nd- order processing alone most likely mediates depth from large disparities. The work presented in this dissertation demonstrates for the first time that quantitative depth is perceived from *both* 1st- and 2nd- order stimuli. These results also highlight the difference in resolution resulting from each type of stimulus in the fused range. While acknowledging the difficulty in evaluating depth percepts at large disparities, the results reported here confirm that only qualitative depth is perceived beyond the fusion limit, and depth is mediated by 2nd- order stereopsis only.

Broadly, these results coincide with Ogle's perceptual dichotomy but the transition from patent to qualitative percepts occurs at the point where observers begin to lose fusion, not within the diplopic range as argued by Ogle (1952b). These results are consistent with widely accepted filter-rectify-filter models of 2nd- order processing and given the physiological data, it is likely that this processing occurs within two streams, which share a common neural substrate.

Bibliography

Anzai, A., Ohzawa, I., & Freeman, R. D. (1999). Neural mechanisms for encoding binocular disparity: receptive field position vs. phase. *Journal of Neurophysiology, 82*(2), 874-890.

Badcock, D. R., & Schor, C. M. (1985). Depth-increment detection function for individual spatial channels. *Journal of the Optical Society of America, 2*(7), 1211–1216.

Barlow, H., Blakemore, C., & Pettigrew, J. D. (1967). The neural mechanism of binocular depth discrimination. *J Physiol (Lond), 193*(1), 327-342.

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

Blakemore, C., & Campbell, F. W. (1969). Adaptation to spatial stimuli. *Journal of Physiology-London, 200*(1), 11-17.

Blakemore, C., Nachmias, J., & Sutton, P. (1970). Perceived spatial frequency shift – evidence for frequency selective neurones in human brain. *Journal of Physiology- London, 210*(3), 727-734.

Breitmeyer, B., & Ogmen, H. (2000). Recent models and findings in visual backward masking:A comparison, review and update. *Perception and Psychophysics, 62*, 1572–1595.

Buckthought, A., & Wilson, H. R. (2007). Interaction between binocular rivalry and depth in plaid patterns. *Vision Research, 47*, 2543–2556.

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to visibility of gratings. *Journal of Physiology- London, 197*(3), 551-566.

Carter, B. E., & Henning, G. B. (1971). Detection of gratings in narrow band visual noise. *Journal of Physiology – London, 219*(2), 355.

Cavanagh, P., & Mather, G. (1989). Motion: the long and short of it. *Spatial Vision, 4*, 103-129.

Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America, 5*, 1986–2007.

Chubb, C., & Sperling, G. (1989). Two motion perception mechanisms revealed through distance driven reversal of apparent motion. *Proc. Natl. Acad. Sci. USA, 86*(1), 2985-2989.

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

Dakin, S. C., & Mareschal, I. (2000). Sensitivity to contrast modula- tion depends on carrier spatial frequency and orientation. *Vision Research, 40*, 311-329.

DeAngelis, G. C., Ohzawa, I., & Freeman, R. D. (1995). Neuronal mechanisms underlying stereopsis: How do simple cells in the visual cortex encode binocular disparity? . *Perception, 24*, 3-31.

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

Derrington, A. M., & Henning, G. B. (1993). Vision – linear and nonlinear mechanisms in pattern vision. *Current Biology, 3*(11), 800-803.
Edwards, M., Pope, D. R., & Schor, C. M. (1998). Luminance contrast and spatial-frequency tuning of the transient-vergence system. *Vision Research, 38*(5), 705–717.
Edwards, M., Pope, D. R., & Schor, C. M. (1999). Orientation tuning of the transient-stereopsis system. *Vision Research, 39*(16), 2717–2727.
Edwards, M., Pope, D. R., & Schor, C. M. (2000). First- and second-order processing in transient stereopsis. *Vision Research, 40*(19), 2645–2651.
Felton, T. B., Richards, W., & Smith, R. A. (1972). Disparity processing of spatial frequencies in man. *Journal of Physiology – London, 225*, 349-362.
Ferster, D. (1981). A comparison of binocular depth mechanisms in areas 17 and 18 of the cat visual corex. *Journal of Physiology, 311*, 623–655.
Fischer, B., & Krueger, J. (1979). Disparity tuning and binocularity of single neurons in the cat visual cortex. *Exp Brain Res, 35*, 1-8.
Fleet, D. J., Wagner, H., & Heeger, D. J. (1996). Neural encoding of binocular disparity: Energy models, position shifts and phase shifts. *Vision Research, 36*(12), 1839-1857.
Foley, J. M., Applebaum, T. H., & Richards, W. A. (1975). Stereopsis with large disparities: discrimination and depth magnitude. *Vision Research, 15*, 417-421.
Freeman, R. D., & Ohzawa, I. (1990). On the neurophysiological organization of binocular vision. *Vision Research, 30*, 1661-1676.
Fukuda, K., Wilcox, L. M., Allison, R. S., & Howard, I. P. (2008). Comparison of depth percepts created by binocular disparity, Panum's limiting case, and monoptic depth. *Journal of Vision Suppl, 8*(6), 1086.
Fukuda, K., Wilcox, L. M., Allison, R. S., & Howard, I. P. (2009). A reevaluation of the tolerance to vertical misalignment in stereopsis. *Journal of Vision, 9*(2), 1-8.
Graham, N. (1972). Spatial frequency channels in human visual system- effects of luminance and pattern drift rate. *Vision Research, 12*(1), 53-68.
Harwerth, R. S., Fredenburg, P. M., & Smith, E. L. r. (2003). Temporal integration for stereoscopic vision. *Vision Research, 43*(5), 505-517.
Harwerth, R. S., & Rawlings, S. C. (1977). Viewing time and stereoscopic thresholds with random-dot stereograms. *American Journal of Optometry & Physiological Optics, 54*, 452-457.
Helmholtz, H. v. (1910). *Physiological optics* (Vol. (Translation by JPC Southall from the 3rd German edition of Handbuch der Physiologischen Optik)). New York: Dover.
Henning, G. B., Hertz, B. G., & Broadbent, D. E. (1975). Some experiments bearing on hypothesis that visual system analyses of spatial patterns in independent bands of spatial frequency. *Vision Research, 15*(8-9), 877-897.
Henning, G. B., Hertz, B. G., & Hinton, J. L. (1981). Effects of different hypothetical detection mechanisms on the shape of spatial-frequency filters inferred from masking experiments. I. Noise masks. *Journal of the Optical Society of America, 71*(5), 574-581.
Hering, E. (1861). *Beitrage zur Physiologie (p.291)*. Leipzig: W. Engelmann.

Hering, E. (1942). *Spatial Sense and Movements of the Eye*: Translated by C.A. Radde (The American Academy of Optometry, Baltimore, Maryland).
Hess, R. F., & Wilcox, L. M. (1994). Linear and non-linear filtering in stereopsis. *Vision Research, 18*(18), 2431-2438.
Hess, R. F., & Wilcox, L. M. (2006). Stereo dynamics are not scale-dependent. *Vision Research, 46*, 1911-1923.
Hess, R. F., & Wilcox, L. M. (2008). The transient nature of 2nd-order stereopsis. *Vision Research, 48*(11), 1327–1334.
Howard, I. P. (1996). Alhazen's neglected discoveries of visual phenomena. *Perception, 25*(10), 1203 – 1217
Howard, I. P., & Rogers, B. J. (1995). *Binocular Vision and Stereopsis*: Oxford University Press.
Howard, I. P., & Rogers, B. J. (2002). *Seeing in Depth* (Vol. 2). Toronto: I. Porteous.
Howard, I. P., & Wade, N. J. (1996). Ptolemy's contributions to the geometry of binocular vision. *Perception, 25*, 1189-1201.
Howarth, P. A. (2011). The geometric horopter. *Vision Research, 51*(4), 397-399.
Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology, 148*, 574-591.
Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol (Lond) 160*, 106-154.
Hubel, D. H., & Wiesel, T. N. (1973). A re-examination of stereoscopic mechanisms in the cat. *J Physiol (Lond), 232*, 29-30.
Jones, R. (1977). Anomalies of disparity detection in human visual system. *Journal of Physiology, 264*(3), 621-640.
Jones, R. (1980). Fusional vergence: Sustained and transient components. *American Journal of Optometry & Physiological Optics, 57*(9), 640-644.
Joshua, D. E., & Bishop, P. O. (1970). Binocular single vision and depth discrimination. Receptive field disparities for central and peripheral vision and binocular interaction on peripheral single units in cat striate cortex. *Exp Brain Res, 10*, 389-416.
Kaye, M. (1978). Stereopsis without binocular correlation. *Vision Research, 18*(8), 1013–1022.
Kovács, I., & Fehér, A. (1997). Non-fourier information in bandpass noise patterns. *Vision Research, 37*, 1167-1177.
Kulikowski, J. J. (1978). Limit of single vision in stereopsis depends on contour sharpness. *Nature, 276*, 126-127.
Langley, K., Fleet, D. J., & Hibbard, P. B. (1999). Stereopsis from contrast envelopes. *Vision Res, 39*, 2313-2324.
LeVay, S., & Voigt, T. (1988). Ocular dominance and disparity coding in cat visual cortex. . *Vis Neurosci, 1*, 395-414.
Lin, L., & Wilson, H. R. (1995). Stereoscopic integration of Fourier and non-Fourier patterns. *Investigative Ophthalmology and Visual Science Suppl., 36*(4), S364.
Liss, P. (1968). Does backward masking by visual noise stop visual processing? *Perception and Psychophysics, 4*, 328–340.

Lu, Z. L., & Sperling, G. (1995). The functional architecture of human visual motion perception. *Vision Research, 35*, 2697-2722.
Lugtigheid, A. J., Wilcox, L. M., Allison, R. S., & Howard, I. P. (2014). Vergence eye movements are not essential for stereoscopic depth. *Proceedings of the Royal Society B, 281*.
Maffei, L., & Fiorentini, A. (1977). Spatial frequency rows in the striate visual cortex. *Vision Research, 17*, 257-264.
Marr, D., & Poggio, T. (1979). A computational theory of human stereo vision. *Proceedings of the Royal Society of London, 204*, 301–328.
Maske, R., Yamane, S., & Bishop, P. O. (1984). Binocular simple cells for local stereopsis: Comparison of receptive field organizations for the two eyes. *Vision Research, 24*, 1921-1929.
Maunsell, J. H. R., & Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interaction and sensitivity to binocular disparity. *Journal of Neurophysiology, 49*, 1148-1167.
Mayhew, J. E. W., & Frisby, J. P. (1976). Rivalrous texture stereograms. *Nature, 264*, 53–56.
Mayhew, J. E. W., & Frisby, J. P. (1979a). Convergent disparity discriminations in narrow-band-filtered random-dot stereograms. *Vision Research, 19*, 63-71.
Mayhew, J. E. W., & Frisby, J. P. (1979b). Surfaces with steep variations in depth pose difficulties for orientationally tuned disparity filters. *Perception, 8*, 691-698.
Mayhew, J. E. W., & Frisby, J. P. (1981). Psychophysical and computational studies towards a theory of human stereopsis. *Artificial Intelligence, 17*(1-3), 349-385.
McKee, S. P., Levi, D. M., & Bowne, S. F. (1990). The imprecision of stereopsis. *Vision Research, 30*(11), 1763–1779.
McKee, S. P., & Taylor, D. G. (2010). The precision of binocular and monocular depth judgments in natural settings. *Journal of Vision, 10*(10), 1-13.
McKee, S. P., Verghese, P., & Farell, B. (2004). What is the depth of a sinusoidal grating? *J Vis, 4*(7), 524-538.
McKee, S. P., Verghese, P., & Farell, B. (2005). Stereo sensitivity depends on stereo matching. *J Vis, 5*(10), 783-792.
McKee, S. P., Verghese, P., Ma-Wyatt, A., & Petrov, Y. (2007). The wallpaper illusion explained. *Journal of Vision, 7*(14), 11.
Mitchell, D. E. (1966). Retinal disparity and diplopia. *Vision Research, 6*, 441-451.
Mitchell, D. E. (1969). Qualitative depth localitation with diplopic images of dissimilar shape. *Vision Research, 9*, 991-994.
Mitchell, D. E. (1970). Properties of stimuli eliciting vergence eye movements and stereopsis. *Vision Research, 10*, 145-162.
Mitchell, D. E., & O'Hagan, S. (1972). Accuracy of stereoscopic localization of small line segments that differ in size or orientation for the two eyes. *Vision Research, 12*(3), 437-454.
Nachmias, J. (1989). Contrast modulated maskers – test of a late nonlinearity hypothesis. *Vision Research, 29*(1), 137-142.

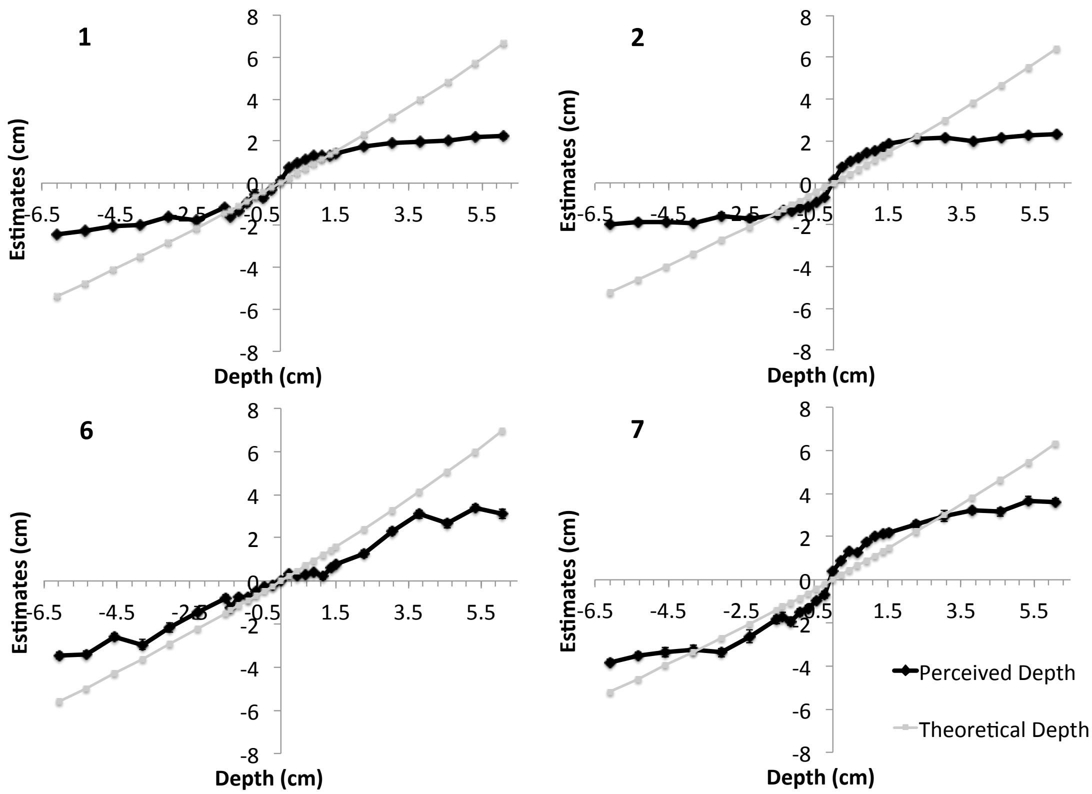
Nachmias, J., & Rogowitz, B. E. (1983). Masking by spatially-modulated gratings. *Vision Research, 23*(12), 1621-1629.
Ogle, K. N. (1952a). Disparity limits of stereopsis. *American Archives of Ophthalmology, 48*(1), 50-60.
Ogle, K. N. (1952b). On the limits of stereoscopic vision. *Journal of Experimental Psychology, 44*(4), 253–259.
Ogle, K. N. (1953). Precision and validity of stereoscopic depth perception from double images. *Journal of the Optical Society of America, 43*(10), 906-913.
Ogle, K. N. (1964a). *The optical space sense* (Vol. 4). London: Academic Press.
Ogle, K. N. (1964b). *Researches in Binocular Vision*: Hafner.
Ogle, K. N., & Wakefield, J. M. (1967). Stereoscopic depth and binocular rivalry. *Vision Research, 7*(1), 89-98.
Ogle, K. N., & Weil, M. P. (1958). Stereoscopic vision and the duration of the stimulus. *American Archives of Ophthalmology, 59*, 4-17.
Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1997). Encoding of binocular disparity by complex cells in the cat's visual cortex. *Journal of Neurophysiology, 77*, 2879-2909.
Ohzawa, I., DeAngelis, G. C., & Freeman, R. F. (1990). Stereoscopic Depth Discrimination in the Visual Cortex: Neurons Ideally Suited as Disparity Detectors. *Science, 249*(4972), 1037-1041.
Panum, P. (1858). *Physiologische untersuchungen über das Sehen mit zwei Augen*. Kiel: Schwerssche Buchhandlung.
Pettigrew, J. D., Nikara, T., & Bishop, P. O. (1968). Binocular interaction on single units in cat striate cortex: Simultaneous stimulation by single moving slit with receptive fields in correspondence. *Experimental Brain Research, 6*, 391–410.
Poggio, G. F. (1995). Mechanisms of stereopsis in monkey visual cortex. *Cerebral Cortex, 3*, 193-204.
Poggio, G. F., & Fischer, B. (1977). Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkeys. *Journal of Neurophysiology, 40*, 1392-1407.
Poggio, G. F., Gonzalez, F., & Krause, F. (1988). Stereoscopic mechanisms in monkey visual cortex: binocular correlation and disparity selectivity. *Journal of Neurophysiology, 8*, 4531-4550.
Poggio, G. F., & Talbot, W. H. (1981). Neural mechanisms of static and dy- namic stereopsis in foveal striate cortex of rhesus monkeys. *Journal of Physiology-London, 315*, 469-492.
Prince, S. J. D., Cumming, B. G., & Parker, A. J. (2002). Range and mechanism of encoding of horizontal disparity in macaque V1. *Journal of Neurophysiology, 87*(1), 209–221.
Read, J. C. A., Parker, A. J., & Cumming, B. G. (2002). A simple model accounts for the response of disparity-tuned VI neurons to anti-correlated images. *Visual Neuroscience, 19*, 735-753.
Reading, R. W. (1970). The threshold of distance discrimination for objects located outside Panum's area. *American Journal of Optometry, 47*, 99-105.

Richards, W. (1971). Anomalous stereoscopic depth perception. *Journal of the Optical Society of America, 61*, 410-414.
Richards, W., & Foley, J. M. (1974). Effect of luminance and contrast on processing large disparities. *Journal of the Optical Society of America, 64*(12), 1703-1705.
Richards, W., & Kaye, M. G. (1974). Local versus global stereopsis: Two mechanisms? *Vision Research, 14*, 1345-1347.
Roy, J. P., Komatsu, H., & Wurtz, R. H. (1992). Disparity sensitivity of neurons in monkey extrastriate area MST. *Journal of Neuroscience, 12*, 2478-2492.
Sato, T. (1983). Depth seen with subjective contours. *Japanese Psychological Research, 25*(4), 213-221.
Sato, T., & Nishida, S. (1994). Does an envelope detecting mechanism mediate stereopsis for band-limited stimuli? *Investigative Ophthalmology and Visual Science (Suppl.), 35*, 1916.
Schofield, A. J. (2000). What does second-order vision see in an image? . *Perception, 29*, 1071–1086.
Schofield, A. J., & Georgeson, M. A. (1999). Sensitivity to modulations in luminance and contrast in visual white noise: separate mechanisms with similar behavior. *Vision Research, 39*, 2697-2716.
Schofield, A. J., & Georgeson, M. A. (2003). Sensitivity to contrast modulation: the spatial frequency dependence of second-order vision. *Vision Research, 43*, 243-259.
Schor, C. M., Edwards, M., & Pope, D. R. (1998). Spatial-frequency and contrast tuning of the transient-stereopsis system. *Vision Research, 38*(20), 3057–3068.
Schor, C. M., Edwards, M., & Sato, M. (2001). Envelope size tuning for stereo-depth perception of small and large disparities. *Vision Research, 41*(20), 2555–2567.
Schor, C. M., & Wood, I. (1983). Disparity range for local stereopsis as a function of luminance spatial frequency. *Vision Research, 23*(12), 1649–1654.
Schor, C. M., Wood, I., & Ogawa, J. (1984). Spatial tuning of static and dynamic local stereopsis. *Vision Research, 24*, 573-578.
Shimono, K. (1984). Evidence for the subsystems in stereopsis- fine and coarse stereopsis. *Japanese Psychological Research, 26*(3), 166-172.
Siderov, J., & Harwerth, R. S. (1993). Precision of Stereoscopic Depth Perception from Double Images. *Vision Research, 33*(1553-1560).
Smallman, H. S., & Macleod, D. I. A. (1994). Size- disparity correlation in stereopsis at contrast threshold. *Journal of the Optical Society of America, 11*, 2169–2183.
Sperling, G. (1963). A model for visual memory tasks. *Human Factors, 5*, 19–31.
Stromeyer, C. F., & Julesz, B. (1972). Spatial frequency masking in vision- critical bands and spread of masking. *Journal of the Optical Society of America, 62*(10), 1221.
Tanaka, H., & Ohzawa, I. (2006). Neural basis for stereopsis from second-order contrast cues. *Journal of Neuroscience, 26*(16), 4370–4382.
Thompson, P. G., & Movshon, J. A. (1978). Storage of spatially specific threshold elevations. *Perception, 7*(1), 65-73.

Treisman, A. (1962). Binocular rivalry and stereoscopic depth perception. *Quarterly Journal of Experimental Psychology, 14*, 23-37.
Westheimer, G. (1979). Cooperative neural processes involved in stereoscopic acuity. *Experimental Brain Research, 36*(3), 585-597.
Westheimer, G., & Mitchell, D. E. (1969). The sensory stimulus for disjunctive eye movements. *Vision Research, 9*(7), 749-755.
Westheimer, G., & Tanzman, I. J. (1956). Qualitative depth localization with diplopic images. *Journal of the Optical Society of America, 46*(2), 116–117.
Wheatstone, C. (1838). Contributions to the Physiology of Vision. Part the First. On Some Remarkable, and Hitherto Unobserved, Phenomena of Binocular Vision. *Philosophical Transactions of the Royal Society of London, 128*, 371–394.
Wichmann, F. A., & Hill, N. J. (2001a). The psychometric function: I. Fitting, Sampling, and Goodness of fit. *Perception and Psychophysics, 63*(8), 1293-1313.
Wichmann, F. A., & Hill, N. J. (2001b). The psychometric function: II.
Bootstrap-based confidence intervals and sampling. *Perception and Psychophysics, 63*(8), 1314–1329.
Wilcox, L. M., & Allison, R. S. (2009). Coarse-fine dichotomies in human stereopsis. *Vision Research, 49*, 2653–2665.
Wilcox, L. M., Harris, J. M., & McKee, S. P. (2007). The role of binocular stereopsis in monoptic depth perception. *Vision Research, 47*, 2367–2377.
Wilcox, L. M., & Hess, R. F. (1995). Dmax for stereopsis depends on size, not spatial frequency content. *Vision Res, 35*(8), 1061-1069.
Wilcox, L. M., & Hess, R. F. (1996). Is the site of non-linear filtering in stereopsis before or after binocular combination. *Vision Research, 36*, 391-399.
Wilcox, L. M., & Hess, R. F. (1997). Scale selection for second-order(non-linear) stereopsis. *Vision Research, 37*(21), 2981-2992.
Wilcox, L. M., & Hess, R. F. (1998). When stereopsis does not improve with increasing contrast. *Vision Res, 38*(23), 3671-3679.
Wilson, H. R. (1999). *Non-Fourier Cortical Processes in Texture, Form, and Motion Perception. Models of Cortical Circuits*.
Wilson, H. R., Ferrera, V. P., & Yo, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Visual Neuroscience, 9*, 79-97.
Zhou, Y. X., & Baker, C. L. (1993). A processing stream in mammalian visual-cortex neurons for non-Fourier responses. *Science, 261*(5117), 98-101.
Ziegler, L. R., & Hess, R. H. (1997). Depth perception during diplopia is direct. *Perception, 26*(1225–1230).

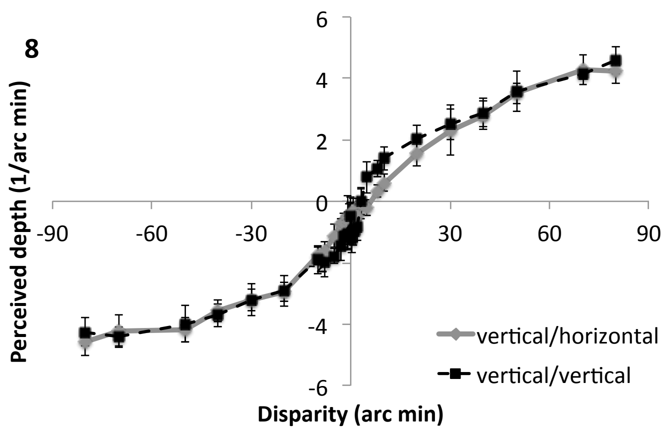
Appendix A.

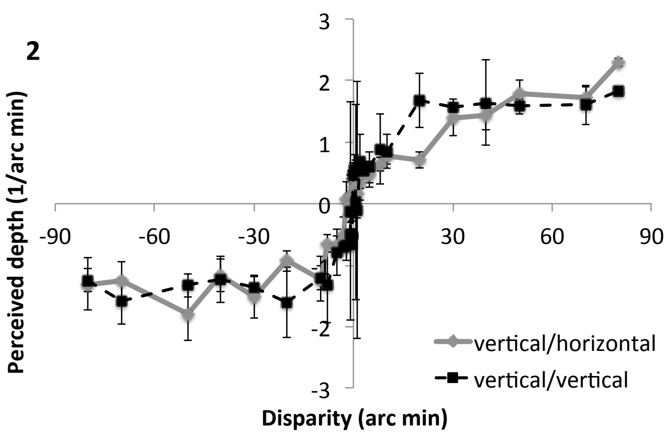
The individual results from Experiment 3.2. Observers estimated the amount of perceived depth between a zero disparity reference and a simple line stimulus. Estimation results are shown for four observers (subjects 1, 2, 6, 7) as a function of actual depth (cm). The grey curves represent the theoretical depth.



Appendix B

A Comparison of depth estimation results when uncorrelated stimuli consisted of images that were vertically oriented in one eye and horizontally oriented in the other eye (grey curve) or when both images were vertically oriented (black curve). Error bars represent 95% confidence intervals.





Appendix C

The results of Experiment 4.2 showing psychometric functions for four observers. The black and grey curves represent responses in the correlated and uncorrelated conditions respectively in the fused, overlapping and diplopic ranges. When fitting the functions, the data for test disparities in the uncrossed direction relative to the target were flipped so that all responses were in the crossed direction. Then the Weibull function was used to fit data above the 50% point (in the crossed direction only).