First- and second-order binocular matching in stereopsis:

Psychophysics and modeling

by

Athena Buckthought

A thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

Department of Psychology

Carleton University
Ottawa, Ontario
September 2004

Your file Votre référence
ISBN: 0-612-97814-1
Our file Notre référence
ISBN: 0-612-97814-1

Abstract

The contribution of first- and second-order binocular matching was investigated in a stereoscopic depth discrimination task. This was done by comparing stereoscopic depth discrimination performance with a fixed contrast window (Experiment 1) to that with a shifted contrast window (Experiment 2). The carrier was a 2 cpd or 8 cpd filtered noise pattern. With the fixed contrast window stereoscopic performance was limited to relatively small disparities, reflecting the disparity limit of first-order binocular matching. As expected, the disparity limit for stereoscopic performance occurred at smaller disparities for the 8 cpd compared to 2 cpd patterns. Compared to the fixed window, stereoscopic performance with the shifted window extended to larger disparities and the differences between the 8 cpd and 2 cpd patterns were reduced. The shifted wide cosine window (Experiment 3) provided the most compelling evidence for binocular second-order matching, as determined in simulations of the binocular energy model. In Experiment 4 binocular first- and second-order matching was investigated using a display in which the disparity of the edges of the pattern was inconsistent with the disparity of the carrier. Results from Experiments 1-4 were consistent with a model of second-order processing with a first stage of filters, followed by a non-linearity and a second-stage of filters at a lower spatial frequency. In Experiment 5 depth discrimination was measured for oriented bandpass filtered noise patterns using filters which varied in mean orientation (θ). The upper disparity limit for

depth discrimination performance increased as the stimulus pattern was progressively more oblique and was inversely proportional to the sine of the angle θ. This indicated that the upper disparity limit in isotropic patterns in Experiments 1-4 was increased by the most horizontal components in the image. These results were consistent with simulations of the binocular energy model (Qian & Zhu, 1997) with oriented receptive fields and assuming that the outputs from complex cells at all orientations are combined using linear summation. This modeling resembled that for motion detection (Bischof & Di Lollo, 1991; Prince et al., 2001), suggesting that stereopsis and motion use similar mechanisms for detection and integration of orientation information.

Keywords: binocular, stereopsis, depth perception, motion perception, second-order, depth discrimination.

Acknowledgements

I would like to thank André Vincent, Filippo Speranza, James Tam, Edmund Rodrique and Sheryl Thingvold and other colleagues at the Communications Research Centre Canada for assistance throughout this project. I thank my co-supervisor John Logan for invaluable assistance throughout my PhD program, for generously helping me with the completion of my thesis work and taking over as supervisor in the absence of Lew Stelmach. Above all I thank my supervisor Lew for enormous encouragement and support, for being a mentor and for providing a stimulating environment for research. Unfortunately Lew passed away during the completion of this thesis so I dedicate this work to his memory.

Table of Contents

Abstract..... ii

Acknowledgements..... iv

Table of Contents..... v

List of Figures..... vii

List of Appendices..... ix

General Introduction 1

First-Order Processing (Binocular Energy Model)..... 5

Binocular Second-Order Processing..... 7

The Half-Shifted Window..... 13

The Detection and Integration of Orientation Information..... 15

General Methods for Experiments 1-5..... 18

Introduction to Experiment 1..... 29

Results 29

Discussion..... 31

Introduction to Experiment 2..... 31

Results..... 32

Discussion..... 33

Introduction to Experiment 3..... 34

Results..... 37

Discussion..... 38

Introduction to Experiment 4..... 45

Results 49

Discussion..... 51

Introduction to Experiment 5..... 55

Results 61

Discussion..... 62

Motion and Stereoscopic Vision..... 70

General Discussion..... 74

Appendix A..... 79

References..... 109

List of Figures

1 Stimuli at low & high spatial frequency 120

2 Stimulus display used in all experiments 121

3 Fixed, shifted & half-shifted window displays 122

4 Hypothesis for binocular matching of oriented components..... 123

5 Oriented filtered noise patterns (Experiment 5)..... 124

6 Experiment 1: shifted Difference of Gaussian..... 126

7 Experiment 1: fixed window hard-edge patterns..... 127

8 Experiment 2: shifted window hard-edge patterns..... 128

9 Experiment 3: fixed window wide cosine patterns..... 129

10 Experiment 3: shifted window wide cosine patterns..... 130

11 Experiment 3: shifted window narrow cosine patterns 131

12 Experiment 4: 2 cpd fixed, shifted & half-shifted (hard-edge)..... 132

13 Experiment 4: 8 cpd fixed, shifted & half-shifted (hard-edge)..... 133

14 Experiment 4: 2 cpd fixed, shifted & half-shifted (cosine)..... 134

15 Experiment 4: 8 cpd fixed, shifted & half-shifted (cosine)..... 135

16 Experiment 5: 2 cpd oriented fixed window..... 136

17 Experiment 5: 8 cpd oriented fixed window..... 137

18 Experiment 5: 2 cpd oriented shifted window..... 138

19 Experiment 5: 8 cpd oriented shifted window 139

20 Follow-up Study to Experiment 5..... 140

A1 Binocular energy model: (A) phase shift & (B) position shift..... 141

A2 Kernels for the phase shift model..... 142

A3 Kernels for the position shift model..... 143

A4 Example simulations for 8 cpd patterns at two disparities..... 144

A5 Simulations (phase shift): 2 cpd & 8 cpd fixed window..... 145

A6 Simulations for uncorrelated noise, narrow & wide cosine..... 146

A7 Simulations for correlated noise, narrow & wide cosine 147

A8 Binocular energy model with oriented kernels..... 148

A9 Simulations (phase shift): 2 cpd orientations 90°, 80°, 70°, 60°..... 149

A10 Simulations (phase shift): 2 cpd orientations 50°, 40°, 30°..... 150

A11 Simulations (phase shift): 8 cpd orientations 90°, 80°, 70°, 60°..... 151

A12 Simulations (phase shift): 8 cpd orientations 50°, 40°, 30° 152

A13 Simulations (phase shift) with oriented kernels compared to predictions. 153

A14 Simulations (position shift) with oriented kernels compared to predictions..... 154

List of Appendices

Appendix A. Simulations of the Binocular Energy Model.....79

First- and second-order binocular matching in stereoscopic depth discrimination

Stereoscopic depth perception depends upon the computation of disparity (i.e. the shift between image features in the left and right eye) at different spatial scales. The importance of spatial scale in stereoscopic depth perception has been confirmed, but the results of different studies have varied according to the class of stimulus. The starting point of my thesis research was an attempt to reconcile the discrepant results in the literature involving pedestal depth discrimination. In this task, stereoscopic depth discrimination thresholds are measured at different reference depths (pedestal disparities); thresholds are generally expected to rise gradually with increasing pedestal disparity as the task became more difficult. Using filtered noise patterns (Figure 1A), Smallman and MacLeod found that stereoscopic thresholds exhibited a rapid rise at modest pedestal disparities. However, depth discrimination varied with spatial frequency, with depth discrimination performance extending to larger disparities for the low compared to high spatial frequency patterns (Smallman & MacLeod, 1997). By comparison, using Difference of Gaussian patterns (Figure 1D), Siderov and Harwerth (1993) found that stereoscopic depth discrimination performance extended towards much larger pedestal disparities. Furthermore, in the study by Siderov and Harwerth depth discrimination performance did not vary with the spatial frequency of the stimulus.

One possible explanation for the discrepant results of Smallman and MacLeod (1997) and Siderov and Harwerth (1993) is that their experiments tapped into different types of stereoscopic processing (first- vs. second-order). Until about a decade ago, it was widely assumed that stereoscopic processing in the human visual system relies on a set of linear spatial frequency tuned channels (Blakemore & Campbell, 1969; Campbell & Robson, 1968; Julesz & Miller, 1975; Mansfield & Parker, 1993; Watson & Robson, 1981). Stereoscopic processing involves an initial filtering stage followed by some binocular comparison procedure (termed binocular matching) to extract the disparity information (Marr & Poggio, 1979). In simple terms, the image is filtered into spatial frequency bands, and disparity is computed in each band. Based on this type of model, stereoscopic depth perception should depend upon the underlying spatial frequency in the image (Julesz & Miller, 1975; Mansfield & Parker, 1993). Starting in the 1990s it was proposed that these filtering operations were not the only input to the binocular matching stage, but in addition a disparity signal may be provided by the stimulus envelope or contrast window (i.e. contrast modulation) (Hess & Wilcox, 1994; Wilcox & Hess, 1995, 1996, 1997, 1998; Zhou & Baker, 1993, 1994). Stereoscopic processing based on the contrast window is referred to as second-order, while stereoscopic processing based on linear filtering is referred to as first-order (Baker, 1999; Chubb, Olzak & Derrington, 2001; Clifford, Freedman & Vaina, 1998; Edwards, Pope & Schor, 2000; Langley, Fleet & Hibbard, 1998, 1999).

Evidence for binocular second-order processing came from a series of studies with Gabor patterns (sine wave patterns within a contrast window); the most important result was that the upper disparity limit for depth perception depended on the contrast window size and not the carrier spatial frequency (Wilcox & Hess, 1995). These results could not be readily explained with a model of stereopsis based on linear filtering and instead non-linear operations were required to extract the signal from the contrast window (Hess & Wilcox, 1994; Wilcox & Hess, 1995, 1996, 1997, 1998).

Further evidence for second-order processing in stereoscopic vision comes from studies showing that observers can perceive depth in images in which the left and right stereopairs have opposite polarity (Pope, Edwards & Schor, 1999), different spatial frequencies (Langley et al., 1998), orthogonal orientations (Schor et al., 2001), uncorrelated random carriers (Wilcox & Hess, 1996) and envelopes of different sizes (Schor et al., 2001). For these types of stimuli it is the shifted contrast envelope or window which is being binocularly matched and not the features within the envelope. The assumption which was made in these studies is that depth perception could not be based on binocular first-order matching since the carrier cannot be matched and therefore must be based upon binocular second-order matching (Pope et al., 1999; Schor, Edwards & Sato, 2001; Wilcox & Hess, 1996).

Based on these ideas, I hypothesize that the second-order pathway involves matching the contrast window of the stimulus and can support depth

perception at disparities where first-order luminance based matching fails. Given the experiments cited above, we can assume that large disparities provide a test case for evaluating the beneficial effects of binocular second-order matching. In Experiments 1-3 I evaluated three different shifted contrast windows (hard-edge, wide cosine and narrow cosine) in extending stereoscopic depth perception to large disparities. These results were consistent with the hypothesis that a shifted contrast window can extend binocular depth perception to larger disparities.

Thus a possible explanation for the discrepant results of Smallman and MacLeod (1997) and Siderov and Harwerth (1993) follows from models of first- and second-order processing. Smallman and MacLeod used filtered noise stimulus within a fixed hard-edge window, as shown in Figure 2A. Binocular disparity is produced by shifting the filtered noise carrier within the fixed hard-edge window. In this type of pattern the only source of depth information is provided by binocular first-order matching of the luminance profile of the pattern (Fleet, Wagner & Heeger, 1996; Prince & Eagle, 2000b; Qian & Zhu, 1997). The prediction from a model based on first-order processing is that the upper disparity limit for stereoscopic performance would be expected to occur at a relatively smaller disparity for higher than for lower spatial frequency patterns, explaining the spatial frequency effect in the results of Smallman and MacLeod.

The Difference of Gaussian (DOG) stimulus used by Siderov and Harwerth (1993) is modulated by a shifted contrast window. Examples of DOG stimuli are shown in Figure 1D. The luminance profile of the DOG is produced by subtracting a broad Gaussian profile from a narrow one to produce a difference of two Gaussian functions; the DOG is sometimes referred to as the "Mexican hat" function and is used to create receptive field profiles (Schor & Wood, 1983). The luminance profile of the DOG function is very similar to a sine wave with a shifted contrast window. Binocular matching in these patterns may be based upon the disparity information in the sinusoidal carrier or in the shifted contrast window. In these patterns second-order binocular matching of the window may have extended stereoscopic performance to larger disparities thus making performance equivalent for the low and high spatial frequency patterns. By comparison, in the fixed window filtered noise stimulus used by Smallman and MacLeod (1997) performance could only be based on binocular matching of the first-order pattern, explaining the poor performance at larger disparities and the prominent effects of spatial frequency. At this point it is useful to consider in further detail models of first-order processing, followed by models of second-order processing in order to motivate the experiments in this thesis.

First-Order Processing (Binocular Energy Model)

The first-order pathway has been modeled using a binocular energy model, that is, a model of binocular complex cells in primary visual cortex which

uses linear filtering to binocularly match luminance features independently in different spatial frequency bands (Fleet et al., 1996; Ohzawa, DeAngelis & Freeman, 1997; Qian & Zhu, 1997; Zhu & Qian, 1996). The receptive fields act as linear filters, with large receptive fields processing low spatial frequencies while smaller receptive fields process high spatial frequencies. The image is filtered in spatial frequency bands and binocular disparity is computed separately in each band. Figure A1 (Appendix) shows a schematic diagram of the binocular energy model and illustrates the computation of disparity by binocular complex cells.

The prevailing model of binocular complex cells is that they have receptive fields covering the same position in both eyes but with differing monocular phases in each input (Ohzawa et al., 1997; Prince & Eagle, 1999; Qian & Zhu, 1997). The largest phase difference which can be represented using differences in monocular phase is equal to a half-cycle of spatial frequency. Thus this model predicts that disparities in vertically oriented, narrow band stimuli may only be detected at up to a half-cycle of spatial frequency and that optimal detection will occur at one-quarter of a cycle. Thus, a key aspect of the binocular energy model is that it predicts the upper disparity limit for depth discrimination to equal approximately a half-cycle of the underlying spatial frequency of the pattern; this half-cycle limit is referred to as the size-disparity correlation (Prince & Eagle, 1999; Qian & Zhu, 1997; Smallman & MacLeod,

1994, 1997). The binocular energy model is described in further detail in Appendix A and this particular version is shown in Figure A1(A).

Insert Figures 1 & 2 about here

Binocular Second-Order Processing

The binocular second-order pathway includes a non-linear sequence of filtering, full-wave rectification and further filtering at a lower spatial frequency in order to extract the contrast window or outline of the stimulus, which are not detected in a first-order linear model (Edwards et al., 2000; McKee, Verghese & Farell, 2004; Sutter, Sperling & Chubb, 1995; Wilson & Kim, 1994). Full-wave rectification, by definition, is mathematically equivalent to taking the absolute value of the filtered response. The stage of full-wave rectification is a non-linearity which is essential to extract the contrast window; however, many other non-linear transforms can be used, such as half-wave rectification or squaring (in which all luminance values in the image are squared) (McKee et al., 2004). In a contrast modulated image this has the effect of producing higher luminance at the centre of the image pattern where the contrast is higher and lower luminance at the edges where contrast is lower. This effectively extracts the contrast window since the resulting image can be binocularly matched in a conventional first-order model at a low spatial frequency which depends upon the width of the original pattern.

The filtered noise patterns shown in Figure 1 are examples of patterns that have variations in luminance as well as contrast since they have been produced from a filtered noise carrier and a contrast window (i.e. contrast modulation or contrast envelope) which drops the contrast of the pattern to zero at the edges. In these images, the contrast window is a hard-edge contrast window (Figure 1A) or is a cosine window (Figure 1B)[1]. Thus a model of second-order processing can extract the hard-edge or cosine window in these patterns.

In general, since second-order binocular matching of the contrast window is performed at a low spatial frequency the disparity limit for depth discrimination based on second-order binocular matching may be larger than that for first-order matching. Furthermore, because stereoscopic performance in the second-order pathway is determined only by the shift of the contrast window, it should not depend upon carrier spatial frequency (Clifford et al., 1998; Wilcox & Hess, 1995). Consequently, the output of the second-order pathway combines with the first-order pathway to extend stereoscopic performance to larger disparities (Edwards et al., 2000; McKee et al., 2004; Schor et al., 2001; Wilcox & Hess, 1995).

Recent work has considered the role of second-order binocular

---

[1] In a hard-edge pattern the contrast drops from maximum to zero over the space of one pixel at the edge of the pattern (Figure 2 and 1A). This is also referred to as rectangular or box windowing. In a cosine window pattern the contrast drops gradually following a cosine function over the space of many pixels (Figure 1B).

matching at large disparities (McKee et al., 2004). McKee et al. tested pedestal depth discrimination with a sine wave carrier and a hard-edge shifted contrast window. It was found that second-order matching of the shifted contrast window extended depth discrimination performance to larger disparities outside the range supported by the first-order luminance based system. The results of McKee et al. were consistent with a model of second-order processing with a first stage of filters, followed by a non-linearity and a second-stage of filters at a lower spatial frequency. One possible interpretation of these results is that there are separate first- and second-order pathways. Alternatively, there may be a single processing pathway which performs both first- and second-order processing, with first-order processing carried out by the first stage of filters and second-order processing carried out by the second stage (McKee et al., 2004).

In Experiment I replicated the discrepant results of Smallman and MacLeod (1997) and Siderov and Harwerth (1993). However, before it is possible to conclude that the differences in these results can be attributed to the effects of contrast windowing and first- and second-order binocular matching, it is necessary to rule out alternative possibilities. In particular, it is necessary to investigate the effects of fixed and shifted windowing using comparable stimuli. Clearly, filtered noise stimuli are quite different in appearance and construction compared to DOG stimuli, as can be noted by comparing Figure 1A (filtered noise) to 1D (Difference of Gaussian). The most

apparent difference is that the filtered noise stimuli are patterns consisting of a two-dimensional (2D) noise carrier whereas the DOG stimuli consist of a one-dimensional (1D) sine wave carrier; in both cases the carrier is modulated by a contrast window. Thus after replicating the results with these two stimulus patterns, the next goal of my thesis was to use a comparable set of stimuli to investigate the effects of spatial frequency and contrast windowing on stereoscopic performance. The stimuli I used for this purpose were 2D filtered noise patterns, similar to those used by Smallman and MacLeod (1997).

I hypothesized that applying a shifted contrast window to filtered noise patterns would make stereoscopic performance more similar for low and high spatial frequency patterns and extend performance to larger disparities. That is, stereoscopic performance with shifted window filtered noise patterns should be more similar to that for DOG patterns. In general, applying a shifted contrast window would make it possible to evaluate the beneficial contribution of second-order contrast window-based matching that could support depth perception at disparities where a first-order luminance based system fails. An example of a shifted contrast window applied to filtered noise is shown in Figure 2B.

Results from Experiment 2 confirmed the expectation that applying a shifted contrast window to filtered noise patterns would reduce the effect of spatial frequency and make it possible to perform the depth discrimination task at larger disparities than with a fixed contrast window. This confirmed that the

shifted contrast window in the DOG stimulus was an important factor extending depth discrimination performance to larger disparities. For the fixed window condition, as stimulus disparity was increased spatial frequency had a prominent effect on stereoscopic performance. Furthermore, performance deteriorated rapidly as stimulus disparity was increased. By comparison, in the shifted window condition spatial frequency had a negligible effect on stereoscopic performance and performance did not deteriorate as stimulus disparity was increased.

The hard-edge window of Experiment 2 (see Figure 1A) contains substantial first-order energy at the edges of the pattern which arises from the abrupt transition between the modulated pattern within the window and the flat gray surround. As this first-order energy is an undesirable attribute of the hard-edge window which can potentially be used to binocularly match the edges of the shifted window, it is important to create a stimulus that does not have a sharp transition in contrast at the edges. A large set of methods from auditory and visual signal processing have been developed to deal with these abrupt transitions in luminance, since problems arising from abrupt transitions in signal energy are common in telecommunications (Hartmann, 1997; Oppenheim & Schafer, 1989). The common solution is to use a ramp or some other gradual transition between the "off" state and "on" state, which are the gray background and the modulated 2D filtered noise carrier pattern, respectively. I used a contrast modulation computed from a cosine function applied to the 2D filtered

noise carrier pattern which completely eliminates residual first-order energy caused by the edges of the contrast window. These cosine windowed stimuli are shown in Figures 1B & 1C. The width of the entire cosine contrast modulated pattern was also manipulated and the rationale for this is discussed below.

In Experiment 3 I further evaluated the contribution of second-order binocular matching in extending stereoscopic performance to larger disparities by using shifted contrast windows. I used shifted wide and narrow cosine windows applied to filtered noise patterns. Baseline performance for binocular first-order matching was provided by the wide cosine fixed window condition in which stereoscopic performance was based predominantly upon binocular matching of the luminance profile of the pattern (i.e. the filtered noise carrier). The wide cosine shifted window was particularly important in that it could provide compelling evidence for the contribution of second-order binocular matching. Computer modeling described in Appendix A showed that a wide cosine contrast window is an improvement over a narrow cosine contrast window because it did not support first-order luminance based binocular matching. That is, although it is important that the cosine patterns have smooth edges, the wide cosine window was better at excluding unwanted first-order energy. Results with narrow cosine window patterns were also relevant because narrow shifted window patterns had been used in previous research

and modeling of the binocular energy function (Prince & Eagle, 2000a, 2000b; Wilcox & Hess, 1995).

Experiments 1-3 plus associated computer simulations described in Appendix A form the core of this thesis. The next two experiments showed directions for further research which I believe are important for elucidating the contribution of first-order and second-order binocular matching in stereoscopic depth perception. Experiment 4 explored how the key variables interact in determining stereoscopic performance while Experiment 5 studied the role of orientation of the filtered noise pattern.

The Half-Shifted Window

In Experiment 4 I compared depth discrimination performance on the fixed and shifted window displays to a third display, the half-shifted window display illustrated in Figure 2C. This display was derived from the fixed window display in that the filtered noise carrier was shifted within the outer edges of the display in the same manner. However, the outer edges of the display only had half the disparity shift compared to the filtered noise carrier. A schematic illustration showing the manner in which the fixed, shifted and half-shifted displays were produced is given in Figure 3. Again, the carrier noise pattern was bandpass filtered at either 2 or 8 cycles per degree (cpd). Examples of these two spatial frequencies are illustrated in Figure 1A.

Insert Figure 3 about here

In Experiment 4 depth discrimination performance for the fixed, shifted and half-shifted window displays was compared to study the contribution of binocular first- and second-order matching. The relative importance of the disparity of the contrast window compared to the disparity of the carrier pattern should be apparent from the results of the half-shifted display. If stereoscopic performance in the half-shifted display depended predominantly on the disparity of the carrier then we would expect performance to match that of the fixed window display. Alternatively if stereoscopic performance in the half-shifted display depended predominantly on the contrast window then we would expect performance to match that of the shifted window display. In fact the results fell approximately midway between these predictions in that stereoscopic performance for the half-shifted window was intermediate to that for the fixed window and shifted window. These results were expected based on the hypothesis that both first- and second-order binocular matching contribute to stereoscopic performance.

I repeated the experiments with the fixed, shifted and half-shifted window using cosine window displays in which the high spatial frequency artefacts at the edges had been removed. Using cosine window displays I was able to reproduce the key trends in the results with hard-edge window patterns

although the absolute level of stereoscopic performance was lower with cosine than with hard-edge displays.

Collectively, the results of Experiments 1-4 were consistent with a model of second-order processing with a first stage of filters, followed by a non-linearity and a second-stage of filters at a lower spatial frequency. The simplest interpretation is that there may be one processing pathway which performs both first- and second-order processing, with first-order processing carried out by the first stage of filters and second-order processing carried out by the second stage (McKee et al., 2004).

The Detection and Integration of Orientation Information

The binocular energy model uses linear filtering to binocularly match luminance features at all orientations in an image (Ohzawa et al.,1997). The image is filtered at each orientation and binocular matching is carried out between corresponding points in the luminance profiles of the left and right eye. (Farell, 2003; Morgan & Castet, 1997; Qian & Zhu, 1997). As shown in Figure 4 I hypothesized that as the orientation of the stimulus pattern becomes progressively more oblique, binocular matching proceeds over a greater distance along the horizontal axis in the image. This would occur because peaks in the horizontal cycle of the pattern are being binocularly matched. Under this hypothesis, the disparity limit for depth discrimination performance would be expected to increase as orientation in the stimulus is progressively more oblique (Farell, 2003; Morgan & Castet, 1997; Qian & Zhu, 1997).

Although this hypothesis is illustrated for sine wave patterns, it is expected to generalize to arbitrary images which do not have evenly spaced luminance peaks. This was studied in Experiment 5 using oriented filtered noise patterns with a range of orientations.

Oriented filtered noise patterns used in Experiment 5 were produced using filters which varied in mean orientation (θ) passed by the filter (see Figure 5B-5H). Based on the hypothesis proposed in this thesis, it was expected that the upper disparity limit for stereoscopic depth perception would be inversely proportional to the sine of the angle θ from horizontal for these oriented patterns. This was studied in a depth discrimination task; it was predicted that depth discrimination performance would extend to larger disparities as orientation in the image is progressively more oblique.

Insert Figures 4 & 5 about here

To summarize and recap, the goal of Experiment 5 was to explore binocular matching of oriented filtered noise patterns with orientation $\theta < 90°$. As expected, depth discrimination performance extended to larger disparities as θ departed from 90°. Observers performed the pedestal depth discrimination task using both fixed and shifted window displays similar to those used in Experiments 1-4. Consistent with earlier results, depth

discrimination performance extended to larger disparities for the shifted compared to fixed window displays.

In Appendix A simulations showed that the results of these psychophysical studies were consistent with a version of the binocular energy model with binocular matching at all orientations in an image. The particular implementation chosen resembled that used for low-level motion perception (Bischof & Di Lollo, 1990, 1991; Prince et al., 2001), providing evidence that motion and stereopsis use similar mechanisms for the detection and integration of orientation information. The processing of orientation information in motion perception is similar to that in stereopsis: motion can be perceived over greater distances as orientation in the image becomes oblique. An implication of this work is that a common set of computational models may be used in further studies of stereopsis and motion (e.g. Hess, Baker & Wilcox, 1999). Furthermore, the similarity in processing between stereopsis and motion has implications for domains which require the integration of motion and stereo information.

General Methods for Experiments 1-5

Observers

The same four observers participated in all experiments. Observers 1-3 were naïve as to the purpose of the experiments. Observer 4 was the author. All observers had good visual acuity and stereoscopic vision.

Training Procedure. Before commencing experimental sessions, all observers were given extensive training on the pedestal discrimination task with trial-by-trial feedback. All the training sessions were completed before the experimental sessions. During the training session, each psychophysical observer performed the depth discrimination task at gradually increasing pedestal disparities (reference depths) with unlimited exposure duration until depth discrimination performance at large pedestal disparities was stable from session to session. Training consisted of 10-15 runs per stimulus condition and spanned a number of days. Each run consisted of 85-300 trials used to obtain a single threshold measurement, using the same adaptive staircase procedure as in sessions for data collection (see details below). Based upon observers' verbal reports, these training sessions were required in order to consistently perceive depth in the stimulus patterns at the exposure duration of 180 ms.

Training is rarely discussed in the literature on stereoscopic vision. A review and meta-analysis of perceptual learning studies (Fine and Jacobs, 2002) only makes reference to one study involving learning in stereopsis, the

study of Fendick and Westheimer (1983). An exception to this was Mayhew and Frisby (1976) who acknowledged the difficulties that observers have with filtered noise patterns. The topic of visual learning in stereoscopic vision was not a part of this thesis; however all subjects including observer 4 (the author) required extensive training to perform the depth discrimination task. The requirement for training seems to be specific to filtered noise patterns; training was not required for other stimulus patterns tested in our research lab including the DOG patterns used in Experiment 1.

Display

All stimuli were generated using a Cambridge Research Systems (CRS) Visual Stimulus Generator (VSG) 2/5 video board and displayed on a high-performance 29" video monitor (EDL 6127) with peak luminance at 85 cd/m^2 (1136 x 851 pixels, 100 Hz frame rate). The monitor was calibrated using the CRS OptiCal system to achieve a linear gray scale. Stereoscopic separation of left and right images was achieved using liquid crystal shutter glasses manufactured by IMAX Corporation, synchronized to the monitor frame rate using an infrared emitter. In the open state the transmission of the shutters was 30%, reducing the effective luminance of the screen accordingly. The viewing distance was 114 cm.

The background luminance was 35 cd/m^2. A zero-disparity frame drawn around the display area (5.2 degrees high and 6.8 degrees wide) and a central Nonius fixation marker (0.35 degrees high x 0.25 degrees wide) were

continuously visible during the experiments. The room was illuminated with fluorescent lights. The light reflected from the blank screen of the computer monitor was 2.7 cd/m^2.

Stimuli

Manner of Producing Disparity. Crossed disparity is defined as the disparity shift of an object nearer the observer than the horopter so that visual lines from the images intersect nearer than the Vieth-Muller circle (Howard & Rogers, 1995). Uncrossed disparity is defined as the shift of an object beyond the horopter; visual lines from the images intersect beyond the Vieth-Muller circle. Crossed disparities are produced by shifting the left stereoview to the right and the right stereoview to the left. Uncrossed disparities are produced by shifting the left stereoview to the left and the right stereoview to the right (Howard & Rogers, 1995). In all cases sub-pixel interpolation is used to create the disparity shift.

Difference of Gaussian (DOG) Patterns. The Difference of Gaussian (DOG) stimulus was generated using the technique described in Schor and Wood (1983) and Siderov and Harwerth (1993). Two spatial frequencies were used, 2 cycles per degree of visual angle (cpd) and 8 cpd. In Equation 1, f represents the spatial frequency in cpd.

$$\mathrm{DOG}(x) = 3\exp(-x^2 f^2 \cdot 15.0234) - 2\exp(-x^2 f^2 \cdot 6.6771) \qquad (1)$$

The constants (15.0234 and 6.6771) yield a DOG with the desired 1.75 octave

bandwidth at half-height, and x represents the horizontal coordinate on the computer monitor in degrees of visual angle. Note that the DOG stimulus has shifted contrast windowing. Disparity was produced by shifting the DOG pattern with sub-pixel accuracy. Figure 1D shows examples of DOG stimuli at the high and low spatial frequency. As shown in the figure the DOG stimuli were truncated at the top and bottom and had a height of 1 degree of visual angle.

Fixed Hard-Edge Window. The fixed window noise patterns were similar to those used by Smallman and MacLeod (1994, 1997). They consisted of isotropic spatially filtered random-dot patterns with a center spatial frequency of 2 cpd or 8 cpd. To avoid ringing in the filter response and to limit the bandwidth of the filtered patterns to $\pm$0.5 cpd, a Kaiser window was applied to the filter kernels (with β=9). Following Kovacs and Fehrer (1997), spurious contrast modulation components were removed from the bandpass patterns. Michelson contrast for the 2 cpd and 8 cpd filtered noise patterns was 80% (calculated as in Howard & Rogers, 1995, pg. 352). In all image patterns, the mean luminance was adjusted if necessary following filtering to match the background grey level of the screen.

Oriented Filtered Noise Patterns. The oriented filtered noise patterns used in Experiment 5 had the same $\pm$0.5 cpd bandwidth as the isotropic filtered patterns but oriented kernels were used with mean orientations of θ equal to

90°, 60°, 30° and 0° as shown in Figure 5. In all cases the orientation bandwidth was equal to 30°.

The width and height of the filtered noise pattern was 3 by 1 degrees of visual angle. In the fixed window patterns, disparity was produced by shifting the noise pattern within the fixed window with sub-pixel accuracy. The edges of the pattern remained fixed at zero disparity (i.e. fixed window). Stimulus width was constant at all magnitudes of disparity. This meant that at one edge of the filtered noise pattern some pixels were removed from the display whereas at the other edge of the pattern new pixels were added. As a result the fixed window stimulus had monocular zones in those regions where the new pixels were drawn. The key aspect of the fixed window stimulus was that only the filtered noise carrier was shifted as disparity was varied and the position of the window remained unchanged. The fixed window stimulus is shown in Figures 2A & 3A.

Shifted Hard-Edge Window. The shifted window noise patterns were produced by shifting the entire pattern. The shifted window stimulus is shown in Figure 2B & Figure 3B. Because the entire pattern was shifted, both the edges of the window and the filtered noise carrier had disparity.

Half-Shifted Hard-Edge Window. The half-shifted window display was derived from the fixed window display. The disparity of the carrier was identical to that in the fixed window display, while the disparity of the contrast window was half that of the carrier. As in the other displays, stimulus width was

constant at all magnitudes of disparity. As a result, the monocular zones were half as large as those in the fixed window display. The key aspect of the half-shifted window display was that stereoscopic performance could be based on either the disparity of the carrier or on the disparity of the contrast window, which were related to each other by a factor of a half. The half-shifted window stimulus is shown in Figure 2C & 3C.

Wide Cosine Window. Cosine window versions of the fixed, shifted and half-shifted displays were produced by applying a cosine contrast modulation to the hard-edge patterns. These are referred to as wide fixed cosine, shifted cosine and half-shifted cosine window displays. The width and height of the wide cosine window displays were 2.5 by 0.75 degrees of visual angle at half height of the cosine contrast modulated edge. Examples of wide cosine window patterns are shown in Figure 1B.

Narrow Cosine Window. Narrow cosine window displays were the same as the wide cosine window displays but had narrower stimulus widths (measured at half height of the cosine contrast modulated edge) equal to 1.4 degrees of visual angle. The vertical height of the stimulus patterns was the same as for the wide cosine window patterns. Examples of narrow cosine window patterns are shown in Figure 1C.

Note that because both wide and narrow cosine window patterns were produced using a cosine function (of differing spatial frequency), this meant that the contrast dropped more abruptly at the edge of the narrow compared to

wide patterns. We verified in further testing with Observers 2 & 4 that in the narrow patterns the shape of the contrast window was not an important factor in determining depth discrimination performance. The other window shapes we used to run tests at this narrow width were hard-edge and Gaussian (computed as in the 1.6 cpd Gabor patterns in Prince & Eagle, 1999).

Design and Procedure

Experimental Trials. At the start of each trial, the observer fixated the Nonius cross and initiated stimulus presentation with a button press when the two halves of the cross appeared to be aligned. Exposure duration was 180 ms. The display consisted of two patterns centered on the screen, displayed one above the other, separated vertically by 0.2 degrees of visual angle (see Figure 2). The top stimulus was always displayed at a pedestal disparity (reference depth) and the bottom stimulus was displayed at a slightly different disparity: the pedestal disparity plus-or-minus a relative disparity. On each trial, subjects indicated with a button press (using the "f" or "b" key on the keyboard) whether the bottom pattern was in-front or behind the top pattern. No feedback was given concerning response accuracy. A staircase algorithm based on PEST (Taylor & Creelman, 1967) was used to estimate the relative disparity threshold at 75% correct. An interleaved staircase procedure was used which consisted of two simultaneous staircases, one for crossed and one for uncrossed pedestal disparities, run in a single session. The two staircases served to measure stereothresholds independently for the crossed disparity

pedestal and uncrossed disparity pedestal. The magnitude of pedestal disparity was constant within a session. Each run consisted of 85-300 trials used to obtain two thresholds for crossed and uncrossed disparities. The larger number of trials was necessary in cases where the observer's responses had higher variability, typically at the largest pedestal disparities. This is typical in adaptive staircase procedures, since the number of trials to converge upon the threshold depends upon the variability in the observer's responses.

The use of interleaved staircases is preferable because observers may inadvertently fixate not on the fixation plane offered by experimental conditions but rather much nearer the stimuli (Howard & Rogers, 1995; Smallman & MacLeod, 1997). This strategy would have turned the large standing disparities in the stimuli into smaller ones, ones small enough for the disparity processing at a given spatial frequency to handle. Simultaneous testing of both crossed and uncrossed disparities makes this strategy less likely to systematically affect results.

Observers were tested at pedestal disparities ranging from 0 to 30 minutes. The spatial frequency of all the patterns was varied at two levels: 2 and 8 cycles per degree (cpd). Subjects were tested at randomized combinations of pedestal disparity and spatial frequency.

Follow-up to Experiment 5: Dmax (Disparity Limit for Depth Perception)

In the follow-up study to Experiment 5, the upper disparity limit for depth perception (Dmax) was measured. It was important to measure Dmax in order

to provide a more precise link between psychophysical data and the simulations of the binocular energy model. The simulations can be used to predict values of Dmax directly but predict stereoscopic thresholds in a depth discrimination task only indirectly. Tsai and Victor (2003) showed that depth discrimination thresholds can be predicted from the binocular energy model by computing a population response for a large set of complex cells tuned to a range of disparities. To decode the disparity represented by the population activity, the population response of the complex cells is compared to a set of templates, one for each disparity. The disparity that is represented by the population activity is taken to be the one that corresponds to a template that minimizes the mismatch. In an approximate way, the modeling shows that thresholds rise (and depth discrimination performance starts to deteriorate) at a value of pedestal disparity close to the half-cycle limit of spatial frequency (Tsai & Victor, 2003). Smallman and MacLeod (1997) also confirmed the same relationship between thresholds in pedestal depth discrimination and predictions from the binocular energy model using a different type of modeling of the response of a population of complex cells tuned to different disparities. In contrast, there is a more direct relationship between the binocular energy model simulations and Dmax: Dmax is predicted as the value of disparity at which the model response drops to zero with increasing disparity (see Appendix A). Based upon these considerations, the simulation results can be

used to quantitatively predict values of Dmax but can only give approximate predictions for depth discrimination thresholds.

Procedure for Dmax. The procedure for measuring the upper disparity limit for depth perception (Dmax) involved using the method of adjustment. In this procedure, a zero disparity reference stimulus and test disparity stimulus were presented on each trial. The observer reported whether the test stimulus appeared to have depth. The depth of the test stimulus was increased until the upper disparity limit for stereopsis was reached (Dmax). Dmax was measured separately for crossed and uncrossed disparities; since no systematic differences were found in these two estimates the average was used as the value for Dmax. These follow-up tests were carried out using fixed window oriented filtered noise patterns at orientations 90°, 80°, 70°, 60°, 50°, 40°, 30° with images similar to those shown in Figure 5. In all cases the orientation bandwidth was 30°. The images were tested at three spatial frequencies: 2 cpd, 4 cpd & 8 cpd. Also isotropic filtered noise patterns similar to the fixed window stimuli shown in Figure 2A were used at the three spatial frequencies. All other aspects of the stereoscopic display and experimental procedure were the same as those used for the pedestal depth discrimination experiments.

Analysis and Presentation of Results

Data are displayed as results for individual subjects on separate graphs with standard error of the estimate calculated from typically 2-6 estimates for each stereo-threshold (more estimates were used for the larger pedestal

disparities because variability was higher). In the field of visual psychophysics, it is common practice to use 2-3 observers and to include the author in this group. Each observer normally serves in thousands of experimental trials and results are presented for each observer independently, with some index of variability. In this thesis I reported results from 4 observers. Each stereothreshold in this thesis was calculated based on approximately 85-300 trials using adaptive methods. A larger number of trials were necessary at larger pedestal disparities, because of the greater variability in the observer's responses. This is typical in adaptive methods since the staircase takes longer to converge upon the threshold if there is more variability in the responses. Because results were presented by individual observer, it was not possible to conduct an analysis of variance (ANOVA). Variability was conveyed by the standard error bars ($\sigma/\sqrt{N}$), where σ is the standard deviation of the 2-6 estimates and N is the number of estimates. Usually these values were smaller than the symbol size representing the mean. Error bars were plotted on all graphs in this thesis except in cases where the error bars were smaller than the symbol size. The low variability was largely a consequence of extensive training prior to actual data collection, as explained earlier. In many cases, the largest error bars were for stereothresholds measured at the largest pedestal disparities (reference depths). This was expected since at these large disparities the depth discrimination task was most difficult.

Experiment 1: Shifted DOG & Fixed Window Hard-Edge Noise

The goal of Experiment 1 was to replicate the results in depth discrimination at pedestal disparities observed with filtered noise stimuli (Smallman & MacLeod, 1997) and with DOG stimuli (Siderov & Harwerth, 1993). Based on previous results I expected that depth discrimination performance with filtered noise patterns would deteriorate at moderate pedestal disparities and the point at which performance would degrade would occur at smaller disparities for high compared to low spatial frequency patterns. By comparison, with DOG patterns stereoscopic performance would be expected to extend to much larger disparities and would not vary with the spatial frequency of the patterns. A further goal was to obtain baseline data for comparison with results obtained with shifted patterns in Experiments 2 and 3. Observers were required to perform the pedestal depth discrimination task with two types of patterns: a) fixed hard-edge filtered noise patterns as illustrated in Figures 1A and 2A, and b) shifted DOG patterns (Figure 1D). Two spatial frequencies were used, 2 cycles per degree (cpd) and 8 cpd, as illustrated in Figures 1A & 1D.

Results

Results are shown in Figure 6 for the DOG and in Figure 7 for the fixed hard-edge filtered noise patterns. Data are shown in separate panels for the four observers. The X-axis represents the pedestal disparity of the reference

pattern while the Y-axis represents the threshold binocular disparity for discrimination of the depth of the reference and test patterns. Each point in the figure was calculated by averaging two to six threshold measurements, each obtained from separate PEST runs. The larger number of threshold measurements was used at the largest pedestal disparities. Error bars represent the standard error of the estimate. Each graph shows results for the 2 and the 8 cpd patterns. Note that some data values were almost identical causing the graphing software to place one symbol on top of the other symbol, making it appear that only one value is represented (e.g. Observer 3, at a pedestal of 30 minutes). Due to limitations of the graphing software this is a problem that may occur in any of the data graphs.

Insert Figures 6 & 7 about here

For the DOG stimulus (Figure 6), performance extended to large disparities and was equivalent for the low and high spatial frequency patterns. By comparison, for the filtered noise stimulus (Figure 7) performance exhibited a dramatic decline at relatively small disparities and depended on spatial frequency. These results replicated previous findings (Siderov & Harwerth, 1993; Smallman & MacLeod, 1997).

Discussion

The point at which depth discrimination begins to degrade in Figure 7 can be estimated as 18-25 minutes for the 2 cpd patterns and 15-20 minutes for the 8 cpd patterns. Each of these disparity values is estimated from the data graph and corresponds to the point on each curve where the function rises from its baseline value. As expected, these trends indicate that depth discrimination performance extended to much larger disparities for the 2 cpd patterns than for the 8 cpd patterns. This was consistent with the notion that spatial frequency was important for the fixed window (Figure 7) but not for the shifted window patterns (Figure 6), a hypothesis that was evaluated further in Experiments 2-3. It was important to replicate these discrepant results as they have implications for the controversy surrounding the size-disparity correlation hypothesis. Smallman and MacLeod (1994, 1997) and Prince and Eagle (1999, 2000a, 2000b) discussed the possibility that these discrepancies may have occurred because of the use of different procedures in previous studies and the failure to control for effects of vergence or other eye movements. A contribution of the present thesis was to confirm the existence of these differences using the same procedures but in a different group of subjects.

Experiment 2: Shifted Window, Hard-Edge Noise

The failure in stereoscopic performance at relatively small disparities observed with fixed window patterns (Figure 7) presumably could be attributed

to the disparity limit on binocular matching of luminance defined features, as predicted by the binocular energy model (Fleet et al., 1996; Prince & Eagle, 2000b; Qian & Zhu, 1997). In Experiment 2 I evaluated the potentially beneficial contribution of a second-order matching process to depth perception at large disparities. Depth discrimination performance was measured for patterns with a hard-edge shifted contrast window as illustrated in Figure 2B. I hypothesized that the shifted contrast window would extend stereoscopic performance to larger disparities and make performance equivalent for low and high spatial frequency patterns. Observers performed the pedestal depth discrimination task with 2 cpd and 8 cpd filtered noise patterns as in Experiment 1. These two spatial frequencies are illustrated in Figure 1A.

Results

Results are shown in Figure 8 with the stereothresholds plotted against pedestal disparity. Figure 8 follows the same format as Figures 6 & 7. The results are shown in separate panels for the four observers. With shifted hard-edge filtered noise patterns performance extended to much larger disparities, compared to the fixed window patterns, as seen by comparing Figures 7 and 8. Another effect of the shifted window was to reduce or eliminate the differences between the 2 and 8 cpd patterns. For two of the four observers (observers 1 & 3), a residual effect of spatial frequency was evident at the largest pedestal disparities. Overall stereoscopic performance with the shifted hard-edge filtered noise patterns was very similar to that observed with DOG patterns in

Experiment 1 (Figure 6), highlighting the importance of the shifted contrast window in these patterns.

Insert Figure 8 about here

Discussion

The results of this experiment confirmed the hypothesis that second-order binocular matching using a shifted window mediated depth perception at disparities where a luminance based system failed. The results of Experiment 2 provide evidence that the spatial frequency effects observed with fixed window patterns may be reduced or eliminated by applying a shifted contrast window to the patterns. However, this raises an important issue: it is important to investigate if stereoscopic performance may be equivalent for all shifted window patterns. The reason why fixed window patterns were used in previous studies (Smallman & MacLeod, 1994, 1997) was to eliminate all monocular cues to depth, thus providing a stimulus that would be processed with first-order binocular matching. This was considered to be an important experimental control. In Experiment 3 the relevant stimulus parameters for shifted window patterns were varied in order to further clarify this issue.

Experiment 3: Shifted Window, Wide & Narrow Cosine

In Experiment 3 I evaluated the contribution of binocular second-order matching using wide and narrow shifted cosine window patterns. Again, baseline performance for binocular first-order matching was provided by the fixed cosine window condition in which stereoscopic performance would be based predominantly upon binocular matching of the luminance profile of the pattern (i.e. the filtered noise carrier). Effects of second-order binocular matching were evaluated in shifted cosine window conditions for wide and narrow patterns. As expected, stereoscopic performance was more similar across the two spatial frequencies (2 and 8 cpd) in the shifted compared to the fixed window patterns. The most important effect to note is that shifted windowing for the cosine stimulus had a beneficial effect on stereoscopic performance compared to the baseline, cosine fixed window condition.

My decision to vary the width of the cosine window patterns at two levels (narrow 1.4° and wide 2.5° visual angle) in Experiment 3 relates to computer simulations of the binocular energy model. Following Prince and Eagle (2000b) I wanted to verify that the binocular energy model could support front/back discrimination with filtered noise patterns that were uncorrelated between the left and right views, in which disparity was conveyed only by the shifted cosine window. It would not be expected that the binocular energy model could discriminate the depth of patterns in which the left- and right-eye views were uncorrelated patterns, because there is no first-order energy in such patterns –

a prerequisite for binocular energy model discrimination. Surprisingly, I found that the binocular energy model could indeed make front/back discriminations with such patterns. In detail, the results of these simulations indicated that depth discrimination performance should be possible over a wide range of disparities, extending up to approximately 30 minutes, at a carrier spatial frequency of 8 cpd.

However I found that the ability of the model to make front/back discriminations with these uncorrelated patterns depended on the stimulus width. The width of stimulus pattern needed to be less than the width of the underlying kernel in the computer simulation. That is, if the stimulus pattern was narrower than the kernel width, then the binocular energy model could perform the discrimination. Conversely, if the stimulus pattern was wider than the kernel width, then the binocular energy model could not perform the discrimination. In detail, the results of these particular simulations indicated that depth discrimination performance should not be possible over a range of disparities up to approximately 30 minutes, again at a carrier spatial frequency of 8 cpd; the output of the binocular energy model averaged to zero over this entire disparity range. Neither Prince and Eagle nor I have a satisfactory explanation for the effect of stimulus width on the simulation results. Nonetheless, the simulation results highlight the importance of using both narrow and wide stimulus widths for cosine windows in Experiment 3. The results of these computer simulations are explained in further detail in

Appendix A. Based on these computer simulations the wide cosine contrast modulated patterns provided a more compelling demonstration of the effects of second-order binocular matching than the narrow patterns.

The exact values for the width of the shifted cosine window patterns were also chosen based on pilot studies in which observers were tested with a range of widths. For wider widths (2.5° visual angle) there was a consistent spatial frequency effect for the observers tested in the pilot study, with depth discrimination performance extending to larger disparities for the low compared to high spatial frequency patterns. This spatial frequency effect was increasingly prominent for wider displays. As I reduced the width of the shifted cosine window display, the spatial frequency effects were gradually reduced. For the narrow width (1.4° visual angle) which has been used in this thesis, the spatial frequency effect was greatly reduced or nonexistent for the observers tested in the pilot study. Therefore the particular widths which were used in Experiment 3 were chosen not only on the basis of the computer simulation results, but also on the basis of the spatial frequency effects in psychophysical testing. The narrow width, in particular, is the width at which the spatial frequency effects are greatly reduced and stereoscopic performance for the high spatial frequency patterns is optimal.

Observers performed the pedestal depth discrimination task with 2 cpd & 8 cpd filtered noise as in Experiments 1-2. The narrow and wide cosine

window patterns are illustrated for these spatial frequencies in Figures 1B & 1C. The wide cosine fixed window served as a baseline condition.

Results

Results are shown in Figures 10 and 11 for wide and narrow cosine window patterns respectively. These results should be compared with those for the fixed window wide cosine patterns which provided a baseline condition (Figure 9). The figures follow the same format as previous figures and plot depth discrimination thresholds against pedestal disparity. There were two effects evident for both the narrow and wide cosine shifted window patterns. Stereoscopic performance extended to larger disparities with the cosine shifted window patterns (both wide and narrow) than with fixed window patterns. This is apparent from comparing the results in Figures 10 and 11 to Figure 9. This was particularly important for the wide shifted cosine window because it provided a compelling demonstration of second-order binocular matching in stereoscopic depth perception. Additionally stereoscopic performance was more similar for the 2 and 8 cpd conditions in shifted than with fixed window patterns. The residual effect of spatial frequency was more pronounced for the wide cosine than for the narrow cosine patterns. This is evident in that the curves for the two spatial frequencies are separated more in Figure 10 than in Figure 11. For all four observers, depth discrimination performance differed for the low and high spatial frequency wide cosine patterns in Figure 10. However, for the narrow cosine patterns a spatial frequency effect was found only for

observer 1, as shown in Figure 11.

Insert Figures 9, 10 & 11 about here

Discussion

Depth discrimination performance was better for the shifted wide cosine (Figure 10) compared to the cosine fixed window patterns (Figure 9), for all four observers. The largest difference was for observer 3 but these effects were apparent for all four observers. This was a key result, since modeling of the binocular energy function demonstrated that the wide cosine window cannot be binocularly matched in a first-order model. These simulations of binocular first-order matching gave virtually identical results for the fixed window patterns and shifted wide cosine patterns. Based on these simulations, stereoscopic performance should have been the same for the fixed window patterns and shifted wide cosine patterns if binocular first-order matching determined performance. The natural conclusion is that binocular second-order matching enhanced stereoscopic performance in the shifted window wide cosine condition. For the shifted wide cosine, performance reflected binocular second-order matching of the contrast window as well as binocular first-order matching of the filtered noise carrier.

In developing the theoretical framework for the present research in terms of first-order and second-order systems, I wanted to rule out the

possibility that the first-order binocular energy model could explain all of the results with shifted window stimuli. This was particularly important because Prince and Eagle (1999, 2000a, 2000b) accounted for much of their own research with shifted stimuli and the research of Wilcox and Hess (1996) using a first-order binocular energy model. In particular, they showed that a first-order model could account for stereoscopic performance with shifted Gaussian modulated noise in which the noise carrier was uncorrelated between the left and the right views. The Gaussian contrast modulation used by Prince and Eagle (2000b) was the same width as the narrow cosine window used in this thesis. Prince and Eagle (2000b) also used a first-order binocular energy model to explain putative second-order effects in studies using Gabor patterns, which were sine wave patterns with shifted or fixed contrast windows. Using the first-order model it was possible to explain a) the oscillations in accuracy of performance with increasing disparity for different envelope sizes of Gabor stimuli (Prince & Eagle 1999, 2000a), b) stereoscopic performance at much larger disparities for Gabor stimuli compared to filtered noise patterns and c) the result that the upper disparity limit for stereopsis (Dmax) for Gabor stimuli increased with envelope size (Wilcox & Hess, 1995).

Depth discrimination performance was poorer for the cosine window display compared to the hard-edge window display. This was most evident for the 8 cpd fixed window display though it was also evident for the shifted display. Compare Figure 9 (fixed cosine window) to Figure 7 (fixed hard-edge

window). There was a large difference in depth discrimination performance for the 8 cpd patterns, particularly for observers 1, 2 & 4. The difference for the 2 cpd pattern was more subtle (observers 2, 3) or non-existent (observers 1,4). Next the effects of applying a cosine window to the shifted patterns can be noted by comparing Figure 8 (shifted window hard-edge) and Figure 10 (shifted window cosine). Again, there was a difference in depth discrimination performance for the 8 cpd patterns (particularly for observers 1, 2 & 4), which was not noticeable for the 2 cpd patterns. From these effects with fixed and shifted window patterns it can be concluded that smoothing the edges of the hard-edge patterns by applying a cosine contrast modulation must have eliminated information useful in binocular matching. The fixed and shifted window cosine patterns were smoothed on both the horizontal and vertical edges, as illustrated in Figure 1B. It was verified with additional pilot studies that smoothing on both the horizontal and vertical edges had an impact on depth discrimination performance.

As a first possible explanation for these effects, it was important to rule out that binocular first-order matching at the edges of the patterns could account for these effects. This might occur because the hard-edge contrast window results in an extra disparity signal at spatial frequencies both higher and lower than the carrier spatial frequency. In particular first-order matching at the edges at a lower spatial frequency could explain these edge effects. In order to rule out this possibility, I tested one person (Observer 2) at lower

contrast values (40%, 30%, 10% and 5% Michelson contrast) using fixed window and shifted window hard-edge patterns. At all values of contrast, including the lowest values, depth discrimination performance was much better for the shifted window compared to fixed window patterns. If the edge effects were due to first-order binocular matching, then at lower values of contrast the edge effects should have been relatively less important compared to the disparity signal at the carrier spatial frequency. Based on the dependence of depth discrimination performance on contrast in these images, it did not seem plausible that the edge effects were due to binocular first-order matching.

The simplest explanation for these effects was that the outline of both the fixed and shifted patterns could be binocularly matched using second-order matching. Recall that models of second-order processing involve three processing levels: an initial filtering stage, a non-linearity (half- or full-wave rectification) and then a second filtering stage. The centre spatial frequencies of the two filtering stages differ, with the second stage typically being lower than the first one by one to two octaves (Edwards et al., 2000; Langley et al., 1999; Schor et al., 2001; Sutter et al., 1995; Wilcox & Hess, 1995; Wilson & Kim, 1994). Consequently, such second-order models are sensitive to local variation in contrast which are present in spatially localized stimuli (Edwards et al., 2000). The result that stereoscopic performance was lower for the cosine window compared to hard-edge patterns indicated that the stimulus widths that

were used in this thesis still allowed the outline of the stimulus patterns to be binocularly matched using second-order processing.

In previous research it was proposed that in order to produce a "pure" first-order stimulus a stimulus display could be used which is spatially extended, preferably as wide as possible (Edwards et al., 2000). In order to test this idea, I carried out further pilot studies using a much wider extended stereoscopic display with a width equal to 12.0 degrees of visual angle at a viewing distance of 114 cm. Stereoscopic depth perception with fixed window displays at this width was approximately the same as for the fixed window wide cosine patterns used in this thesis. These additional tests served to confirm that the fixed window wide cosine patterns used in this thesis were the stimulus patterns which most closely corresponded to a "pure" first-order stimulus, because smoothing the edges of the pattern reduced the impact of second-order binocular matching of the stimulus outline. Thus the fixed window wide cosine pattern was the most appropriate stimulus to estimate depth discrimination performance based upon first-order luminance based matching. Interestingly, second-order binocular matching of the outline of the stimulus was an important factor for both the fixed and shifted window patterns.

Presumably the hard edge provided a discontinuity which compared to the cosine edge could be more readily detected by second-order processing. Consistent with this, the hard-edge pattern appears distinct from the background while the cosine pattern blends into the background, as seen in

Figures 1A & 1B. The hard-edge contrast windows activated second-order filters at a range of different spatial frequencies so the cosine contrast windows were preferable in order to activate second-order filters at a particular spatial frequency.

Using cosine window patterns allows us to determine estimates for the spatial frequency tuning of the second-stage filters in the second-order processing pathway. Based on the results with the 8 cpd wide cosine and hard-edge shifted window patterns we may conclude that the width of the cosine window was too wide for the carrier spatial frequency of 8 cpd and the width needed to be reduced for optimal performance. Indeed, this was found to be true since stereoscopic performance for the 8 cpd carrier improved with the narrow cosine window and became equivalent to stereoscopic performance for the 2 cpd carrier with the wide cosine. For an 8 cpd filtered noise carrier the optimal stimulus width was the width of the narrow cosine (1.4 degrees) while for the 2 cpd carrier the optimal width was that for the wide cosine (2.5 degrees). Using these widths we can calculate the spatial frequency tuning of the second-stage filters in the second-order processing pathway as follows. The first-stage 8 cpd filters feed into second-stage filters tuned to a spatial frequency of 1/(1.4)=0.71 cpd, while 2 cpd filters feed into 1/(2.5)=0.40 cpd filters. These results are consistent with prevailing models of the second-order pathway in which first-stage filters are linked to second-stage filters at a much lower spatial frequency (Edwards et al., 2000; Langley et al., 1999; Schor et al.,

2001; Sutter et al., 1995; Wilcox & Hess, 1995; Wilson & Kim, 1994). Further experiments could be carried out at a range of carrier spatial frequencies and stimulus widths in order to determine if there is a systematic relationship between the spatial frequency tuning of first- and second-stage filters.

The model for second-order processing used here is similar to models of second-order processing in the texture literature which also employ two stages of filtering; cosine contrast modulations are used to isolate second-order filters at a particular spatial frequency (Kingdom & Keeble, 1996; Motoyoshi & Nishida, 2001; Prins & Kingdom, 2002; Zhou & Baker, 1993, 1994).

As discussed above, the fixed cosine window can be considered to be the stimulus which is most representative of binocular first-order matching. This means that the results in Figure 9 can be compared with the half-cycle limits predicted by the size-disparity correlation (Prince & Eagle, 1999; Smallman & MacLeod, 1997). The point at which depth discrimination performance begins to degrade in Figure 9 occurred at disparity values of approximately 0-3 and 11-23 minutes, respectively. Empirically the point at which depth discrimination performance degrades is estimated from each curve in Figure 9 and corresponds to the point on the curve where the function rises from its baseline value. The half- cycle limit for the 2 cpd and 8 cpd patterns was expected to equal 15 minutes and 3.75 minutes respectively. While stereoscopic performance varied with the spatial frequency of the pattern in fixed window displays it appeared that stereoscopic performance appeared to extend to

slightly larger disparities than expected (Prince & Eagle, 1999; Smallman & MacLeod, 1997). Based on modeling work, it is expected that depth discrimination performance should not be possible at pedestal disparities above the disparity limit predicted by the size-disparity correlation (Tsai & Victor, 2003). However, the relationship between the point at which depth discrimination performance degrades and the half-cycle limit is only approximate. This issue will be discussed again in the context of Experiment 5.

Experiment 4: The Half-Shifted Window

In Experiment 4 I introduced a new display referred to as the half-shifted window, shown in Figure 2C. In the half-shifted display the hard-edge or cosine contrast modulated window was shifted by half the disparity of the filtered noise carrier. In Experiment 4 I compared stereoscopic performance in the half-shifted display with the fixed and shifted displays at two spatial frequencies (2 and 8 cpd) and with two window shapes (hard-edge and cosine contrast modulated). The half-shifted display made it possible to assess the relative importance of these factors in determining stereoscopic performance in the pedestal discrimination task.

Figure 3 provides more information about the half-shifted display by comparing it to the construction of the more traditional fixed and shifted displays. In the fixed window display binocular disparity was produced by shifting the carrier within the fixed edges of the display. This produced

monocular zones that were uncorrelated between the left and right views of the pattern ensuring that the width of the pattern was constant at all disparities. In Figure 3 the monocular zones are depicted with grey regions. In the shifted window display, binocular disparity was produced by shifting the entire pattern, so that the edges had disparity and there were no monocular zones. The half-shifted window display was derived from the fixed and the shifted display types in that the carrier pattern was shifted by the full disparity while the outer edges of the pattern were shifted by half this disparity. The carrier noise pattern was bandpass filtered at either 2 or 8 cpd. Examples of these two spatial frequencies are illustrated in Figure 1A.

I hypothesized that the relative importance of the disparity of the contrast window compared to the disparity of the carrier pattern would be readily apparent in the results for the half-shifted display. If stereoscopic performance in the half-shifted display depended predominantly on the disparity of the carrier then we would expect performance to match that of the fixed window display at the corresponding 100% disparity of the carrier. Conversely, if stereoscopic performance in the half-shifted display depended predominantly on the contrast window then we would expect that performance would match that of the shifted window display, at the corresponding 50% disparity of the shifted window. These effects would be expected to be affected by the shape of the window (hard-edge or cosine contrast modulated), but the overall trends with a cosine window should be the same as for the hard-edge patterns. For

example, we would expect that the absolute level of performance would be lower with cosine than with hard-edge displays and the effects of windowing at 8 cpd would be more evident for the cosine windowed condition and less evident for the hard-edge windowed condition, in agreement with the effects of cosine windowing in Experiment 3.

As it turns out results fell approximately midway between these predictions in that that stereoscopic performance in the half-shifted condition did not match that of the fixed condition, nor were they equivalent to those in the shifted condition at 50% of the nominal disparity value. Stereoscopic performance for the half-shifted window was intermediate to that for the fixed and shifted window patterns. For cosine window patterns stereoscopic performance for the half-shifted window was again intermediate to that for the fixed and shifted window patterns but the absolute levels of performance were lower. As in the case of hard-edge patterns this trend was more apparent for high spatial frequency patterns. These results were to be expected based upon results from Experiments 1-3, since the interpretation is that both first- and second-order binocular matching contributed to stereoscopic performance.

To summarize, in Experiment 4 fixed, shifted, and half-shifted window displays were used to study the contribution of first- and second-order binocular matching. The fixed and shifted window conditions from Experiments 1-2 provided benchmark levels of stereoscopic performance. Theoretically the fixed window display can be considered to reflect binocular first-order matching

while the shifted window can be considered to reflect both first-and second-order binocular matching. The half-shifted display was designed to determine the relative contribution of first and second-order matching by further comparing the effects of spatial frequency, window shift and window shape on stereoscopic performance. On the basis of these assumptions the results suggested that neither first- nor second-order matching dominated processing in the half-shifted condition. Effects of both the carrier and the window were evident in the results. The shifted window caused an improvement in stereoscopic performance reflecting the beneficial role of binocular second-order matching of the edges of the pattern. The carrier limited the extent of this improvement, reflecting the role of binocular first-order matching of the carrier pattern.

The relevance of Experiment 4 was to show that in the presence of two inconsistent disparity signals (one from the carrier and one from the window), both signals made a contribution to stereoscopic performance, leading to the conclusion that both first- and second-order binocular matching contributed to depth discrimination performance at large disparities. A simpler alternative, that only second-order binocular matching contributes to depth discrimination performance for any pattern where the window is shifted, is not supported by the results of this thesis. Instead I found support for the more complex alternative that both binocular first-order and second-order matching made a contribution in determining stereoscopic performance at larger disparities for

shifted window patterns. These results will be discussed in terms of first- and second-order processing pathways. The results were consistent with a model of second-order processing with a first stage of filters, followed by a non-linearity and a second-stage of filters at a lower spatial frequency. However, it is possible that there is a single processing pathway which performs both first- and second-order processing, with first-order processing carried out by the first stage of filters and second-order processing carried out by the second stage.

To summarize, observers performed the pedestal depth discrimination task as in Experiments 1-3 using the half-shifted window display at the low and high spatial frequencies and cosine window versions of these displays.

Results

Results with the half-shifted window patterns are shown in Figures 12 and 13 for low and high spatial frequency patterns respectively. The X-axis represents the pedestal disparity while the Y-axis represents the threshold binocular disparity for depth discrimination. Separate panels show results for the four observers. Note that in Figures 12-15 the data for the fixed and shifted window patterns has been replotted from the earlier figures since this is data from Experiments 1-3. For the low spatial frequency patterns (Figure 12) depth discrimination performance did not differ across the three conditions, except for observer 3. For observer 3, performance in the half-shifted condition was intermediate to the fixed and shifted window conditions. For the high spatial

frequency patterns (Figure 13) performance in the half-shifted condition was intermediate to the fixed and shifted conditions for all four observers.

The results with cosine window versions of the half-shifted window display are shown in Figures 14-15. These graphs follow the same format as Figures 12-13. As in the case of the hard-edge patterns, performance in the half-shifted condition was intermediate to the fixed and shifted conditions. Again this trend was more apparent for high spatial frequency patterns (Figure 15) than low spatial frequency patterns (Figure 14). For the low spatial frequency patterns this trend was only apparent for observer 3. Generally stereoscopic performance was worse for cosine window patterns compared to hard-edge patterns, as expected given the results of Experiment 3.

Insert Figures 12, 13, 14 & 15 about here

These results may be interpreted as follows. The half-shifted condition can be compared with the fixed and shifted conditions which set a lower and upper bound on performance. In the fixed window condition disparity thresholds exhibited a rapid rise with increasing pedestal disparity. This point at which performance degraded occurred at smaller disparities for the high compared to low spatial frequency patterns. In the shifted window condition stereoscopic performance extended to larger disparities compared to the fixed window condition. For this condition performance reflected binocular second-

order matching of the contrast window as well as binocular first-order matching of the filtered noise carrier. In the half-shifted window, stereoscopic performance still varied with spatial frequency, but extended to larger disparities compared to the fixed window condition, indicating that both first- and second-order processing contributed.

Discussion

From the results it was evident that performance in the half-shifted window patterns was intermediate to that for the fixed and shifted window patterns. That is, performance was not equivalent to that in the shifted condition even though the half-shifted window provided a clear disparity signal. This indicated that binocular matching of the contrast envelope enhanced performance relative to the fixed window condition, but the degree of improvement was sub-optimal. We may conclude that the output from the binocular second-order pathway contributed an enhancement to depth discrimination performance but did not dominate performance. Since there was a prominent effect of spatial frequency in the half-shifted window patterns it was apparent that both first- and second-order binocular matching contributed to depth discrimination at large disparities. Overall, the conclusions of Experiment 4 were consistent with Experiments 1-3.

The key point of the results from Experiments 1-4 was that the second-order system extended stereoscopic performance to larger disparities. However, one scenario that could be readily dismissed was that the second-

order system could entirely override the first-order system. If this was the case, then stereoscopic performance for low and high spatial frequency patterns would have been the same for all types of shifted windows. Furthermore, it is possible to dismiss the possibility that the extraction of the contrast window precedes binocular matching of the carrier. Evidence for this processing scheme was found in studies of stereoscopic performance near the horopter (i.e., zero disparity) with uncorrelated noise patterns (Wilcox & Hess, 1996). However, near the horopter the contributions of the first-order and second-order systems were likely more difficult to tease apart as the performance of both systems was optimal at these small disparities. In pilot studies I found that depth discrimination was not possible with uncorrelated shifted window patterns if the pedestal disparity was greater than zero. Evidently, a more complex scheme was necessary to describe first- and second-order pathways.

Recent work also found that second-order binocular matching extended depth discrimination performance to larger disparities using sine wave patterns in a shifted hard-edge contrast window (McKee et al., 2004). Binocular matching of the edges of the sine wave pattern was clearly important in accounting for the results of this study, but could have involved either first- or second-order processing. McKee et al. ruled out first-order binocular matching in accounting for stereoscopic performance at large disparities since the contribution of edge matching was still present even with sine wave patterns presented with very low (5%) contrast. McKee et al. proposed that the type of

stereo mechanism required to explain the edge matching either must respond to a substantial spatial region (e.g. be tuned to a very low spatial frequency) or must operate functionally in a different way from the first-order mechanisms (e.g. be able to influence the responses at distant locations), but probably needs both properties. Thus McKee et al.'s results are consistent with those of this thesis in proposing that binocular first-order matching cannot account for all of the effects observed with shifted window patterns. McKee et al. proposed that binocular first- and second-order processing operate in parallel and combine their outputs (McKee et al., 2004). Collectively, the results of this thesis and the study of McKee et al. provide strong evidence for the importance of binocular first- and second-order matching in stereoscopic depth perception.

A direction for future research would be to study and characterize the mapping and pooling of first-order and second-order systems in binocular vision (Farell, Li & McKee, 2004a, 2004b; McKee et al., 2003, 2004). This work could draw upon the rich set of studies of second-order effects in motion and texture perception (e.g. Chubb et al., 2001). The general model would need to account for results showing that a contrast window was ineffective in supporting depth perception when the carrier differed in spatial frequency in the two eyes, or when there was no overlapping spatial frequency content in the two views (Edwards et al., 2000; Mayhew & Frisby, 1976). Edwards et al. proposed that contour information from the first-order and second-order paths was extracted separately and that stereoscopic performance was based on the

combined output of the two paths. This proposal is similar to schemes suggested in the texture discrimination and segregation literature (e.g. Sutter et al., 1995) and in the motion literature (e.g., Boulton & Baker, 1993; Clifford et al., 1998; Sato, 1998) in which perception is based on the combined outputs of first- and second-order pathways.

In agreement with previous research, the results of Experiments 1-4 provide evidence for first- and second-order pathways. However, the most parsimonious interpretation which fits all previous results is that the first- and second-order pathways are not separate but instead function as a single pathway which performs both first- and second-order processing. Compatible with this, the results of Experiments 1-4 showed that depth discrimination for the low and high spatial frequency patterns was affected by the width of the patterns. For the shifted window wide cosine patterns there was a prominent effect of spatial frequency. This indicated that a model of second-order processing with a first stage of filters, followed by a non-linearity and a second-stage of filters at a lower spatial frequency gave a satisfactory account of these results. The first-stage filters serve to perform first-order processing so that there was no evidence from these experiments of an extra parallel processing pathway. The spatial frequency effects for the shifted window wide cosine patterns are therefore consistent with a two-stage filtering model which includes both first- and second-order processing.

Experiments 5: Oriented Filtered Noise

One feature of the results in Experiments 1-4 was unexpected. While the point at which depth discrimination performance degraded varied with the spatial frequency of the pattern in fixed window displays the magnitude of this variation was not numerically consistent with the half-cycle limit predicted by the size-disparity correlation (Prince & Eagle, 1999; Smallman & MacLeod, 1997). That is, while the half- cycle limit for the 2 cpd and 8 cpd patterns was expected to equal 15 minutes and 3.75 minutes respectively, the point at which depth discrimination performance begins to degrade in Figure 7 occurred at disparity values of approximately 18-25 and 15-20 minutes, respectively. With a cosine window the values were approximately 0-3 and 11-23 minutes, respectively. Empirically the point at which depth discrimination performance degrades is estimated from curves in Figures 7 and 9 and corresponds to the point on each curve where the function rises from its baseline value.

Based on the model of depth discrimination of Tsai and Victor (2003) we might expect that depth discrimination performance should not be possible at pedestal disparities above the disparity limit predicted by the size-disparity correlation hypothesis. Of the four values observed, three were above their theoretically expected limits; that is depth discrimination could be performed at disparities above those predicted numerically by the half-cycle limit. Of course the point at which depth discrimination performance is degraded gives only a rough estimate of the disparity limit for depth perception. A goal of Experiment

5 was to ascertain with greater precision whether stereoscopic depth perception is in fact possible at disparities above the half-cycle limit of spatial frequency. In particular Experiment 5 explored the possibility that the half-cycle limit for the hard-edge and cosine windowed patterns was exceeded because the stimulus pattern was isotropically filtered.

As shown in Figure 5A, the isotropic filtered noise patterns were created using filters which passed all orientations equally, within a band of spatial frequencies between the upper and lower cut-off spatial frequencies. However, this ideal bandpass filter may also be used to create oriented (anisotropic) filtered noise patterns by limiting the filter to a band of orientations but using the same upper and lower cut-off spatial frequencies, as shown in Figures 5B-5H. In Experiment 5 I investigated the hypothesis that anisotropically filtered noise patterns might have a different disparity limit compared to the isotropically filtered patterns used in Experiments 1-4. Under this hypothesis depth discrimination would extend to larger disparities for oblique patterns.

The binocular energy model uses linear filtering to match luminance features at all orientations in an image (Ohzawa et al.,1997). The image is filtered and binocular matching is carried out between corresponding points in the luminance profiles of the left and right eye (Farell, 2003; Morgan & Castet, 1997; Qian & Zhu, 1997). As the orientation of the filter departs progressively from $\theta = 90°$ and the filtered image takes on the appearance of an oblique pattern (Figure 4), binocular matching might proceed over a greater distance

along the horizontal axis in the image. This would occur because peaks in the horizontal cycle of the pattern are being binocularly matched. As a result the upper disparity limit for depth perception should extend towards larger disparities as orientation in the stimulus pattern becomes progressively more oblique (Farell, 2003; Morgan & Castet, 1997; Qian & Zhu, 1997). Indeed, the results of Experiment 5 indicated that the orientation of the spatial frequency filter had the expected effect on stereoscopic performance in the pedestal discrimination task: depth discrimination performance was possible at larger disparities as the pattern became more oblique.

Oriented bandpass filtered noise patterns used in Experiment 5 were produced using filters which varied in central spatial frequency and mean orientation (θ). Figure 5A illustrates the two-dimensional Fourier domain of an image showing an ideal bandpass filter with cut-off spatial frequencies F_L and F_H. A particular spatial frequency component had a vertical spatial frequency V_s and a horizontal spatial frequency U_s. A key feature of the isotropically filtered stimulus was that the horizontal spatial frequency U_s could be lower than the lower cut-off spatial frequency (F_L) passed by the filter[2] and therefore also lower than the centre spatial frequency passed by the filter. I hypothesized that the upper disparity limit for stereoscopic depth perception would be inversely proportional to the sine of the angle θ.

[2]For example, 8 cpd patterns in Experiment 1 were created with a filter with centre spatial frequency 8 cpd and lower and upper cut-off frequencies 7.5 and 8.5 cpd. A particular 8 cpd oriented component with orientation 60° had vertical spatial frequency 8*cos(60°)=4 cpd and horizontal spatial frequency 8*sin(60°)= 6.93 cpd which was lower than the lower cut-off spatial frequency and centre spatial frequency.

Intuitively, as illustrated in Figure 4 for sine wave gratings, binocular matching may proceed over a greater distance horizontally as the sine wave takes on a more oblique appearance because the horizontal period (λ_{hor}) increases in inverse proportion to the sine of the angle (Farell, 2003; Morgan & Castet, 1997; Qian & Zhu, 1997). Based on this hypothesis, the upper disparity limit for depth perception is predicted to equal to a half-cycle of horizontal spatial frequency, instead of a half-cycle of spatial frequency from the size-disparity correlation hypothesis (Prince & Eagle, 1999; Smallman & MacLeod, 1994, 1997). Figure 4C illustrates the calculation of the half-cycle of horizontal spatial frequency for oriented sine waves based on trigonometry. Figure 4 shows only sine wave gratings for clarity of exposition but the hypothesis is expected to generalize to arbitrary images which do not contain evenly spaced peaks.

As a further rationale for the Experiment 5, note that low-level motion detection is modeled using detection of perpendicular phase shifts (Adelson & Bergen, 1985; Bischof & Di Lollo, 1990, 1991; Prince et al., 2001; Simoncelli & Heeger, 1988; van Santen & Sperling, 1985). Each oriented motion detector responds to motion in a direction perpendicular to its orientation (Adelson & Bergen, 1985; Bischof & Di Lollo, 1991; van Santen & Sperling, 1985). It has been shown in the literature on motion perception using filtered random-dot kinematograms that coherent motion can be reliably perceived up to a limit know as D_{max} which corresponds to the half-cycle displacement limit of the

spatial frequency of the stimulus. Beyond this limit coherent motion can no longer be seen reliably. Interestingly, D_{max} increases as the orientation of the image filter departs from 90° and the image takes on the appearance of progressively more oblique bands, like those illustrated in Figures 4 & 5 (Bischof & Di Lollo, 1991; Prince et al., 2001). This occurs even though the oblique orientations in the image are not perpendicular to the direction of motion. Indeed, the oblique orientations allow D_{max} to be extended beyond the half-cycle displacement limit of the nominal lowest spatial frequency in the image (Bischof & Di Lollo, 1990, 1991; Prince et al., 2001).

I hypothesized that the models that have been applied to motion perception may also be applied to stereoscopic depth perception. In Appendix A simulations of the binocular energy model (Fleet et al., 1996; Qian & Zhu, 1997) followed motion perception modeling by Bischof and DiLollo (1991) which incorporated detection of perpendicular shifts at all orientations in an image. This gave a possible explanation for the results of Experiment 5 which showed that depth discrimination performance extended to progressively larger disparities as orientation in the image patterns was progressively more oblique. Furthermore, it gave an explanation for stereoscopic performance in Experiments 1-4 at large disparities, above the half-cycle limit. In the isotropic filtered patterns the oblique components had the effect of extending depth discrimination performance to larger disparities. The results of Experiment 5

were useful in providing evidence that stereopsis and motion use similar mechanisms for detection and integration of oriented components in an image.

To recap and summarize, observers performed the pedestal depth discrimination task as in Experiments 1-4 with oriented filtered noise patterns. Filtered noise patterns similar to those shown in Figure 5 were used, with mean orientation θ passed by the filter equal to 90° (vertically filtered), 60°, 30° and 0° (horizontally filtered). These were bandpass filtered noise patterns filtered at a low or high spatial frequency (2 cpd or 8 cpd). Under the hypothesis proposed in this thesis the results should indicate that the upper disparity limit for depth perception varies with the orientation of image components and is inversely proportional to the sine of the angle (θ) of orientation from horizontal. Observers performed the pedestal depth discrimination task using both fixed and shifted window displays similar to those used in Experiments 1-4. The point at which depth discrimination performance degrades was used as an estimate to compare with the theoretical prediction for the disparity limit for depth perception. Consistent with expectations, the results of Experiment 5 confirmed that depth discrimination performance extended to larger disparities for shifted compared to fixed window patterns for oriented noise patterns.

Empirical testing and modeling was performed using isotropic as well as anisotropic, oriented filtered noise patterns. A model similar to that used in motion perception was used to successfully model the results, suggesting that stereopsis and motion use similar processing for the detection and integration

of information across different orientations in the image (Bischof & Di Lollo, 1990, 1991; Prince et al., 2001). The revised model including anisotropic processing was applied to the results of Experiments 1-5 in order to show that the modified model was capable of accounting for results both with isotropic and with anisotropic patterns.

Results

Results for the oriented filtered noise patterns are shown in Figure 16-19, with the stereoscopic thresholds plotted against pedestal disparity. Individual curves on the graphs represent data from filtered noise patterns with orientation θ equal to 90° (vertically filtered), 60°, 30° or 0° (horizontally filtered). Note that since these are filtered noise patterns the value of orientation indicates mean orientation passed by the filter. With fixed window hard-edge filtered noise patterns shown in Figures 16-17, depth discrimination performance extended to larger disparities for the horizontally filtered compared to vertically filtered patterns, with data for the other orientations falling in between. Compare the position of the curves with open circles (vertically filtered, θ=90°) to the curve with filled diamonds (horizontally filtered, θ=0°). These differences were much greater for the 8 cpd patterns than for the 2 cpd patterns. This is evident in that there is a greater separation between the curves in Figure 17 compared to Figure 16. For both the 2 cpd and 8 cpd patterns the differences were most apparent at the largest pedestal disparities where the curves were most widely separated. Overall, the results provided

evidence for the hypothesis that depth discrimination performance varies with orientation and extends to larger disparities for oblique orientations.

Insert Figures 16, 17, 18 & 19 about here

The same trends were evident in the data for the shifted window patterns in Figures 18-19. In Figure 19 showing the data for the 8 cpd patterns, note that the position of the curves with open circles (vertically filtered, θ=90°) is again higher than the curve with filled diamonds (horizontally filtered, θ=0°), and the other orientations fall in between these curves. For the 2 cpd patterns, however, there was no difference in depth discrimination performance for filtered noise at different orientations (Figure 18). For both the 2 cpd and 8 cpd patterns depth discrimination performance was better overall for the shifted compared to fixed window patterns. Compare the curves in Figure 18 to Figure 16 (fixed versus shifted 2 cpd patterns) and Figure 19 to Figure 17 (fixed versus shifted 8 cpd patterns). As in Experiments 1-3, the shifted window extended stereoscopic performance to larger disparities. Furthermore another effect of the shifted window was to reduce the differences between the low and high spatial frequency patterns.

Discussion

Compatible with the hypothesis proposed in this thesis, depth discrimination performance extended to larger disparities for oblique

orientations. The most obvious differences were in the high spatial frequency patterns, in particular for fixed window patterns. By comparison, for the low spatial frequency patterns depth discrimination performance did not show as much variation with orientation. This might be expected since the disparity range which was tested represented a much larger range in terms of phase disparities (as opposed to disparities in minutes), for the high compared to low spatial frequency patterns. For example, disparities ranging from 3.75 minutes to 15 minutes are equivalent to phase disparities of 0.5 to 2 cycles for the 8 cpd patterns, but are equivalent to phase disparities of only 0.125 to 0.5 cycles for the 2 cpd patterns. The results of the simulations of the binocular energy model give some support for this explanation, as discussed below.

Simulations of the binocular energy model (Fleet et al., 1996; Qian & Zhu, 1997) were performed with oriented filtering and an implementation chosen to be similar to typical motion perception models which incorporated detection of perpendicular shifts at all orientations in an image. Simulations were carried out on image patterns similar to the filtered noise patterns used in Experiment 5 with filtering using kernels at different orientations. Simulations are shown in Figures A9-A12. The simulations are important in illustrating the predicted variation of the upper disparity limit for stereoscopic performance with spatial frequency and orientation of the kernel. For each spatial frequency pattern the disparity limit gradually increased as the orientation of the kernel was progressively more oblique, a pattern which fits the psychophysical results.

Figure A13 compares two sets of theoretical predictions of the upper disparity limits for stereoscopic depth discrimination with the values obtained in the simulations of the binocular energy model. The Y-axis shows the upper disparity limit in minutes of visual angle for the binocular energy model response (filled squares), compared to the prediction of a half-cycle of horizontal spatial frequency (filled circles) and the prediction of a half-cycle of spatial frequency from the size-disparity correlation (dashed line). The X-axis shows the orientation of the kernel. The upper disparity limit for the model was obtained from the simulation output curves shown in Figures A9-A12. The disparity values from the simulations gave an excellent fit to the theoretical predictions, although values were in some cases slightly below the predicted values. The most important trend to note is that the upper disparity limit is inversely proportional to the sine of the angle of the orientation from horizontal. This validates the use of the binocular energy model as appropriate for the theoretical predictions in this thesis. The hypothesis illustrated in Figure 4 used essentially a geometrical argument to calculate the variation of the upper disparity limit with orientation. The simulations showed that the increase in disparity limit for oblique orientations is a consequence of a processing stage in the binocular energy model which uses detection of perpendicular shifts for disparity computations.

The results of Experiment 5 for oriented filtered noise patterns can be compared with the predictions of the simulations of the binocular energy model

in Figures A13 & A14. As an example, compare the simulation results in Figure A14 from the position shift model to the psychophysical data. In Figure A14, the disparity limits for depth perception for the 8 cpd oriented patterns were equal to 5.2, 5.9 and 8.6 minutes for orientations equal to 90°, 60° and 30°. These values should be compared against the curves in Figure 17; the point at which depth discrimination performance degrades is estimated from each curve as the point at which the curve rises from its baseline value. It should be emphasized that these estimates are approximate. For these high spatial frequency patterns, the predicted disparity limits gave a reasonably close fit to point at which performance degrades on the psychophysical curves.

For the low spatial frequency patterns, depth discrimination performance was much more similar for patterns with different orientations. For example, the curves in Figure 16 (fixed window) are close to one another and the curves in Figure 18 (shifted window) lie on top of one another. These results were consistent with modelling of the binocular energy function since the disparity limits for these low spatial frequency patterns were predicted to be high enough that good performance in the stereoscopic depth discrimination task should be possible at the largest pedestal disparities. In detail, the upper disparity limits for the 2 cpd patterns from Figure A14 were predicted to be 21.3, 24.6 and 35.5 minutes for orientations 90°, 60° and 30°. An examination of Figure 16 shows that the depth discrimination curves were consistent with these disparity limits. That is, the point at which depth discrimination performances degrades is

approximated by these disparity limits. The curves in Figure 18 are for shifted patterns and we may conclude that second-order matching extended stereoscopic performance to larger disparities.

Up to this point, the upper disparity limits for depth perception obtained from the binocular energy model simulations were compared to the point at which depth discrimination performance degrades, estimated from the depth discrimination curves; each of these points is estimated from a depth discrimination curve as the point at which the curve rises from its baseline values. An alternative method to relate the binocular energy model simulations to psychophysical results is to measure the upper disparity limit for depth perception more directly in a different psychophysical task. Therefore, as a follow-up study to Experiment 5, I used a different stereoscopic depth perception task to obtain an independent estimate for the upper disparity limit for depth perception for the oriented filtered noise patterns. In this task, which uses the method of adjustment, the observer reported whether the stimulus pattern has depth and the depth of the stimulus pattern was gradually increased until depth was no longer perceived. The upper disparity limit obtained in this way is referred to as Dmax and can be compared directly with the upper disparity limit predicted from the simulations of the binocular energy model. Thus the use of the method of adjustment to measure Dmax provides an important test of the fit between the binocular energy model simulations and

the psychophysical results. In particular it serves to verify that the binocular energy model is appropriate to provide a fit to the psychophysical data.

In Figure 20, the values of Dmax for oriented filtered noise patterns at orientations ranging from 90° (vertically filtered) to 30° are compared to predictions from the simulations. In Figure 20, the X-axis shows the orientation of the filtered noise patterns and the Y-axis shows the value of Dmax in minutes. Separate curves are drawn for the values of Dmax obtained at spatial frequencies of 2 cpd, 4 cpd & 8 cpd (filled circles, filled triangles, filled squares). The predicted values of Dmax obtained from the simulations are plotted with dashed lines (three dashed lines in each panel correspond to the 2 cpd, 4 cpd & 8 cpd patterns). The values of Dmax for isotropic filtered noise patterns are also plotted on each panel for the three spatial frequencies (open circles, open triangles, open squares, labelled "iso").

Several trends are evident from inspecting these psychophysical data and making a comparison to modelling results. First, inspection of the psychophysical data reveals a prominent effect of spatial frequency for these images, which is consistent with our earlier results with fixed window patterns. The extra spatial frequency of 4 cpd was used in this study because there was such a large difference between the Dmax values for the 2 cpd and 8cpd patterns. This fits with my earlier interpretations that the limit for luminance-based matching for the 8 cpd patterns occurs at very small disparities. Next, comparing the Dmax values to the simulation results reveals that the binocular

energy model simulations gave a good fit to these psychophysical results since the data points fall close to the dashed lines. Generally, the values of Dmax were slightly above the values predicted from the simulations; the most important trend is that the slope of the lines is the same indicating that for all these curves the values of Dmax increased as orientation was progressively more oblique. The advantage of using the method of adjustment is that the values of Dmax obtained with this method provide a direct measure of the upper disparity limit for depth perception which allows for comparison to the simulation results. These results, together with the results of the depth discrimination studies, provide support for the hypothesis that the upper disparity limit for depth discrimination increases as the orientation in the image is progressively more oblique.

Note that in order to calculate each point on the three curves shown in Figure 20 the model response from Figure A14 was summed over orientation; for example, to obtain the value for an orientation of 80° the model response was summed over orientations of 90°, 80° and 70°. This was done to reflect the orientation bandwidth of 30°, which was the same for all the oriented patterns. Summing the model response over orientation in this way resulted in a small change in the slope of the three curves in Figure 20.

The version of the binocular energy model used to produce the curves in Figure 20 was the position shift model, with the disparity range of the complex cells set to four times the range specified by the half-cycle limit of spatial

frequency. For example, for the simulations for the 2 cpd patterns the disparities represented by the complex cells ranged from 60 minutes crossed disparity to 60 minutes uncrossed disparity. Since there is a close fit between the modelling and the psychophysical data with the parameters in the simulations set in this way, it can be concluded that a disparity representation outside of the half-cycle limit of spatial frequency is necessary in order to account for the results with filtered noise patterns. This indicates that the phase shift model by itself cannot account for all of the psychophysical results in this thesis since the results at large disparities exceed the disparity range for this model. This is consistent with other studies which have proposed that both the phase shift model and position shift model may be important in modeling stereoscopic depth perception because depth perception extends to large disparities for some classes of stimuli (Fleet et al., 1996; Prince & Eagle, 1999, 2000a, 2000b; Qian & Zhu, 1997; Smallman & MacLeod, 1997).

Insert Figure 20 about here

In summary, the results of Experiment 5 and the follow-up study were consistent with a version of the binocular energy model using oriented filtering and assuming that the outputs from complex cells at all orientations are combined using linear summation. Based on these two studies, the disparity limit for depth discrimination performance increased as orientation was

progressively more oblique, which is consistent with the hypothesis that the disparity limit is equal to a half-cycle of horizontal spatial frequency. This provides an explanation for the results of Experiment 1-5 with isotropic filtered noise patterns in which depth discrimination was possible at large disparities, much larger than the half-cycle limit for the lowest spatial frequency in these patterns. In short, the oblique or most horizontal orientations in the filtered noise patterns extended the upper disparity limit for depth discrimination performance.

Motion and Stereoscopic Vision

Experiment 5 was a preliminary study which provides some directions for further research which may investigate whether stereopsis and motion use similar processing for the detection and integration of orientation information. I have carried out further studies (Buckthought & Stelmach, 2004) in stereoscopic depth perception with filtered noise patterns, systematically varying orientation bandwidth, spatial frequency bandwidth, and lower cut-off spatial frequency, using the same filtering parameters as Bischof & Di Lollo (1991) in order to perform a comparison of the data to this study. As a next step, it would be important to verify these results by studying stereopsis and motion in the same group of subjects.

In fact there are many similarities between stereo and motion, involving either similarity in the computational models that have been applied to both domains, or similarity at the perceptual level. Some implications of the work in

this thesis involve domains which require the integration of motion and stereo information, which will be briefly touched upon here.

To start with computational models, motion and stereo both must solve an initial correspondence problem. When we watch a sequence of images in which objects appear at successive positions, this gives a convincing impression of motion. The visual system matches the corresponding points in successive images to perceive apparent motion. In stereo, the correspondence problem is similar: the visual system must match corresponding points in the left eye and right eye image in order to compute disparity and perceive depth. In this thesis the solution used to solve the correspondence problem for stereopsis lies in the binocular energy model, which is similar to motion energy models (Adelson & Bergen, 1985; van Santen & Sperling, 1985). For both stereo and motion this is the filtering stage of low-level processing, which is integrated with other stages to describe higher-level processing (Simoncelli & Heeger, 1998; Wilson & Kim, 1994; Yuille & Grzywacz, 1998).

Neurons in middle temporal area MT of the macaque monkey are selective for motion as well as disparity (DeAngelis & Newsome, 1999; Simoncelli & Heeger, 1998). The functional significance of interactions between motion and disparity in area MT was demonstrated in studies that showed that perception of motion transparency was dependent on disparity information (Bradley, Qian & Anderson, 1995; Lappe, 1996; Qian, Anderson &

Adelson, 1994). It was previously demonstrated that the response of MT neurons to a random-dot pattern moving in its preferred direction was greatly reduced by the presentation of a second transparent dot pattern moving in the opposite direction. Interestingly, these reductions in the response occurred only when the disparity between preferred and anti-preferred patterns was within a certain range of each other. When the stimuli were clearly separated in depth, no reductions in the neuronal response were observed (Lappe, 1996).

Based on these results, it appears that the interaction between disparity and motion in area MT served to separate motion signals that originated from different depth planes (Lappe, 1996). However, when motion signals were presented on similar depth planes, an averaging of motion signals occurred. The importance of disparity in processing of motion was also demonstrated in studies in which pattern motion in plaid stimuli could only be observed when plaid components were presented on a common depth plane (Kwas, von Grunau & Dube, 1995). From the results in the present thesis, one could ask whether these interactions between disparity and motion occur for the first time at the level of the binocular energy model processing, or at higher-level processing stages. In the simplest scenario, if the binocular energy model is modified so that neurons are selective for both motion and disparity, then we would observe effects such as upper limits for disparity and motion processing that were similar for neurons at the same orientation and spatial frequency (Qian, 1994; Simoncelli & Heeger, 1998; Wilson & Kim, 1994). However, if

these interactions between disparity and motion occur for the first time at a higher stage of processing for motion and disparity computation then such effects would not be observed.

An example of the ecological significance in the similarity between stereopsis and motion may lie in how the cues of motion parallax and disparity are integrated together to result in the perception of depth. Motion parallax refers to the relative motion of the images of objects at different distances caused by the motion of the observer with respect to the objects. For an object at a given distance and a given motion of the observer, the extent of motion parallax between the object and a second object is proportional to the depth between the objects (Howard & Rogers, 1995).

Binocular disparity and motion parallax are cues that usually covary together in depth perception, since both cues are a consequence of perspective. The simultaneous difference in perspective is the basis of binocular stereopsis and the change of perspective over time is the basis of motion parallax. Perspective viewing offers a source of information about the relative depth and three-dimensional structure of objects and surfaces as well as information about the absolute distance to the surface (Howard & Rogers, 1995). Given that disparity and motion parallax covary in depth perception, it would be expected that the processing of these two cues may both demonstrate the effects of orientation and spatial frequency illustrated in this thesis. In previous studies thresholds for perceiving three-dimensional

structure defined by either disparity or motion were measured for curved surfaces or surfaces with disparity corrugations (Howard & Rogers, 1995). It would be interesting to systematically vary spatial frequency and orientation bandwidth in a study of the relative importance of binocular disparity and motion parallax as cues for veridical depth perception (Howard & Rogers, 1995). In summary, the work of this thesis provides some direction for future research in stereo and motion integration.

General Discussion

The point at which depth discrimination performance degraded in fixed window displays reflected the limit of first-order luminance-based binocular matching (Fleet et al., 1996, Qian & Zhu, 1997). In Experiments 1-3 the contribution of second-order binocular matching was evaluated using shifted contrast windows. It was found that stereoscopic performance extended to larger disparities with shifted compared to fixed contrast windows.

In particular, the shifted window wide cosine patterns provided the most compelling evidence for second-order matching because they could not be binocularly matched by the first-order system, as determined in computer modeling of the binocular energy function. The other two types of shifted windows (hard-edge and narrow cosine) yielded comparable results to those obtained using the cosine window, but the effects could not be attributed unambiguously to second-order binocular matching because the disparity of

these stimuli could be detected by the first-order luminance-based system. Depth discrimination performance for the wide cosine shifted window reflected both binocular first-order matching of the carrier pattern and second-order matching of the contrast window. The improvement in performance over the fixed window condition showed the contribution of binocular second-order matching.

In developing the theoretical framework in terms of first-order and second-order processing, I wanted to rule out the possibility that the first-order binocular energy model could explain all of the results with shifted window stimuli. This was particularly important because Prince and Eagle (1999, 2000a, 2000b) accounted for much of the previous work in stereoscopic vision using a first-order binocular energy model.

As a comment on the spatial frequency effects in the present results, in the fixed window conditions (Figure 7 & Figure 9) stereoscopic performance for the 2 cpd and 8 cpd patterns was possible at disparities much larger than the half-cycle limit values of 3.75' and 15'. Based upon the modeling by Tsai and Victor (2003) the point at which depth discrimination performance degraded on these tasks should have been approximated by these half-cycle limits. One possible explanation for this result is pooling of disparity computations across orientation in these images. This possibility was investigated in Experiment 5, in which the depth discrimination was studied with oriented filtered noise patterns (Bischof & DiLollo, 1991; Fleet et al., 1996; Qian & Zhu, 1997). The

results of Experiment 5 indicated that depth discrimination performance extended to larger disparities for oblique orientations. This provided an explanation for the results of Experiments 1-3, in which depth discrimination performance was possible at disparities much larger than the half-cycle limit of spatial frequency. The oblique orientations in the filtered noise patterns had extended depth discrimination performance to larger disparities.

These psychophysical results along with the modeling of the binocular energy function provided clarification in the controversy surrounding the size-disparity correlation hypothesis. One controversy pertained to the discrepant results that a spatial frequency effect was observed with filtered noise patterns (Smallman & MacLeod, 1994, 1997) but was not observed with DOG patterns (Schor & Wood, 1983; Siderov & Harwerth, 1993). It was possible to reconcile these results in terms of the different contrast windows used in these patterns. Applying a shifted contrast window to the noise patterns made stereoscopic performance extend to larger disparities and made it much more similar to that for the DOG patterns. As another contribution to this issue, Experiment 5 showed that stereoscopic performance also extends towards larger disparities for isotropic filtered noise patterns as a result of pooling across orientation of the disparity computations. Much of the controversy has resulted from the assumption that stereoscopic depth perception should always be limited to a range of disparities up to a half-cycle of spatial frequency for band-limited stimuli such as sine wave gratings and filtered noise patterns, regardless of

orientation bandwidth (Prince & Eagle, 1999, 2000a, 2000b; Smallman & MacLeod, 1994, 1997). A contribution of the present work is the development of an appropriate model for detection and integration of orientation information which reconciles the results for different classes of stimulus patterns. The results from the follow-up study shown in Figure 20 indicated that the position shift model with an extended disparity representation may be necessary in order to account for stereoscopic perception at large disparities with the filtered noise patterns. In particular, it is evident that the binocular energy model makes an appropriate fit in predicting the variation of the upper disparity limit for depth perception with orientation in the image; however, the best fit is obtained when the position shift model is used with binocular complex cells representing a large disparity range.

What is the ecological significance to the human visual system of binocular matching of second-order features? The present work provides an answer to this question in showing that a key contribution of second-order binocular matching of contrast windows is to extend stereopsis to larger disparities than can be supported using only a first-order system. This was evident in my research at disparities where the first-order system failed. In naturalistic situations the first-order system may fail because of binocular false matches. Binocular false matches refer to erroneous matches between image features in one eye with image features in the other eye when the visual system attempts to compute binocular disparity. This could arise with textured

or repetitive patterns that have many false matches or are ambiguous in depth. Binocular false matches also occur in scenes with overlapping objects at different depths; semi-occluded regions are areas in the scene which can only be perceived by one eye. These semi-occluded regions can not be binocularly matched with features perceived by the other eye to compute disparity. My research also showed that second-order matching is not an override on the first-order system, but an enhancement to stereoscopic depth perception. Both first- and second-order processing are important in determining the final stereoscopic depth percept.

Appendix A

In this thesis, a particular model of stereoscopic vision has been chosen to model psychophysical results. There are, however, a large number of models of stereoscopic vision in the psychophysical and computer vision literature. Of particular interest are stereoscopic correspondence algorithms, which model the first stages of disparity computation. Recently, Scharstein and Szeliski (2002) reviewed all the classes of stereo correspondence algorithms in the computer vision literature. Scharstein and Szeliski quantitatively analyzed the performance of a large number of current stereo correspondence algorithms based on their ability to produce a disparity map for a set of standard images. Using their results, it is possible to determine which algorithms are most promising as models for stereoscopic vision.

A brief review will be performed here of stereo correspondence algorithms to provide the rationale for the particular model that was used in this thesis and to situate it within the framework of existing models. However, it is beyond the scope of this review to describe these stereo correspondence algorithms in technical detail; further technical details are available from the papers cited by Scharstein and Szeliski.

Stereo vision correspondence algorithms can be classified based on the underlying assumptions about the physical world and the image formation process. For example, different algorithms use different methods to determine that points in the left and right images match, i.e., that they are projections of

the same scene point. Equally important are assumptions about the world or scene geometry and the visual appearance of objects. Starting from the fact that the physical world consists of piecewise-smooth surfaces, algorithms have built-in smoothness assumptions without which the correspondence problem would be underconstrained. For example, the smoothness assumption could be used to constrain disparity values across a surface in the image to change gradually and would reject any abrupt changes in disparity (Scharstein & Szeliski, 2002).

An important issue in understanding an algorithm is the representation used internally and output externally by the algorithm. It is preferable that an algorithm can produce a dense disparity map with a disparity estimate at each pixel, where disparity is the difference in location of corresponding features seen by the left and right eyes. Algorithms which only compute disparity for certain features or sparse locations in the image were excluded from consideration as models that could be used in this thesis because they limit any further analysis which can be performed for surface or shape representation (Scharstein & Szeliski, 2002).

Scharstein and Szeliski developed a taxonomy and categorization scheme for stereo matching algorithms based on the observation that stereo algorithms generally perform (subsets of) the following four steps: (1) matching cost computation; (2) cost (support) aggregation; (3) disparity computation and optimization; and (4) disparity refinement, which is an extra step to improve the

disparity estimate.

These processing stages are slightly different in each stereo correspondence algorithm. Here the traditional sum-of-squared-differences (SSD) algorithm will be described, with an explanation in terms of these processing stages. This serves to give a more concrete example of a typical stereo correspondence algorithm and to show that it fits into this framework. In the SSD algorithm: (1) the matching cost computed at each image point is the squared difference of intensity values at a given disparity; this means that the SSD describes the likelihood that a particular disparity value is correct at that image point; (2) aggregation is done by summing matching cost over square windows with constant disparity; this means that the disparity computation is smoothed or averaged over a local area on the image; (3) disparities are computed by selecting the minimal (winning) aggregated value at each pixel; this means that at each pixel the "winning" disparity is chosen using a winner-take-all approach (Scharstein & Szeliski, 2002). Note that step (4) is not described here because it is not used in the SSD alogorithm.

Most stereo correspondence algorithms can be categorized according to how they perform each of these four steps. These four steps are the algorithmic "building blocks" from which a large set of existing algorithms can be constructed (Scharstein & Szeliski, 2002). Most important is the distinction between local and global algorithms for disparity computation. For example, in local algorithms the disparity computation at a given point depends only on

intensity values within a finite window and usually smoothness assumptions are incorporated at the second stage of processing (referred to as aggregating support). This means that pixels are only assigned particular disparity values if they are consistent with adjacent disparity values, to be consistent with the assumption that most surfaces are smooth and vary slowly in depth. The SSD algorithm follows this processing sequence and is a local algorithm (Scharstein & Szeliski, 2002).

In local methods, the emphasis is on the first two processing stages (i.e., matching cost computation and cost aggregation). Computing the final disparities is the simplest stage since it involves only choosing at each pixel the disparity associated with the minimum cost value. Thus these methods use a local "winner-take-all" (WTA) method to choose the disparity value at each pixel (Scharstein & Szeliski, 2002).

In contrast, global methods perform the most important work during the disparity computation phase and often skip the aggregation step. Many global methods are formulated in an "energy-minimization" framework. Once the global energy has been defined, a variety of algorithms can be used to find a local minimum; global algorithms make explicit smoothness assumptions and then solve an optimization problem. Such algorithms typically do not perform an aggregation step, but rather seek a disparity assignment (step 3) that minimizes a global cost function that combines data (step 1) and smoothness terms. In other words, all global algorithms contain a function which solves an

optimization problem in order to find a global solution to produce a disparity map across the image. The main distinction between these different global algorithms is the minimization procedure used (e.g. simulated annealing, probabilistic (mean-field) diffusion or graph cuts) (for a review of these algorithms, see Scharstein & Szeliski, 2002 and articles cited therein).

Based upon the quantitative comparison of a large number of stereo correspondence algorithms using a testbed, Scharstein and Szeliski came to the conclusion that the global optimization algorithms gave the best results in all regions of the images, since they were able to process areas without texture, as well as discontinuities and boundaries (e.g., Birchfield & Tomasi, 1999; Boykov, Veksler & Zabih, 2001; Kolmogorov & Zabih, 2001; Lin & Tomasi, 2002). However, there are problems with the global algorithms that make them unsuitable for the modeling work in this thesis. In this thesis, the use of global constraints in modeling stereoscopic vision is not desirable since it is more useful to try to account parsimoniously for the psychophysical data with a minimum of assumptions. To start with, it is preferable to use a local algorithm in order to account for the psychophysical data; following this analysis with a local algorithm, then a global algorithm can be used if necessary. For this reason global stereo correspondence algorithms were excluded as models to be used in this thesis.

According to the quantitative evaluation performed by Scharstein and Szeliski, local algorithms do not perform as well as global algorithms since they

do poorly in areas without texture or near discontinuities (e.g. Hirschmuller, 2001; Muhlmann et al., 2001; Veksler, 2001). The problem with these algorithms is that discontinuities are processed incorrectly if they lie inside the local processing window which computes disparities in stage (1) because the discontinuities may be erroneously smoothed or averaged over. For this reason, many local algorithms incorporate additional algorithmic stages, such as an "adaptive" or "sliding" window for disparity computation which changes its processing area depending on the disparity values, instead of a fixed window. Discontinuities which fall within the adaptive window are handled correctly and are not averaged over because the window shifts its position when they are encountered (e.g. Birchfield & Tomasi, 1998). However, this restricts the use of local algorithms for modeling because an adaptive window is not plausible as part of a model of the human visual system.

The binocular energy function, which was used for modeling work in this thesis, has the advantage of being neurophysiologically plausible since it is based upon binocular complex cells in primary visual cortex. It is a local algorithm, however, because it combines steps (1) and (2) and uses a matching cost that is based on a support region, as it performs disparity computations over a local area of the stimulus which are similar to a normalized cross-correlation (which is mathematically equivalent to linear filtering, followed by disparity computation). The binocular energy model can compute disparities in areas with and without texture in an image.

Furthermore, the output of the simulations can be used to produce a disparity map and for many images it correctly produces sharp depth edges at the position of discontinuities in the image and does not smooth over these discontinuities (Qian & Zhu, 1997). Although I did not use the disparity map output of the binocular energy model as the output which would be predictive of stereoscopic perception, it can be demonstrated mathematically that the binocular energy model gives an output which averages to zero whenever it also fails to compute a disparity map (Qian & Zhu, 1997). Thus, although the binocular energy model is a local stereo correspondence algorithm, it does not suffer from the same limitations as many other local algorithms. The approach used in this thesis is similar to other modeling based on modifications to the binocular energy function in order to fit psychophysical data (Bridge, Cumming & Parker, 2001; Chen, Wang & Qian, 2001; Gray et al., 1998; Prince & Eagle, 2000a, 2000b; Read, Parker & Cumming, 2002; Tsai & Victor, 2003). However, as an improvement to this previous modeling work using the binocular energy function, I incorporated the use of oriented receptive fields for the binocular complex cells and the modeling of first- and second-order processing.

In this appendix I describe two sets of simulations of the binocular energy model carried out using the same methods as Qian and Zhu (1997) and Fleet et al. (1996). The purpose of the first set of simulations was to evaluate a key assumption underlying the rationale for Experiments 1-3, namely that the

binocular energy model could compute the disparity of shifted narrow cosine contrast windows, but could not compute the disparity of shifted wide cosine contrast windows. By showing that the binocular energy model could not account for performance with wide cosine contrast windows, it was possible to attribute the results from these patterns to second-order binocular matching. The results of the psychophysics experiments together with the modeling results provided evidence that both first- and second-order binocular matching contributed to stereoscopic performance.

The second set of simulations showed that the binocular energy model could be generalized to include filtering at all orientations in the image. The purpose of this second set of simulations was to show that the results of Experiments 1-5 were consistent with a version of the binocular energy model similar to quadrature models of motion perception (Adelson & Bergen, 1985; Bischof & Di Lollo, 1990, 1991; Prince et al., 2001; Simoncelli & Heeger, 1988; van Santen & Sperling, 1985). The results of these simulations gave a possible explanation for the results in Experiment 5, that the upper disparity limit for stereoscopic depth perception increased as orientation in the image was progressively more oblique. These simulations also gave a possible explanation for results of Experiments 1-3 indicating that depth discrimination performance was possible at disparities larger than the half-cycle limit of spatial frequency. Results of both sets of simulations are presented in Figures A4-A7 & Figures A9-A14.

The binocular energy model was introduced in neurophysiological work (Ohzawa, et al.,1997) and was developed formally by Qian and Zhu (1997) and Fleet et al. (1996). The model describes how complex cells compute the binocular disparity of images, as shown in Figure A1. At the input, the model consists of four monocular cells organized in pairs. The output of each monocular pair is summed to get a binocular simple cell response. At the next stage, the output of each simple cell is squared and the two simple cell responses are summed to get the complex cell response. Conceivably, a single binocular simple cell could compute disparity using inputs from only one pair of monocular cells. This type of detector would, however, be sensitive to the absolute position of stimulus features within the receptive field of the monocular cells. Thus, simple cells are not reliable disparity detectors. To create a reliable detector, two simple cells are used, each tuned to the same disparity, but offset relative to each other by 90 degrees of phase.

This quadrature arrangement of simple cells is produced by appropriately setting the phases of the monocular pairs. Residual sensitivity to the absolute position of stimulus features can be further reduced by pooling across spatially neighbouring complex cells with the same disparity tuning, as implemented here and by Qian and Zhu (1997).

The binocular energy model has two variants: phase shift and position shift. For phase shift, disparity selective complex cells combine the outputs of monocular cells at the same retinal positions in the left and right eyes.

Binocular disparity is represented by the relative shift of the on-off subregions within the monocular receptive fields (Figure A1(A)). For position shift, disparity selective cells combine the outputs of monocular cells at different retinal positions in the left and right eyes. Binocular disparity is represented by the relative shift of the entire monocular receptive fields. The organization of the on-off subregions is the same in the left and the right eyes (Figure A1(B)).

Insert Figure A1 about here

Equations 2, 3, and 4 show how to calculate the output of the binocular energy model from one pair of quadrature simple cells. The symbols in the equations match those in the schematic diagram of the model, illustrated graphically in Figure A1(A).

$$S_1 = L_1 + R_1 \tag{2}$$

$$S_2 = L_2 + R_2 \tag{3}$$

$$C = (L_1 + R_1)^2 + (L_2 + R_2)^2 \tag{4}$$

Symbols L_1, R_1, L_2, and R_2 are the monocular responses, S_1 and S_2 are the simple cell responses and C is the complex cell response. L_1, R_1, L_2, and R_2 are calculated by the convolution of the kernels for the phase shift or the position shift representations at each location in the input image. The

equations for the individual kernels can be derived from the Gabor function in equation (5)

$$k(x) = \exp(-\frac{(x + offset)^2}{2\sigma^2}) \cos(\omega(x + offset) + \varphi) \quad (5)$$

where x is the location in the image, σ is the standard deviation of the Gaussian envelope, ω is the spatial frequency and φ the phase of the receptive field. To generate kernels for the phase representation the offset was set to zero and the φ parameter was varied for different disparities. To generate kernels for the position shift representation the φ parameter was set to produce either an even or an odd kernel and the offset parameter was varied for different disparities. In the simulations σ was set to 0.0975 degrees of visual angle for the 8 cpd kernel and was scaled proportionately for all other spatial frequencies of the kernel. Kernels for the phase shift and position shift variants of the binocular energy model are illustrated in Figures A2 and A3, respectively. The fifth kernel was tuned to zero disparity and there were four uncrossed and four crossed disparity kernels.

In further detail, the kernels for the phase shift model shown in Figure A2 were produced as follows. The nine simple cells shown in the left panel (S1) had the following monocular phases: (-6π/8, 2π/8), (-5π/8,π/8), (-4π/8, 0), (-3π/8, - π /8), (-2π/8, -2π/8), (-π/8, -3π/8), (0, -4π/8), (π/8, -5π/8), (2π/8, -6π/8). The 9 simple cells shown in the right panel (S2) were in quadrature (offset by 90 degrees) from those in the left panel. The X-axis in the Figure is in units of

pixels. A value of 64 pixels per degree of visual angle was used in the simulations.

Similarly, the kernels for the position shift variant shown in Figure A3 were produced as follows. In the left panel (S1) the phase of the kernel was 0. In the right panel (S2) the phase was $\pi/2$. Disparity was produced by shifting the position of the kernel along the X-axis. The position shift model represented a range of disparities that was four times greater than the phase shift model in Figure A2. That is, each kernel in this figure had disparity tuning equal to four times the disparity tuning of the corresponding kernel in Figure A2.

The maximum disparity that can be represented using a phase shift model is equal to a half-cycle of the underlying spatial frequency of the kernel. As discussed earlier, this limit is referred to as the size-disparity correlation. This limit occurs because the monocular cell receptive fields in the phase shift model can have a difference in phase which is equal to at most a half-cycle. By comparison, for a position shift model the range of disparities can be set to an arbitrarily large value. In the present simulations this was set arbitrarily to four times the range of the phase shift model. All simulations were run with crossed disparity patterns.

For the first set of simulations, I ran each simulation with kernels centred at three different spatial frequencies and summed the three responses. The spatial frequencies of the kernels were 1) the spatial frequency of the stimulus pattern, 2) one octave below the spatial frequency of the stimulus pattern and

3) one-half octave above the spatial frequency of the stimulus pattern. For example, for the 8 cpd stimulus pattern, the kernels were set to: 8, 4 and 11.3 cpd.

Insert Figures A2 & A3 about here

I convolved every point in the image with the 9 kernels illustrated in Figures A2 and A3 and after appropriately summing and squaring the outputs of these kernels following equations 2-5, I obtained the output of the nine complex cells. Note that although Figure A1 shows vertically oriented monocular cell receptive fields, only isotropic filtering (and not vertical filtering) was used in the first set of simulations. For each of the nine complex cells the spatially adjacent cell responses were pooled to reduce the monocular phase dependence and noise in the response (Qian & Zhu, 1997). The spatial pooling was done using a Gaussian smoothing function, so that the response of the complex cell may be increased or decreased slightly depending on whether the responses of complex cells at adjacent positions in the image were higher or lower. The Gaussian smoothing function used for these simulations had a standard deviation of 4 pixels and size 21 x 21 pixels. Next, for every point in the image the sum of the responses of the four crossed disparity complex cells gave the total crossed energy while the response of the four uncrossed disparity complex cells gave the total uncrossed energy. The difference

between the total crossed and the total uncrossed energy was used as the model output. Finally, for a given stimulus pattern the response for a vertical column of points at the centre of the image was averaged to obtain the final estimate of binocular energy at the given disparity of the stimulus pattern. This calculation was performed ten times for each stimulus pattern and for each stimulus disparity value. The code was written in Matlab. In summary, the simulation yielded an estimate of the crossed minus the uncrossed binocular energy for a given stimulus pattern across a range of binocular disparities of the stimulus pattern.

Examples of two simulation runs are shown in Figure A4 (separately in panels A & B). These examples illustrate the output of the phase shift version of the binocular energy model for a fixed window hard-edge pattern similar to that used in Experiment 1. The spatial frequency of the pattern was 8 cpd in both simulations. In the first simulation shown in Figure A4(A) the pattern had a disparity of 3 minutes crossed disparity. The 3D surface plots in Figure A4(A) show the corresponding crossed and uncrossed binocular energy and the difference between the crossed and uncrossed energy calculated for each point in the image. Note that at this disparity of 3 minutes the binocular energy for the crossed disparity exceeded the binocular energy for the uncrossed disparity, as the surface plot is located in the positive region of the graph. In the second simulation shown in Figure A4(B), the pattern had a disparity of 4 minutes crossed disparity. Note that at this disparity of 4 minutes the binocular

energy for the crossed disparity was approximately equivalent to the binocular energy for the uncrossed disparity, which yielded a surface plot that straddled a value of zero in the difference graph. I concluded that at this slightly larger value of disparity the binocular energy model failed to compute the disparity of this image. The average output of 10 simulations like that shown in Figure A4 was used to calculate the binocular energy function at each particular disparity, spatial frequency and window condition. Although the entire 3D surface plot is shown in Figure A4, I used a slice along the Y-axis at the central X position of this surface as an estimate which could be predictive of stereoscopic performance.

Insert Figure A4 about here

First I simulated the fixed window hard-edge patterns used in Experiment 1 with the phase version of the binocular energy model. The results of an average over ten simulations are shown in Figure A5 for the 8 cpd and the 2 cpd filtered noise patterns at crossed disparities. Consider Figure A5(B), illustrating the simulation for the 8 cpd pattern which had crossed disparity. At a disparity of zero the crossed minus uncrossed energy canceled to zero indicating that the binocular energy model detected neither crossed nor uncrossed disparity in the stimulus pattern. For disparities up to about 5 minutes the crossed energy dominated the output of the response which

indicated that the binocular energy model could compute the disparity for this image over this disparity range. The 5 minute disparity point in this curve could be used as an estimate of the disparity limit for stereoscopic performance with the 8 cpd fixed window hard-edge pattern. Beyond about 5 minutes of disparity the model exhibited a "ringing" response; that is, the model cycled between correct and incorrect disparity responses. The amplitude of this ringing gradually dropped to zero as disparity was increased. The ringing was expected due to the cyclical nature of the resulting filtered noise image because the first stage of processing in the binocular energy model was linear filtering (Fleet et al., 1996; Qian & Zhu, 1997); the spacing of the luminance peaks in the filtered image was determined by the spatial frequency of the kernel. At progressively larger disparities the binocular energy model output cycled between correct and incorrect responses as successive luminance peaks in the left and right images were binocularly matched.

Insert Figure A5 about here

Figure A5(A) illustrates the simulation for the 2 cpd fixed hard-edge pattern which had crossed disparity. For this pattern the crossed energy dominated the response up to a disparity limit of approximately 20 minutes. These disparity limits were consistent with the size-disparity correlation hypothesis and were in general agreement with the observed trends in the

results with fixed window hard-edge patterns (Experiment 1, Figure 7). In further simulations I confirmed that the position shift and phase shift versions of the binocular energy model yielded approximately the same results with the fixed window hard-edge pattern. Based on these simulations I concluded that the binocular energy model provided a good estimate of the upper disparity limit on stereoscopic performance with the fixed window hard-edge pattern.

The first set of simulations pertained to the narrow and wide cosine shifted window patterns. These were important because I wished to show that the binocular energy model could not represent the disparity of wide shifted cosine window patterns at large disparities. Prince and Eagle (2000b) had shown using uncorrelated filtered noise and the position shift model that it was not necessary to invoke second-order binocular matching to explain stereoscopic performance with shifted Gaussian windowed patterns. Note that these patterns had the same width as the narrow cosine window patterns used in this thesis. Indeed, they reported that the first-order binocular energy model predicted an extended disparity limit for these narrow shifted patterns. It was important to show with simulation results that the width of these shifted patterns was the most important factor and that window shape (Gaussian versus cosine) was not relevant. Indeed I found that the results of simulations were the same for these narrow shifted patterns regardless of window shape. In agreement with Prince and Eagle (2000b) I found that the crossed binocular energy of shifted narrow uncorrelated patterns extended to large disparities, up

to about 35 minutes for narrow cosine patterns, predicting that stereoscopic performance with these patterns should be possible for disparities greatly exceeding the half-cycle limit expected from the size-disparity hypothesis. In detail, these simulations were carried out with a narrow cosine window, uncorrelated random noise pattern filtered at a center spatial frequency of 8 cpd. The same simulations were also carried out for narrow Gaussian patterns used by Prince and Eagle (2000b) and confirmed that the results were the same. In these uncorrelated patterns the left and right view were independently generated filtered noise patterns and only the contrast modulation window conveyed the binocular disparity. The average of ten simulations is shown in Figure A6(A). The results of simulations with a narrow Gaussian shifted window were identical and are not shown.

Compare Figures A5(B) and A6(A). Whereas the disparity limit in Figure A5 reflected a disparity consistent with the size-disparity correlation hypothesis, the disparity limit in Figure A6 was much larger.

Insert Figure A6 about here

Based solely on the results of the simulation results with shifted narrow cosine window patterns shown in Figure A6(A) we might conclude that the binocular energy model could detect the disparity of other shifted windowed patterns such as the wide patterns used in Experiments 2-3. In order to test this

possibility I ran simulations with an uncorrelated shifted window wide cosine pattern. The average of ten simulation runs is shown in Figure A6(B). From these simulations it appears that the binocular energy model could not detect the disparity for the uncorrelated wide cosine window. The crossed minus uncrossed energy was close to zero.

I expected that further simulations with correlated patterns would confirm the trends with uncorrelated noise patterns. This can be appreciated by comparing Figures A7 and A6. As with uncorrelated narrow cosine window patterns (Figure A6(A)), the model output for correlated patterns (Figure A7(A)) exhibited crossed binocular energy over a large range of disparities and exceeded the limit predicted by the size-disparity correlation hypothesis. Thus we can conclude that while the binocular energy model could compute the disparity of a narrow cosine window pattern over a large range of disparities, it could compute the disparity of a wide cosine window pattern (Figure A7(B)) over a much smaller range of disparities. In the simulation performance only extended to a disparity of 4 minutes for the 8 cpd pattern. This was consistent with the size-disparity correlation hypothesis in that the disparity limit was near the 8 cpd half-cycle limit of spatial frequency. These patterns were identical to the wide cosine window patterns used in Experiment 3. The simulations shown here suggest that stereoscopic performance with a wide cosine window pattern should be similar to that observed with a fixed window hard-edge pattern. Based on these simulations stereoscopic performance for both the wide cosine

and fixed window hard-edge patterns should be limited to a disparity range predicted by the size-disparity correlation. Compare the output of the binocular energy model for the simulations shown in Figures A5(B) & A7(B). In both cases the upper disparity limit was approximately 5 minutes. Since the experimental results showed that the disparity limit for stereoscopic performance with shifted wide cosine window patterns was much greater than for the fixed window condition (compare Figures 7 & 10) one could conclude that second-order binocular matching contributed to performance at large disparities in the wide cosine window condition. Based upon further analysis of results from the psychophysics experiments, both first- and second-order binocular matching appear to have contributed to stereoscopic performance in the wide cosine window condition.

Insert Figure A7 about here

The simulations showed that the width of the stimulus patterns was an important variable determining whether the pattern can be binocularly matched in a first-order model since narrow patterns could be binocularly matched regardless of shape (i.e. Gaussian or cosine). In particular, it was possible for patterns to be binocularly matched in a first-order model if the width of the stimulus pattern was less than the width of the underlying kernel in the computer simulation. That is, if the stimulus pattern was narrower than the

kernel width then the binocular energy model could perform the discrimination. Conversely, if the stimulus pattern was wider than the kernel width then the binocular energy model could not perform the discrimination. Note that simulations with the phase shift version of the binocular energy model were always limited by the size-disparity correlation hypothesis and thus could not be used to account for depth discrimination performance at large disparities with shifted window patterns under any conditions.

The second set of simulations of the binocular energy model involved oriented filtering at all orientations in the image. Up to this point modeling was carried out using isotropic filtering only but here the modeling was generalized to include oriented kernels in order to filter the image at all orientations. The purpose of the simulations was to show that the results of Experiments 1-5 were consistent with a version of the binocular energy model (phase shift variant) similar to quadrature models of motion perception (Adelson & Bergen, 1985; Bischof & Di Lollo, 1990, 1991; Prince et al., 2001; Simoncelli & Heeger, 1998; van Santen & Sperling, 1985). These simulations gave a possible explanation for the results of Experiment 5 which indicated that the disparity limit for depth discrimination increased as the orientation of the filtered noise patterns was progressively more oblique. Further, the simulations gave an explanation for the results of Experiments 1-3 indicating that depth discrimination performance was possible at disparities larger than the half-cycle limit of spatial frequency. These simulations showed that the results of the

psychophysical experiments were consistent with a particular implementation of the binocular energy model (phase shift variant) which detected perpendicular shifts at different orientations in an image.

Figure A8 illustrates that both the position shift and phase shift models could be generalized to include filtering using oblique kernels at all orientations in an image. In the case of the phase shift variant of the binocular energy model illustrated in Figure A8(A), the oriented kernels were given internal phase shifts in order to detect disparities perpendicular to kernel orientation. The simulations were carried out so that each oblique kernel was given a number of different internal phase shifts to detect a range of disparity shifts. These internal phase shifts were perpendicular to the kernel orientation. The simulations were carried out the same way as the first set of simulations, with the exception that spatial pooling over a slightly larger area of the stimulus display was necessary to produce the effect of orientation. A Gaussian spatial pooling function was used for this purpose, with a standard deviation of 64 pixels and size of 161 x 161 pixels. Furthermore, it was important to use a large enough processing kernel in these simulations; the size of the processing kernel varied with spatial frequency and was larger for the low spatial frequency kernels. In all other respects the simulations were similar to those shown in Figure A4. This gave an upper disparity limit for depth perception which increased as the orientation of the kernel was progressively more oblique. Simulations were run in order to assess whether the data from the

psychophysical experiments in the present studies were consistent with this modeling.

Insert Figure A8 about here

In the case of the position shift variant illustrated in Figure A8(B), the oblique kernels were shifted in a direction perpendicular to the orientation of the kernel. In this variant of the binocular energy model each oblique kernel was shifted using a range of perpendicular shifts, with larger shifts equivalent to larger disparity shifts. Simulations based on the position shift variant would therefore be expected to give similar results as the phase shift variant if the complex cells were set to have the same disparity range. The phase shift variant has the disadvantage that the maximum disparity that can be represented using the largest internal phase shift in the receptive fields is equal to a half-cycle of spatial frequency of the kernel while the position shift variant does not have this limitation. Therefore, it was important to also run simulations with the position shift variant in order to model stereoscopic performance at large disparities. Simulations with the position shift variant were conducted with the range of disparities arbitrarily set to four times the range used with the phase shift variant. However, the simulations using the phase shift variant may be considered to be the most important because they involve the implementation most similar to quadrature models of motion

perception (Adelson & Bergen, 1985; Bischof & DiLollo, 1990, 1991; Prince et al., 2001; Simoncelli & Heeger, 1988; van Santen & Sperling, 1985). The assumption which is being made in choosing this particular model is that it is important to show that stereoscopic depth perception can be modeled using a model bearing similarity to a motion perception model. Of course further work involving psychophysics and modeling would be required to validate this approach and fully test the similarities between the domains of stereopsis and motion.

Insert Figures A9-A12 about here

Simulations of the phase shift variant of the binocular energy model with oriented kernels are shown in Figures A9-A12. As an example, consider the top panel in Figure A9 illustrating the simulation for the 2 cpd pattern with a 90° (vertical) kernel. At a disparity of zero the crossed minus uncrossed energy cancelled to zero indicating that the binocular energy model detected neither crossed nor uncrossed disparity in the stimulus pattern. For disparities up to about 15 minutes the crossed energy dominated the output of the response. More precisely, the point at which the curve crossed zero was equal to 15.4 minutes while the peak response occurred at 6.86 minutes, as determined by curve fitting (see details below). Beyond 15.4 minutes of disparity the model continued to cycle between correct (positive peaks) and incorrect (negative

peak) responses and then settled to an average of zero. This was expected due to the cyclical nature of the pattern following the oriented bandpass filtering performed by the oriented kernels as discussed above for the first set of simulations. The 15.4 minute disparity point on this curve could be used as an estimate of the disparity limit for stereoscopic performance obtained with the vertical kernel which can be compared to the predicted half-cycle limit of 15 minutes. Similarly the 6.86 minutes disparity point (the peak response) should be compared with the predicted quarter-cycle value of 7.5 minutes. Note that all the other curves obtained with different kernel orientations are very similar in appearance to the simulation curves for a vertically oriented kernel except for slight shifts in the peak and zero response.

The upper disparity limits were measured from the simulation output curves (Figures A9-A12) using the following method. Each simulation output curve was fit to a four-parameter damped sine wave curve shown in Equation 6, using non-linear regression and using the significance level of the F-test as evidence that the regression curve gave an appropriate fit.

$$f(x) = a\exp(-\frac{x}{d})\sin(2\pi x / b + c) \qquad (6)$$

It was possible to fit all simulation curves (Figures A9-A12) in this way and measure the zero crossing following the largest initial peak in the simulation curve. In Figure A13 all the zero response disparity values found in this way

are summarized for different kernel orientations for the 2 cpd & 8 cpd filtered noise patterns.

Insert Figures A13 & A14 about here

Figure A13 compares two predictions of the upper disparity limits for stereoscopic depth discrimination with the values obtained in the simulations of the binocular energy model. The Y-axis shows the upper disparity limit for the binocular energy model response (filled squares, labeled "Model"), compared to the prediction of a half-cycle of horizontal spatial frequency (filled circles, labeled "Hor Phase") and a prediction of a half-cycle of spatial frequency from the size-disparity correlation (dashed line, labeled "Half-Cycle"). The X-axis shows the orientation of the kernel (in degrees from horizontal) used to run the simulation. As explained above, the upper disparity limit is the disparity value at which simulation output went to zero following the largest initial peak response. The disparity values from the simulations gave an excellent fit to the predicted half-cycle of horizontal spatial frequency although in some cases these values fell slightly below the predicted values (in particular for the 2 cpd images). The most important trend to note is that the upper disparity limit was inversely proportional to the sine of the angle of the orientation from horizontal. That is, the two curves (filled circles, filled squares) are always parallel across all values of kernel orientation. This validated the use of the binocular energy

model as appropriate for giving a fit to the hypothesis that the upper disparity limit for depth discrimination is inversely proportional to the sine of the angle of orientation from horizontal in the image. These results are equivalent to the results obtained in motion perception modeling: the outputs of motion perception modeling give values of D_{max} that are inversely proportional to the sine of the angle of orientation from horizontal (Bischof & Di Lollo, 1990,1991). If the simulation outputs are added at different orientations, then it is possible to predict the disparity limit values for the isotropic filtered noise patterns.

Figure A14 follows the same format as Figure A13 and also compares two predictions for the upper disparity limits for stereoscopic depth discrimination with the values obtained in simulations of the binocular energy model. However, in this case the simulations were carried out using the position shift model with the range of disparities represented by the complex cells set to four times the range used in the phase shift model. The model output (plotted with squares, labeled "Model") was consistently above the prediction from a half-cycle of horizontal spatial frequency (plotted with circles, labeled "Hor Phase"). However, for both spatial frequencies the curve plotted with squares is still parallel to the curve plotted with circles. This illustrates that the position shift model could represent disparities significantly above the disparity limits from the phase shift model, but the upper disparity limit was still inversely proportional to the sine of the angle of the orientation from horizontal. Therefore, even if the simulations were run with complex cell tuning set to a

larger range of disparities, the disparity limits for stereoscopic performance were still consistent with the most important trends obtained in Experiments 1-5.

In interpreting these simulation results I expected that the disparity limit for stereoscopic performance in isotropic filtered noise patterns would be predicted by the combined output of the simulations at all orientations in the image in analogy with motion perception modeling (Bischof & Di Lollo, 1990,1991). The resulting disparity limits could be compared with the depth discrimination curves (Figures 7 & 9) and should be compared to the point on the curve where the function rises from its baseline value which is the point that depth discrimination performance degrades (for example 18-25 and 15-20 minutes for the curves in Figure 7).

In the follow-up study to Experiment 5, simulations of the binocular energy model were used to provide a fit to the upper disparity limit for depth perception, Dmax, for oriented filtered noise patterns. In Figure 20, the values of Dmax for oriented filtered noise patterns at orientations ranging from 90° (vertically filtered) to 30° are compared to predictions from the simulations. In Figure 20, the X-axis shows the orientation of the filtered noise patterns and the Y-axis shows the value of Dmax in minutes. Separate curves are drawn for the values of Dmax obtained at spatial frequencies equal to 2 cpd, 4 cpd, 8 cpd (filled circles, filled triangles, filled squares). Dmax values are also plotted on each panel for isotropic filtered noise patterns at the three spatial frequencies

(open circles, open triangles, open squares). The predicted values of Dmax obtained from the simulations for the oriented patterns are plotted with dashed lines (three dashed lines in each panel correspond to the 2 cpd, 4 cpd & 8 cpd patterns).

In order to calculate these predicted curves, the binocular energy model response from Figure A14 was summed over orientation; for example, to obtain the value for an orientation of 80° the model response was summed over orientations 90°, 80° & 70°. This was done to reflect the 30° orientation bandwidth of the oriented filtered noise patterns. Summing over orientation in this way results in a small change in slope of these three curves. The values of Dmax fall close to the dashed lines, indicating that the binocular energy model simulations gave a good fit to these psychophysical results. The most important trend observed here is that the slope of the lines is the same indicating that for all these curves the values of Dmax increased for oblique orientations. This provides a further illustration of the manner in which the binocular energy model simulations provide a fit to psychophysical data.

In summary, the simulations gave a possible explanation for the results of Experiments 1-5 with filtered noise patterns, in which depth discrimination was possible at large disparities, much larger than the half-cycle limit for the lowest spatial frequency in these patterns. On the basis of these simulations, it is possible to conclude that the oblique or most horizontal orientations in the

filtered noise patterns extended the upper disparity limit for depth discrimination.

References

Adelson, E.H. & Bergen, J.R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A, 2(2)*, 284-299.

Baker, C.L. Jr. (1999). Central neural mechanisms for detecting second-order motion. *Current Opinion in Neurobiology, 9(4)*, 461-6.

Birchfield, S. & Tomasi, C. (1998). Depth discontinuities by pixel-to-pixel stereo. In: *International Conference on Computer Vision (ICCV)*, 1073-1080.

Birchfield, S. & Tomasi, C. (1999). Multiway cut for stereo and motion with slanted surfaces. In: *International Conference on Computer Vision (ICCV)*, 489-495.

Bischof, W.F. & Di Lollo, V. (1990). Perception of directional sampled motion in relation to displacement and spatial frequency: Evidence for a unitary motion system. *Vision Research, 30*, 1341-1362.

Bischof, W.F. & Di Lollo, V. (1991). On the half-cycle displacement limit of sampled directional motion. *Vision Research, 31*, 649-660.

Blakemore, C. & Campbell, F.W. (1969). On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology, 203*, 237-260.

Boulton, J. C. & Baker, C. L. (1993). Different parameters control motion perception above and below a critical density. *Vision Research, 33*, 1803-1811.

Boykov, Y., Veksler, O. & Zabih, R. (2001). Fast approximate energy minimization via graph cuts. *IEEE Transactions on Image Processing (TRAMI), 23(11)*, 1222-1239.

Bradley, D., Qian, N. & Anderson, R. (1995). Integration of motion and stereopsis in middle temporal cortical area of macaques. *Nature (London), 373*, 609-611.

Bridge, H., Cumming, B.G. & Parker, A.J. (2001). Modeling V1 neuronal responses to orientation disparity. *Visual Neuroscience, 18*, 879-891.

Buckthought, A. & Stelmach, L.B. (2004). Binocular matching of oriented components in stereopsis: Psychophysics and modelling. Vision Sciences Society, April 30-May 5, Sarasota.

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

Chen, Y., Wang, Y. & Qian, N. (2001). Modeling V1 disparity tuning to time-varying stimuli. *Journal of Neurophysiology, 86*, 143-155.

Chubb, C., Olzak, L. & Derrington, A. (Eds.) (2001). Second-order processes in vision. *Journal of the Optical Society of America A, 18*, 2175-2370.

Clifford , C.W.G, Freedman, J.N. & Vaina, L.M. (1998). First- and second-order motion perception in Gabor micropattern stimuli: psychophysics and computer modeling. *Cognitive Brain Research, 6*, 263-271.

DeAngelis, G.C. & Newsome, W.T. (1999). Organization of disparity-selective neurons in macaque area MT. *Journal of Neuroscience, 19(4)*, 1398-

1415.

Edwards, M., Pope, D. R. & Schor, C. M. (1999). Orientation tuning of the transient-stereopsis system. *Vision Research, 39*, 2717-2727.

Edwards, M., Pope, D. R. & Schor, C. M. (2000). First- and second-order processing in transient stereopsis. *Vision Research, 40*, 2645-2651.

Edwards, M. & Schor, C.M. (1999). Depth aliasing by the transient-stereopsis system. *Vision Research, 39*, 4333-4340.

Farell, B. (2003). Detecting disparity in two-dimensional patterns. *Vision Research, 43*, 1009-1026.

Farell, B., Li, S. & McKee, S.P. (2004a). Disparity increment thresholds for gratings. *Journal of Vision, 4*, 156-168.

Farell, B., Li, S. & McKee, S.P. (2004b). Coarse scales, fine scales and their interactions in stereo vision. *Journal of Vision, 4*, 488-499.

Fendick, M., & Westheimer, G. (1983). Effects of practice and the separation of test targets on foveal and peripheral stereoacuity. *Vision Research, 23*, 145-150.

Fine, I. & Jacobs, R.A. (2002). Comparing perceptual tasks: a review. *Journal of Vision, 2(2)*, 190-203.

Fleet, D. J., Wagner, H. & Heeger, D. J. (1996). Encoding of binocular disparity: Energy models, position shifts and phase shifts. *Vision Research, 36*, 1839-1858.

Gray, M.S., Pouget, A., Zemel, R.S., Nowlan, S.J. & Sejnowski, T.J. (1998). Reliable disparity estimation through selective integration. *Visual Neuroscience, 15,* 511-528.

Hartmann, W.M. (1997). *Signals, Sound and Sensation.* Woodbury, NY: AIP Press.

Hess R.F., Baker C.L. Jr. & Wilcox L.M. (1999). Comparison of motion and stereopsis: linear and nonlinear performance. *Journal of the Optical Society of America A, 16(5)*, 987-94.

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

Hirshmuller, H. (2001). Improvements in real-time correlation-based stereo vision. In*: IEEE Workshop on Stereo and Multi Baseline Vision, International Journal of Computer Vision (IJCV).*

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

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

Julesz, B. & Miller, J. (1975). Independent spatial-frequency tuned channels in binocular fusion and rivalry. *Perception, 4*, 125-143.

Kingdom, F.A.A. & Keeble, D.R.T. (1996). A linear systems approach to the detection of both abrupt and smooth spatial variations in orientation-defined textures. *Vision Research, 36*, 409-420.

Kolmogorov, V. & Zabih, R. (2001). Computing visual correspondence with occlusions using graph cuts. In: *International Conference on Computer Vision (ICCV), volume II*, 508-515.

Kovacs, I. & Feher, A. (1997). Non-Fourier information in bandpass noise patterns. *Vision Research, 37*, 1167-1175.

Kwas, M., von Grunau, M. & Dube, S. (1995). The effects of motion adaptation and disparity on motion integration. *Canadian Journal of Experimental Psychology, 49*, 80-92.

Langley, K., Fleet, D.J. & Hibbard, P.B. (1998). Linear and non-linear transparencies in binocular vision. *Proceedings of the Royal Society of London B, 265*, 1837-1845.

Langley, K., Fleet, D.J. & Hibbard, P.B. (1999). Stereopsis from contrast envelopes. *Vision Research, 39*, 2313-2324.

Lappe, M. (1996). Functional consequences of an integration of motion and stereopsis in area MT of monkey extrastriate visual cortex. *Neural Computation, 8,* 1449-61.

Lin, M. & Tomasi, C. (2002). Surfaces with occlusions from layered stereo. Technical Report, Stanford University.

Mansfield, J.S. & Parker, A.J. (1993). An orientation-tuned component in the contrast masking of stereopsis. *Vision Research, 33*, 1535-1544.

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

Mayhew, J. E. & Frisby, J. P. (1979). Convergent disparity discriminations in narrow-band-filtered random-dot stereograms. *Vision Research, 19*, 63-71.

McKee, S.P., Farell, B. & Verghese, P. (2003). The cost of resolving stereo ambiguity. Vision Sciences Society, Sarasota, Florida, May 9-14.

McKee, S. P., Verghese, P. & Farell, B. (2002). Edges and gratings: Interactions between 1st and 2nd order stereo systems [Abstract]. *Journal of Vision, 2(10),* 74a, http://journalofvision.org/2/10/74/, doi:10.1167/2.10.74.

McKee, S. P., Verghese, P. & Farell, B. (2004). What is the depth of a sinusoidal grating? *Journal of Vision, 4*, 524-538.

Morgan, M.J. & Castet, E. (1997). The aperture problem in stereopsis. *Vision Research, 18*, 2737-2744.

Motoyoshi, I. & Nishida, S. (2001). Visual response saturation to orientation contrast in the perception of texture boundary. *Journal of the Optical Society of America, A*, 2209-2219.

Muhlmann, K., Maier, D., Hesser, J. & Manner, R. (2001). Calculating dense disparity maps from color stereo images, an efficient implementation. In: *IEEE Workshop on Stereo and Multi-baseline Vision, International Journal of Computer Vision (IJCV).*

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.

Oppenheim, A.V. & Schafer, R.W. (1989). *Discrete-Time Signal Processing*. Englewood Cliffs, New Jersey: Prentice-Hall.

Prince, S.J.D. & Eagle, R.A. (1999). Size-disparity correlation in human binocular depth perception. *Proceedings of the Royal Society of London B, 266*, 1361-65.

Prince, S.J.D. & Eagle, R.A. (2000a). Stereo correspondence in one-dimensional Gabor stimuli. *Vision Research, 40*, 913-924.

Prince, S.J.D. & Eagle, R.A. (2000b). Weighted directional energy model of human stereo correspondence. *Vision Research, 40*, 1143-55.

Prince, S.J.D., Eagle, R.A. & Rogers, B.J. (1998). Contrast masking reveals spatial frequency channels in stereopsis. *Perception, 27*, 1345-1357.

Prince, S.J.D., Offen, S., Cumming, B.G. & Eagle, R.A. (2001). The integration of orientation information in the motion correspondence problem. *Perception, 30*, 367-380.

Prins, N. & Kingdom, F.A.A. (2002). Orientation- and frequency-modulated textures at low depths of modulation are processed by off-orientation and off-frequency texture mechanisms. *Vision Research, 42*, 705-713.

Qian, N. (1994). Computing stereo disparity and motion with known binocular cell properties. *Neural Computation, 6*, 390-404.

Qian, N. (1997). Binocular disparity and the perception of depth. *Neuron, 18,* 59-68.

Qian, N. & Zhu, Y. (1997). Physiological computation of binocular disparity. *Vision Research, 37*, 1811-27.

Read, J.C.A., Parker, A.J. & Cumming, B.G. (2002). A simple model accounts for the response of disparity-tuned V1 neurons to anti-correlated images. *Visual Neuroscience, 19*, 735-53.

Sato, T. (1998). D-sub(max): Relations to low- and high-level motion processes. Watanabe, T. (Ed). *High-level motion processing: Computational, neurobiological, and psychophysical perspectives*. (pp. 115-151). Cambridge, MA: MIT Press .

Scharstein, D. & Szeliski, R. (2002). A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision, 47(1),* 7-42.

Schor, C.M., Edwards, M. & Pope, D.R. (1998). Spatial-frequency and contrast tuning of the transient-stereopsis system. *Vision Research, 38*, 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*, 2555-2567.

Schor, C. & Wood, I. (1983). Disparity range for local stereopsis as a function of luminance spatial frequency. *Vision Research, 23*, 1649-1654.

Siderov, J. & Harwerth, R.S. (1993). Precision of stereoscopic depth perception from double images. *Vision Research, 33*, 1553-1560.

Simoncelli, E.P. & Heeger, D.J. (1998). A model of neuronal responses in visual area MT. *Vision Research, 38*, 743-761.

Smallman, H.S. & MacLeod, D.I.A. (1994). Size-disparity correlation in stereopsis at contrast threshold. *Journal of the Optical Society of America A, 11*, 2169-2182.

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

Stelmach, L.B. & Buckthought, A. (submitted). First- and second-order processing in stereoscopic depth discrimination. *Vision Research*.

Sutter, A., Sperling, G. & Chubb, C. (1995). Measuring the spatial frequency selectivity of second-order texture mechanisms. *Vision Research, 35,* 915-924 .

Taylor, M. M. & Creelman, C. D. (1967). PEST: Efficient estimates on probability functions. *Journal of the Acoustical Society of America, 41*, 782-787.

Tsai, J.J. & Victor, J.D. (2003). Reading a population code: a multi-scale neural model for representing binocular disparity. *Vision Research, 43*, 445-466.

van Ee, R. & Erkelens, C.J. (1996). Temporal aspects of binocular slant perception. *Vision Research, 36*, 43-51.

van Santen, J.P.H. & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A, 2*, 300-321.

Veksler, O. (2001). Stereo matching by compact windows via minimum ratio cycle. In: *International Conference on Computer Vision (ICCV), volume I*, 540-547.

Watson, A.B. & Robson, J.G. (1981). Discrimination a threshold: Labelled detectors in human vision. *Vision Research, 21*, 1115-1122.

Wilcox, L.M. & Hess, R.F. (1995). Dmax for stereopsis depends on size, not spatial frequency content. *Vision Research, 35*, 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*, 2981-2992.

Wilcox, L.M. & Hess, R.F. (1998). When stereopsis does not improve with increasing contrast. *Vision Research, 38*, 3671-3679.

Wilson, H.R. & Kim, J. (1994). A model for motion coherence and transparency. *Visual Neuroscience, 11*, 1205-20.

Yuille, A.L. & Grzywacz, N.M. (1998). A theoretical framework for visual motion. Watanabe, T. (Ed). *High-level motion processing: Computational,*

neurobiological, and psychophysical perspectives. (pp. 187-211). Cambridge, MA: MIT Press.

Zhou, Y.-X. & Baker, C.L. Jr. (1993). A processing stream in mammalian visual cortex neurons for non-Fourier responses. *Science, 261*, 98-101.

Zhou, Y.-X. & Baker, C.L. Jr. (1994). Envelope-responsive neurons in Area 17 and 18 of cat. *Journal of Neurophysiology, 72*, 2134-2150.

Zhu, Y. D. & Qian, N. (1996). Binocular receptive field models, disparity tuning, and characteristic disparity. *Neural Computation, 8*, 1611-41.

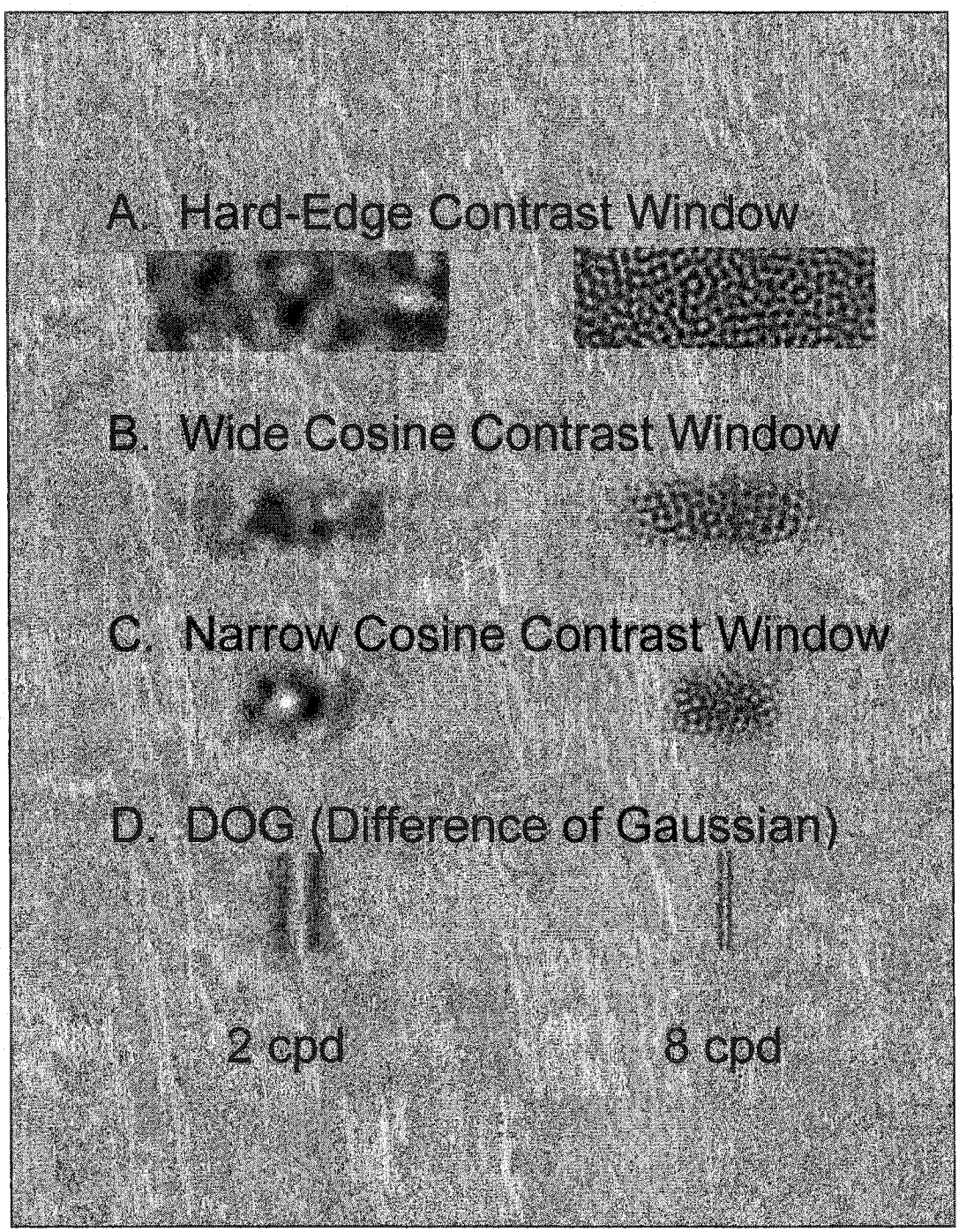

Figure 1. Four types of stimuli with different types of contrast modulation (contrast windows) at the low & high spatial frequency (2 cpd, 8 cpd). A. Hard-edge contrast window. B, C. Wide and narrow cosine contrast window (i.e. cosine contrast modulation). D. Difference of Gaussian (DOG).

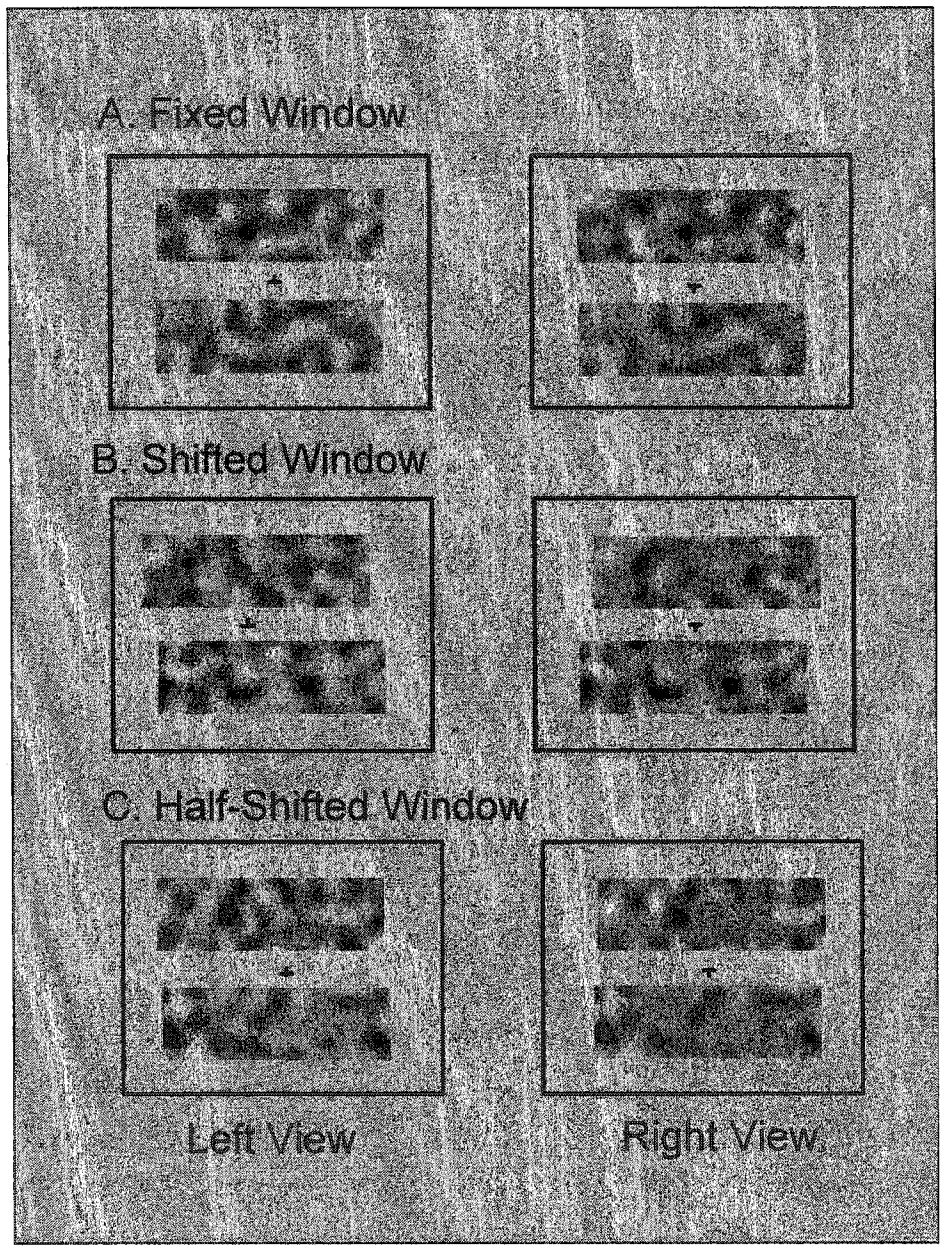

Figure 2. The stimulus display used in all experiments. A zero-disparity frame was drawn around the reference and test stimulus patterns. The reference and test patterns were above and below the Nonius marker, respectively. The figure shows the three types of windows: A. Fixed, B. Shifted and C. Half-shifted for the hard-edge contrast modulation. These three window types define the manner of producing disparity.

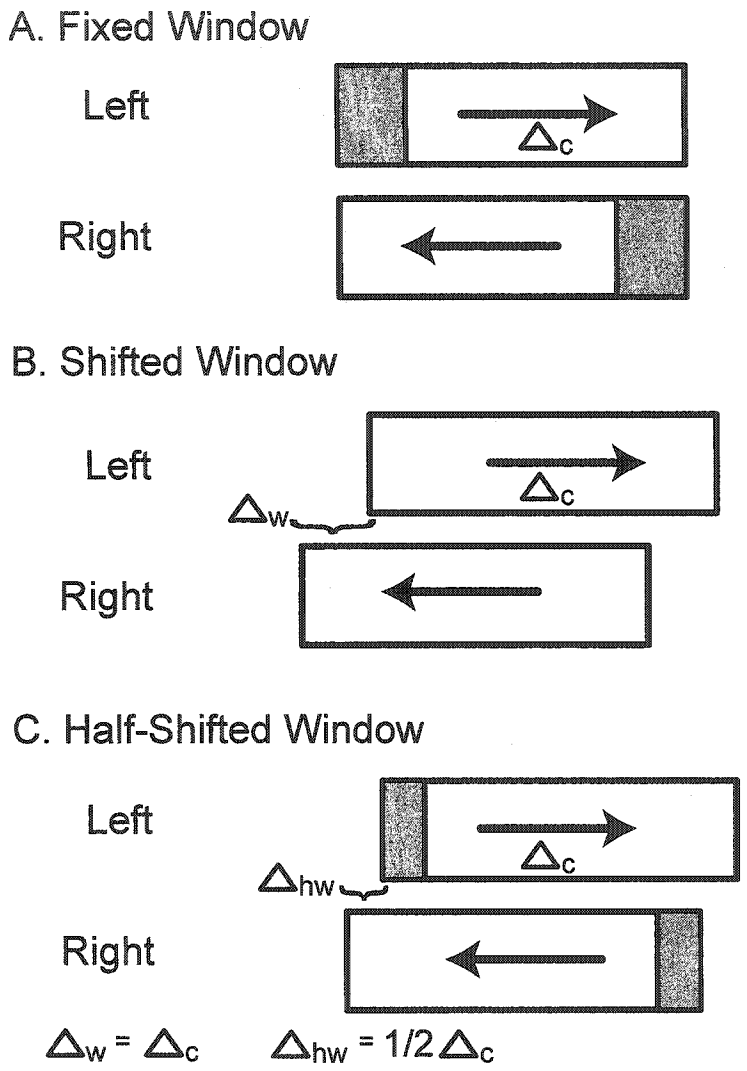

Figure 3. Schematic illustrating how the three stereoscopic displays were produced. The disparity of the carrier noise pattern (Δc) was the same across all three display types. Disparity (Δc) was produced by shifting the noise pattern with sub-pixel accuracy. The disparity of the contrast window or outline of the pattern differed across the three display types. A. Fixed window. The contrast window had zero disparity (Δw =0). The width of the pattern was constant at all disparities. Monocular zones that were uncorrelated between the left and right stereo views are shown in grey. B. Shifted window. The contrast window had the same disparity as the carrier (Δw = Δc). There were no monocular zones. C. Half-shifted window. The disparity of the contrast window was half that of the carrier (Δhw = ½ Δc). The monocular zones were half as large as those in the fixed window display.

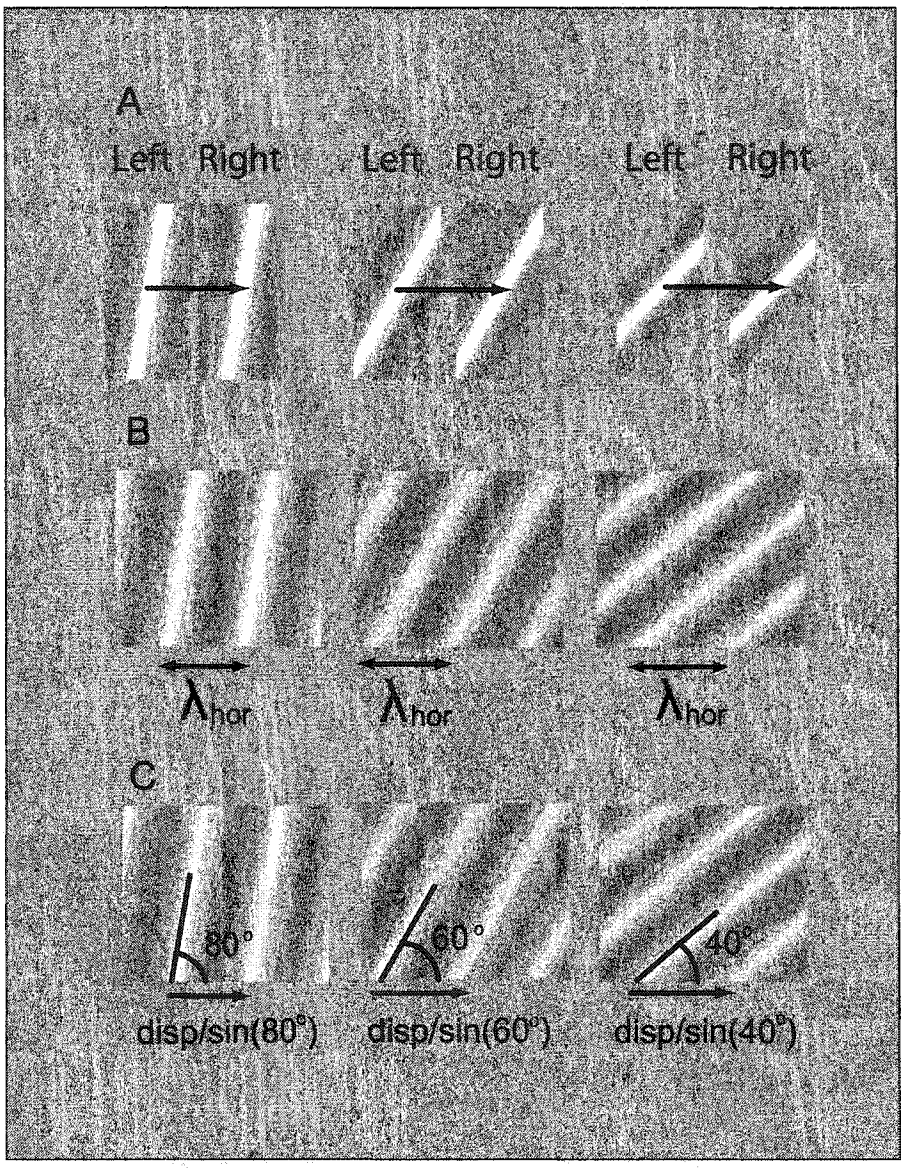

Figure 4. Hypothesis for binocular matching of oriented components. A. Left & right stereopairs for sine wave patterns at three orientations (80°, 60°, 40°). As the orientation is progressively more oblique binocular matching proceeds over a greater distance along the horizontal axis in the image. Sine wave patterns with the same orientation are shown again in B, with images that include more cycles of the sine waves. The horizontal period λ_{hor} increases as the orientation is progressively more oblique and horizontal spatial frequency decreases. C. Since the disparity limit is predicted to equal a constant fraction of the horizontal period the disparity limit is inversely proportional to the sine of the angle from horizontal. The disparity limit is predicted to equal a half-cycle of horizontal spatial frequency.

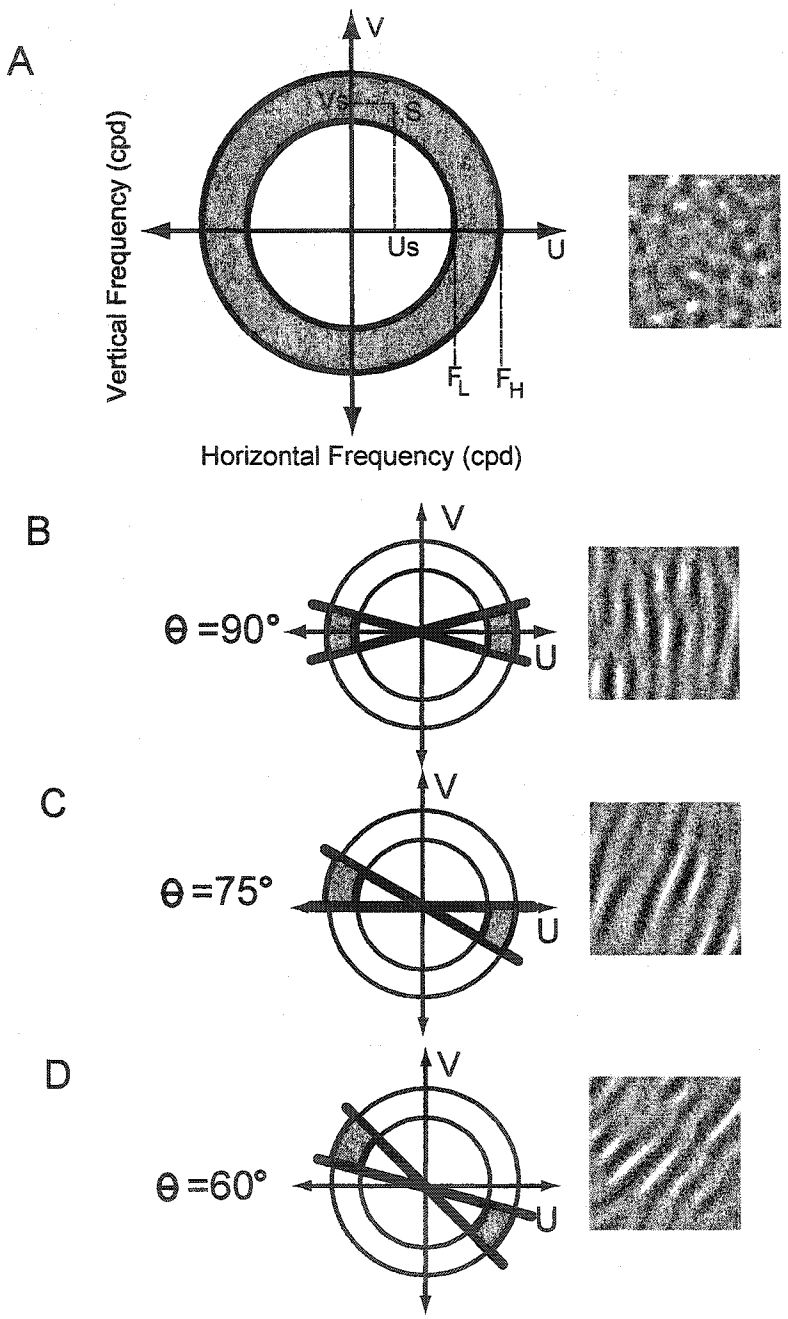

Figure 5. Oriented filtered noise patterns. A. Two-dimensional Fourier domain of an image showing an ideal isotropic bandpass filter with cut-off spatial frequencies F_L and F_H. A particular spatial frequency component has vertical & horizontal spatial frequency V_s & U_s and in this case, U_s is below F_L. In panels B-D all filters had the same orientation bandwidth equal to 30° but different mean centre orientations (θ) passed by the filter equal to 90°, 75°, 60°.

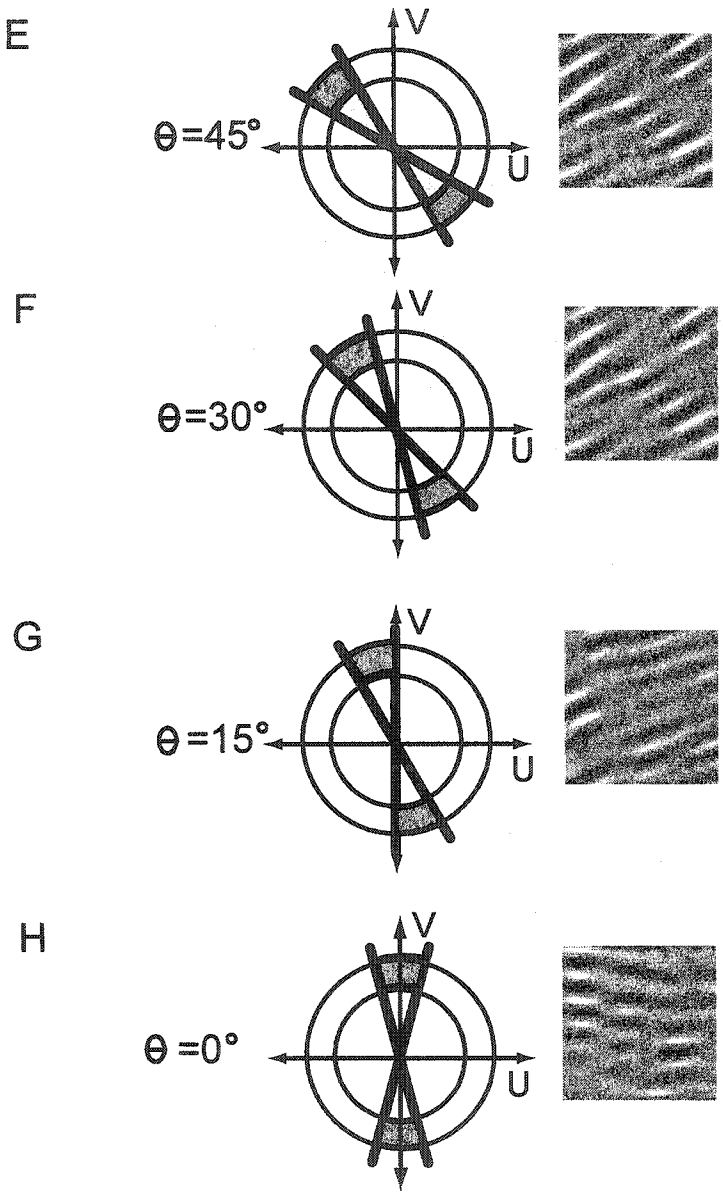

Figure 5 (cont'd). Oriented filtered noise patterns. In panels E-H all filters had the same orientation bandwidth equal to 30° but different mean centre orientations (θ) passed by the filter equal to 45°, 30°, 15° and 0°.

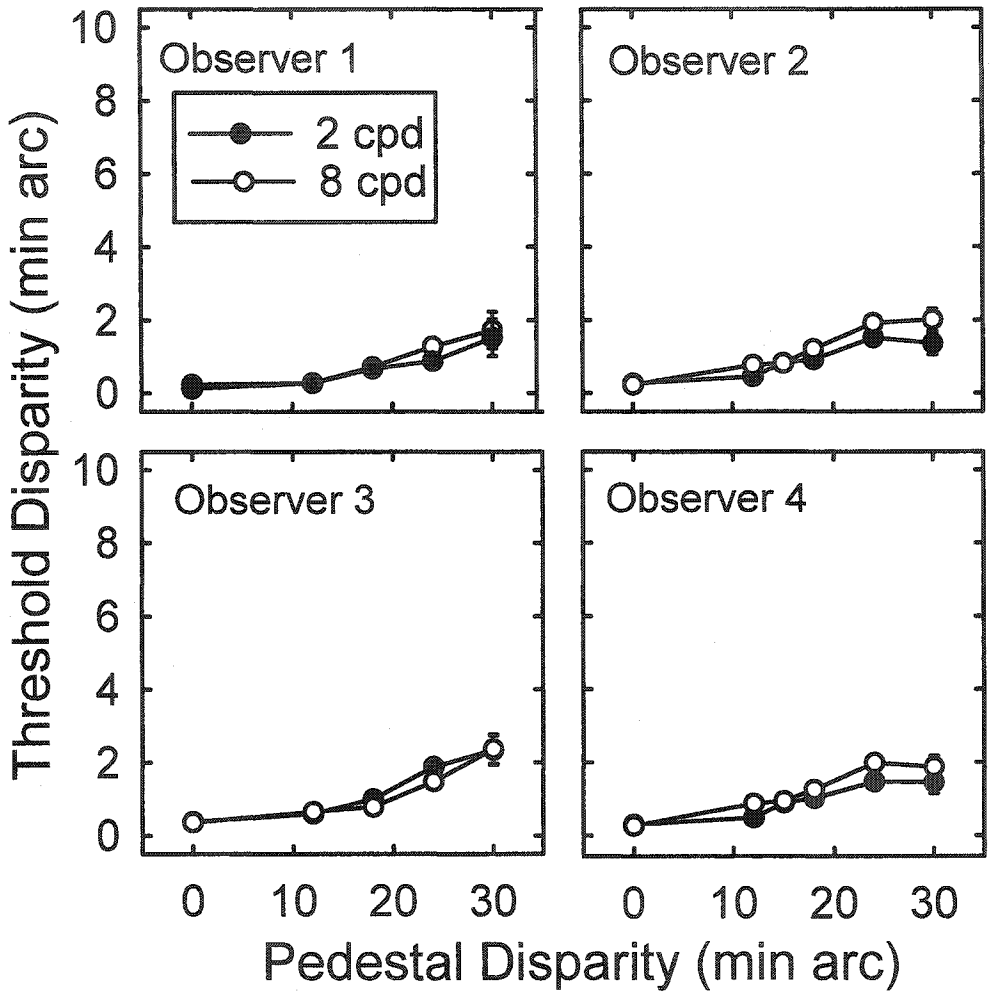

Figure 6. Results of Experiment 1 for the shifted window Difference of Gaussian (DOG) patterns shown in separate panels for four observers. Error bars show the standard error of the mean (except where it is smaller than the symbol size). X-axis shows the pedestal disparity of the reference pattern. Y-axis shows the depth discrimination threshold for the test pattern measured at each value of pedestal disparity. Separate curves are drawn for the 2 cpd & 8 cpd patterns. Note that some data values were almost identical causing the graphing software to place one symbol on top of the other symbol, making it appear that only one value is represented (e.g. Observer 3, at a pedestal of 30 minutes). Due to limitations of the graphing software this is a problem that may occur in any of the data graphs.

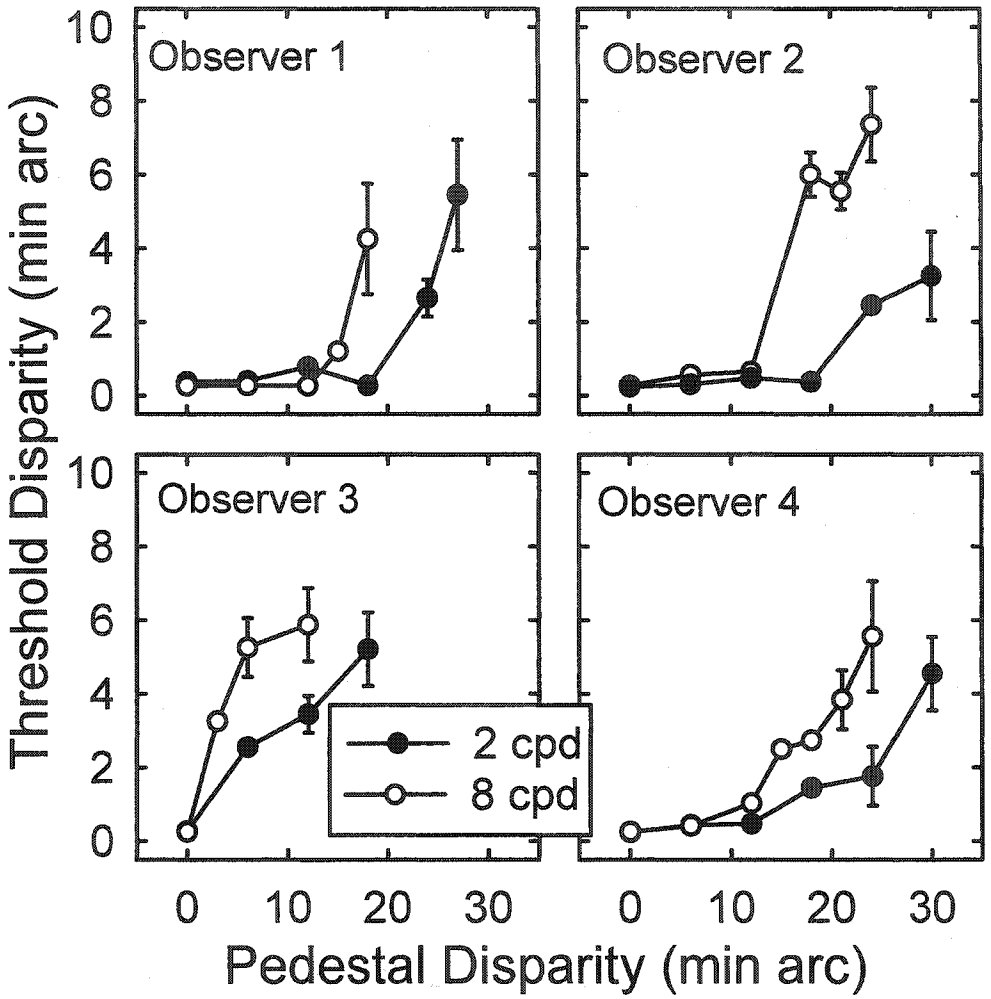

Figure 7. Results of Experiment 1 for fixed window hard-edge filtered noise patterns. Follows the same format as Figure 6.

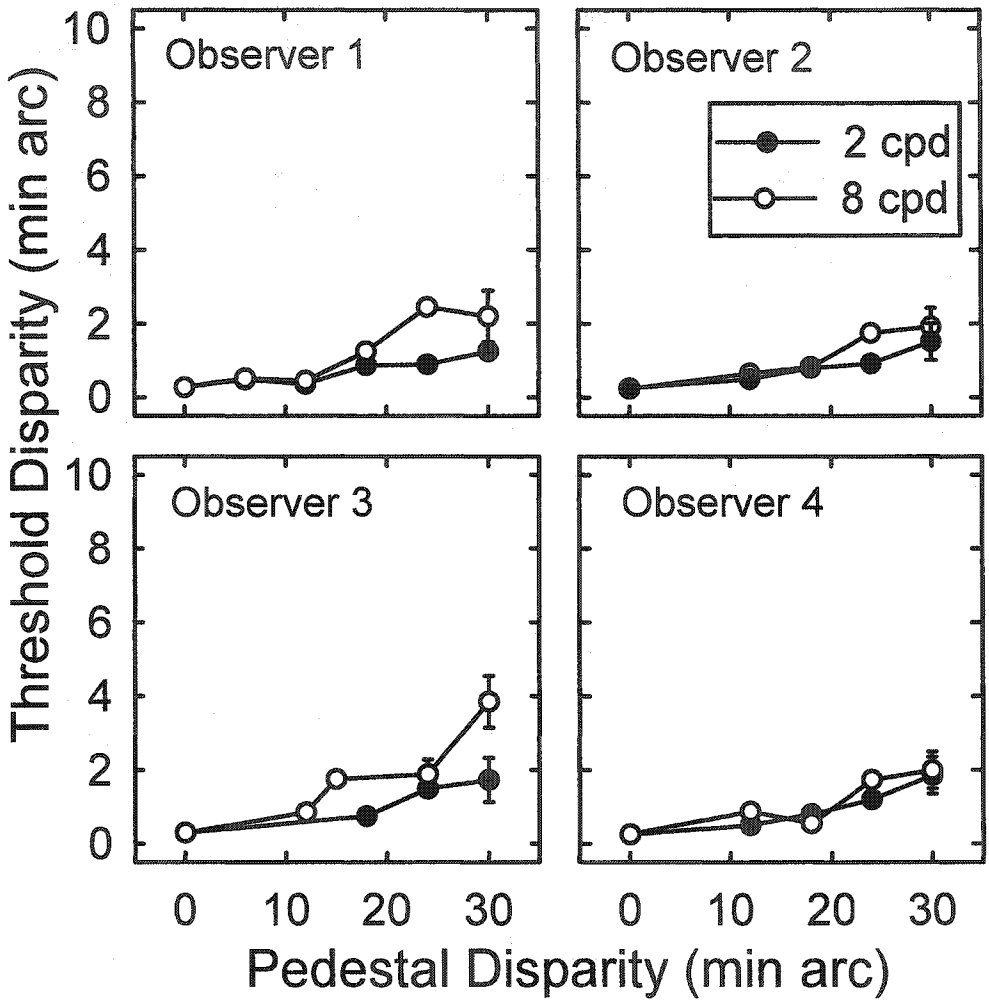

Figure 8. Results of Experiment 2 for the shifted window hard-edge patterns. Follows the same format as Figure 6.

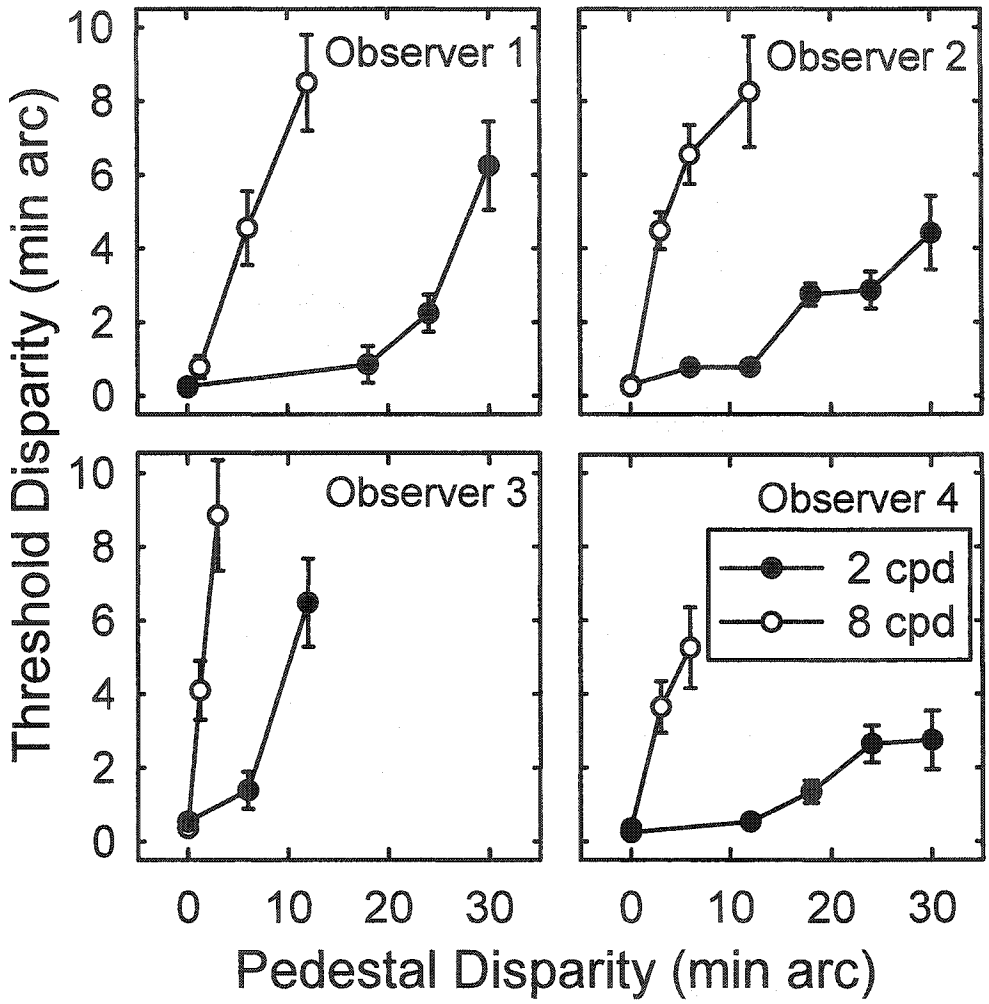

Figure 9. Results of Experiment 3 for the fixed window wide cosine patterns. Follows the same format as Figure 6.

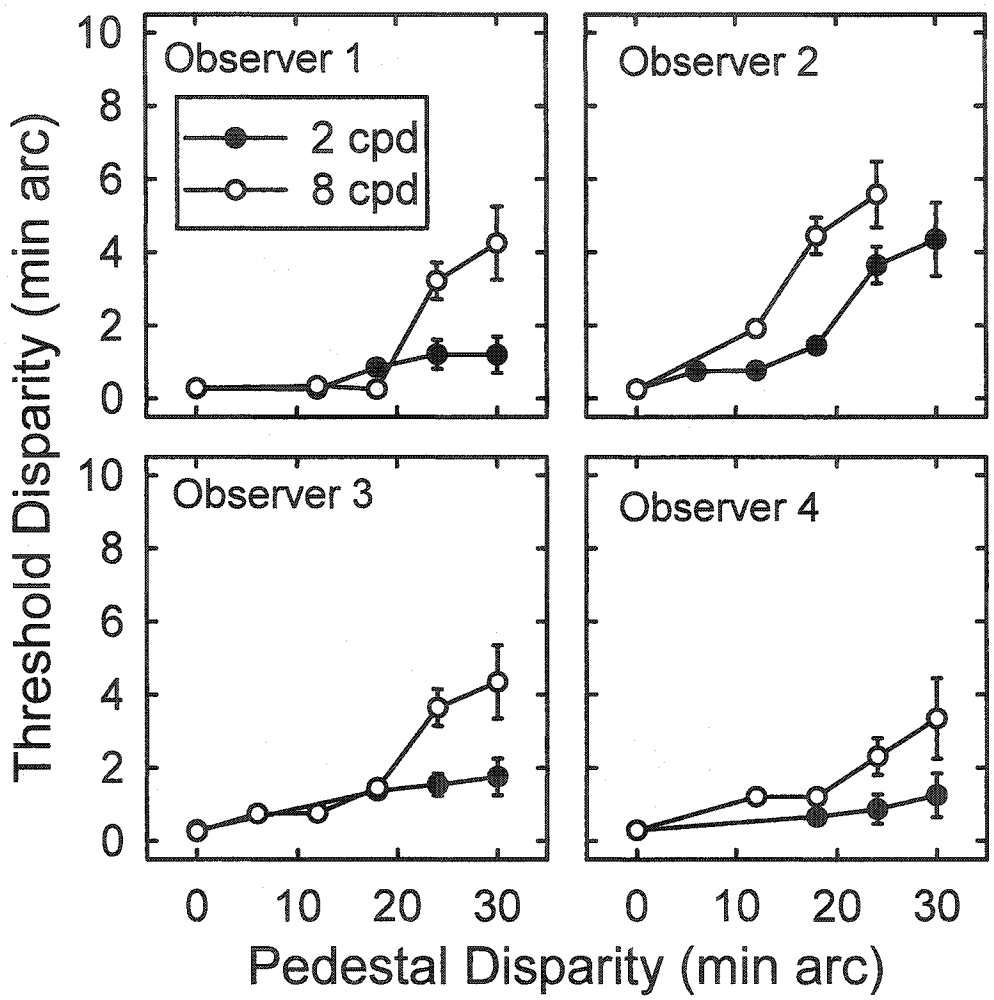

Figure 10. Results of Experiment 3 for the shifted window wide cosine patterns. Follows the same format as Figure 6.

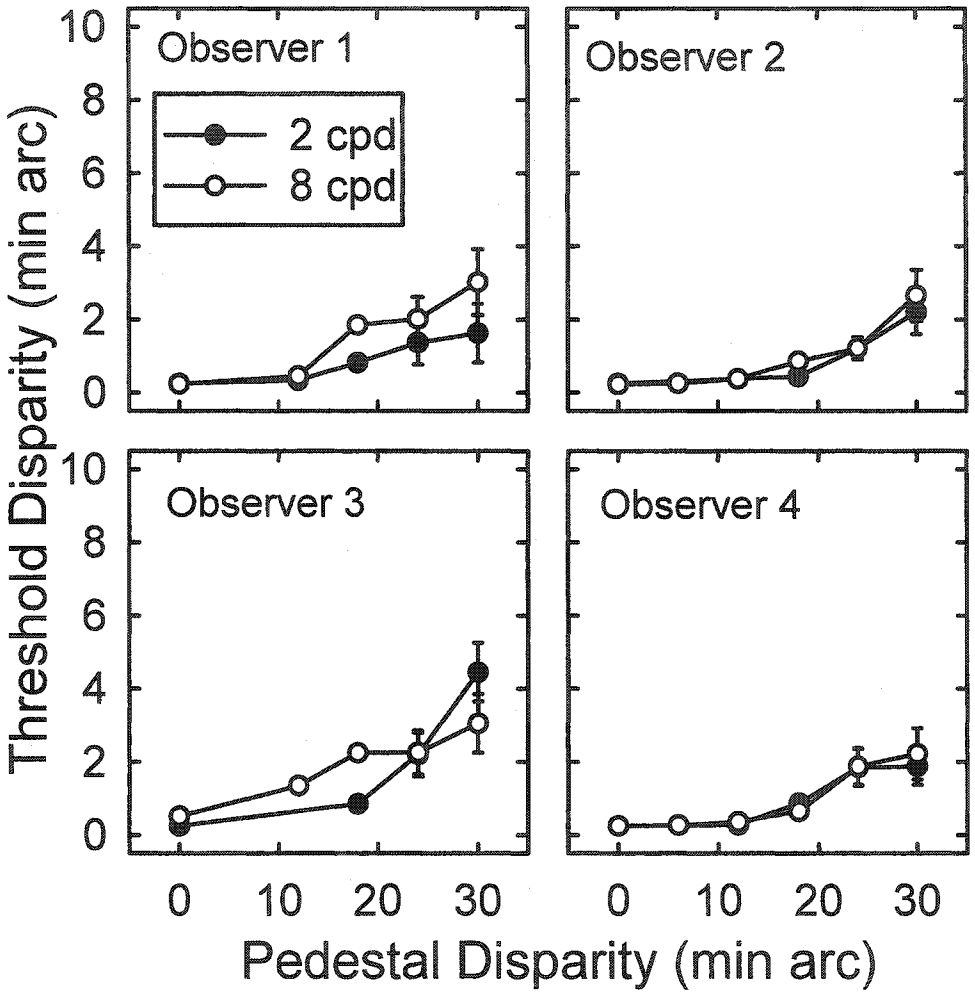

Figure 11. Results of Experiment 3 for the shifted window narrow cosine patterns. Follows the same format as Figure 6.

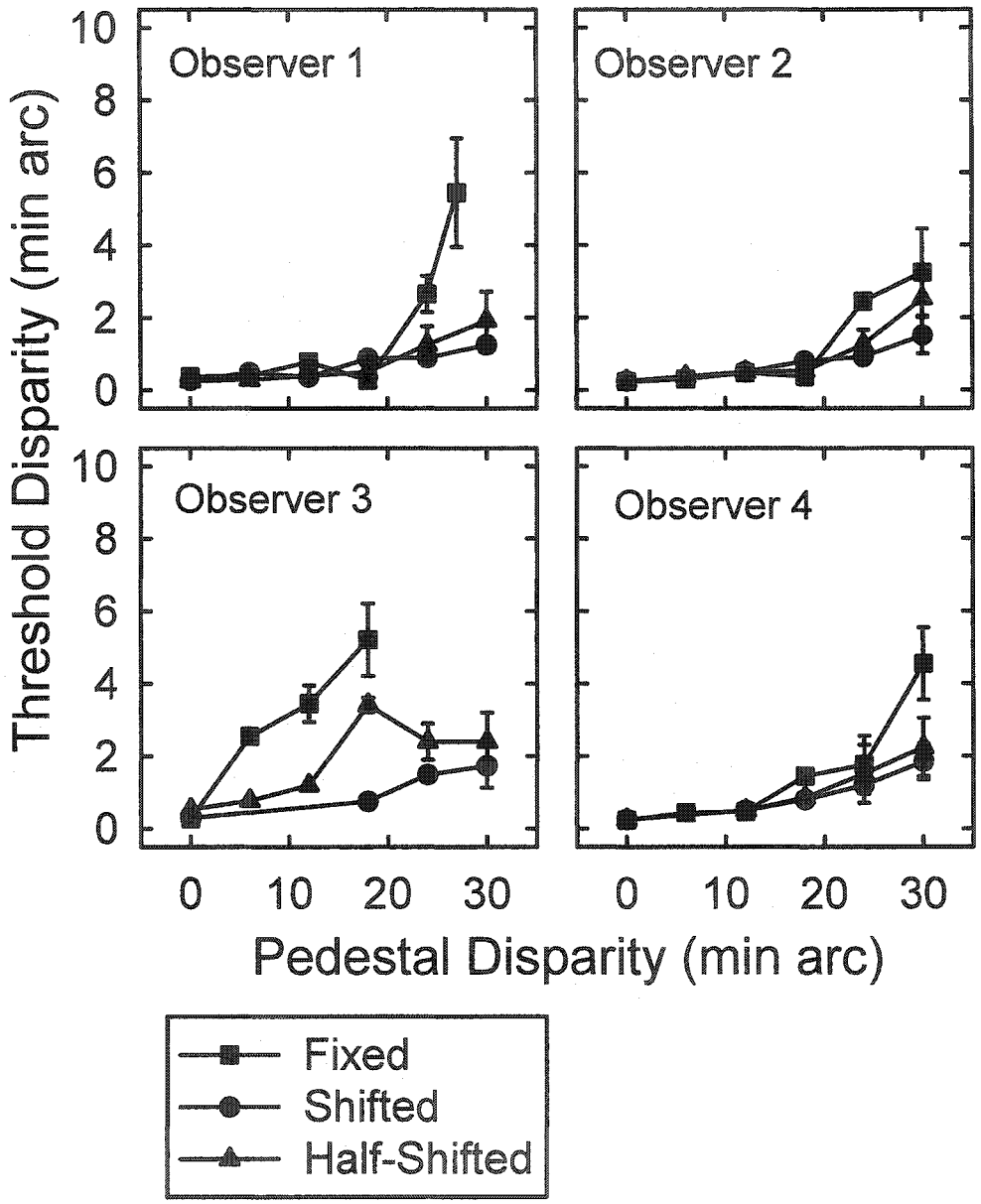

Figure 12. Results of Experiment 4 for 2 cpd fixed, shifted and half-shifted window (hard-edge patterns). Error bars show the standard error of the mean (except where it is smaller than the symbol size). X-axis shows the pedestal disparity of the reference pattern. Y-axis shows the depth discrimination threshold at each pedestal disparity. Note that the data for the fixed and shifted patterns are the same as in Figures 7-8.

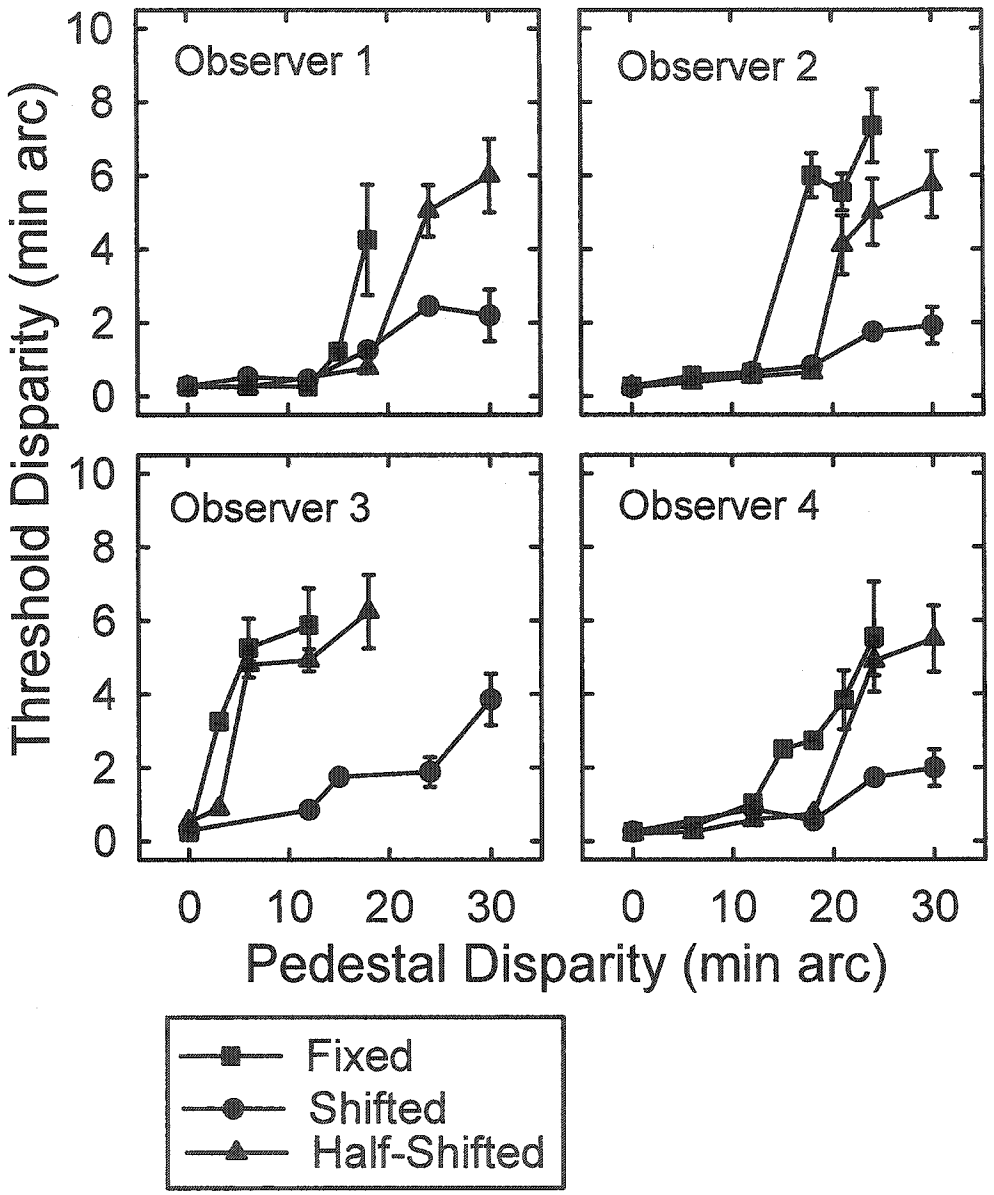

Figure 13. Results of Experiment 4 for 8 cpd fixed, shifted & half-shifted window (hard-edge patterns). Follows the same format as Figure 12. Note that the data for the fixed and shifted patterns are the same as in Figures 7-8.

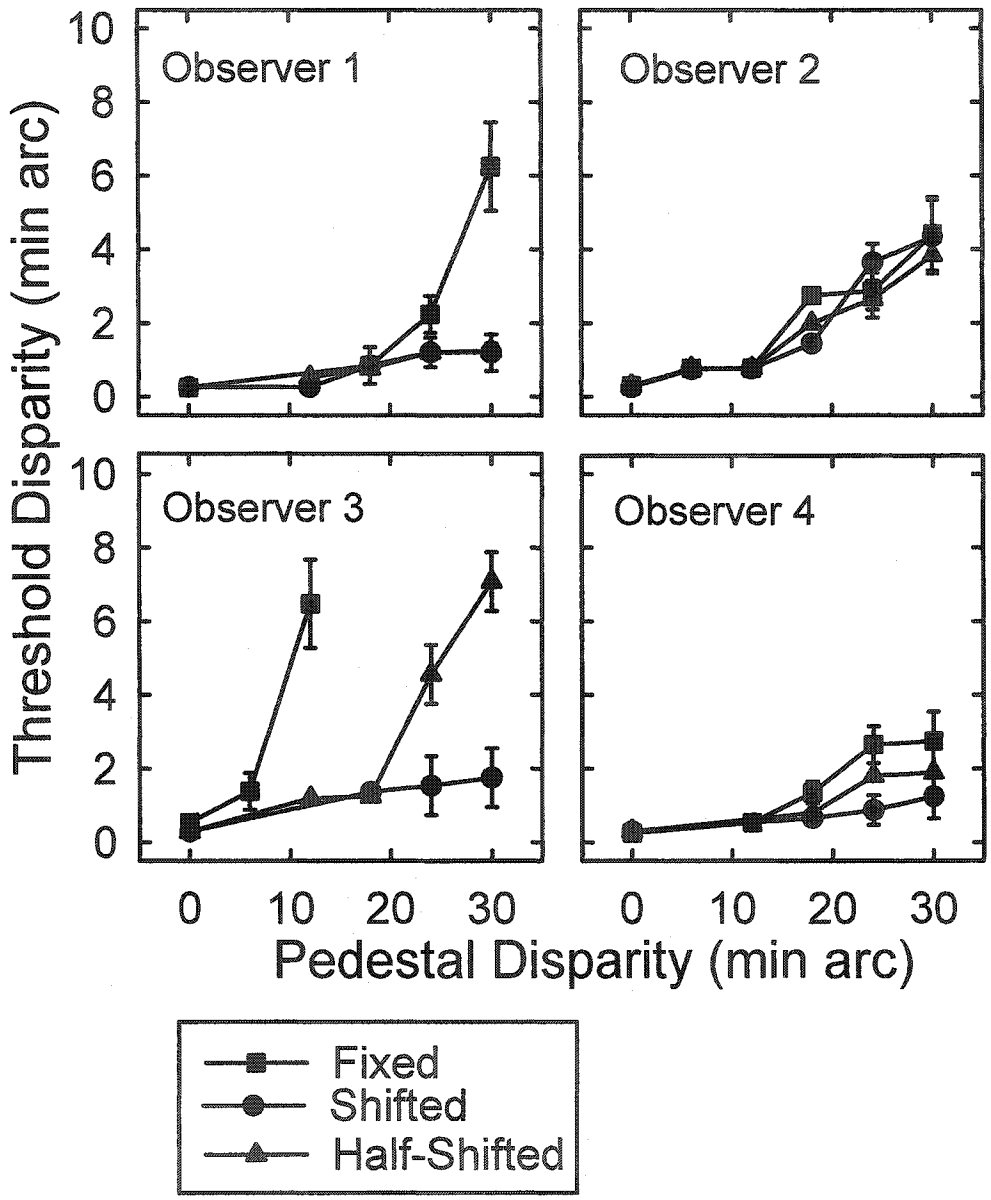

Figure 14. Results of Experiment 4 for 2 cpd fixed, shifted & half-shifted window (cosine window patterns). Follows the same format as Figure 12. Note that the data for the fixed and shifted patterns are the same as in Figures 9-10.

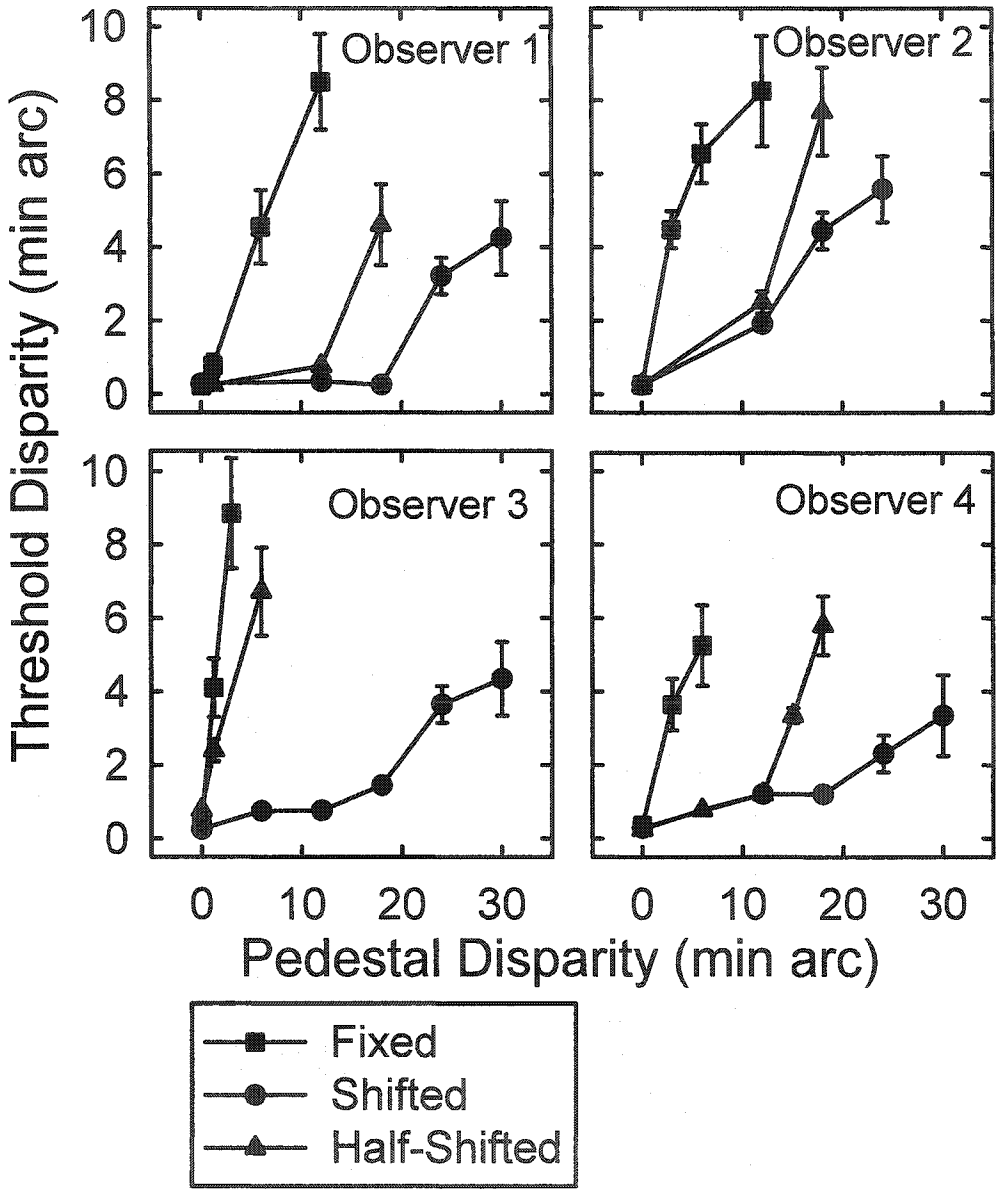

Figure 15. Results of Experiment 4 for 8 cpd fixed, shifted and half-shifted window (cosine window patterns). Follows the same format as Figure 12. Note that the data for the fixed and shifted patterns are the same as in Figures 9-10.

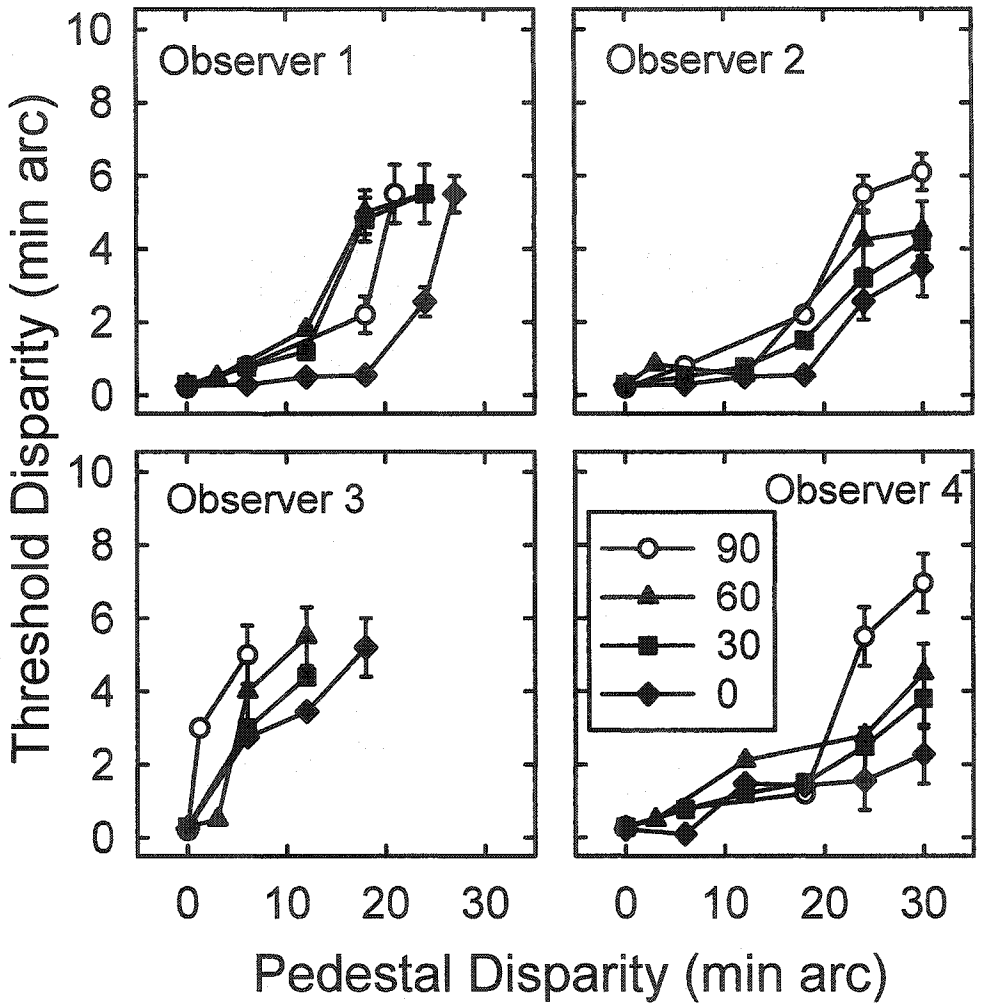

Figure 16. Results of Experiment 5 for 2 cpd fixed window oriented filtered noise patterns. Individual curves were plotted for filter orientations θ equal to 90° (vertically filtered), 60°, 30° and 0° (horizontally filtered). Error bars show the standard error of the mean (except where it is smaller than the symbol size). X-axis shows the pedestal disparity of the reference pattern. Y-axis shows the depth discrimination threshold at each pedestal disparity.

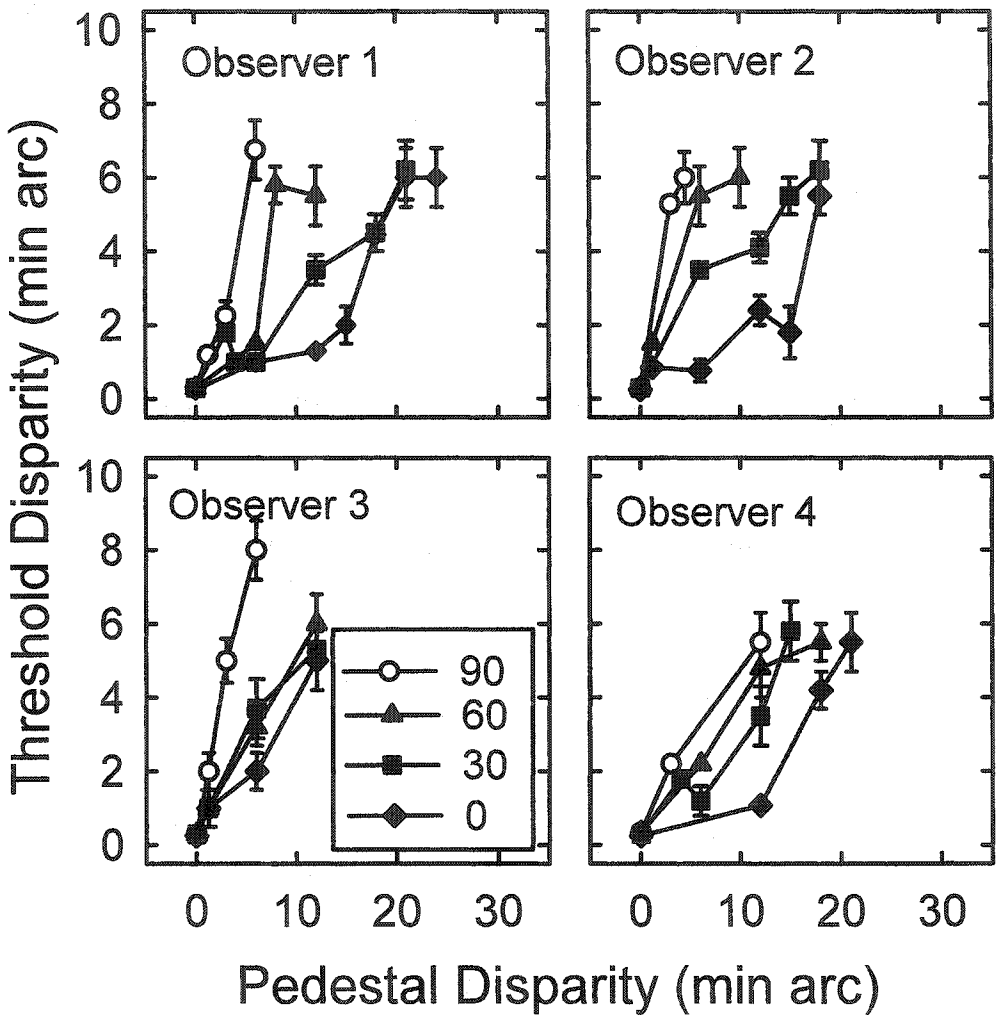

Figure 17. Results of Experiment 5 for 8 cpd fixed window oriented filtered noise patterns. Individual curves are for orientations 90° (vertically filtered), 60°, 30° and 0° (horizontally filtered). Follows the same format as Figure 16.

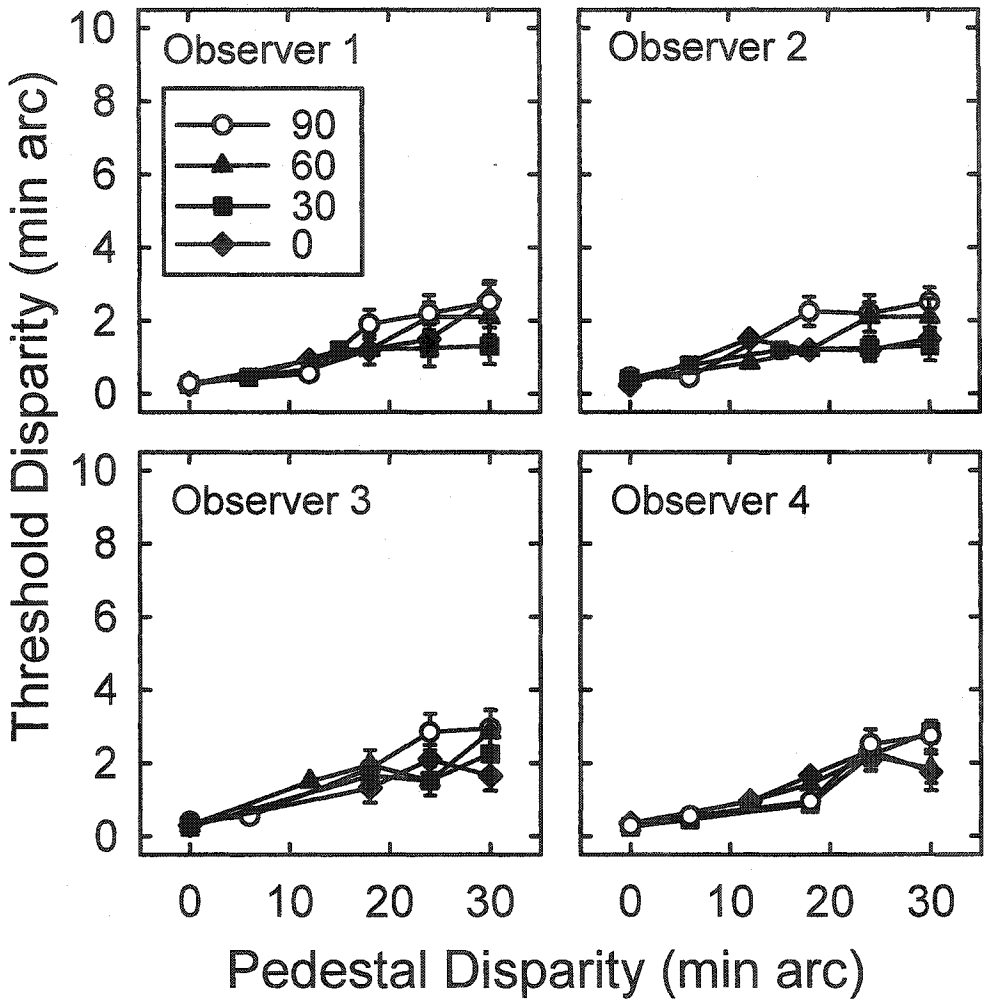

Figure 18. Results of Experiment 5 for 2 cpd shifted window oriented filtered noise patterns. Individual curves are for orientations 90° (vertically filtered), 60°, 30° and 0° (horizontally filtered). Follows the same format as Figures 16.

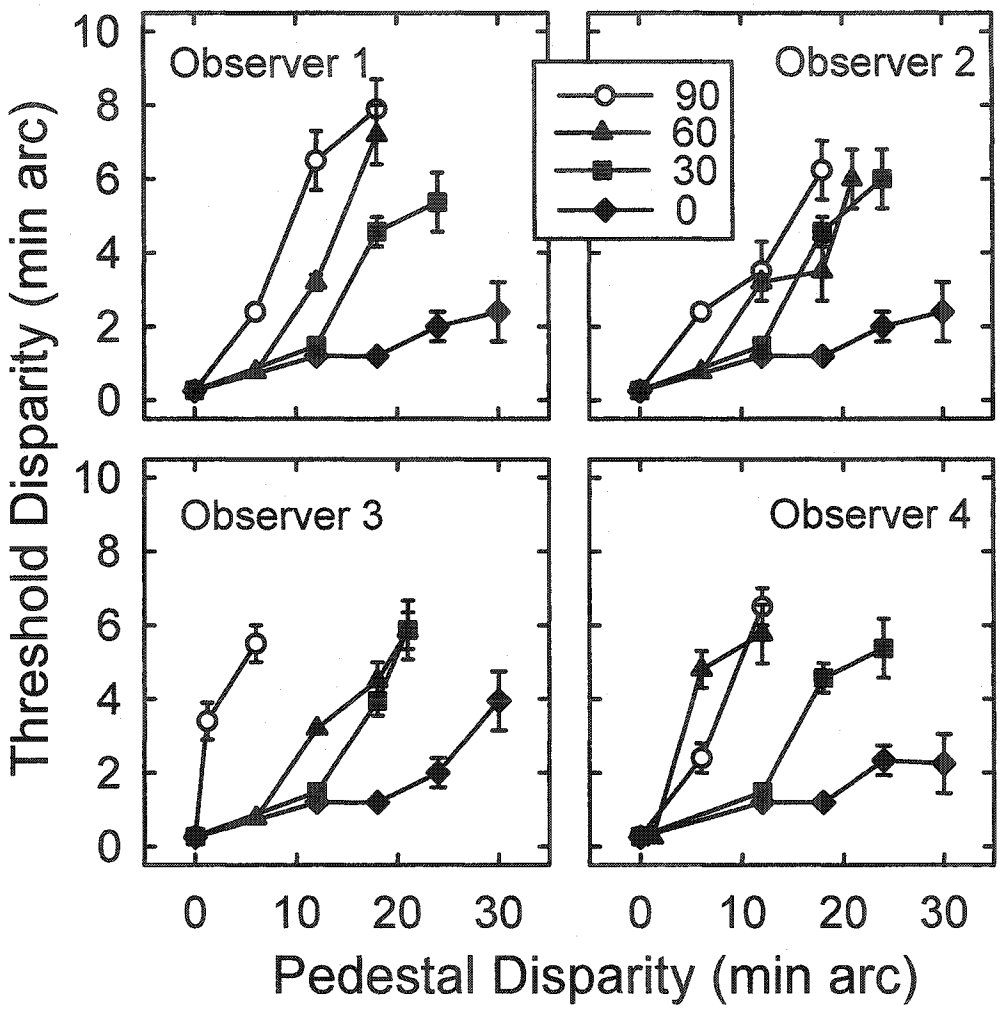

Figure 19. Results of Experiment 5 for 8 cpd shifted window oriented filtered noise patterns. Individual curves are for orientations 90° (vertically filtered), 60°, 30° and 0° (horizontally filtered). Follows the same format as Figure 16.

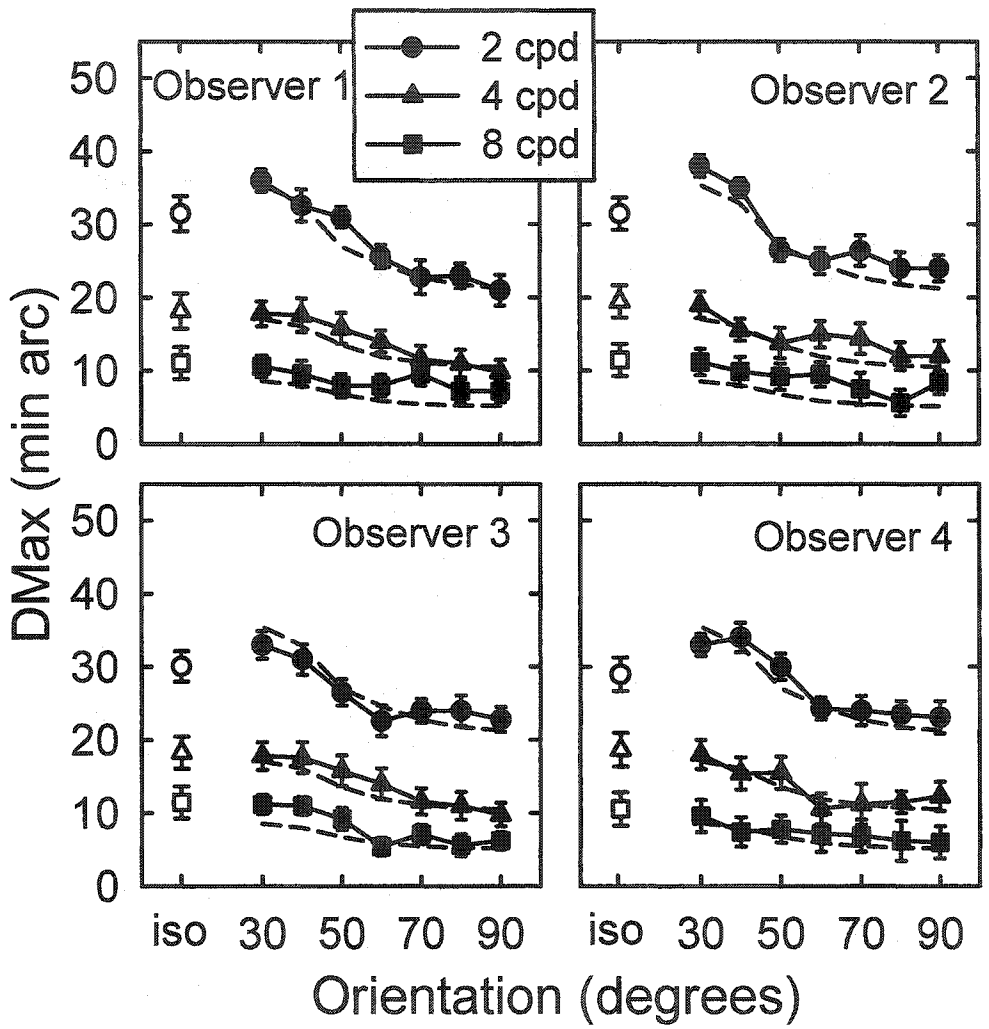

Figure 20. Results of follow-up to Experiment 5 comparing Dmax (upper disparity limit for depth perception) to predictions from binocular energy model simulations. Data and simulation results are for fixed window oriented filtered noise at orientations 90° (vertically filtered), 80°, 70°, 60°, 50°, 40°, & 30°. Separate curves are drawn for 2 cpd, 4 cpd & 8 cpd filtered noise (filled circles, filled triangles, filled squares). A data point is also plotted for isotropic filtered noise patterns for each spatial frequency (open circles, open triangles, open squares, labeled "iso" on each panel). On each panel, the predictions from the simulations for these three spatial frequencies are plotted with dashed lines for comparison with the psychophysics results.

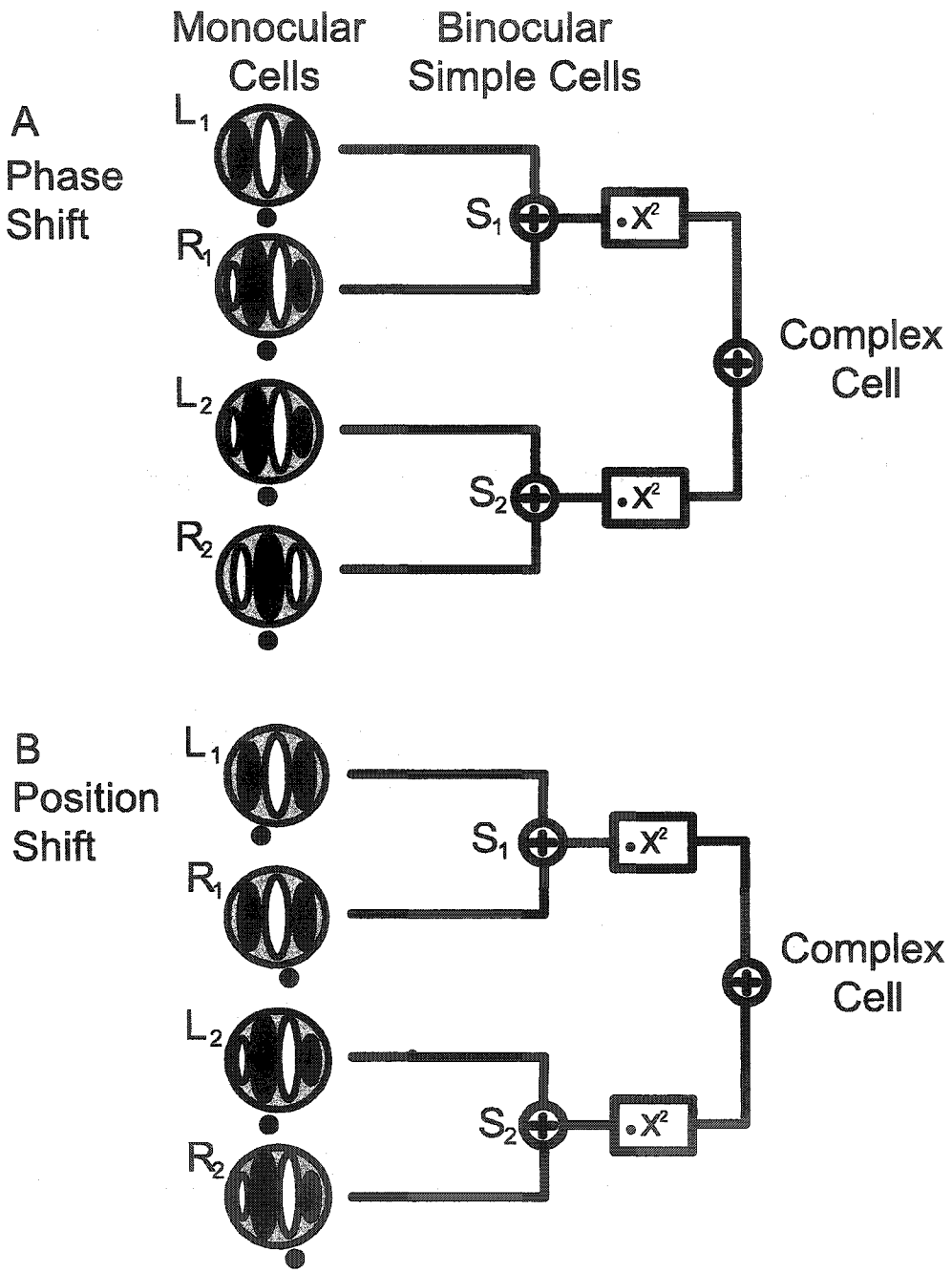

Figure A1. Binocular energy model: (A) phase shift and (B) position shift.

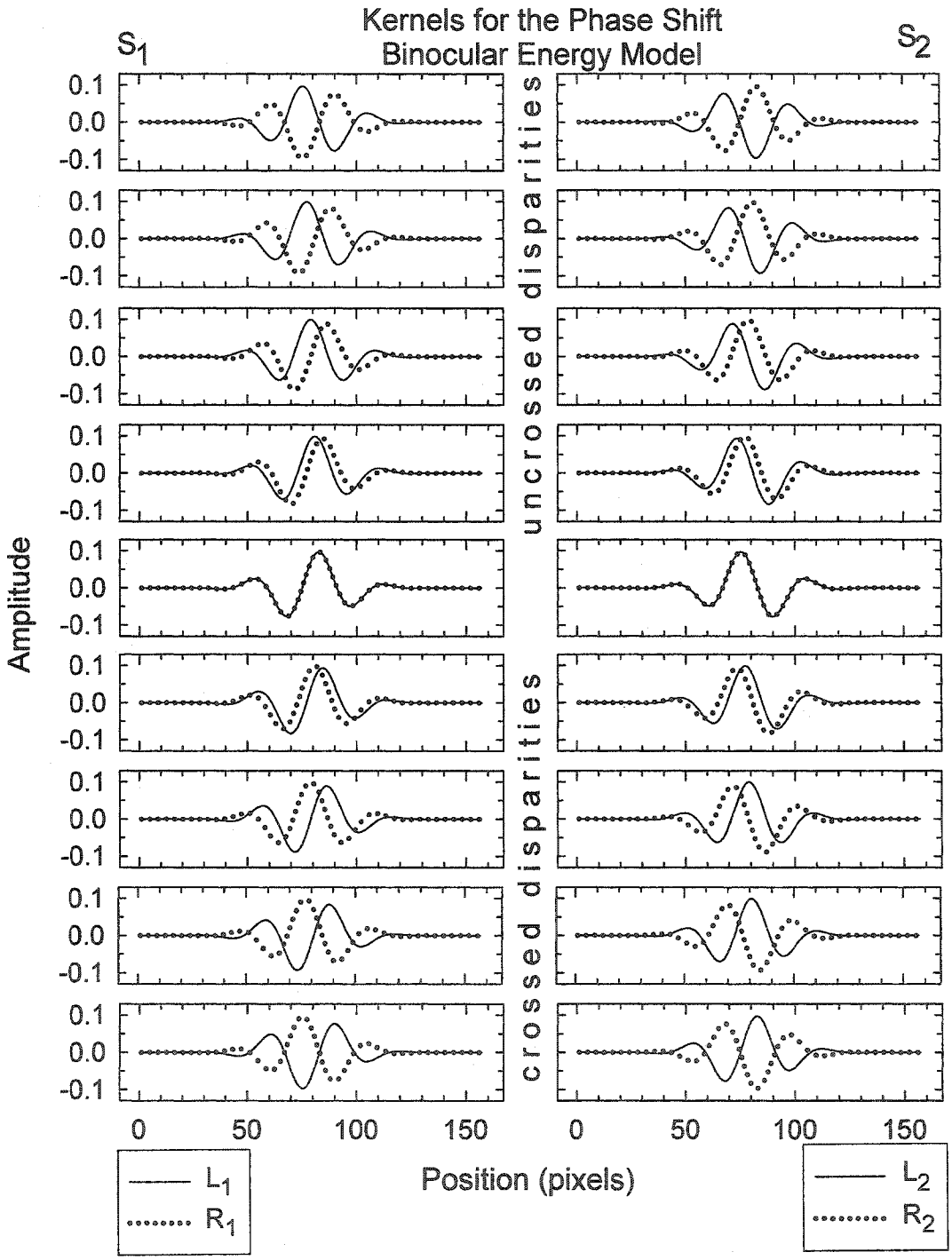

Figure A2. The kernels for the phase shift version of binocular energy model

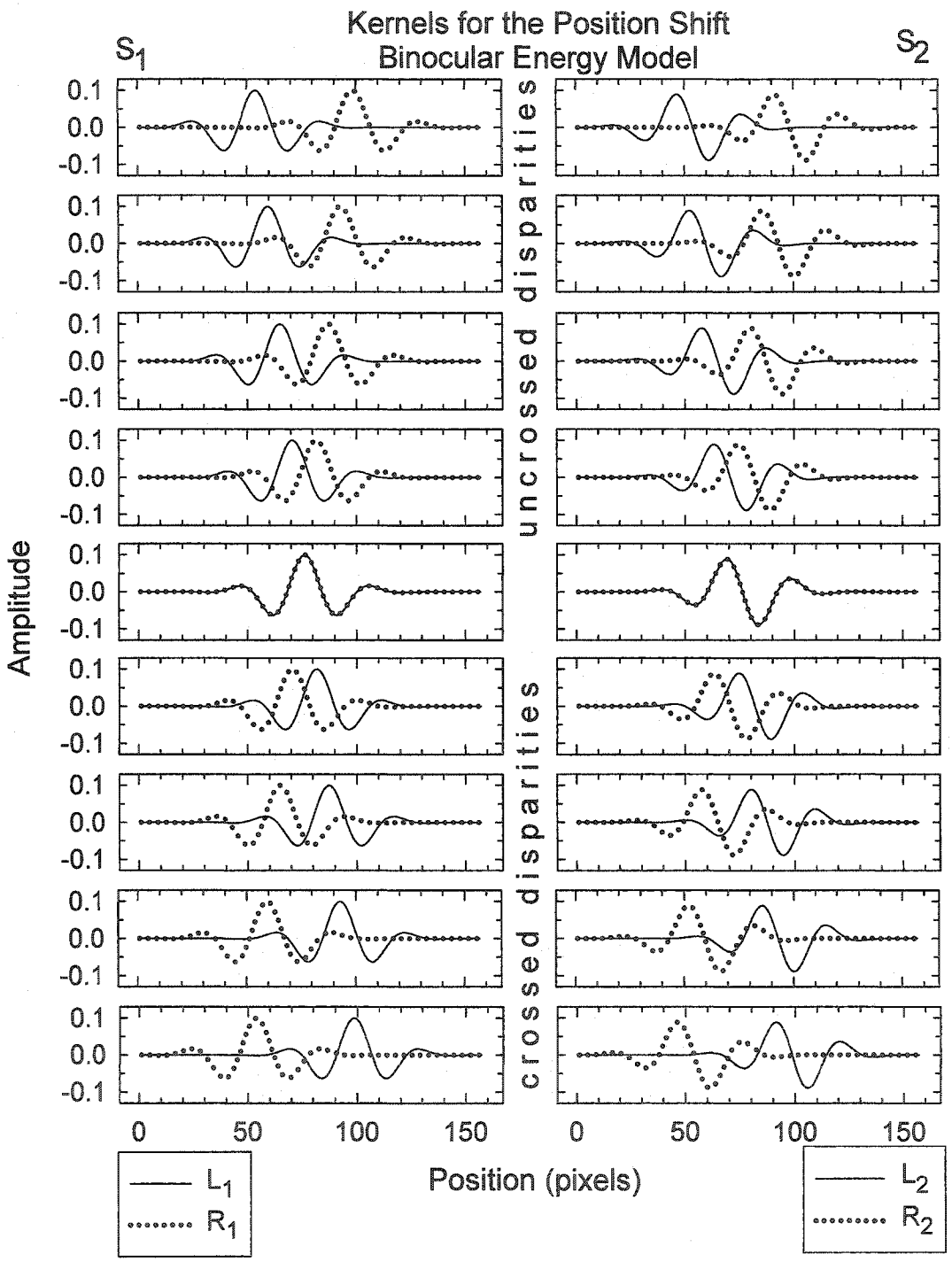

Figure A3. The kernels for the position shift version of binocular energy model.

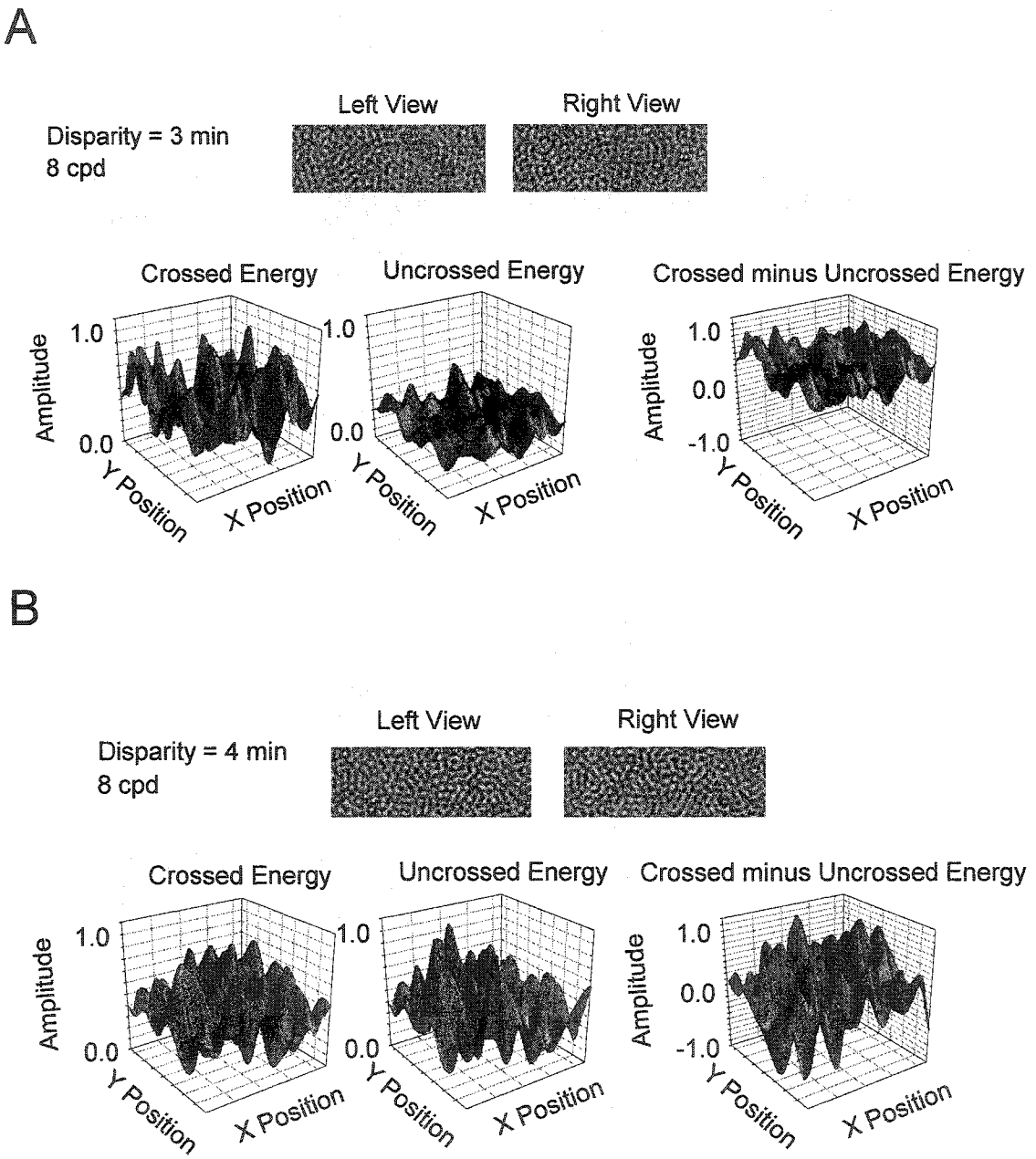

Figure A4. Simulation of the binocular energy model (phase shift version) for the fixed hard-edge window patterns at 8 cpd at disparity of (A) 3 minutes and (B) 4 minutes. For each disparity the corresponding crossed and uncrossed binocular energy is shown in the left and centre panels. The difference between the crossed and uncrossed energy is shown in the right panel.

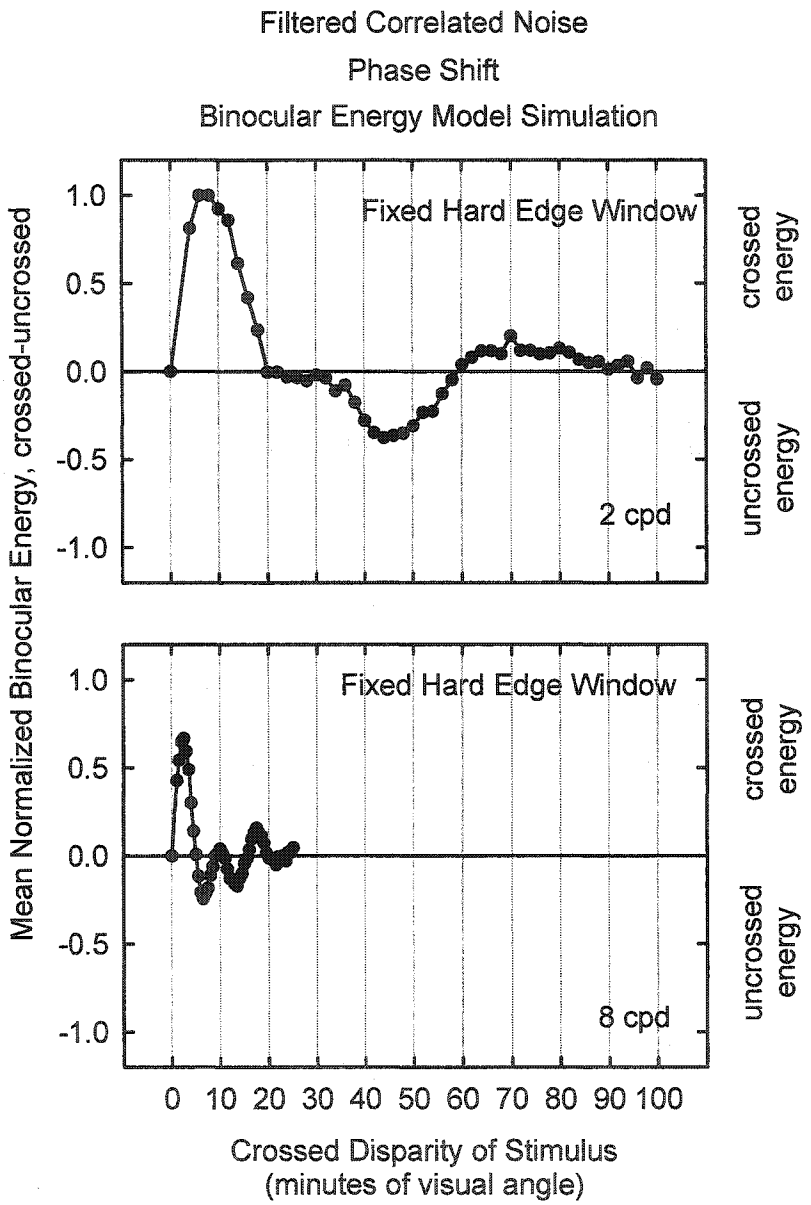

Figure A5. Simulations of the binocular energy model (phase shift version) for the fixed window hard-edge patterns used in Experiment 1. Simulations are shown for the 2 cpd & 8 cpd patterns.

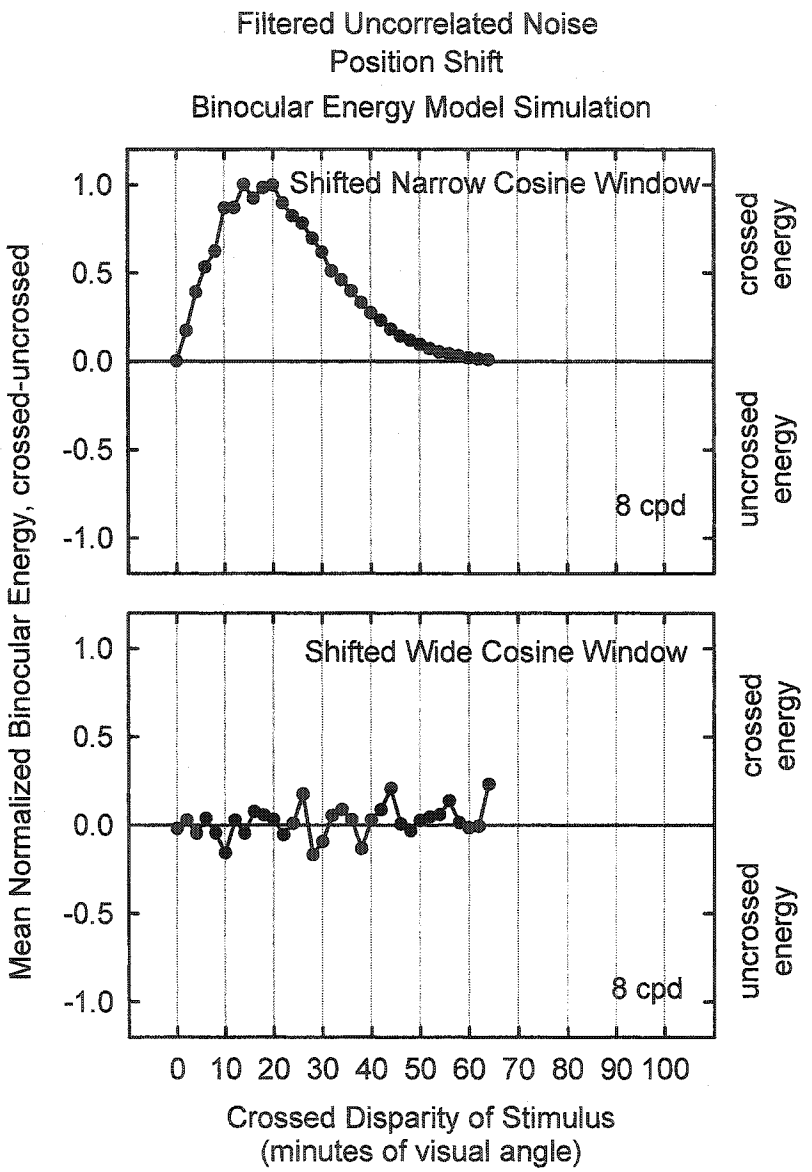

Figure A6. Simulations of the binocular energy model (position shift version) for uncorrelated filtered noise stimuli with a shifted narrow cosine contrast window (top panel) and a shifted wide cosine contrast window (bottom panel). Both simulations were for 8 cpd patterns. Follows the same format as Figure A5.

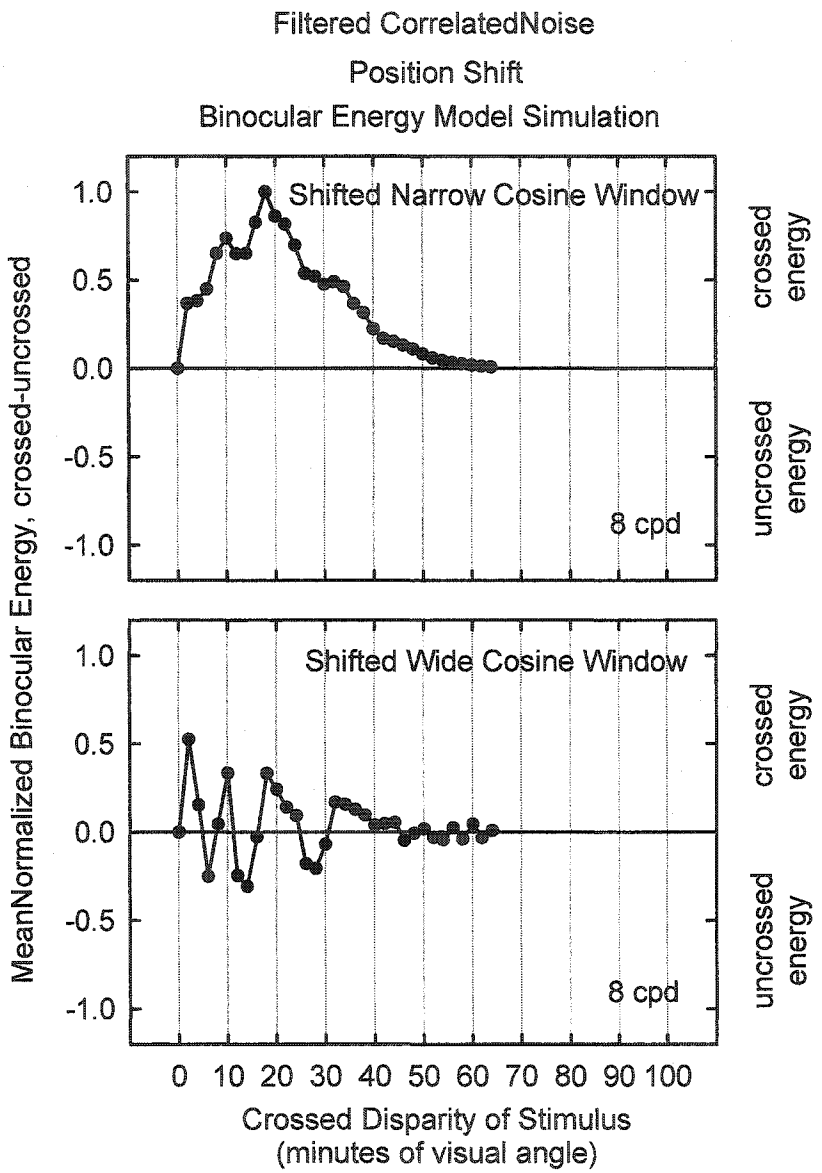

Figure A7. Simulations of the binocular energy model (position shift version) for correlated filtered noise stimuli with a shifted narrow cosine contrast window (top panel) and a shifted wide cosine contrast window (bottom panel). Both simulations were for 8 cpd patterns. Follows the same format as Figures A5-A6.

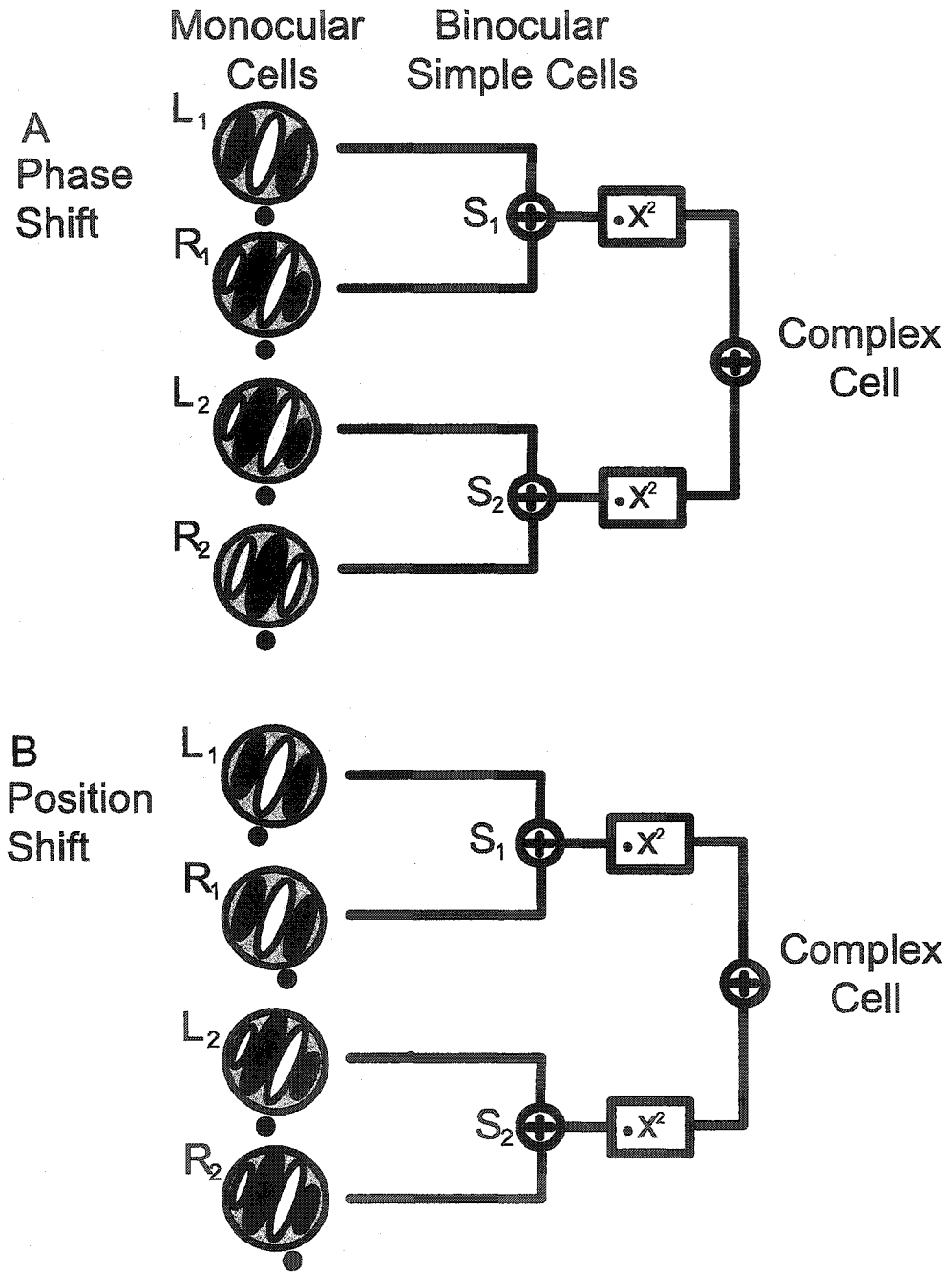

Figure A8. The binocular energy model: (A) phase shift; and (B) position shift versions with oblique oriented kernels.

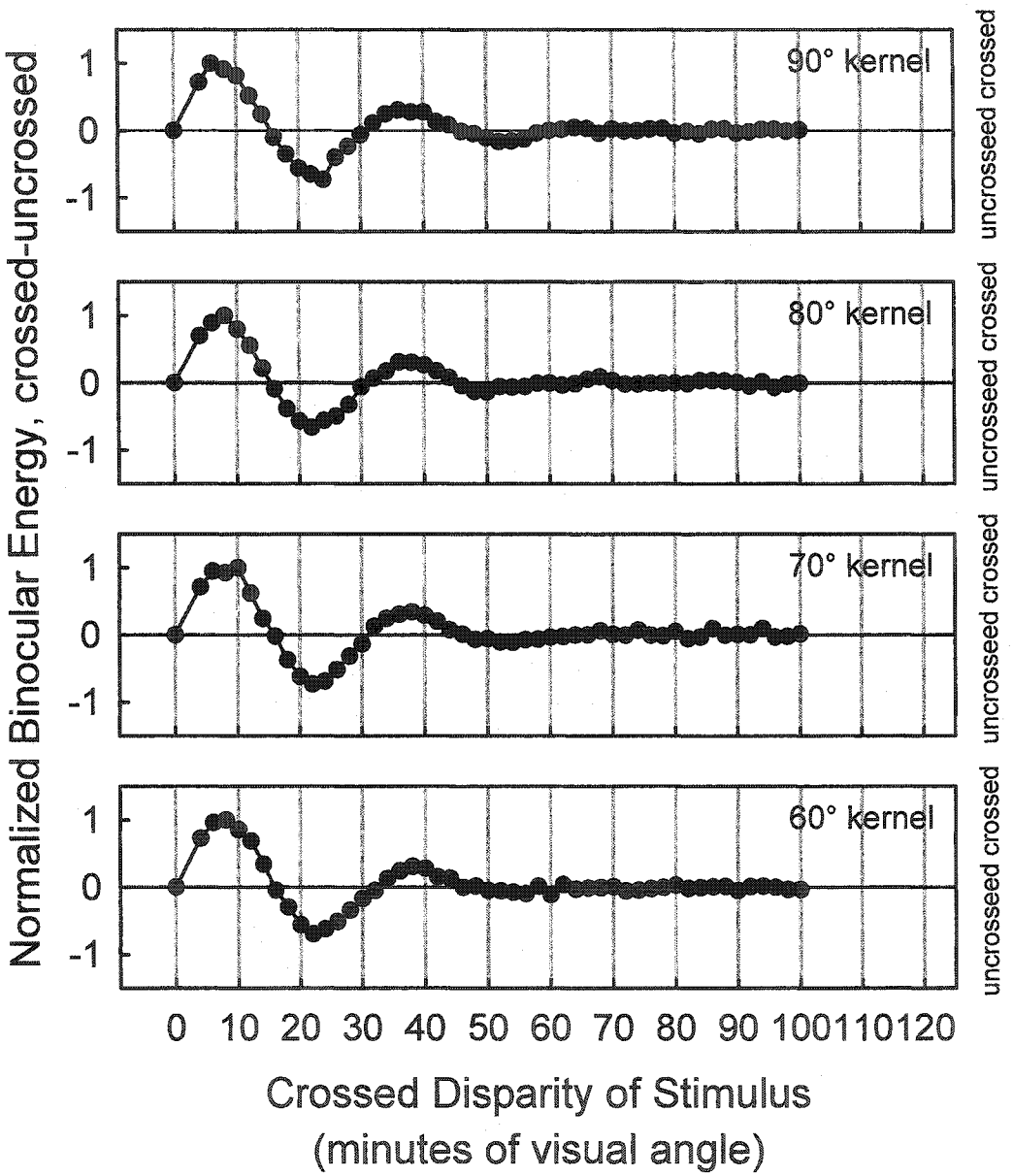

Figure A9. Simulations of the binocular energy model (phase shift variant) with oriented kernels for the fixed window oriented filtered noise patterns as in Figure 5. Simulations are shown for 2 cpd filtered noise patterns with kernels oriented at 90°, 80°, 70° and 60°.

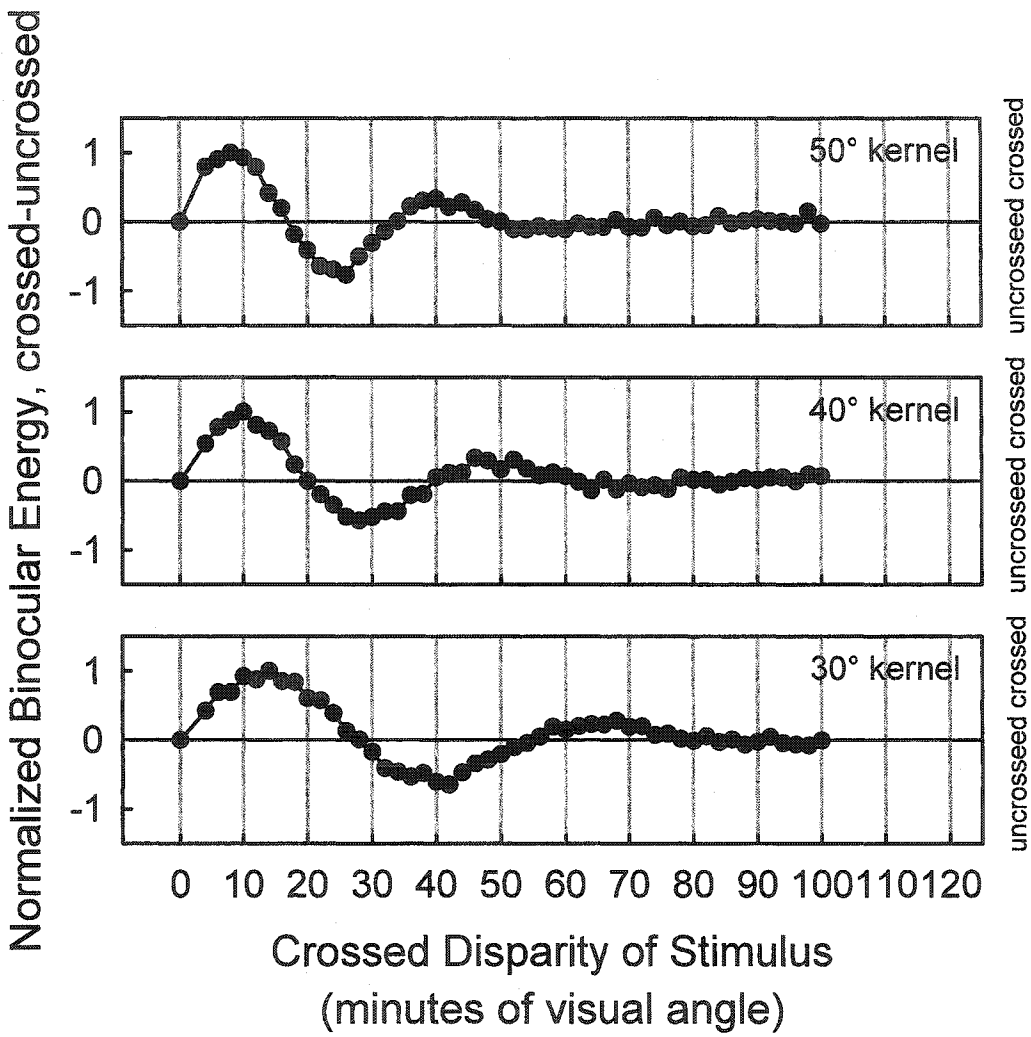

Figure A10. Simulations are the same as those in Figure A9, but for 2 cpd fixed window oriented filtered noise patterns and kernels oriented at 50°, 40° and 30°.

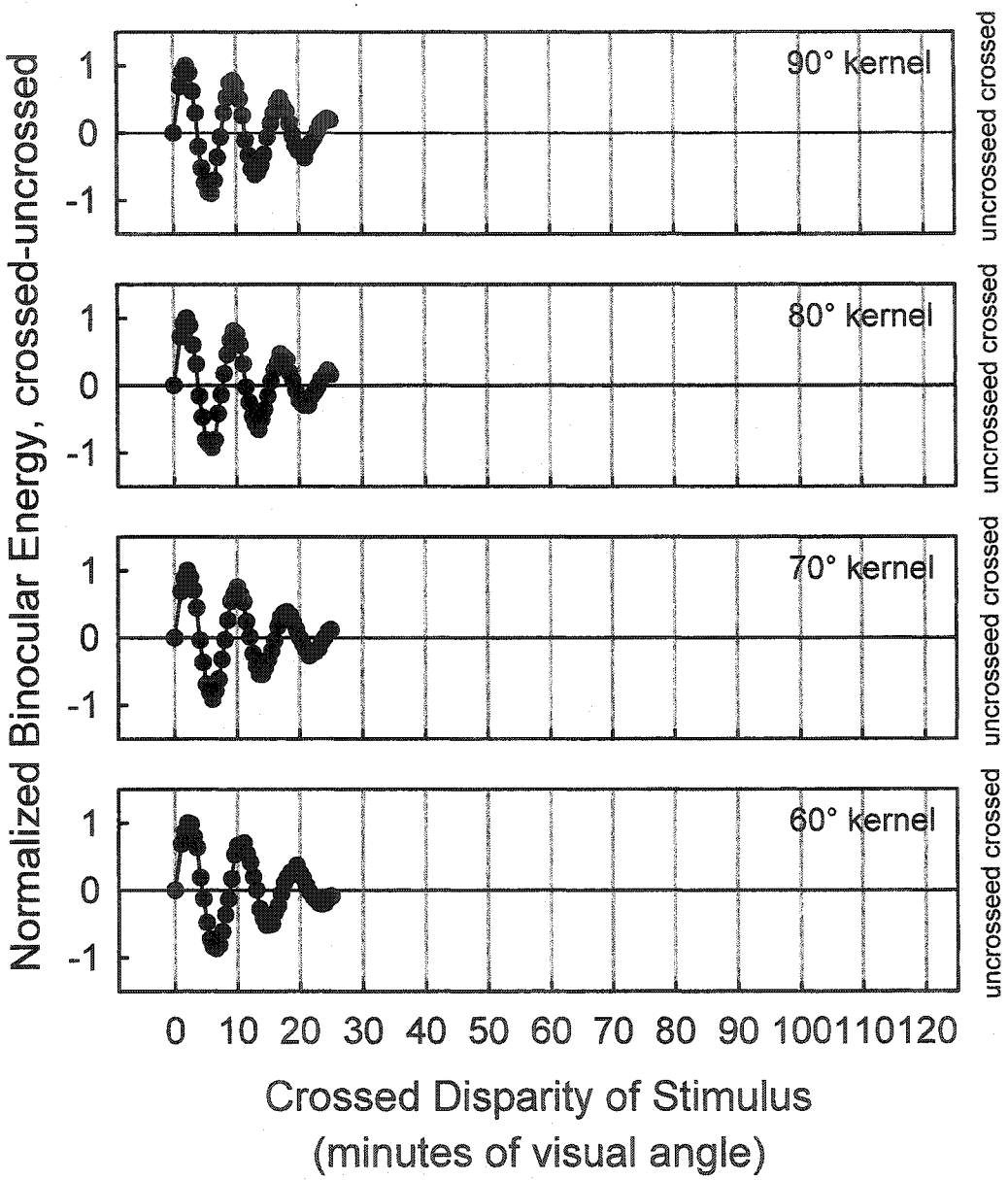

Figure A11. Simulations are the same as those in Figures A9-A10 but for 8 cpd fixed window oriented filtered noise patterns and kernels oriented at 90°, 80°, 70° and 60°.

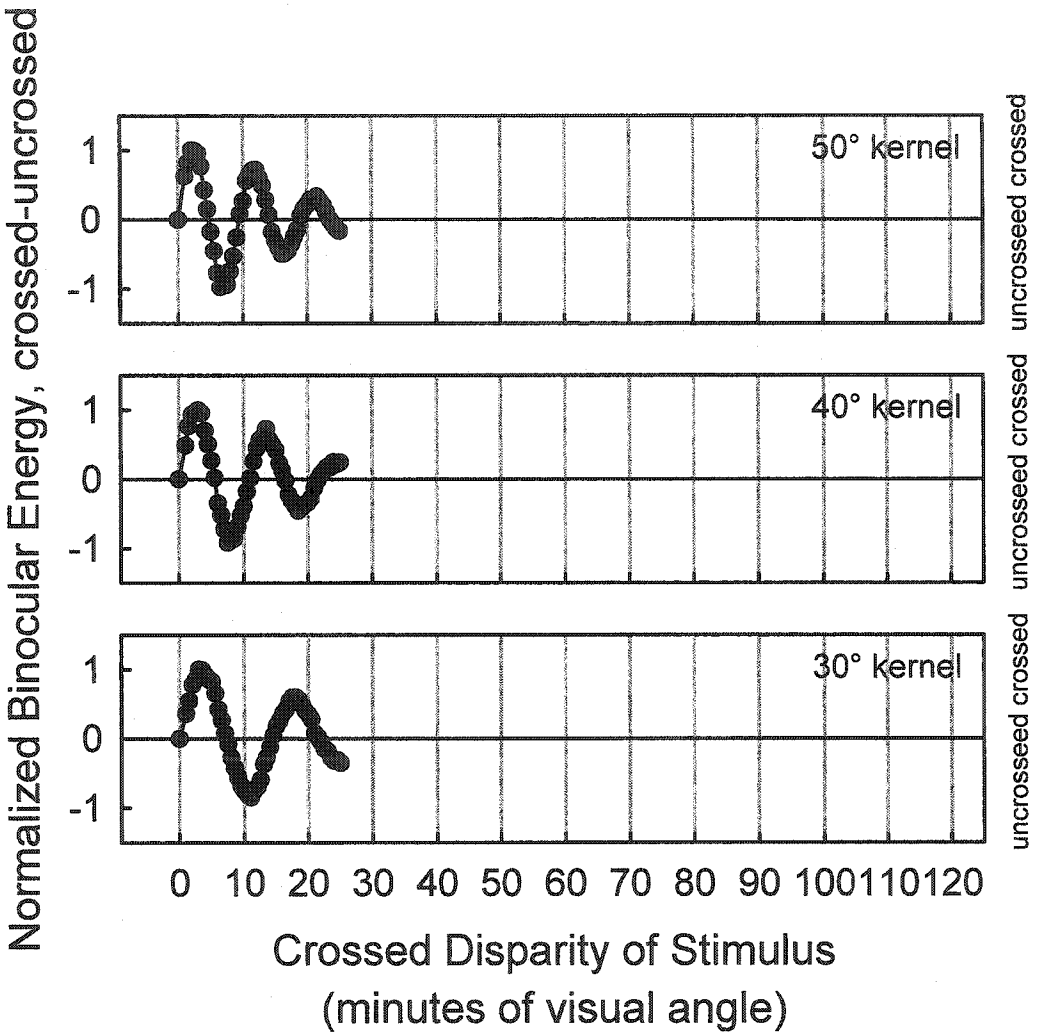

Figure A12. Simulations are the same as those in Figure A9-A11 but for 8 cpd fixed window oriented filtered noise patterns and kernels oriented at 50°, 40° and 30°.

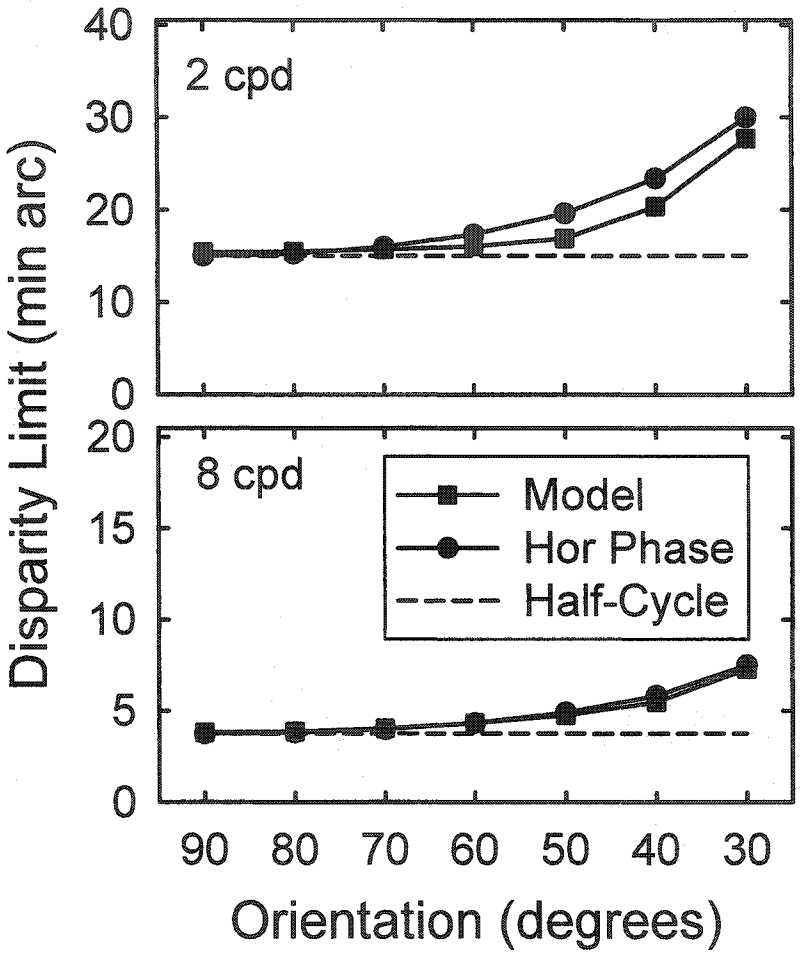

Figure A13. Simulations (binocular energy model, phase shift variant) as a function of kernel orientation are compared to theoretical predictions. Separate panels are for 2 cpd & 8 cpd images. X-axis shows the orientation of the kernel in degrees from horizontal (values from 90° to 30°). Y-axis shows the upper disparity limit in minutes. The binocular energy model output (squares) is compared to the prediction of a half-cycle of horizontal spatial frequency (circles) and a half-cycle of spatial frequency from the size-disparity correlation (dashed line). The upper disparity limit from the binocular energy model output was obtained from the curves in Figures A9-A12 using curve fitting. The half-cycle of horizontal spatial frequency was calculated by dividing the nominal value of disparity (i.e. half-cycle of spatial frequency) by the sine of the angle from horizontal as shown in Figure 4.

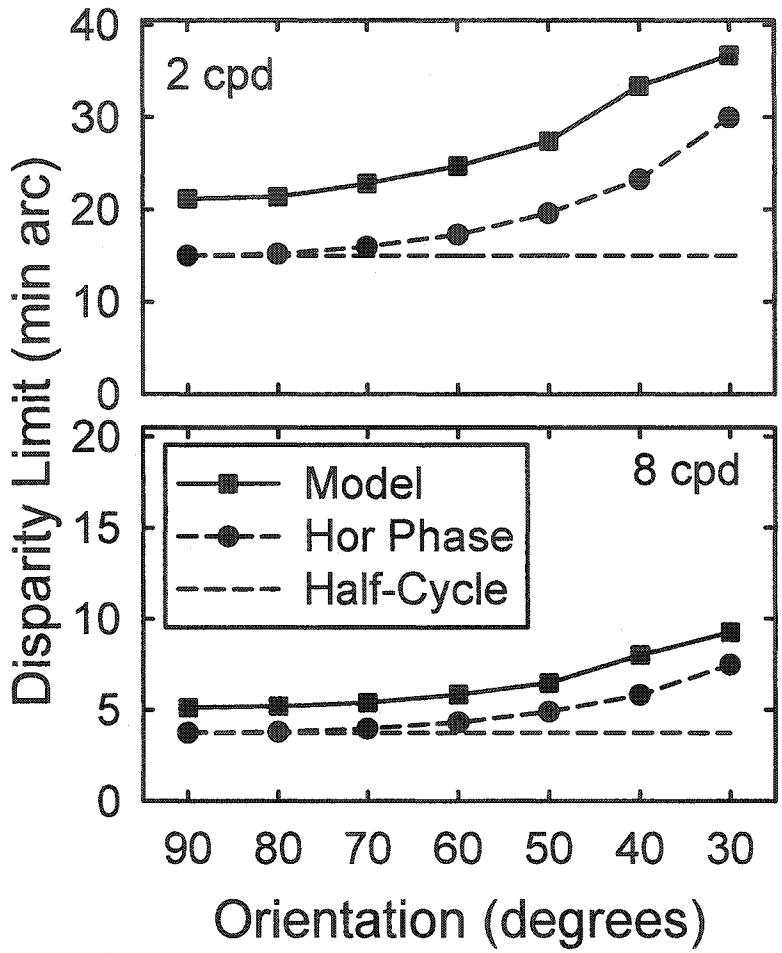

Figure A14. Same as Figure A13 except that modeling results are for the position shift variant of the binocular energy model. The disparity range of the complex cells was set to four times the range used in the phase shift model. Follows the same format as Figure A13.