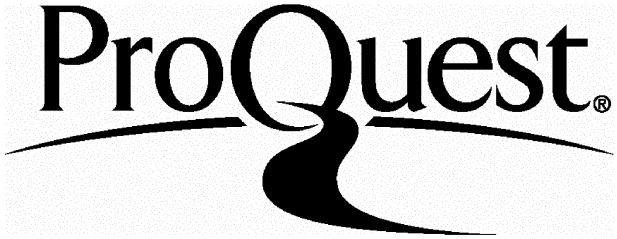
Linear and non-linear mechanisms in the perception of stereoscopic slant and transparency

Paul B. Hibbard

University College London

August 3, 1997

Thesis Submitted for the Degree of Doctor of Philosophy



Abstract

This thesis explored the role of contrast disparities in stereopsis, and the nature of the encoding of surface slant from stereoscopic cues. Contrast disparities may be defined as interocular differences in the position of image regions corresponding in terms of image contrast, rather than luminance. It was found that, for simple plaid stimuli, stereoscopic slant thresholds could be predicted from disparities in the plaid's components. Further, the perceived slant of grating and plaid stimuli was found to be underestimated, with the degree of underestimation for plaids depending on the orientation of their component gratings. These results may be explained in terms of the orientation and spatial frequency disparities in the Fourier components of the stimuli, and are consistent with the notion that orientation disparities provide the primary cue to stereoscopic slant (Rogers and Graham, 1983).

For plaids with orthogonal components, differing in contrast and spatial frequency, stereoscopic transparency was observed. Transparency was also observed in stimuli for which depth was defined by contrast modulation disparities. Transparency was only perceived for crossed disparities of the contrast modulation, such that the modulation appeared to lie in front of its carrier. This asymmetry was not evident if additional luminance disparities were introduced to the image. These results support the view that stereopsis has access to independent, linear and nonlinear channels (Hess and Wilcox, 1994). However, it was found that adaptation to the carrier of a contrast modulated stimulus increased the minimum contrast at which contrast disparities could be detected, suggesting that significant nonlinearities in stereopsis are preceded by a stage of linear filtering.

These results were explained using a model in which luminance and contrast disparities are processed by independent linear and nonlinear mechanisms,s haring a common linear filtering stage early in processing.

Acknowledgements

I would like to thank my supervisor, Keith Langley, for his help, advice and support throughout the thesis. Oliver Braddick and David Fleet also provided many helpful comments and ideas.

I am grateful to many others at UCL for their encouragement and support. I would particularly like to thank Colin Clifford, Tom Hartley, Stevie Sackin, Peter Howell, Alan Johnston, Mark Gardner, Nico Preston, John Draper, Kate Bradford, and the Psychology Allstars . I would also like to thank David Green and George Houghton, for interesting me in Psychology as an undergraduate.

Finally, I would like to thank Isabel, Jo, Paul(s), Carol, Ben, Mikie, Arshad, Garreth, Helen and the Loughborough posse for their friendship throughout.

This thesis is for my parents.

Contents

List of Figures

List of Tables

1. Introduction

1 General Introduction

One of the more interesting features of many visual systems is their ability to infer depth information from differences in the images projected to the left and right eyes. These differences are introduced by the two eyes viewing the same scene from different locations. A different geometric relationship will exist between objects in the visual scene, and the two retinae. This geometric relationship will determine the mapping of the three dimensional coordinate frame to the image formed on each retina. It follows that the two binocular images will not be identical. The differences between the two images are known as binocular disparities, and may be specified in terms of the differences in either the optic arrays subtended by each point to the two eyes, or the positions of the projections of corresponding image points.

The disparity associated with a point is determined by the position of the point relative to the observer, and the observer's viewing geometry. If a viewer fixates a

point, such that both eyes are pointed directly towards it, the point will have zero disparity. Other points will have disparities determined in part by their distance in depth away from this fixation point. Disparities thus provide a source of information relating to the distance to points in the image. Wheatstone (1838) first demonstrated that depth may be perceived when two slightly different photographs or drawings are presented to the two eyes.

To determine the binocular disparity associated with a point, it is necessary to establish which points in the two retinal images correspond to the same point in three-dimensional space. This is known as the correspondence problem (Marr and Poggio, 1979). Once correspondence has been determined, it is possible to compare the positions of points in the two images, and to infer the binocular disparity. In the stereograms used by Wheatstone, features were visible monocularly which were thought to have been used to solve the correspondence problem (Sherrington, 1906). However, Julesz (1960) demonstrated that extensive pre-processing of monocular images is not required to solve the correspondence problem. Julesz devised a stimulus known as a random dot stereogram, consisting of a binocular pair of images, both of which consist of random visual noise. The noise in the two images of a random dot stereogram is identical, except that some points are shifted between left and right images, so as to produce a binocular disparity. Depth is perceived in these stereograms consistent with these disparity cues. These stimuli are interesting due to the ambiguity inherent in the correspondence problem. For any given point in the left image, there will be a number of points in the right image with the same luminance, to which the point may be matched. These stimuli demonstrate that the correspondence problem may be solved in the presence of point-wise ambiguity in left- and right-eye matches.

Once correspondence has been achieved, it is possible to determine binocular disparity, and from this to infer depth. Depth supplies information relating to the

distance to objects, and to their shape. This thesis addresses two questions related to the representation of surface shape on the basis of stereoscopic cues. The first question is what types of mechanisms are used to infer surface shape from stereoscopic disparity. While it is in principle possible to determine the depth of individual points from horizontal disparities, it has been suggested that the analysis of higher order properties of shape, such as depth discontinuities, and surface orientation and curvature, exploits corresponding higher order properties of image disparities (e.g. Rogers and Graham, 1983; Gillam, Flagg and Finlay, 1984; Stevens and Brookes, 1987; Brookes and Stevens, 1989; Rogers and Cagenello, 1989). The second question relates to the types of monocular information which may be used for binocular matching to establish correspondence. Conventional models of stereopsis (e.g. Marr and Poggio, 1979; Grimson, 1980; Mayhew and Frisby, 1980), developed in light of the correspondence problem in random dot stereograms, match points on the basis of their luminance. However, it has been suggested that contrast envelopes may present another source of monocular information which may be used for binocular matching, and which may support depth perception (Liu, Schor and Ramachandran, 1992; Sato and Nishida, 1993; 1994; Hess and Wilcox, 1994; Wilcox and Hess, 1995; 1996).

In this chapter, the theoretical and empirical backgrounds of these issues are discussed. In section 2, *The geometry of binocular image projection*, the relationships between three dimensional shape, depth, and binocular disparity are discussed. Section 3, *Stereoscopic slant*, presents evidence that, in the representation of surface slant, stereospsis is able to exploit binocular differences other than the positional disparities associated with individual points. Section 4, *Nonlinear channels in stereopsis*, is concerned with the suggestion that contrast envelopes may serve as primitives for stereospsis. Finally Section 5, *Overview*, outlines the goals of the thesis.

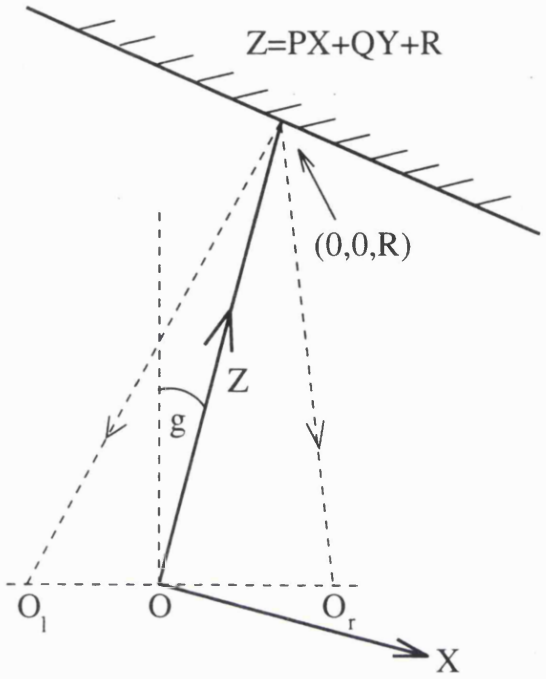


Figure 1.1: *The coordinate system used. For details, see text.*

2 The geometry of binocular image projection

Disparities occur in binocular image pairs as a result of the projection of a single visual scene to two image planes, horizontally separated in space. An analysis of how points in three-dimensional space project independently to the two retinae may be used to understand these disparities. Longuet-Higgins and Prazdny (1981) analysed the general problem of the projection of a moving three-dimensional scene to a single image plane; Mayhew and Longuet-Higgins (1982) applied this analysis to the specific problem of binocular vision. The geometry presented below is based on this latter analysis.

Figure 1.1 represents the geometry of binocular viewing. The problem addressed

is that of the disparities introduced as a result of the projection of the three-dimensional world coordinate system XYZ to the two-dimensional left- and right-image coordinate systems, xy and $x'y'$. The world coordinate system used has an origin at O, the midpoint of the line joining the optical centres of the eyes (O_l and O_r). The Z axis is defined by the line joining O to the fixation point, $(0, 0, R)$. The distance R, in the direction of the Z axis, gives the fixation distance. The X axis lies in the plane containing the line O_lO_r and the fixation point; the Y axis is normal to this plane. The gaze angle, g, is given by the angle between the X axis and the line O_lO_r. The interocular distance O_lO_r is given by I. The following analysis refers to the projection of the plane $Z = PX + QY + R$, which refers to any plane passing through the fixation point. It is assumed that the fixation distance R is sufficiently larger than the interocular distance I that terms in the second order of $\frac{I}{R}$ may be ignored. The horizontal and vertical disparity of each point in the scene is given by $(d_h, d_v) = (x' - x, y' - y)$, which describes the shift in the projected image of the point between left and right images. This disparity is given by:

$$d_h = \left[(P\cos g + \sin g)x + Qy\cos g + (\cos g - P\sin g)x^2 - Qxy\sin g\right]\frac{I}{R} \quad (1)$$

$$d_v = \left[y\sin g + (\cos g - P\sin g)xy - Qy^2\sin g\right]\frac{I}{R} \quad (2)$$

Assuming g is small, then $\sin g$ and $\cos g$ may be replaced by g and 1, respectively, and Pg and Qg may be neglected. (1) and (2) then become:

$$d_h = \left[(Px + Qy + gx + x^2\right]\frac{I}{R} \quad (3)$$

$$d_v = [gy + xy]\frac{I}{R} \quad (4)$$

For the approximations given in equations (3) and (4), both horizontal and vertical

disparities vary spatially. However, only horizontal disparities are affected by surface shape, which is here determined by P and Q. Slant about a vertical axis will introduce a horizontal gradient of horizontal disparity. Similarly, slant about a horizontal axis will introduce a vertical gradient of horizontal disparity. Both horizontal and vertical disparities are scaled by $\frac{I}{R}$, the ratio of the interocular distance to the viewing distance.

This analysis suggests a method of computing the unknown parameters P and Q, which determine the slant of the surface about a vertical and a horizontal axis, respectively. Equation (3) shows that slant about a horizontal axis introduces a vertical gradient of horizontal disparity which is directly proportional to the magnitude of slant. A measurement of the vertical gradient of disparity could thus be used to estimate the magnitude of slant about a horizontal axis. This estimate would need to be scaled by the quantity $\frac{I}{R}$; Longuet-Higgins (1982) suggested that this latter quantity may be estimated from vertical disparities. Similarly, equation (3) shows that slant about a vertical axis introduces a horizontal gradient of horizontal disparity, which is again directly proportional to the magnitude of slant. However, disparity also changes horizontally as a result of eccentricity.

A gradient of horizontal disparity may be represented using matrix notation by:

$$\begin{bmatrix} d_h \\ d_v \end{bmatrix} = \begin{bmatrix} G_h & G_v \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{5}$$

where G_h and G_v give the horizontal and vertical components of the gradient, respectively. A vertical gradient of disparity may be described as a horizontal shear between left and right images. Similarly, a horizontal gradient of disparity represents a horizontal expansion or compression of the left image relative to the right image. These transformations are illustrated in figure 1.2. In principle, measure-

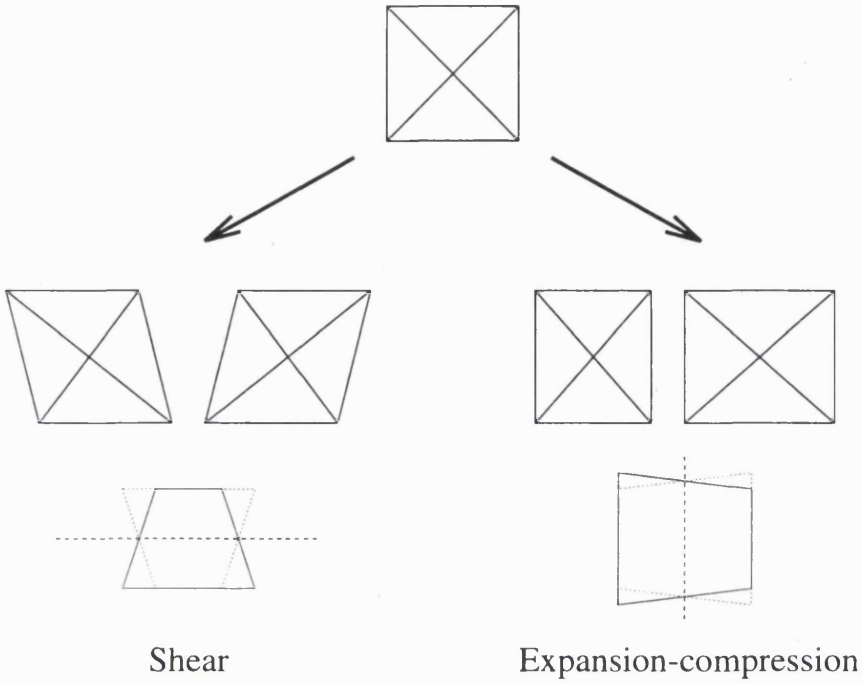


Figure 1.2: *Binocular transformations resulting from gradients of horizontal disparity. A horizontal shear is introduced by slant about a horizontal axis, while a horizontal expansion-compression is introduced by slant about a vertical axis. Perspective figures represent slant about these axes.*

ments of horizontal and vertical gradients of horizontal disparity may be used to estimate surface slant. However, disparity gradients are also introduced by other image transformations, unrelated to slant. These include two dimensional rotation of an image, as may occur if the eyes are torsionally misaligned, and unequal image magnification, or aniseikonia. In general, gradients of disparity may be described as an affine transformation:

$$\begin{bmatrix} d_h \\ d_v \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{6}$$

which may be rewritten:

$$\begin{bmatrix} d_h \\ d_v \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{bmatrix} + \begin{bmatrix} 0 & -\lambda_2 \\ \lambda_2 & 0 \end{bmatrix} + \begin{bmatrix} 0 & \lambda_3 \\ \lambda_3 & 0 \end{bmatrix} + \begin{bmatrix} -\lambda_4 & 0 \\ 0 & \lambda_4 \end{bmatrix} \tag{7}$$

where $\lambda_1 = \frac{1}{2}(a+d)$, $\lambda_2 = \frac{1}{2}(c-b)$, $\lambda_3 = \frac{1}{2}(b+c)$ and $\lambda_4 = \frac{1}{2}(d-a)$. Equation (7) represents an irreducible representation of the affine transformation, in terms of its geometrical invariants (Koenderink and van Doorn, 1975). The first term in (7) represents a uniform expansion of magnitude λ_1. The second term represents a rotation. The final two terms in (7) represent the deformation component. Rotation through a small angle α is given by:

$$\begin{bmatrix} d_h \\ d_v \end{bmatrix} = \begin{bmatrix} 1-\cos\alpha & 1-\sin\alpha \\ 1+\sin\alpha & 1-\cos\alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{8}$$

$$\approx \begin{bmatrix} 0 & -\lambda_2 \\ \lambda_2 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{9}$$

where $\lambda_2 = \sin\alpha$ and the small angle approximation $\cos\alpha \approx 1$ is used.

The deformation component represents a stretching of the image in one direction, accompanied by a compression in the orthogonal direction. Koenderink and van Doorn (1976) showed that the deformation component is directly related to surface slant. This may be seen by substituting the transformation in (5) into (7):

$$\begin{bmatrix} d_h \\ d_v \end{bmatrix} = \begin{bmatrix} G_h & 0 \\ 0 & G_h \end{bmatrix} + \begin{bmatrix} 0 & -G_v \\ G_v & 0 \end{bmatrix} + \begin{bmatrix} G_h & 0 \\ 0 & G_h \end{bmatrix} + \begin{bmatrix} 0 & G_v \\ -G_v & 0 \end{bmatrix} \quad (10)$$

Slant about a horizontal or a vertical axis will introduce a component of rotation, or of dilation, respectively. Both will introduce a component of deformation. In the same way, rotation and dilation of an image will both introduce horizontal and vertical gradients of disparity. The advantage of the representation in equation (10) is that deformation is unaffected by rotation and dilation, and is directly related to surface slant. A mechanism sensitive to the magnitude and direction of deformation would therefore have direct access to the magnitude and direction of surface slant (the magnitude being subject to a scaling factor).

3 Stereoscopic Slant

The geometric analysis presented above describes how disparities are related to surface orientation. This section reviews empirical research into the use of stereoscopic information in the perception of shape. As in the geometric analysis, the main focus of the studies described here is the representation of surface slant. Additionally, only the role of horizontal disparities is discussed; the question of how vertical disparities may be used to scale depth estimates (Longuet-Higgins, 1982; Mayhew and Longuet-Higgins, 1982; Gårding, Porrill, Mayhew and Frisby, 1996; Rogers and Bradshaw, 1996) is not addressed.

Equations (3) and (4) show that, in a local neighbourhood, spatial variations in disparity are related to the orientation of a viewed surface. This analysis assumes that viewed points lie on a continuous planar surface. If, however, there is a depth discontinuity in the image, and neighbouring points lie at different depths, there will be

a corresponding disparity discontinuity in the binocular images. While such discontinuities cause problems for models of stereopsis (Nelson, 1975; Mayhew and Frisby, 1980), they also present a potentially efficient strategy for encoding surface slant (Gillam, Flagg and Finlay, 1984). For planar surfaces, disparities on a surface are redundant once the disparities at its boundaries have been determined, as they will vary linearly between the boundaries. Gillam et al. (1984) found that subjects were able to report the slant of a surface more quickly and accurately if the boundaries of the slanted region were defined by disparity discontinuities. Gillam, Chambers and Russo (1988) provided further examples of how stereoscopic efficiency is improved by the presence of disparity discontinuities. In addition, they reported similar facilitation of the identification of surface slant by the presence of discontinuities in disparity gradients, as occur when two surfaces meet on a common line.

Gillam et al. (1988) argued that, once correspondence has been achieved, additional processing, making use of higher order properties of disparity, is required to recover surface structure (fusion being achieved in their experiments up to several seconds before subjects were able to identify the slant of a surface). Ryan and Gillam (1993) argued that surface slant may be directly estimated from disparity gradients. The representation of slant would not then rely on the prior representation of the horizontal disparities associated with individual points. Prolonged exposure to a slanted surface will cause a subsequently presented frontoparallel surface to appear to slant in the opposite direction (Kohler and Emery, 1947; Bergman and Gibson, 1959; Wenderoth, 1970). Ryan and Gillam (1993) explored this aftereffect in stimuli containing discrete disparate elements. By maintaining the disparities associated with individual elements, but varying their separation, they showed that the strength of the aftereffect depended not on relative disparities themselves, but on the gradient of disparity that existed between points. Fusional limits may be equally dependent on the magnitude of the gradient of disparity. While it has been suggested that fusion is only possible within the range of disparities known as Panum's fusional area (Ogle,

1950), it has been demonstrated that fusion of a point is less likely if other points are present which are close in visual direction, but which lie at a difference distance (Volkmann, 1864; Helmholtz 1909). Burt and Julesz (1980) suggested that fusion is determined by the gradient of disparity, rather than its magnitude. They found that fusion does not occur for disparity gradients greater than around 1. Again, these results suggest that disparity gradients play an important role in stereospsis.

The magnitude of the disparity gradient generated by a slanting surface depends directly on the magnitude of slant, and is unaffected by direction. It follows that a procedure by which slant was represented on the basis of disparity gradients would be unaffected by the direction of slant. Images with equal magnitudes of disparity gradient would be perceived as surfaces with equal slant, regardless of direction. There exists, however, an anisotropy in the perception of slant about horizontal and vertical axes. Surfaces slanting about a horizontal axis have more apparent slant, and exhibit lower slant thresholds, and faster resolution, than do surfaces slanting about a vertical axis (Wallach and Bacon, 1976; Rogers and Graham 1983; Gillam et al., 1984, 1988; Michison and McKee, 1990; Mitchison and Westheimer, 1990; Gillam and Ryan, 1992; Cagenello and Rogers, 1993). It is possible that the visual system interprets vertical disparity gradients as resulting from surface slant more readily than it does horizontal gradients. Rogers and Bradshaw (1994) argued that this may be a deliberate strategy, since horizontal gradients will exist even for frontoparallel surfaces extending into the periphery (as is evident in equation (3)). A horizontal gradient of disparity cannot therefore be taken to be the result of slant about a vertical axis, in the same way that a vertical gradient of disparity may be taken to result from slant about a horizontal axis. This strategy may account for the relatively poor perception of slant about a vertical axis. In addition, Gillam (1968) demonstrated that slant about a vertical axis is more influenced by perspective conflict than is slant about a horizontal axis. For these reasons, Mitchison and McKee (1990) argued that an internal reference may exist for slant about a horizontal

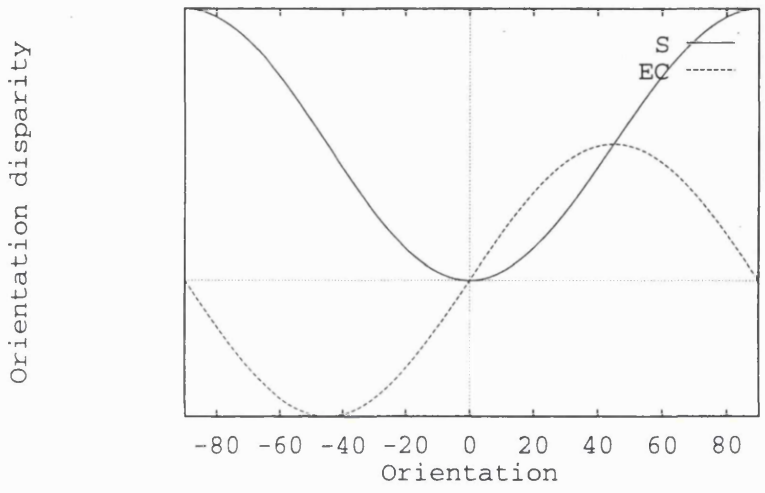


Figure 1.3: *Orientation disparity, plotted against orientation, for equal magnitudes of shear (solid line) and expansion-compression (dashed line). Orientation disparities for the two transformations have equal magnitudes only for orientations of* $\pm 45°$. *Overall, more orientation disparity is introduced by horizontal shear than by expansion-compression.*

axis, that is not available for slant about a vertical axis. This reference would allow for increased sensitivity to slant about a horizontal axis.

As an alternative, Rogers and Graham (1983) suggested that the anisotropy may be explained if slant perception is based on orientation disparities, rather than gradients of positional disparities. Orientation disparities were defined by Blakemore, Fiorentini and Maffei (1972) as interocular differences in the orientation of corresponding images contours. Figure 1.2 shows that vertical and horizontal disparity gradients may be described as a horizontal shear, and a horizontal expansion-compression,

respectively. Both transformations introduce orientation disparities. Orientation disparities for an image contour depend both on the type and magnitude of transformation, and on the orientation of the contour. Here, the orientation of a contour is defined in terms of its absolute (cyclopean) orientation, rather than the individual orientations present in the binocular images. For example, figure 1.2 shows that a horizontal shear does not affect the orientation of horizontal lines, and produces a maximum orientation disparity for vertical lines. For an expansion-compression, orientation disparities are evident for neither horizontal nor vertical lines, and have the greatest magnitude for lines at $\pm 45°$. Figure 1.3 shows how orientation disparity depends on orientation, for both shear and expansion-compression transformations. Integration of these graphs reveals that, on average, orientation disparities are 57% greater for a shear than for an expansion-compression of the same magnitude (Cagenello and Rogers, 1993). If orientation disparities provided the primary cue to surface slant, one would predict that slant about a horizontal axis would be more readily perceived than slant about a vertical axis.

Figure 1.3 shows that shear and expansion-compression introduce equal magnitudes of orientation disparities to lines oriented at $\pm 45°$. Cagenello and Rogers (1993) found that, while the anisotropy between slant about horizontal and vertical axes was observed for random dot stereograms, and stimuli consisting of grids of horizontal and vertical lines, it was not apparent for grids of diagonal lines. These findings provide evidence that the anisotropy in slant thresholds is a result of the role of orientation disparities in the encoding of slant. However, Mitchison and McKee (1990) found an anisotropy in slant thresholds for both grids of horizontal and vertical lines, and grids of diagonal lines. Cagenello and Rogers (1993) argued that this difference was due to the size of the stimuli used in the two studies. Mitchison and McKee used stimuli with a diameter of 0.75 degree, whereas those of Cagenello and Rogers had a diameter of 10.66 degrees. Orban, Vandenbussche and Vogels (1984) demonstrated that line orientation discrimination is based on positional rather than

orientational differences for small (< 2degrees) stimuli. Cagenello and Rogers suggested that orientation disparities may similarly be used as a cue to slant only for sufficiently large stimuli. Gillam and Ryan (1992), however, reported anisotropic slant perception for diagonal lines, even for large stimuli.

Rogers and Graham (1983) showed that there is an analogous anisotropy in the perception of surfaces defined by motion parallax. For horizontal head movements, motion parallax information may be described as a horizontal shear for surfaces slanting about a horizontal axis, and as a horizontal expansion-compression for slant about a vertical axis. For vertical head movements, slant about a horizontal axis results in a vertical expansion-compression, whereas slant about a vertical axis results in a vertical shear. Rogers and Graham found that slant about a horizontal axis was more readily perceived for horizontal head movements. Conversely, slant about a vertical axis was more readily perceived with vertical head movements. These results demonstrate that it is a shear transformation that is more readily perceived than an expansion-compression transformation, rather than the axis of slant *per se* which is responsible for the anisotropy. These results may again be explained in terms of the orientation changes associated with the two transformations, which are unaffected by the directions of the transformations. Finally, the anisotropy is also evident for stereoscopically defined curved surfaces (Rogers and Graham, 1983; Cagenello and Rogers, 1988; Rogers and Cagenello, 1989). The analysis of orientation changes may thus represent a general strategy in the encoding of shape from both motion and stereo cues.

Other evidence has been forwarded to support the role of orientation disparities in stereoscopic slant. von der Heydt, Hänny and Dürsteller (1981) examined binocular pairs of one dimensional noise images. They found that, even when a different noise sample was presented to each eye, so that the stimulus appeared rivalrous, an inclined surface was perceived if there existed a difference in the orientation

of lines between the two eyes. In these stimuli, the luminances of corresponding points in left- and right-eye images are uncorrelated, and it is not possible to match points between images to determine horizontal disparities. Orientation disparities therefore provide the only cue to slant. Ninio (1985) investigated slant perception in line stereograms containing different amounts of orientation and positional disparity. He found that a smooth surface was more likely to be observed in stimuli containing orientation disparities, than in those without.

The role of orientation disparities in stereopsis is also supported by physiological evidence. Blakemore et al. (1972) found binocular cells in cat primary visual cortex that showed a difference in the orientation tuning of their left- and right-eye monocular receptive fields. Blakemore et al. argued that this difference in orientation tuning made the cells ideally suited as orientation disparity detectors, and suggested they may play a role in stereopsis. Nelson, Kato and Bishop (1977) have also described binocular cells in cat visual cortex which were tuned to different orientations in left and right eyes, and which responded to binocular orientation disparities. Hänny, von der Heydt and Poggio (1980) found cells similarly tuned to orientation disparities in the prestriate cortex of monkeys.

Interocular differences in spatial frequency have also been proposed as a cue to slant (Blakemore, 1970; Tyler and Sutter, 1979). Blakemore (1970) showed that slant about a vertical axis is perceived if vertical gratings with a difference in spatial frequency are presented to left and right eyes. This difference in frequency introduces a horizontal gradient of disparity. Slant may result from positional disparities, or from the frequency difference directly. Tyler and Sutter (1979) conducted experiments in which vertical sinusoidal gratings, drifting in opposite directions, were presented to left and right eyes. At slow drift rates, this stimulus appeared to move in depth either towards or away from the observer, depending on the direction of drift. Above a certain drift rate, however, motion in depth was no longer perceived. When this

stimulus was presented with a difference in frequency between left- and right-eye gratings, at a drift rate sufficient to destroy the perception of motion in depth, slant was observed. In addition, it was reported that slant was evident in dichoptic displays of uncorrelated vertical one-dimensional noise, differing between left- and right- eyes in mean spatial frequency. It was proposed that this slant percept was based on spatial frequency differences.

However, evidence exists that would appear to refute the notion that orientation and spatial frequency disparities play a central role in the perception of strereoscopic slant. The first, and most significant evidence against the orientation disparity account of slant perception, is the finding of Cagenello and Rogers (1993) that horizontal axis slant thresholds are smaller for random dot stimuli than for grids of horizontal and vertical lines. As vertical lines provide the greatest orientation disparities, stimuli containing vertical lines would be expected to provide the most powerful cue to horizontal axis slant. Rogers and Bradshaw (1994) argued that interocular spatial frequency differences may not provide a reliable cue to surface slant, due to the frequency differences introduced by eccentricity, even for frontoparallel surfaces. Halpern, Wilson and Blake (1996) have also argued that stereopsis from frequency differences is not robust. Their subjects failed to perceive slant in two-dimensional bandpassed random noise stimuli if the noise had an orientation bandwidth greater than $\pm 14°$; neither did they observe slant in oppositely drifted uncorrelated one-dimensional random noise stimuli with an interocular spatial frequency difference. Halpern et al. argued that slant may be perceived on the basis of positional disparities, even for uncorrelated noise patterns. Stereopsis in this case would be supported by local correlations in images. They used a computer simulation to show that this strategy may produce performance above chance in a slant discrimination task. They argued that the unrobust nature of stereopsis from frequency differences found in their empirical observations may be explained if the process relied on positional disparities measured on the basis of chance local corre-

lations in globally uncorrelated stimuli. Similarly, although the stimuli of von der Heydt et al. (1981) are based on one dimensional, uncorrelated monocular signals, it is possible that sufficient local correlations may have existed to support stereopsis.

Horizontal disparities may also play a role in the representation of surface slant. For a vertical grating, with an interocular difference in orientation or spatial frequency, analysis of positional disparities, and of orientation or frequency disparities, would predict the perception of the same slant. However, an analysis of positional disparities would also predict that, for a sufficiently large difference in orientation or spatial frequency, the surface would appear as a series of slanted patches, separated by horizontal or vertical depth discontinuities, introduced by aliasing. This discontinuous percept has been reported for both orientation (Piggins, 1978) and frequency (Tyler and Sutter, 1979; DeValois and DeValois, 1990; Halpern, Wilson and Blake, 1996) differences, and was described by Tyler and Sutter as being similar in appearance to a Venetian blind. DeValois and DeValois reported bistability between the perception of a Venetian blind, and a single surface, for gratings with an interocular frequency difference. This bistability may be explained in terms of conflicting depth cues generated by positional disparities and frequency disparities. Alternatively, it may represent conflict between local and global solutions of the correspondence problem. In either case, the existence of the Venetian blind percept demonstrates that position disparities play a role in the perception of slanted surfaces.

Another important issue is the question of how surface slant is encoded from disparity cues. Equation (10) demonstrates that disparity gradients may be introduced by rotation and dilation, as well as by horizontal shear and expansion-compression. In addition, orientation disparity is introduced by rotation, and frequency disparity by dilation. Both transformations may equally be represented as combinations of horizontal and vertical positional disparities. A mechanism representing slant on the basis of locally measured positional, orientation or frequency disparities would

be expected to perceive slant given a rotation or dilation between left and right images. Howard and Kaneko (1994; Kaneko and Howard, 1994) found that rotation and dilation of images between the two eyes generated relatively little apparent slant. Howard and Kaneko (1994) suggested that slant about a horizontal axis is derived from the difference between horizontal shear and vertical shear. Similarly, it was suggested that slant about a vertical axis is derived from the difference between horizontal and vertical expansion-compression (Kaneko and Howard, 1994). This strategy would predict that slant would be perceived from vertical shear, or from deformation, which is equivalent to equal but opposite horizontal and vertical shear, but not from rotation, which is equivalent to equal magnitudes of horizontal and vertical shear. Similarly, it is predicted that slant would be perceived from vertical expansion-compression (as is apparent in Ogle's induced effect (Ogle, 1938)), but not dilation. Howard and Kaneko's results contradict those of Gillam and Rogers (1991), who found that slant was perceived from rotation, but not from vertical shear. Howard and Kaneko (1994) argued that this was due to the zero disparity surround used in the latter study. They found that, for stimuli with a black surround, slant *was* perceived for cyclorotated stimuli subtending 10 degrees of visual angle, but not for larger stimuli. For stimuli with a textured, zero disparity surround, slant was perceived for all sizes of stimuli studied. They proposed that, while horizontal shear and expansion compression are measured locally, the equivalent vertical transformations may be measured more globally. This strategy would help to discount torsional misalignment of the eyes.

From a theoretical analysis, Koenderink and van Doorn (1976) proposed that slant may be encoded on the basis of binocular deformation. Equation (10) above demonstrates that deformation is related directly to surface slant. Further, Koenderink and van Doorn suggested that deformation may be encoded via an analysis of angular disparities, defined as the interocular difference in the angle formed between corresponding pairs of image contours. This encoding of slant would be unaffected

by global dilation or rotation of the image between left and right eyes.

Computational models have shown that estimates of surface slant may be obtained from algorithms based on orientation disparities. Wildes (1991) showed that angular disparities, measured from pairs of image contours, may be used to measure slant on the basis of deformation. Jones and Malik (1992) presented an algorithm which estimates slant on the basis of both orientation and spatial frequency differences. These models demonstrate that the theories proposed to account for psychophysical results are able to provide reliable estimates of slant.

4 Nonlinearity in stereopsis

It has been suggested that, in addition to disparities defined by local changes in luminance, stereopsis may also make use of disparities in contrast envelopes. Gabor stimuli have been used to study the role of contrast envelopes in stereopsis (Liu et al., 1992; Sato and Nishida, 1993; 1994; Hess and Wilcox, 1994; Wilcox and Hess, 1993, 1994, 1995, 1996). A Gabor stimulus consists of the product of a sinusoidal grating, and a Gaussian contrast envelope, such that the contrast of the sinusoidal carrier within any region is determined by the magnitude of the Gaussian envelope within that region. For a given carrier frequency, the size of the Gaussian envelope will determine the bandwidth of the stimulus; increasing the size of the envelope decreases the bandwidth of the Gabor. Hess and Wilcox (1994) found that stereoacuity for Gabor stimuli depends primarily on the size of the Gaussian envelope when the stimulus has a bandwidth smaller than 0.5 octaves. For larger bandwidths, stereoacuity is determined by the spatial frequency of the sinusoidal carrier. Moreover, Wilcox and Hess (1995) showed that the upper depth limits of stereopsis are determined primarily by the size of the Gaussian envelope for Gabor stimuli, relatively independently of the carrier frequency. Wilcox and Hess (1996) also investigated contrast

envelope disparities in stimuli in which one-dimensional noise was modulated by a vertical Gabor signal. They found that, although stereoacuity was poor, depth was perceived when the noise was uncorrelated between left and right eyes. Further, they found that stimuli in which the left- and right-eye one dimensional noise signals were orthogonal did not support stereoscopic depth. Liu et al. (1992) presented subjects with Gabor stimuli in which envelope and carrier disparities were manipulated independently. They found that stereoacuity was better when the envelope and the carrier had identical disparities than when only the carrier had a nonzero disparity. In contrast to Wilcox and Hess (1996), they also reported that when presented with binocular Gabor stimuli in which left- and right-eye sinusoidal carriers were orthogonal, subjects perceived depth correctly from the envelope while the carriers appeared rivalrous. Sato and Nishida (1993) presented subjects with second order random dot stereograms, much like those used in some studies of non-Fourier motion. They found that upper limits of disparity were lower with second order stimuli than with conventional random-dot stereograms.

Hess and Wilcox (1994) suggested that these results demonstrate that stereopsis has access to independent linear and nonlinear channels. They proposed that luminance disparities are processed by a disparity mechanism based on linear filters; cortical simple cells have been modelled as such a linear filtering stage (e.g. Hubel and Wiesel, 1962; Campbell, Cooper and Enroth-Cugell, 1969; Movshon, Thompson and Tolhurst, 1978; Ohzawa and Freeman, 1986). A nonlinear operation is however necessary to make contrast envelopes explicit and available to disparity processing. Wilcox and Hess (1996) proposed that complex cells may provide such a nonlinearity (Spitzer and Hochstein, 1985). On the basis of their finding that envelope disparities supported stereopsis when left- and right-eye carrier signals were uncorrelated, Wilcox and Hess (1996) argued that the nonlinearity making the contrast envelope explicit occurs prior to binocular combination.

This two channel model of stereopsis is analogous to those used to explain non-Fourier motion (e.g. Chubb and Sperling, 1988; Victor and Conte, 1992; Wilson, Ferrera and Yo, 1992; Zhou and Baker, 1993; Fleet and Langley, 1994a). Derrington and Badcock (1986) demonstrated that, in a plaid formed from the product of a high frequency (carrier) and low frequency (modulation) sinusoidal grating, transparent motion may be seen if the carrier and modulation move with different velocities. This stimulus contains no Fourier component with the velocity of the modulation, yet motion is nevertheless seen with this velocity. Chubb and Sperling (1988) described this as "non-Fourier" motion, and defined a class of stimuli, which they labelled drift balanced stimuli, for which this idea was extended. They proposed that motion in these stimuli could not readily be understood in terms of their Fourier spectra. Fleet and Langley (1994a) demonstrated that idealisations of many of these stimuli have a relatively simple characterisation in the Fourier frequency domain. However, a simple mechanism relying on velocity estimates from image Fourier components would fail to detect motion in these stimuli. Chubb and Sperling suggested that motion may be perceived following full-wave rectification of the image (after a stage of spatially broadband filtering), which would have the effect of introducing Fourier components with the required velocity. This nonlinearity forms the initial stage of an independent, non-Fourier channel in motion processing, which may account for the perception of motion which does not correspond directly to image Fourier components. Other two channel models have been proposed, in which nonlinearities occur relatively late in processing, after a stage of orientation- and spatial frequency-specific filtering (Wilson et al., 1992; Zhou and Baker, 1993; Fleet and Langley, 1994a).

It is argued here that it is not necessary to propose independent linear and nonlinear channels to understand the above results. Rather, they may be explained in terms of a single disparity processing mechanism. Fleet, Wagner and Heeger (1996) proposed a model of stereopsis in which disparities are processed by a binocular energy

mechanism. The receptive fields of cortical simple cells may be modelled using two dimensional Gabor functions (Jones and Palmer, 1987), with adjacent cells demonstrating phase relationships of 90° or 180° (Foster, Gaska, Marcelja and Pollen, 1983; Liu, Gaska, Jacobson and Pollen, 1992; Palmer and Davis, 1981; Pollen and Ronner, 1981). Gabor functions with a quadrature (90°) phase relationship may be used to compute an energy response (Adelson and Bergen, 1985; Emerson, Bergen and Adelson, 1992; Heeger, 1992). Fleet et al. proposed that complex cells may function as binocular energy neurons, nonlinearly combining responses from quadrature pairs of linear neurons (simple cells) in both left- and right-eyes. Binocular energy neurons in the model are given disparity tuning by altering the left- and right-eye receptive fields of simple cells. The binocular receptive fields are related by position shifts and phase shifts. A position shifted neuron has receptive fields with an identical shape in the left- and right-eye, but which are shifted in position between eyes. For a phase shifted neuron, receptive fields in the left and right eyes occupy the same position, but have a difference in shape, taking the form of a shift of the phase of the sinusoidal component of the Gabor. Additionally, hybrid neurons were proposed, combining a position shift with a phase shift. To provide robust and reliable estimates, the model computes disparity by pooling responses of binocular energy units over position, orientation and scale.

In this model, simple and complex cells do not constitute independent linear and nonlinear channels, but stages in a single channel. The nonlinearities in the model will however mean that its responses will be sensitive to image envelope disparities. This will be true especially when the carrier and envelope have different disparities. Here, a perfect match between left and right signals cannot be achieved with any disparity; matching under these circumstances would be expected to be unrobust, and noise sensitive. This may explain why disparity thresholds, and upper disparity limits, are dependent on envelope size for Gabor stimuli in which envelope and carrier have equal disparities (Hess and Wilcox, 1994; Wilcox and Hess; 1995). It will also

explain why disparity is more easily detected when carrier and envelope have the same disparity, and left and right images are perfectly correlated. In addition, by virtue of its integration across scale and position, this computational strategy is able to detect disparities larger than one half cycle of the spatial frequency components of the stimulus.

It may also be possible to account for stereospsis from contrast modulated uncorrelated noise patterns using a single channel model. Wilcox and Hess (1996) demonstrated that, for a spatial average taken over the entire of extent of the stimulus, their stimuli were uncorrelated. This does not however exclude the possibility that local correlations may have existed. Any such local correlations may have been sufficient to support stereopsis. Using a similar argument, Halpern et al. (1996) showed that the stimuli of Tyler and Sutter (1979) provided sufficient local correlations that a positional disparity based mechanism could support a slant discrimination task above chance. While a mechanism which is dependent on chance local correlations would not be expected to be robust, it should be borne in mind that stereoacuity for the uncorrelated stimuli used by Wilcox and Hess was poor.

As an alternative, it may be possible to explain the detection of contrast envelopes in stereopsis if a single, linear processing channel is preceded by an early nonlinearity. Non-linearities are known to occur early in visual processing, in the responses of X-cells in the cat LGN (Derrington, 1987). Chen, Makous, and Williams (1993) suggested that nonlinearities may occur retinally. It is often assumed that these nonlinearities may take the form of a logarithmic compression of the input signal (Derrington, 1987). These nonlinearities would affect all subsequent visual processing, and thus may not be identified with nonlinearities present in two channel models, proposed to detect non-Fourier image properties. Such nonlinearities would introduce Fourier components which are not present in the original image. These additional components are known as distortion products. Both compressive and ex-

pansive nonlinearities may introduce additional components with the orientation and spatial frequency of image contrast variations. Burton (1973) suggested that the visual system may rely on these distortion products to detect contrast beats. As such, early nonlinearities may play a functional role in the detection of image contrast variations. Burton (1973) found that prolonged exposure to the product of a high frequency (carrier) and a lower frequency (modulation) sinusoidal grating increased contrast detection thresholds for gratings close in orientation and spatial frequency to the modulation. Since there is no Fourier component with this orientation and frequency in the image, it was proposed that adaptation of the relevant channel resulted from the presence of distortion products. Henning, Hertz and Broadbent (1975) found reciprocal masking effects between contrast beats and luminance gratings of equal orientation and frequency. Masking effects were greatest when gratings were 90° out of phase with the distortion product that woud be expected to be introduced by a compressive nonlinearity, suggesting that the masking is not caused by such a compression. Smallman and Harris (1996) suggested that early visual nonlinearities may be expansive rather than compressive. In contrast, Scott-Samuel and Georgeson (1995) provided evidence for an early compressive nonlinearity, introducing a distortion product whose magnitude is influenced by the temporal frequency of the contrast envelope.

If spatial contrast variations, such as contrast beats in plaid stimuli, were detected on the basis of distortion products introduced by an early nonlinearity, cortical simple cells would be expected to respond to beats of the correct orientation and spatial frequency, as well as to luminance gratings. Albrecht and DeValois (1981) found that simple cells did *not* respond to a contrast beat at the orientation and spatial frequency at which the cell would be optimally stimulated by a luminance grating. Derrington (1990) suggested that nonlinear responses of X-cells in cat LGN may be nulled by summation of on- and off-centered neurons, thus preventing cortical cells from responding to distortion products. Zhou and Baker (1993), however,

found a subpopulation of neurons in cat areas 17 and 18 which did respond to both luminance gratings and contrast beats. These neurons showed orientation and spatial frequency tuning for both types of stimulus. However, the optimal beat spatial frequency was always lower than the optimal luminance grating spatial frequency. Additionally, responses to beats showed a marked dependence on the spatial frequency of the carrier grating. These results are not consistent with a simple early distortion product hypothesis. Rather, Zhou and Baker argued that the cell responses recorded resulted from nonlinearities occurring after orientation- and frequency-specific filtering, constituting a distinct pathway whose role may be to analyse spatial and temporal contrast variations.

Derrington and Badcock (1985) provided further evidence that contrast beat detection does not rely on distortion products. In a plaid formed from the sum of two sinusoidal gratings, with similar orientations and spatial frequencies, the contrast of any distortion product will depend on the contrast of the two components. If contrast beats were processed on the basis of this distortion product, then their detectability should depend on the product of the contrasts of the component gratings. Increasing the contrast of one component should reduce the contrast of the other component required to detect the beat. Conversely, Derrington and Badcock found that increasing the contrast of one of the components *increased* the contrast in the other component required. These results would be predicted if beats are detected on the basis of spatial variations in contrast; such variations would become harder to detect at higher contrasts (Legge, 1981).

The above results suggest that early nonlinearities do not contribute significantly to the perception of image contrast envelopes. However, distortion products do produce significant masking and adaptation effects. It is possible further that they may play a functional role in stereopsis. If this were the case, then the perception of depth from contrast envelopes need not rely on a separate nonlinear processing

channel.

5 Overview

The studies discussed in the previous sections demonstrate that stereopsis does not represent the action of a single, uniform mechanism. Rather, the visual system appears to be sensitive to properties of disparity fields that are directly related to important properties of object and surface shape. These include discontinuities in, and gradients of, horizontal disparity, orientation disparities and, more controversially, spatial frequency disparities, [1]. In addition, it has been proposed that stereopsis has access to independent linear and nonlinear channels. This suggestion is made of the grounds that luminance[2] defined binocular disparities is not necessary for the perception of stereoscopic depth.

The aim of this thesis is to explore how luminance and contrast envelope information are used in stereopsis, and how different types of disparity are combined in the representation of depth. Primarily, the thesis explores the relationship between luminance and contrast disparities. In doing so, it addresses the question of whether the two forms of disparity are processed by a single mechanism, or whether separate linear and nonlinear stereoscopic processing channels exist.

Previous studies of contrast envelope stereopsis have used Gabor and contrast modu-

[1] Although outside the scope of this thesis, it is also proposed that stereopsis may be divided into global and local operations (Julesz, 1971; Tyler, 1971,1975,1990). Global stereopsis is associated with stimuli such as random dot stereograms, in which the correspondence problem has to be solved in the face of ambiguous possible binocular matches. Conversely, local processing operates on features for which binocular matching is unambiguous.

[2] Again outside the scope of the current discussion, it should be noted that isoluminant chromatic stereograms can in some situations support depth perception (Tyler and Cavanagh, 1989).

lated random dot stimuli to manipulate luminance and contrast envelope disparities independently. These stimuli have suffered from the problem that, when luminance and contrast disparities are different, they provide different cues to depth. An important observation of Rogers and Graham (1983) was that, for slanted surfaces, the magnitude of orientation disparities may vary independently of the magnitude and direction of surface slant. For a given binocular transformation, the orientation disparity of an image contour will depend on its orientation. In this thesis, this property is used to manipulate luminance and contrast envelope disparities independently, without necessarily introducing any conflict between the two. This methodology is applied to assess the roles of the two sources of disparity information in stereopsis. Further, this manipulation allows for an examination of the roles of positional, orientation and frequency disparities from the two sources of information in the representation of surface slant.

In non-Fourier motion stimuli, a contrast modulation may be observed to move over its carrier (Derrington and Badcock, 1986; Chubb and Sperling, 1988; Fleet and Langley, 1994a). A question central to this thesis is how transparency may similarly be observed in stereopsis. Multiple surfaces, appearing in transparency, may be apparent in random dot stereograms (Akerstrom and Todd, 1988; Weinshall, 1990; Parker, Johnston, Mansfield and Young, 1991; Langley, Fleet and Hibbard, 1995). Here, the question of how transparency may be observed in stimuli containing both luminance and contrast disparities is addressed. Finally, the nature of the processing of contrast disparities is analysed.

2. Grating and plaid slant thresholds

1 Introduction

The experiments presented here address two questions concerning the encoding of stereoscopic slant. The first relates to the types of image disparity that are utilised in slant perception; the second, the types of monocular information which may support this disparity processing. Specifically, the experiments address the roles of orientation, positional and spatial frequency disparities in determining stereoscopic slant thresholds, and the extent to which slant may be inferred from disparities in image contrast envelopes.

Previous studies have demonstrated the importance of orientation disparities in stereoscopic slant (e.g. von der Heydt, Hänny and Dürsteller 1981; Rogers and Graham, 1983; Ninio 1985; Cagenello and Rogers, 1993). Rogers and Graham

(1983) showed that the anisotropy related to the axis of slant may be predicted if orientation disparities represent an important cue to slant. Cagenello and Rogers (1993) demonstrated that this anisotropy is dependent on the orientation of surface contours, which determine the orientation disparity associated with a given binocular transformation. Further, von der Heydt et al. (1981) showed that slant is clearly perceived from binocular images containing lines of random luminance with different orientations in left and right eyes, even if the lines themselves are uncorrelated. Finally, Ninio (1985) demonstrated that orientation disparities are important in the perception of smooth surfaces.

Orientation disparities cannot, however, explain the perception of slant in all situations. Slant may be perceived in stimuli consisting of grids of horizontal and vertical lines, with a horizontal gradient of disparity. This stimulus contains no orientation disparities. Slant perception must therefore make use of other disparity cues. Two possible cues are positional disparity gradients (Ryan and Gillam, 1993) and diffrequencies (Blakemore, 1970; Tyler and Sutter, 1979; Halpern, Patterson and Blake, 1987; Tyler, 1990; Rogers and Bradshaw, 1994; Halpern, Wilson and Blake, 1996). Diffrequencies are defined as interocular differences in spatial frequency. Typically, diffrequencies have been studied with relation to vertical, one dimensional signals (Blakemore, 1970; Tyler and Sutter, 1979; Halpern et al., 1987; Rogers and Bradshaw, 1994). Under these conditions, diffrequencies are associated with slant about a vertical axis (Blakemore, 1970).

For stimuli at orientations other than vertical, diffrequencies may be introduced by slant about other axes. Halpern et al. (1996) showed that slant was perceived in stimuli consisting of sinusoidal gratings oriented at 45°, with an interocular difference in frequency. Here, slant was seen about an axis orthogonal to the orientation of the grating. Additionally, for non-vertical stimuli, slant about a vertical axis will introduce orientation disparities as well as diffrequencies. Figure 2.1 shows

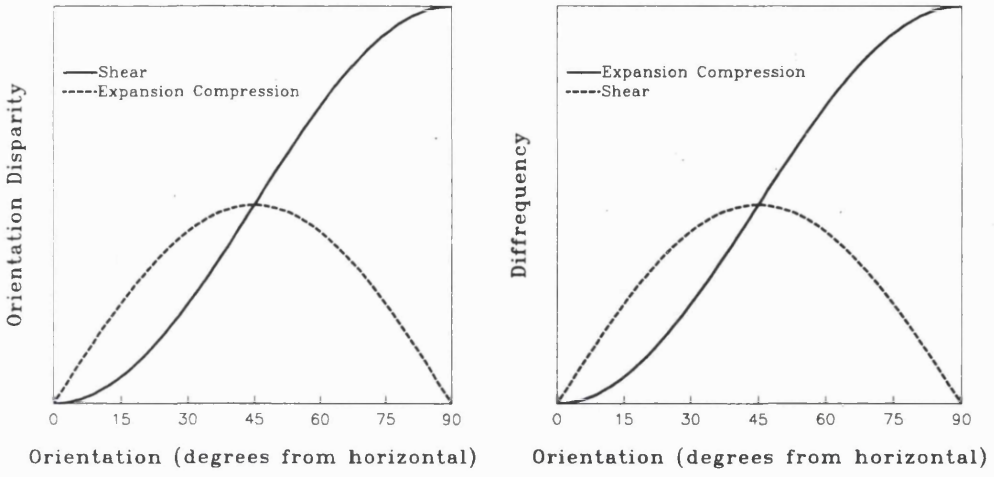


Figure 2.1: **(A)** *Orientation and* **(B)** *spatial frequency disparities produced for an image contour by a constant magnitude of shear or expansion-compression. Disparities depend upon the orientation of the contour. The shear transformation produces maximum orientation differences for vertical gratings, and maximum frequency differences for gratings oriented at* $\pm 45°$*; the reverse is true for an expansion-compression (see appendix A).*

how orientation and frequency disparities vary as a function of orientation. For slant about a horizontal axis, maximum orientation disparities occur for vertical image contours, and maximum diffrequencies for contours at 45°. For slant about a vertical axis, this pattern is reversed. Overall, maximum orientation disparities occur for vertical contours subjected to a vertical disparity gradient, while maximum frequency disparities occur for vertical contours subjected to a horizontal disparity gradient.

In general, orientation disparities will be associated with slant about a horizontal axis, while diffrequencies will be associated with slant about a vertical axis, and

hence horizontal gradients of disparity. Given the unreliability of horizontal gradients of disparity as a cue to slant, Rogers and Bradshaw (1994) argued that diffrequencies may not be used to encode slant. Alternatively, both orientation disparities and diffrequencies may play some role in the encoding of slant. The relatively poor perception of slant about a vertical axis might then arise if sensitivity to orientation disparities is greater than sensitivity to diffrequencies. This poor sensitivity to diffrequency may again be related to the unreliability of horizontal gradients of disparity as a cue to slant.

If diffrequencies do *not* play a role in stereoscopic slant, then slant in stimuli in which there are no orientation disparities must result from gradients of horizontal disparity. These may be represented directly, or implicitly in the disparities associated with individual points. The first question addressed by the experiments presented in this chapter is the extent to which disparities in orientation, spatial frequency and position contribute to the perception of slant.

The second question addressed relates to the potential role played by envelope disparities in stereoscopic slant. Other studies have demonstrated the influence of envelope disparities on the perception of stereoscopic depth. (Liu, Schor and Ramachandran, 1992; Sato and Nishida, 1993; 1994; Hess and Wilcox, 1994; Wilcox and Hess, 1993, 1994, 1995, 1996). These experiments have often employed Gabor stimuli (Liu et al., 1992; Hess and Wilcox, 1994; Wilcox and Hess, 1995), for which envelope and carrier disparities may be manipulated either independently, or simultaneously. When manipulated independently, envelope and carrier will provide conflicting disparity cues. If the carrier and envelope in these stimuli provided independent sources of information, processed by linear and nonlinear stereoscopic channels, respectively, then it might be expected that under these conditions transparency would be observed. In motion, transparency may be observed with related stimuli (e.g. Derrington and Badcock, 1985). When carrier and envelope carry the

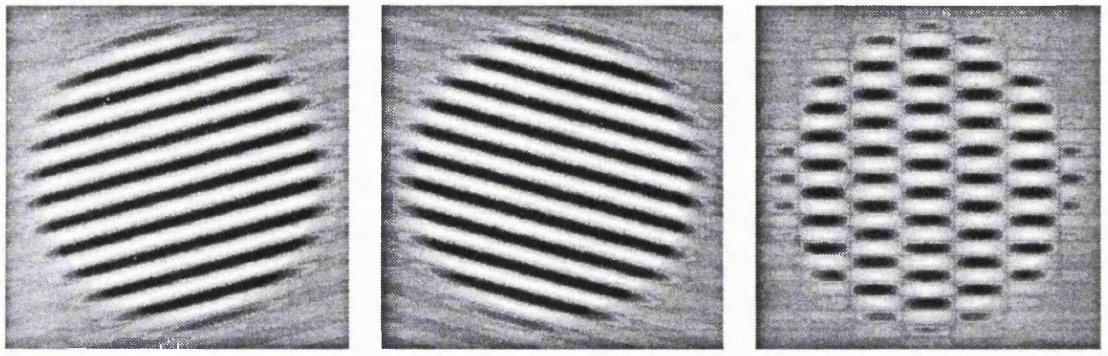
same disparity information, however, it is not clear that the effects of envelope disparity on stereoacuity, and on fusional range, necessarily demonstrate the existence of an independent nonlinear channel. Rather, depth may be encoded by a single disparity processing channel, the effectiveness of which is influenced by changes in both the carrier and envelope of an image. It is intended in this chapter to address the extent to which contrast envelope disparities contribute directly to the perception of slant. Stimuli were chosen so that carrier and envelope disparities could be manipulated independently, while still representing a single surface. This is possible for a slanted surface, since stereoscopic slant thresholds depend on the orientation of stimuli (Cagenello and Rogers, 1993). For stimuli in which carrier and contrast envelope have different orientations, a consideration of the carrier, or of the contrast envelope, will predict different stereoscopic slant thresholds. The stimuli used in these experiments were plaids, formed from the sum of two sinusoidal gratings. The two component gratings had equal magnitudes of spatial frequency and contrast, and were symmetrically oriented about horizontal. This stimulus may be described equivalently as the sum of two components, or as the product of a carrier and a contrast modulation. For relatively small angular separations, this stimulus will be consistently perceived as a horizontal carrier modulated by vertical beats; an example of the stimulus is shown in figure 2.2.

The spatial frequencies of the carrier and beats are given by the mean and half the difference of the component spatial frequencies:

$$\begin{aligned} I(x,y) &= \sin(fy\cos\theta + fx\sin\theta) + \sin(fy\cos\theta - fx\sin\theta) \\ &= \sin(fy\cos\theta)\sin(fx\sin\theta) \end{aligned} \tag{1}$$

where f and $\pm\theta$ refer to the spatial frequency and orientation of the component gratings, respectively.

The binocular transformations of horizontal shear and expansion compression, which



(A) (B) (C)

Figure 2.2: **(A)** *and* **(B)** *Examples of the sinusoidal gratings used.* **(C)** *An example of a plaid used. This plaid was produced by adding the two gratings in* **(A)** *and* **(B)**. *It appears as a horizontal carrier with a vertical contrast modulation.*

are related to surface slant, will introduce disparities to the carrier and contrast envelope, and to the individual component gratings. Orientation and spatial frequency disparities will be greater for the vertical contrast modulation than for the components, which are close to horizontal in orientation. If contrast envelope disparities are perceived directly, then it might be expected that slant in these stimuli would be perceived when component disparities were below threshold. Alternatively, slant thresholds may be determined by component disparities. In this case, slant thresholds would be expected to depend on the orientation of image components.

This chapter presents three experiments. The first examined the effects of contrast and spatial frequency on stereoscopic slant thresholds. The results of the first experiment acted as controls in the later experiments. The second experiment examined slant thresholds for surfaces defined by binocular pairs of individual sinusoidal gratings, as the absolute (cyclopean) orientation of the gratings was varied. This experiment was used to assess the relative importance of position, orientation and

spatial frequency disparity cues in stereoscopic slant. Finally, slant thresholds were measured for plaid patterns. For these stimuli, the contrast beat provides a source of disparity information in addition to that provided by the image Fourier components. By comparing results for the latter two experiments, it was possible to evaluate the extent to which plaid slant thresholds depend on disparities present in the individual plaid components and hence the contribution, if any, of any nonlinear channel to stereoscopic slant thresholds.

2 Methods

2.1 Stimulus generation and display

Stimuli were generated and stored in the RAM of a SUN SPARC 20 Workstation, with 32 Mbytes of RAM. Images were displayed on a single, colour monitor, with a refresh rate of 76 Hz, using an 8 bit display driver. In many of the experiments presented in this and the following chapters, it was important to ensure that nonlinearities in the display did not lead to artefacts in the experimental results. Linearity was ensured by taking luminance measurements from the monitor, and, by using gamma-correction, forming a suitable lookup table. The resulting display used 233 grey scales, and nonlinearities were estimated to be less than 0.1%. The mean luminance of the display was 37.8cdm^{-2}. Image pixels were on a square lattice, and had a width of 0.25mm.Stereoscopic images were presented using a modified Wheatstone stereoscope (figure 2.3), with parallel viewing geometry. The viewing distance was 44cm. At this distance, image pixels subtended a width of 2.0 minutes of visual angle. One possible source of error in this experimental setup may have arisen from differences in pixel geometry between the two halves of the screen. By comparing the sizes of the two monocular images, it was estimated that these differences would

introduce disparities of less than the size of one pixel. However, this could still result in a discriminable disparity, a possibility that should borne in mind.

Stimuli were windowed in software using a circular aperture with a diameter of 4.7 degrees of visual angle, the edges of which were softened with a Gaussian window with a standard deviation of 0.94 degrees. Sub-pixel accuracy was achieved in the stimuli using a standard bilinear interpolation procedure (Georgeson, Freeman and Scott-Samuel, 1996). This procedure allowed for a minimum disparity of 0.5 second of arc. For plaid stimuli, individual components were transformed prior to summation. In subsequent chapters, other stimuli were used which involved both the summation and mulitplication of signals. In all cases, individual signals were transformed prior to combination. All experiments were carried out in a dark room, with the experimental monitor providing the only source of illumination.

2.2 Subjects

Three subjects were used. One was naïve to the purpose of the experiment. All subjects had normal or corrected to normal vision. Unless specified otherwise, these subjects were also used in the experiments presented in the following chapters.

2.3 Procedure

Slant perception has been demonstrated to be affected by cyclovergence (e.g. Rogers, 1992; Swash, Rogers, Bradshaw and Cagenello, 1995). It was important therefore to ensure that the effects of such eye movements were minimised. Initially, an identical Gaussian noise pattern was presented to each eye for 1 s. This presentation time was found to be sufficient for subjects to obtain fusion of the stimulus. As well as minimising the effects of eye movements, the noise stimulus, which appeared as

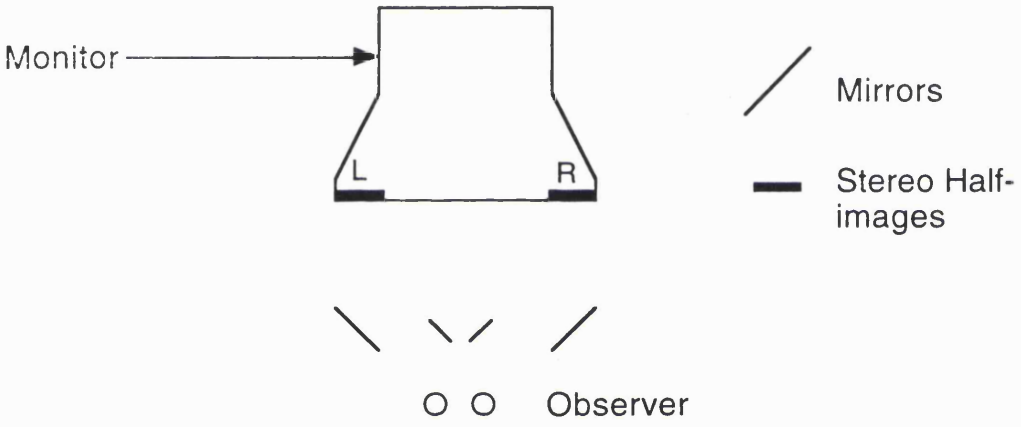


Figure 2.3: *Diagram of the experimental apparatus used.*

a frontoparallel surface, also served as a reference plane for judgements about the slant of the test surfaces, and helped to prevent the buildup of aftereffects. After presentation of the noise stimulus, the screen was blanked for 500ms, after which time the test stimulus was presented, and remained visible until the subject made a response. Subjects were permitted to examine the stimuli, since presentation time is critical to perception of stereoscopic slant (Gillam, Chambers and Russo, 1988). Subjects were however requested to respond as quickly as they could, to minimise cyclovergence during the time course of the trial. Typically, responses were made within 2 seconds. Cyclovergent eye movements are not elicited by horizontal shear or expansion-compression transformations (Rogers, 1992; Howard and Kaneko, 1994). Howard and Kaneko found that, for cyclorotated stimuli subtending a visual angle of 30 degrees, cyclovergent responses to null the cyclorotation took around 10 seconds. They found that slant was not perceived in these stimuli. Slant was however perceived for smaller stimuli, subtending 10 degrees of visual angle, with no constraints on viewing time. The stimuli used in the current study subtended a diameter of less than 10 degrees. Thus, while cyclovergence was not checked for explicilty, for example using the nonius method (Howard and Rogers, 1995), it was concluded that cyclovergent eye movements were unlikely to significantly affect the results.

Test stimuli were single sinusoidal gratings or plaids. Subjects were asked to classify each stimulus as either a "ground plane", with the bottom appearing closer in depth, or a "sky plane" (top closer), for the shear and rotation transformations, or as a "left wall" (left closer) or "right wall" (right closer) for the expansion-compression stimuli. The magnitude of slant was varied between trials using the APE adaptive probit analysis algorithm (Watt and Andrews, 1981). Slant discrimination thresholds were defined as the standard deviation of the psychometric function, as measured by the APE algorithm. Each psychometric function is based on 64 individual trials; data points are defined as the mean of 3 independently measured functions.

3 Experiments

3.1 Contrast and spatial frequency effects on stereoscopic slant thresholds

The first experiment was undertaken to test for effects of spatial frequency and contrast on stereoscopic slant thresholds. Both parameters are known to be important factors in depth discrimination tasks (Schor and Wood, 1983; Halpern and Blake, 1988; Legge and Gu, 1987; Hess and Wilcox, 1994; Kontsevich and Tyler, 1994), but have not been studied with direct relationship to stereoscopic slant.

This experiment was intended as a control for later experiments. As such, the effects of contrast and spatial frequency were not examined for all conditions to be used. Rather, a subset of the stimuli was chosen to ascertain whether slant thresholds appeared to be affected by contrast and spatial frequency, and whether these effects might account for the results of the subsequent experiments. Slant thresholds were measured for vertical sinusoidal gratings, and for plaids whose components were oriented at $\pm 25°$, as the frequency of the gratings, or plaid components, was varied. The spatial frequency of the components was varied to produce contrast modulations (beats) with spatial frequencies between 0.5 cycles/degree and 2.0 cycles/degree. Thresholds were measured for vertical grating stimuli with the same spatial frequency as the plaid beats. Thresholds were measured for both shear and expansion-compression transformations. Slant thresholds were also measured over a wide range of suprathreshold contrasts, for both grating and plaid stimuli. A shear transformation was used, and gratings and plaid components had a spatial frequency of 3.2 cycles/degree.

3.1.1 Results

The results of manipulating spatial frequency, for all subjects and conditions, are shown in figure 2.4. Results are presented in terms of disparity gradients, which describe the spatial change in disparity in the stimuli. These gradients are measured in the units of degrees/degree[1]. Although there was some variability between subjects and conditions, no consistent effects were evident for the conditions measured. For subject KL, thresholds were lower for plaids than for gratings; this was not evident for the other two subjects. The second and third experiments in this chapter explored thresholds for gratings and plaids, respectively. This between subject difference will be discussed in relation to the results of these experiments.

Results from the manipulation of grating and plaid contrast[2] are shown in figure 2.5. The data are plotted on log-log axes. A linear regression on the data gave a mean slope of -0.47 for gratings, and -0.78 for plaids (see figure legend for details). Thus, both stimuli exhibited a similar relationship between contrast and slant thresholds, but the slope for plaids was considerably higher.

3.1.2 Discussion

For both grating and plaid stimuli, thresholds were found to be independent of spatial frequency. Although Schor and Wood (1983) and Kontsevich and Tyler (1994) both found disparity discriminability to be linearly related to spatial frequency for

[1]The geometrically predicted slant of a surface, about a vertical axis, for a horizontal disparity gradient, is given by $\tan^{-1}\left(\frac{g-1}{g+1}\frac{2d}{i}\right)$. Here, g gives the disparity gradient, d the viewing distance, and i the interocular distance. Similarly, a vertical gradient of disparity predicts a surface with a horizontal gradient, and a slant of $\tan^{-1}\left(g\frac{d}{i}\right)$.

[2]Results are plotted in terms of the Michelson contrast of the stimuli, defined as $\frac{L_{max}-L_{min}}{L_{max}+L_{min}}$, where L_{max} and L_{min} are the maximum and minimum luminances of the stimuli, respectively.

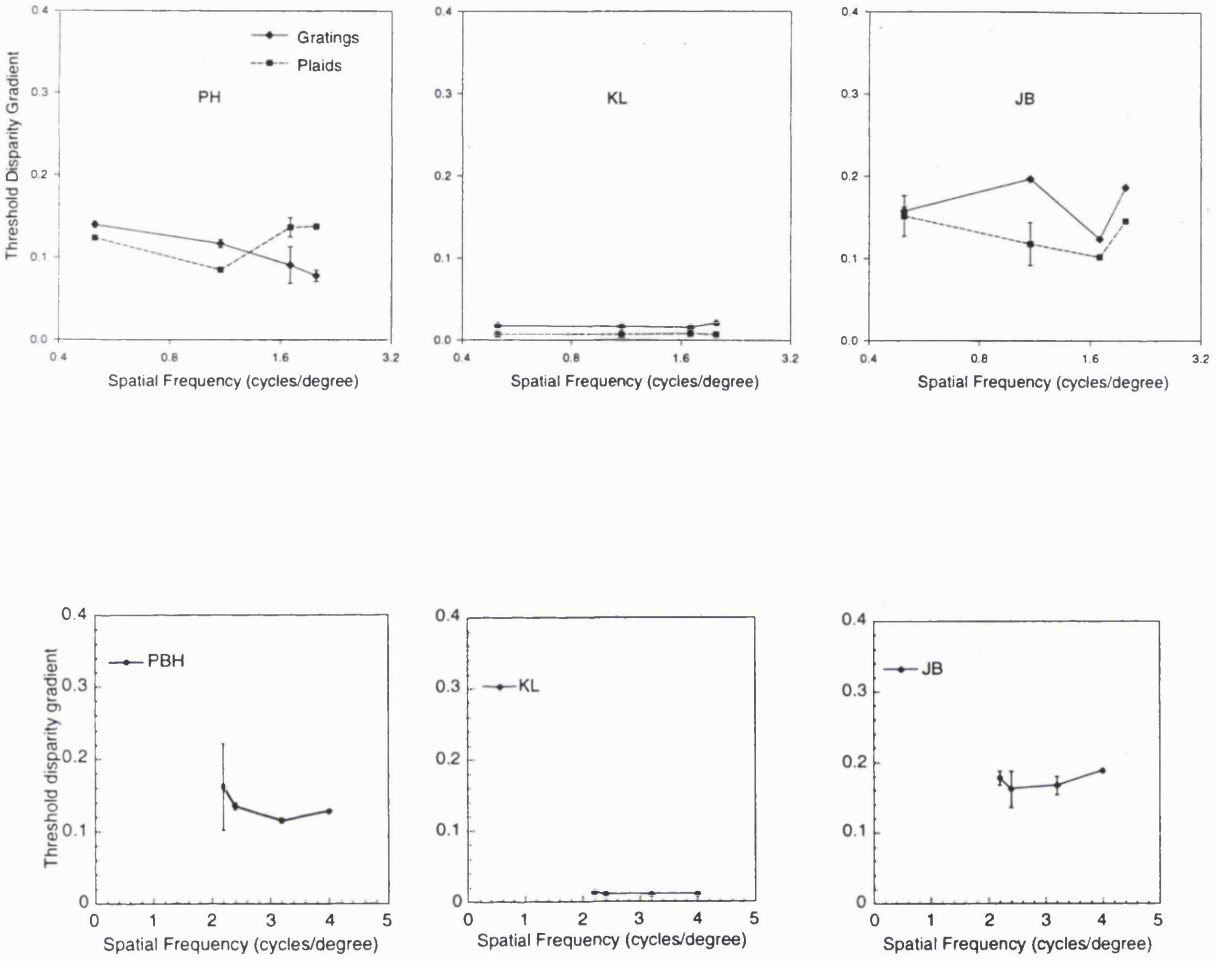


Figure 2.4: Plaid and grating slant (top three graphs) and inclination (bottom three graphs) thresholds plotted against spatial frequency. Only slant thresholds were measured for the plaid stimuli. Error bars in this and all other graphs represent 1 standard error of the mean (note that, for the sample sizes used here, this is equal to 0.577 times the standard deviation).

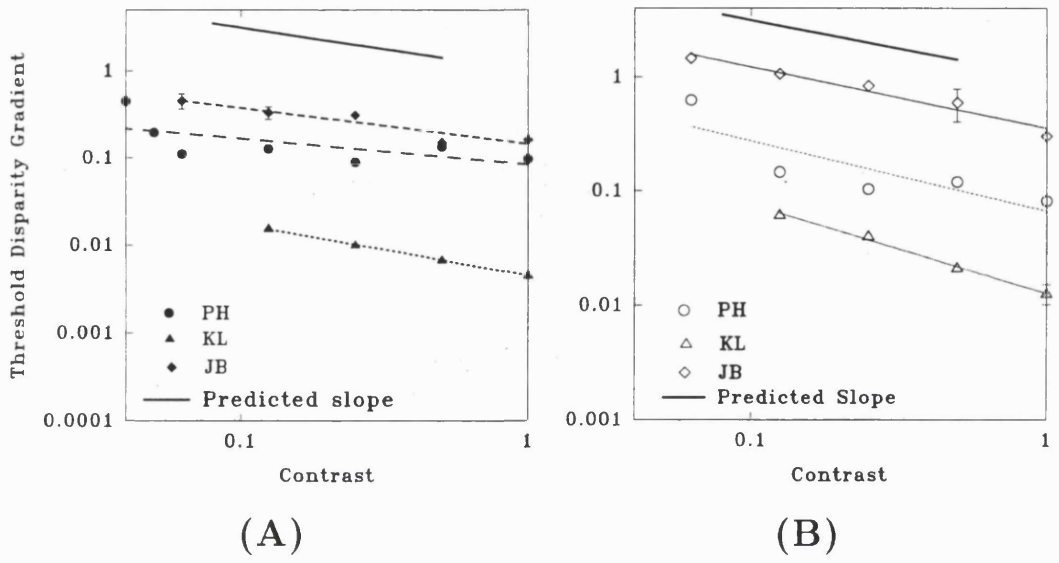


Figure 2.5: *Horizontal axis slant thresholds plotted against contrast.* **(A)** *Gratings* **(B)** *Plaids. The thick lines represent the slope predicted by the expected square-root relationship.*

frequencies below 2.5 cycles/degree, Hess and Wilcox (1994) and Wilcox and Hess (1995) showed these effects to hold only for relatively broadband stimuli. For narrowband stimuli, stereoacuity did not depend on spatial frequency. Since the stimuli used here were narrowband, the lack of dependence of slant discrimination thresholds on spatial frequency is in agreement with similar studies that examined depth discrimination.

Legge and Gu (1987), Hess and Wilcox (1994) and Kontsevich and Tyler (1994) found the discriminability of disparity to be inversely related to the square root of contrast. When the results are plotted on log-log axes, this relationship predicts a line with a slope of -0.5. This result was found here for slanted gratings, but not for plaids. However, if slant thresholds depend on the perceived contrast of the plaids, then this difference may be accounted for. Georgeson and Shackleton (1994) found

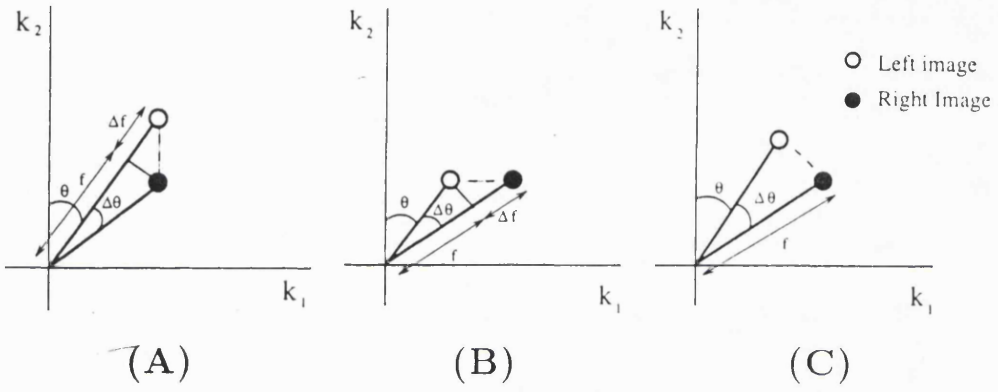


Figure 2.6: *The three stereoscopic transformations used.* (**A**) *Shear* (**B**) *Expansion-compression* (**C**) *Rotation. Each transformation is expressed in the Fourier frequency domain.*

the perceived contrast of plaids to be less than the perceived contrast of a single grating stimulus with an identical Michelson contrast. As the effects of contrast on stereoacuity are greatest at low contrast levels, a reduction in perceived plaid contrast would lead to results showing a greater slope than that predicted from its Michelson contrast. The data were recalibrated for the plaid stimuli, based upon the data of Georgeson and Shackleton, to plot slant thresholds against the predicted perceived contrast of the stimuli. This led to a revised slope of -0.64, which is closer to the expected value of -0.5. This suggests that slant thresholds of plaids were affected by their *perceived* contrast.

3.2 Slant thresholds for sinusoidal gratings

In the second experiment, slant thresholds were measured for surfaces defined by sinusoidal gratings. The orientation of the gratings was manipulated between blocks of trials. The experimental procedure was identical to that used in the first experiment. Grating orientation, relative to horizontal, ranged between 10° and 90°. Sub-

jects ran three blocks of trials, using the three binocular transformations of shear, expansion-compression, and rotation, as shown in figure 2.6. Grating contrast was fixed at 99.8%, and the cyclopean spatial frequency was 3.2 cycles/degree.

3.2.1 Results

Results are presented in figure 2.7. It was not possible to measure thresholds for grating orientations below 10° for subjects PH and KL, or below 15° for subject JB. For orientations below these limits subjects were unable to perceive slant reliably for all ranges of slant tested. For gratings oriented further from horizontal, slant was reported reliably for all three binocular transformations. Similarly, Howard and Kaneko (1994) found slant to be perceived from cyclorotated stimuli, subtending 10 degrees, both with and without a zero disparity reference. Thresholds showed a marked dependence on orientation. For the shear and rotation transformations, slant thresholds increased for orientations close to horizontal. Results were similar for these two conditions, although thresholds were in general slightly lower for shear than for rotation. The similarity between these results may reflect the relative unimportane of the interocular spatial frequency differences that were introduced by the shear transformation, but not by the rotation.

For the expansion-compression, thresholds were smallest for oblique gratings, and rose asymmetrically for larger and smaller angles.

The binocular transformations used introduced either gradients of disparity or, for the rotation, a disparity field which may be approximated as a disparity gradient. These disparity gradients introduced both orientation and spatial frequency disparities. Curves were fitted to the data under the assumption that slant thresholds were determined by the magnitude of the gradient of disparity. Morgan and Castet (1995) demonstrated that, for one dimensional stimuli, stereoacuity is determined not by

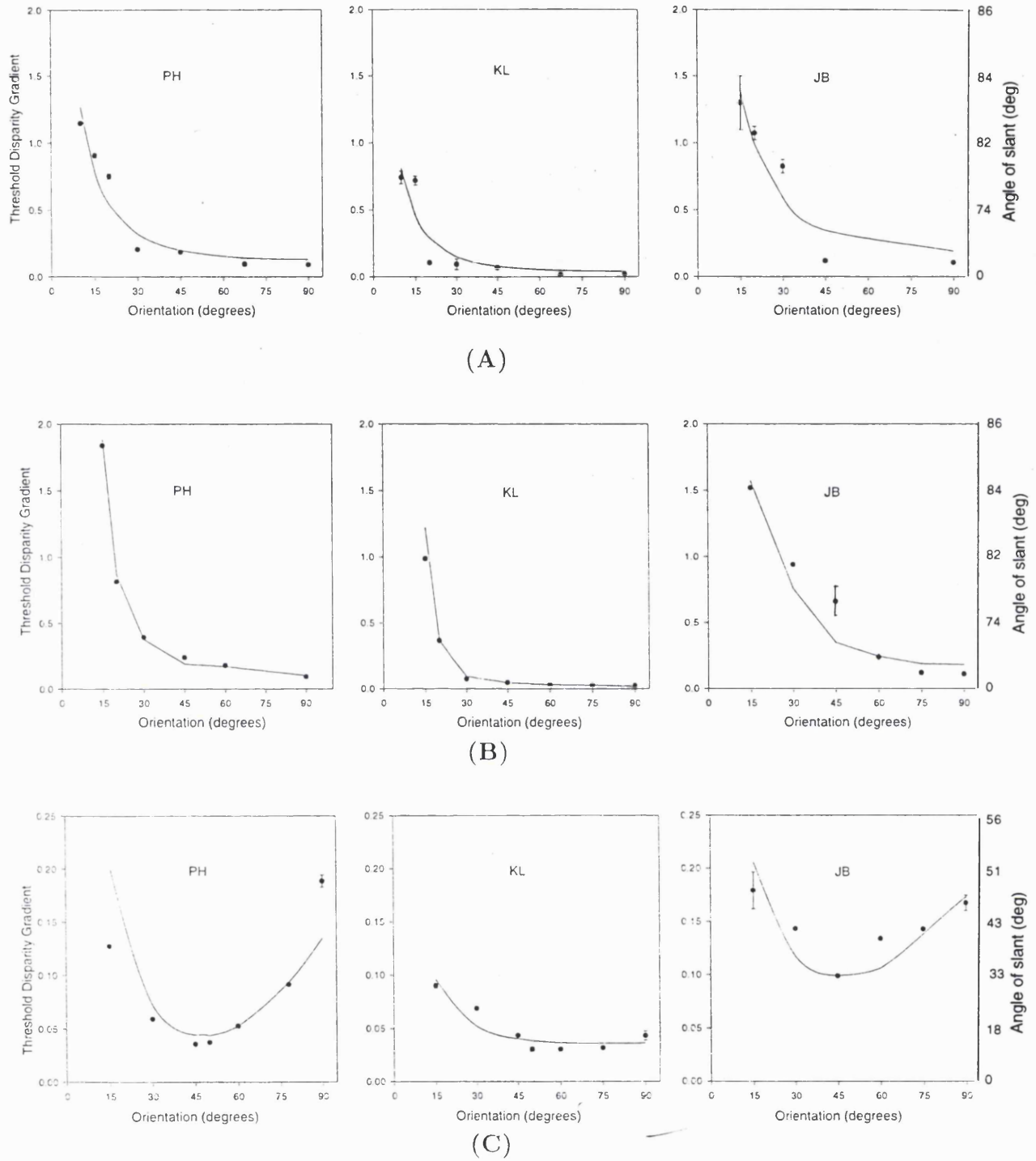


*Figure 2.7 Slant thresholds for grating stimuli plotted as a function of orientation. (**A**) results for shear, (**B**) results for rotation, (**C**) results for expansion-compression. Lines represent the best fitting curve, which was produced by a model involving orientation disparities and (for shear and expansion-compression) spatial frequency disparities*

3.2.2 Discussion

For the shear condition, fitted curves were extrapolated for orientations close to horizontal. It was found that, for orientations where thresholds were not obtainable, predicted threshold disparity gradients were 1.1(KL), 1.8(PH) and 2.1(JB). Burt and Julesz (1980) proposed a disparity gradient of around 1 as an upper limit on the perception of slant. Subjects would not therefor be expected to perform the task required reliably under these circumstances.

The curve fits shown in figure (2.7) demonstrate that slant thresholds may be predicted by assuming that both orientation disparities and spatial frequency disparities are used as cues to slant. This model could account for slant perception for all stimuli, including those for which no orientation disparities existed.

By representing orientation and frequency differences in the Fourier frequency domain, their magnitudes may be directly compared (figure 2.8). For a horizontally periodic stimulus, for example, a spatial frequency change will alter its horizontal frequency, whereas an orientation change will alter its vertical frequency. This representation assumes that sensitivity to differences in spatial frequency is determined by the proportional difference in frequency, such that a greater difference is required as frequency is increased (Bowne, 1990). It also assumes that sensitivity to orientation differences is relatively independent of spatial frequency; Burr and Wijesundra (1991) found that orientation discrimination for high contrast gratings is unaffected by spatial frequency for frequencies above 0.2 cycles/degree. On the basis of the curve fits, the relative sensitivity to orientation disparities and diffrequencies was estimated. Sensitivity to orientation disparities was found to be 1.75 times that of diffrequencies for subject JB, 1.72 for subject KL, and 1.73 for subject PH. On the whole, subjects were more sensitive to binocular differences in orientation than spatial frequency. A similar trend has been found in other ar-

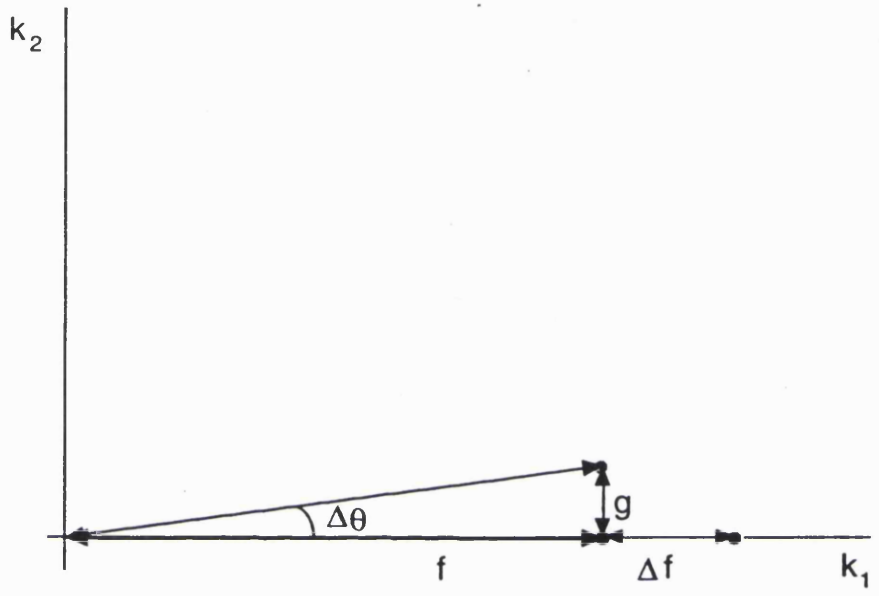


Figure 2.8: *A difference in orientation may be directly compared to a difference in spatial frequency in the Fourier frequency domain. This figure represents a vertical grating, with a spatial frequency of f . Its frequency may be altered by an amount Δf. Alternatively, altering its vertical frequency by an amount g will change the grating's orientation. For the grating with frequency f, if Δf and g represent the smallest discernable changes in spatial frequency and orientation, respectively, then the ratio of sensitivity to orientation and frequency is given by $g : \Delta f$.*

eas of research. For two dimensional spatial discrimination tasks, just noticeable differences in orientation of between 0.5° and 1.14° have been reported (Burbeck and Regan, 1983; Bradley and Skottun, 1984; Heeley and Timney, 1988; Bowne, 1990; Heeley and Buchannon-Smith, 1994), while discrimination on the basis of spatial frequency requires a difference in frequency of between 2% and 5% (Thomas, 1983; Hirsch and Hylton, 1982; Bradley and Skottun, 1984; Mayer and Kim, 1986; Burbeck and Regan, 1983; Bowne, 1990). Judgements of orientation differences are thus approximately 2-3 times as sensitive as judgements of frequency differences. Electrophysiological estimates of spatial frequency bandwidths of cortical simple cells are generally around 6 times greater than corresponding estimates of orientation bandwidths (DeValois, Yund and Hepler, 1982; Blakemore and Nachmias 1971; Movshon and Blakemore, 1973; DeValois, Albrecht and Thorell, 1982; Foster, Gaska and Pollen, 1983; Movshon, Thompson and Tolhurst, 1978). This ratio has been shown to be relatively constant regardless of the actual bandwidths of particular cells (Movshon et al., 1978; DeValois et al, 1982).

It is proposed that the variation in slant thresholds as a result of grating orientation, and of the transformation type, may be due to unequal sensitivity to binocular differences in orientation and spatial frequency. One possible explanation for this would be that binocular cells involved in stereoscopic processing are able to encode orientation differences with greater accuracy than spatial frequency differences. This may then account for the anisotropy in slant thresholds reported (Rogers and Graham, 1983; Cagenello and Rogers, 1993).

There exists an aperture problem in stereopsis similar to that in motion, such that the direction of slant of a single grating stimulus is ambiguous. This might be expected to affect the results presented here. However, despite this inherent ambiguity, Halpern, Wilson and Blake (1996) found that a single grating would appear to slant about an axis orthogonal to its orienation. This result was also found here,

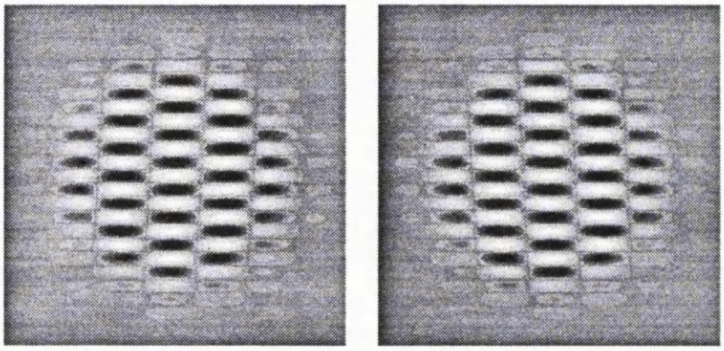


Figure 2.9: *An example of a plaid stimulus, which has been sheared about a horizontal axis between left and right eye views. Cross eyed fusion reveals the plaid to slant about a horizontal axis.*

so subjects were able to respond consistently to the two directions of slant.

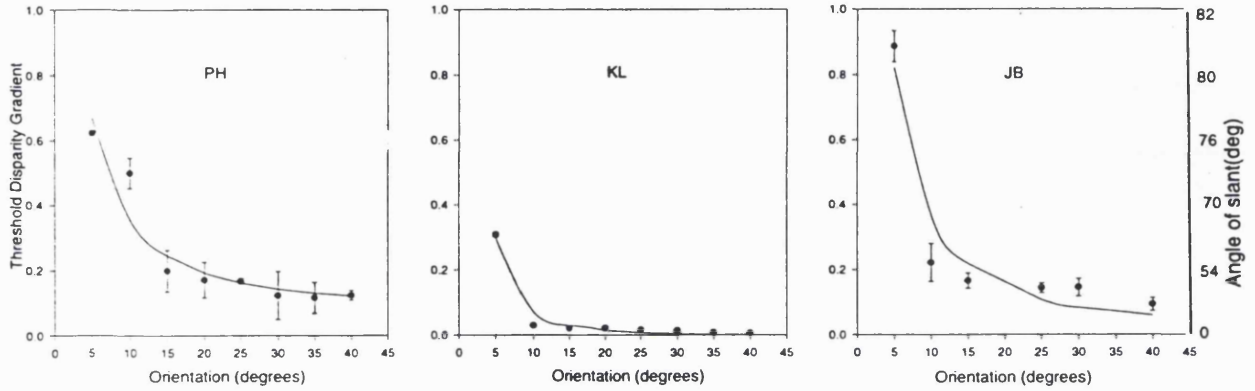
An unlimited presentation time was used for all stimuli, which were visible until subjects made a response. This was done since latencies to perceive slant can depend critically on the axis of slant (Gillam et al., 1988). However, as discussed earlier, it is possible that this methodology may have affected the results obtained. First, it is possible that perceived slant was affected by cyclovergence, which could have occured over the time for which each stimulus was presented. Second, latencies to report slant depend on the orientation of surfaces, allowing the possibility that thresholds would appear artificially high if subjects did not allow sufficient time before responding. However, it was decided that an unlimited response time provided the best solution to the potential problems of cylovergence, and of anisotropic processing times for slants about difference axes.

3.3 Slant thresholds for plaid stimuli

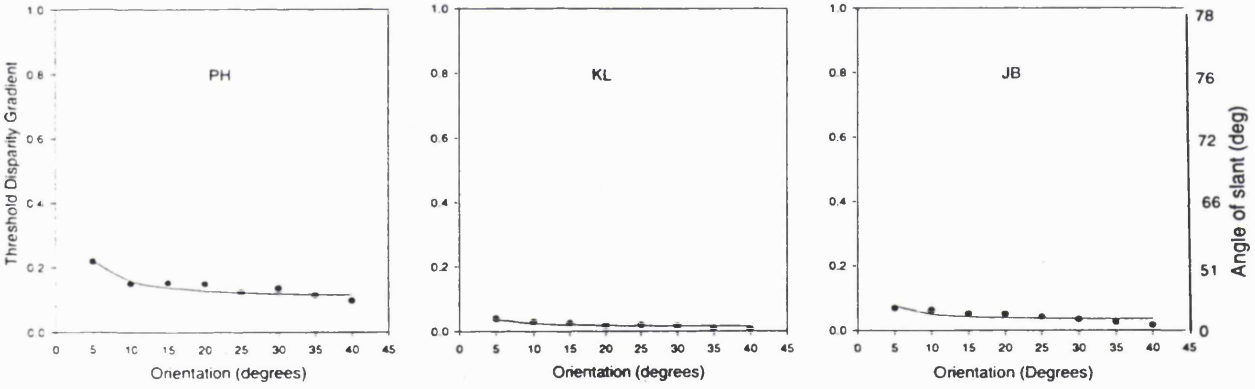
The final experiment measured slant thresholds for plaid stimuli. Component orientation was manipulated to determine whether slant discrimination for plaids could be predicted from the disparities present in individual component gratings. This would predict that plaid slant discrimination thresholds would show a similar dependence on orientation as do thresholds for gratings. The range of orientations studied was limited to between $\pm 5°$ and $\pm 40°$. Small component angular separations lead to beats with low spatial frequencies, thus limiting the number of cycles it is possible to present given the size of the stimulus window. For the plaid stimuli used, at least two full cycles of the contrast modulation were always visible. For all plaid stimuli, the carrier had a spatial frequency of 3.2 cycles/degree (the frequency of the plaid's components); the beat spatial frequencies ranged between 0.279 cycles/degree (for plaids with components at $\pm 5°$), and 2.1cycles/degree (for plaids with components at $\pm 45°$). The plaids were transformed in a manner consistent with surfaces slanting about both horizontal and vertical axes. An example of the stimuli is given in figure 2.9.

If plaid slant thresholds depend on disparities in their individual components, it should be possible to model plaid thresholds on the basis of the data from the second experiment. Curves were fitted to the data using the best fitting model from the previous experiment, the model based upon orientation and spatial frequency disparities. Plaid thresholds were not expected to have the same magnitude as grating thresholds, since two components rather than one were present. Therefore, in modelling the results, thresholds were allowed to be different to the single grating condition, while the relative sensitivity to orientation and spatial frequency disparities found in the models fit in experiment 2 was maintained. The model was fitted to the data for both shear and expansion-compression conditions simultaneously.

3.3.1 Results



(A)



(B)

Figure 2.10: *Slant thresholds for plaid stimuli as a function of component orientation.* **(A)** *Shear* **(B)** *Expansion-Compression. Curves represent predictions of the orientation and spatial frequency disparity model.*

Results for all three subjects are shown in figure 2.10; the results are again further tabulated in appendix A. In each case, slant thresholds were smaller than for a single grating with the orientation of the plaid components. Thresholds did however increase as the orientation of components approached horizontal. This was most pronounced for surfaces with a horizontal axis than for those with a vertical axis; in the latter case, the effects of component orientation were small.

3.3.2 Discussion

The results show that plaid slant thresholds were affected by the orientation of their component gratings. This suggests that plaid slant thresholds were determined by component disparities. However, it should be noted that the effects of orientation were appreciable only with components at $\pm 5°$; at this orientation, thresholds could not be measured for single gratings. Thus, while thresholds for grating and plaid stimuli demonstrated similar trends, a direct comparison of the two results is not possible. Manipulating component orientation also affected other aspects of the stimuli. From equation (1), it can be seen that as the orientation of the plaid components approaches horizontal, the spatial frequency of the beats is reduced. Additionally, decreasing the separation of the plaid components *increases* the perceived contrast of the stimulus (Georgeson and Shackleton, 1994). Experiment 1 showed spatial frequency to have no effect on slant thresholds. Increases in perceived contrast of stimuli *decreased* slant thresholds, suggesting that slant thresholds may decrease with decreasing component separation. It may be concluded that the results found here were not a consequence of the change in beat frequency, or the perceived contrast of the plaids.

Thresholds were lower for plaids than for gratings. This would be expected given that there are two components rather than one present. If disparities in the plaid

beats were made available, slant thresholds may be expected to be unaffected by component separation. This would be expected because the beats would always be vertical, and slant thresholds are determined primarily by the orientation of surface contours (as is evident in the results of the second experiment, and the results presented by Cagenello and Rogers (1993)). The results presented here provide little support for the notion that disparities in contrast beats contributed to the perception of slant.

It is interesting to note that thresholds for the expansion-compression condition were *lower* than those for the shear condition. This is a reversal of the usual anisotropy in stereoscopic slant (Rogers and Graham, 1983). This reversal is predicted if one again assumes that orientation disparities play a major role in determining slant thresholds. Orientation disparities are greater for expansion-compression than for shear for orientations between horizontal and 45°. It is also consistent with results reported by Cagenello and Rogers (1993), who found the anisotropy to be lost if images contained components oriented at ±45°.

4 Conclusions

Slant thresholds for both grating and plaid stimuli can be accounted for by a model in which an analysis of the binocular affine transformation makes use of both orientation disparities and spatial frequency and/or positional disparities. The data suggest that, however they may be encoded, sensitivity to spatial frequency disparities is considerably lower than sensitivity to orientation disparities.

Given the variabilty in the magnitude of thresholds reported between subejcts, it would have been illuminating to have measured thresholds for a larger subject group. However, while caution must be born in mind when drawing conclusions from a

small number of subjects, it is interesting to note that, despite these differences in magnitudes found, consitent trends were found in the data for all subjects.

These data do not support the notion that slant is detected from the contrast envelope in plaids. This is perhaps surprising given other results that suggest a role for a nonlinear channel in stereopsis (Hess and Wilcox, 1994; Sato and Nishida, 1993; Fleet and Langley, 1994b; Hibbard, Langley and Fleet, 1994; Lin and Wilson, 1995). The results are however consistent with results for analogous studies in motion. Welch (1989) found plaid motion discriminability to be determined by the velocities of the individual sinusoidal components rather than the plaid velocity. These results together with those of the current study suggest that for plaid stimuli in which component and envelope together signal a single coherent structure, motion and stereo processing may be based purely on an analysis of changes in the component gratings.

3. Perceived slant in grating and plaid stimuli

1 Introduction

In the previous chapter, it was reported that stereoscopic slant thresholds for grating stimuli were dependent on the orientation of gratings, for both horizontal and vertical axes of slant. Plaid slant thresholds were found to be similarly related to the orientations of the plaid's components. These results were interpreted in terms of a model in which orientation disparities play a critical role in the perception of slant at threshold. This model is consistent with existing literature on slant thresholds (e.g. Rogers and Graham, 1983; Mitchison and McKee, 1990; Cagenello and Rogers, 1993). This chapter explores the perception of slant in gratings and plaids with suprathreshold slant.

Suprathreshold slant perception exhibits the same anisotropy with respect to the axis of slant as do slant thresholds. Latencies to perceive slant are greater for

slant about a vertical axis than for slant about a horizontal axis. (Wallach and Bacon, 1976; Gillam Flagg and Finlay, 1984, Gillam, Chambers and Russo, 1988) Additionally, less slant is typically perceived for vertical than for horizontal axes of slant (Gillam, Flagg and Finlay, 1984; Gillam, Chambers and Russo, 1988; Mitchison and McKee, 1990). Again, this anisotropy is influenced by the orientation of surface contours. Gillam et al. (1984) studied slant perception for surfaces slanting about both horizontal and vertical axes, for a range of surface types. Surfaces were defined by vertical and horizontal lines, and grids formed from diagonal or horizontal and vertical lines. The poorest perception of slant was found for grids of horizontal and vertical lines, slanting about a vertical axis; this stimulus contains no orientation disparities. Gillam et al. (1988) also found anisotropic latencies for slanted random dot stereograms. These results are consistent with slant thresholds measured for similar stimuli by Cagenello and Rogers (1993), and support the notion that both threshold and suprathreshold slant perception are based primarily on an analysis of orientation disparities.

Mitchison and McKee (1990) and Gillam and Ryan (1992) both reported that the anisotropy in slant perception persisted when the orientations of surface elements were manipulated so as to equalise orientation disparities. In stimuli containing lines oriented at $\pm 45°$, slant about horizontal and vertical axes produce equal orientation disparities. However, more slant was perceived in grids of diagonal lines slanting about a horizontal axis than a vertical axis. These results suggest that factors other than orientation disparity contribute to the perception of slant.

Gillam (1968) suggested that conflicting perspective cues may affect stereoscopic slant judgements. It is typical in investigating stereoscopic slant to manipulate disparity cues only, to ensure that perceived slant can only result from stereoscopic information. This has the effect of introducing a conflict between stereoscopic and perspective depth cues. Consider for example a surface defined by horizontal and

vertical lines, slanting about a vertical axis. Horizontal lines will tend to converge as the surface draws away from the observer (linear perspective), while vertical lines will become closer together (texture compression) (Gillam, 1968). If linear perspective and texture gradient are consistent with a frontoparallel surface, perspective cues may conflict with stereoscopic cues indicating a slanting surface. Ryan and Gillam (1994) found that linear perspective in horizontal lines provided the strongest perspective cue, and led to the greatest conflict with stereoscopic cues. Adding horizontal lines to vertical lines presented with a horizontal gradient of disparity will decrease the perceived slant of the figure about a vertical axis (Gillam, 1968). The addition of horizontal lines introduces a conflicting linear perspective cue, without contributing additional disparity information. Similar results have been reported at threshold (McKee, 1983). Mitchison and McKee (1990) found that adding a square border, with no disparity or perspective cues, to stereoscopically slanting stimuli diminished the perceived slant. This had more effect when the surface was slanting about a vertical than a horizontal axis. This is consistent with Gillam's (1968) notion that linear perspective in horizontal lines provides the strongest perspective cue to slant.

Gillam and Ryan (1992) analysed the contributions of orientation disparities and conflicting perspective cues to the perception of stereoscopic slant. They examined slant perception in stimuli defined by horizontal, vertical and diagonal lines, and in grids formed from combinations of these. While most of their results could be explained in terms of interactions between orientation disparities and conflicting perspective cues, they suggested that other factors must also influence perceived slant.

The experiments presented here examined perceived slant in binocularly presented grating and plaid stimuli. These were motivated by three considerations. First, to establish the extent to which the results of the previous chapter apply to supra-

threshold slant. Second, to further explore the effect of the orientation of image contours on slant perception. Third, to examine whether non-Fourier stereoscopic channels may be evident for suprathreshold slant. Studies which have argued for a role for envelope disparities have reported stereoacuity for envelope disparity to be poorer than that for luminance defined disparity (Liu, Schor and Ramachandran, 1992; Hess and Wilcox, 1994, Lin and Wilson, 1995). Envelope disparities may have been undetectable in the threshold stimuli presented in the previous chapter, and may be expected to be evident for greater slants.

Typically, perceived slant in stereoscopic displays is measured by matching to the slant of a real three-dimensional stimulus. This methodology allows a comparison to be made between the geometrically predicted slant of the surface and its perceived slant. The experiments presented here concerned the relative effectiveness of alternative sources of disparity information. As such, the absolute slant perceived was not an issue. Rather, the experiments considered the perceived slant elicited by an equal binocular transformation applied to different stimuli. The experiments presented employed a stereoscopically defined probe stimulus, against which the slant of the test stimuli was compared. While this procedure did not allow the absolute magnitude of perceived slant to be assessed, it had the advantage that the results could be considered solely on the basis of the stereoscopically presented stimuli. Some results of the studies reviewed earlier cannot be accounted for on the basis of stereoscopic or perspective cues, and it may be that some aspect of the comparison task was responsible. By using stereoscopically defined test and probe stimuli, this problem was avoided.

2 Methods

2.1 Subjects

The two experimenters acted as subjects in this experiment. The two subjects had normal vision, and had acted as subjects in the experiments presented in the previous chapter.

2.2 Procedure

The procedure used was identical for all experiments. Each trial consisted of two intervals, in which a test and a probe stimulus were presented. The test stimulus was either a single vertical sinusoidal grating, or a plaid formed from two component gratings. Details are given in Sections 3 and 4 below. The probe stimulus was a plaid composed from three component sinusoidal gratings. These components had a spatial frequency of 2.5 cycles/degree. One of the probe plaid components was vertical, the other two were symmetrically oriented 60° either side of vertical. This stimulus is shown in figure 3.1. This probe stimulus was chosen so as to avoid the aperture problem in stereoscopic slant, related to that in two-dimensional motion (Wallach, 1935). The orientation of a planar surface has two degrees of freedom, which may be described as the slant and tilt of the surface (Stevens, 1983). Halpern, Wilson and Blake (1996) found that a single oblique sinusoidal grating may appear to slant about an axis orthogonal to its orientation. As is the case for two dimensional motion, the aperture problem may be avoided by considering stimuli containing contours at different orientations. A probe stimulus containing gratings at three different orientations was used so that both the slant and the tilt of the stimulus could be uniquely determined, while maintaining a stimulus that was as

similar as possible to the test stimuli. For the single vertical grating stimuli, shear and expansion produced the perception of inclination and slant, respectively, so a direct comparison between probe and grating stimuli was in all cases possible. In contrast to the experiments presented in chapter 2, all stimuli were viewed through a software-generated hard circular window, the edges of which were not softened with a Gaussian contrast envelope. The diameter of the window was 7.9 degrees.

The aim of the experiments was to determine the slant of the probe stimulus that appeared equal to that of the test stimulus. Both intervals of each trial had a duration of 700ms. Both the first and second interval were preceded by the presentation of a random noise stimulus, with zero disparity, for 1000ms. This stimulus served to maintain fusion between trials, and thus to minimise eye-movements and torsional misalignment, which may affect the perceived slant of stimuli (Rogers, 1992; Howard, Ohmi and Sun, 1993; Swash, Rogers, Bradshaw and Cagenello, 1995). Again, while a nonius fixation was not used, each stimulus was presented with a small, central fixation point to aid fusion. The subjects' task was to decide which interval contained the stimulus with the greater slant. The interval containing the test stimulus was varied randomly between trials, to remove the possibility that the results were influenced by the order of presentation of test and probe stimuli. For each block of trials, a single test stimulus was used. Each test stimulus had either a horizontal or a vertical gradient of disparity. The magnitude of the disparity gradient was fixed for a given block of trials; the sign of the disparity gradient was alternated between successive trials. This was done to inhibit depth after-effects (Kohler and Emery, 1947; Bergman and Gibson, 1959; Wenderoth, 1970). For each trial, the probe stimulus had a disparity gradient with the same direction and sign as the test stimulus. Its magnitude was controlled as an independent variable using the APE algorithm (Watt and Andrews, 1981), independently for each sign of disparity gradient. Each block consisted of 128 trials, 64 for each sign of disparity. The APE algorithm was used to fit a psychometric function to the data obtained, independently for each

sign of disparity gradient. This function was used to determine the magnitude of disparity gradient of the probe stimulus that was perceived to have greater slant than the test stimulus on 50% of occasions. This 50% point was taken to represent the slant at which the test and probe stimuli appeared equally slanted.

3 Experiments

3.1 Sinusoidal gratings

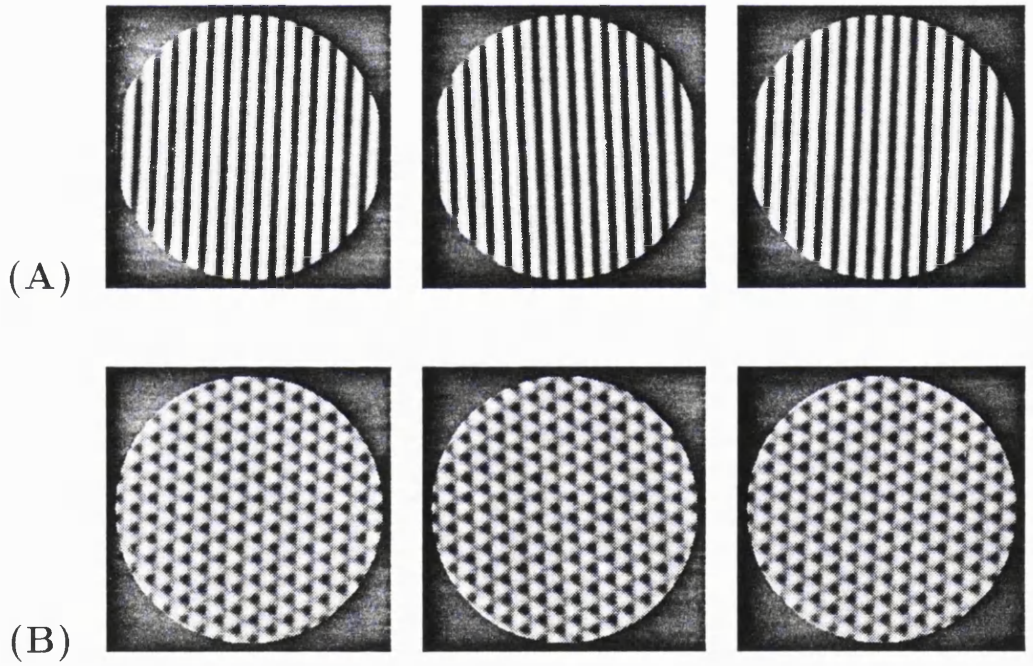


Figure 3.1: *Stimuli used in the first experiment.* **(A)** *A grating stimulus.* **(B)** *The probe stimulus. The stimuli have an equal vertical gradient of disparity.*

The first experiment assessed the perceived slant of grating stimuli over a range of suprathreshold slants. Each block of trials used a vertical sinusoidal grating, with a spatial frequency of 2.5 cycles/degree. The test stimulus had a horizontal or

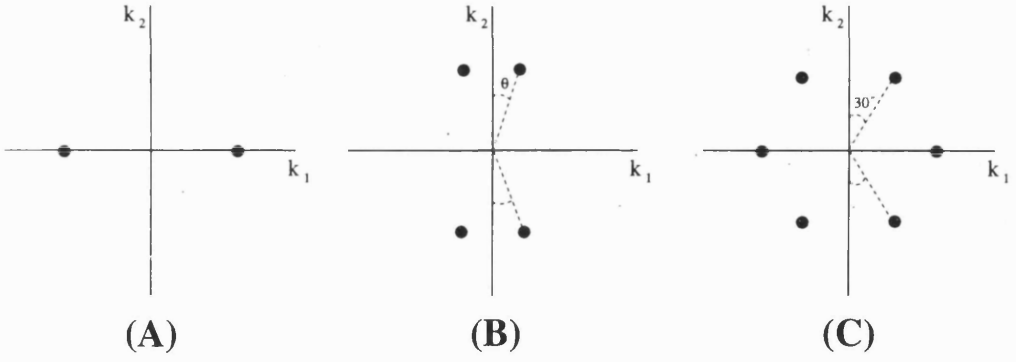
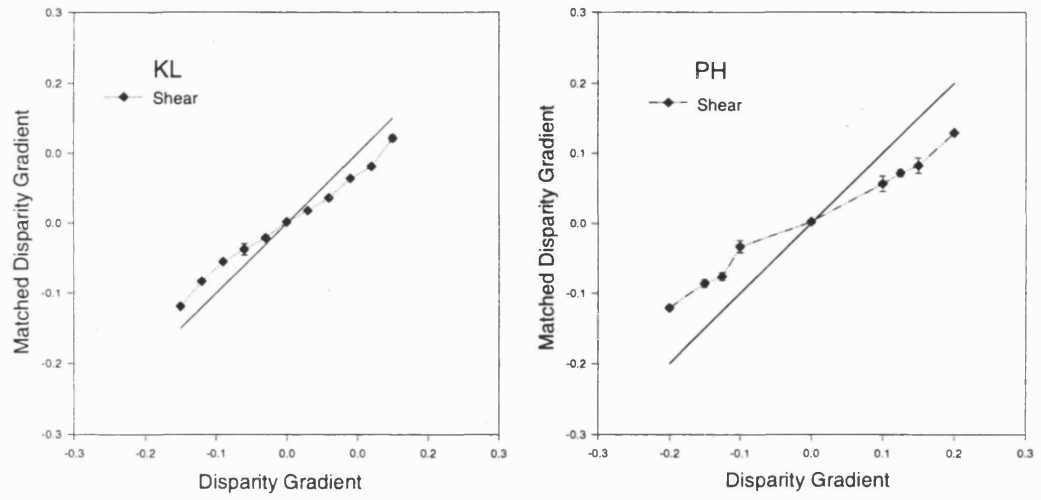


Figure 3.2: *Fourier transforms of the stimuli shown in figures 3.1 and 3.4.* **(A)** *Grating.* **(B)** *Plaid.* **(C)** *Probe. Transforms represent the stimuli prior to binocular transformation.*

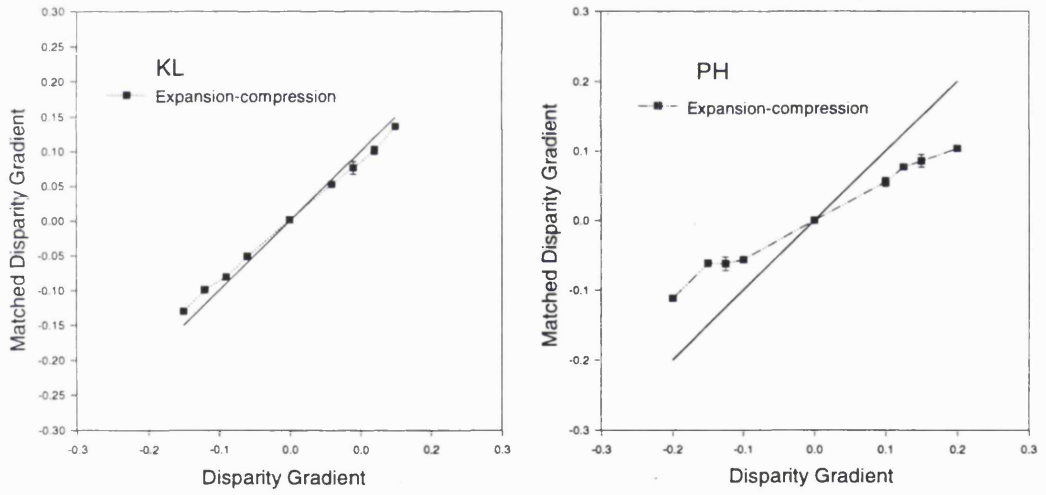
vertical gradient of disparity, the magnitude of which was varied between blocks. The procedure described above was used to assess the magnitude of disparity gradient in the probe stimulus that appeared equally slanted. Figure 3.1 shows the grating and probe stimuli; slant may be perceived in each stimulus. Figures 3.2A and 3.2C show the Fourier transforms of the stimuli. This representation highlights the differences in orientation of the components of each stimulus.

3.1.1 Results

Figure 3.3 shows 50% points for grating stimuli. The 50% point represents the magnitude of disparity gradient in the probe stimulus appearing to have the same slant as the test stimulus. If test and probe stimuli had equal perceived slant for equal disparity gradients, the results would lie on a straight line through the origin, with a slope of unity. This is the solid line in the graphs in figure 3.3. Generally, the grating stimulus appeared equal in slant to a probe stimulus with a smaller disparity gradient. A grating appeared to slant less than a three component probe



(A)



(B)

Figure 3.3: *50% points plotted against magnitude of grating disparity gradient.* **(A)** *Vertical disparity gradient.* **(B)** *Horizontal disparity gradient.*

	R^2	Slope
PH (Shear)	0.98256	0.5368
PH (expansion)	0.97761	0.5819
KL (Shear)	0.9892	0.7268
PH (expansion)	0.9964	0.8778

Table 3.1: *Results of linear regressions on the data shown in figure 3.3. R^2 values, and the slopes of the regression fits are shown. In all cases, the slopes were significantly below 1.0, ($p < 0.05$) showing that slant and inclination are consistently underestimated.*

stimulus with an equal magnitude of disparity gradient. The discrepancy between the actual and perceived slant of test stimuli tended to increase with increasing slant. Linear regressions on these data showed this underestimation to be signficant for both surface types, for both subjects (table 3.1)

3.1.2 Discussion

Grating stimuli appeared less slanted than probe stimuli with equal disparity gradients. This result may be related to the increased orientation disparity contained in the probe stimuli. However, the maximum orientation disparity in each stimulus (that of the vertical component of the probe stimulus) was equal for a given magnitude of slant or inclination. Additionally, slant was perceived in a grating with a horizontal disparity gradient. This stimulus has no orientation disparity, so the perception of slant could not have been due to orientation disparity alone.

A stereoscopically defined probe stimulus was used to minimise the contribution of factors other than disparity to trends apparent in the data. This procedure does

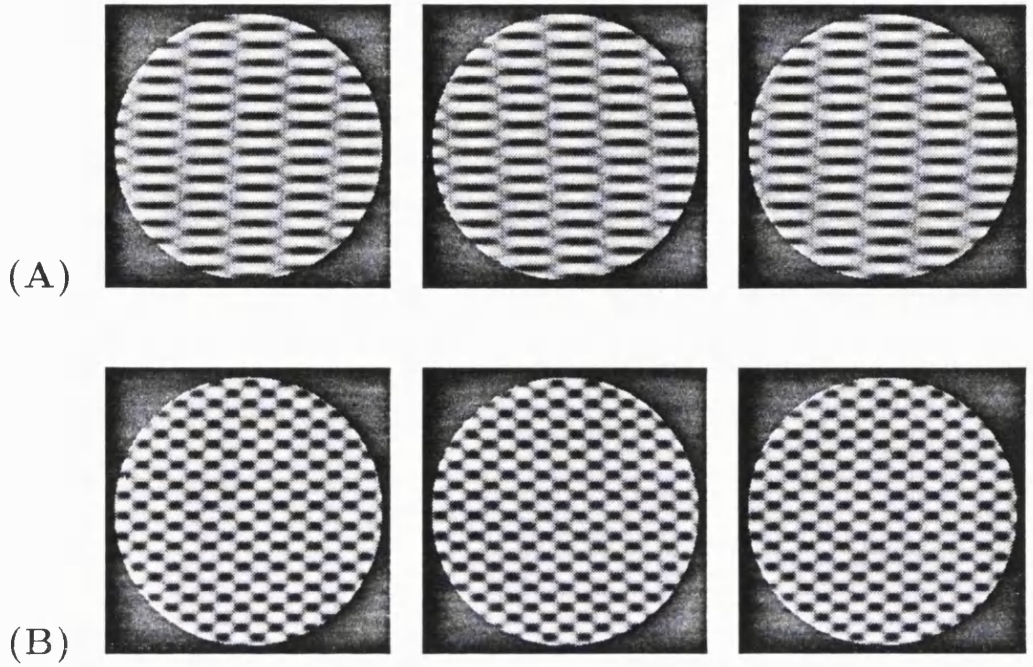


Figure 3.4: *Examples of the plaid stimuli used in the second experiment. In* **(A)** *, the components are oriented at* $\pm 10°$*; in* **(B)** *, they are oriented at* $\pm 30°$*. While both stimuli have an equal vertical gradient of disparity,* **(A)** *should appear less slanted than* **(B)**.

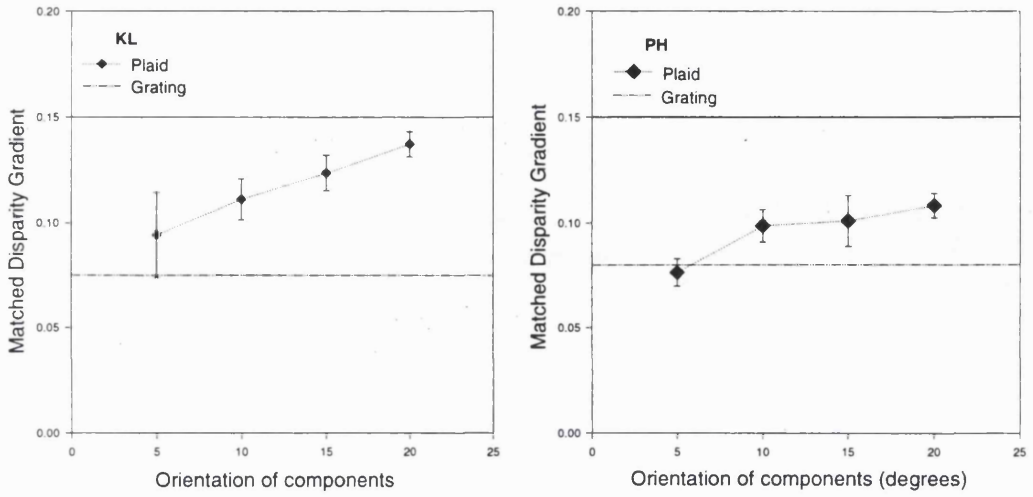
not remove the effects of perspective conflict. However, perspective conflict would have been less for a single vertical grating than for the probe stimulus, which had both oblique and vertical components. While perspective conflict may have affected the perceived slant of the probe stimulus, this would not account for the fact that it appeared *more* slanted than the grating stimulus. While perspective conflict will undoubtedly have affected the perceived slant of both test and probe stimuli, its effects would not predict the results found here.

3.2 Plaids

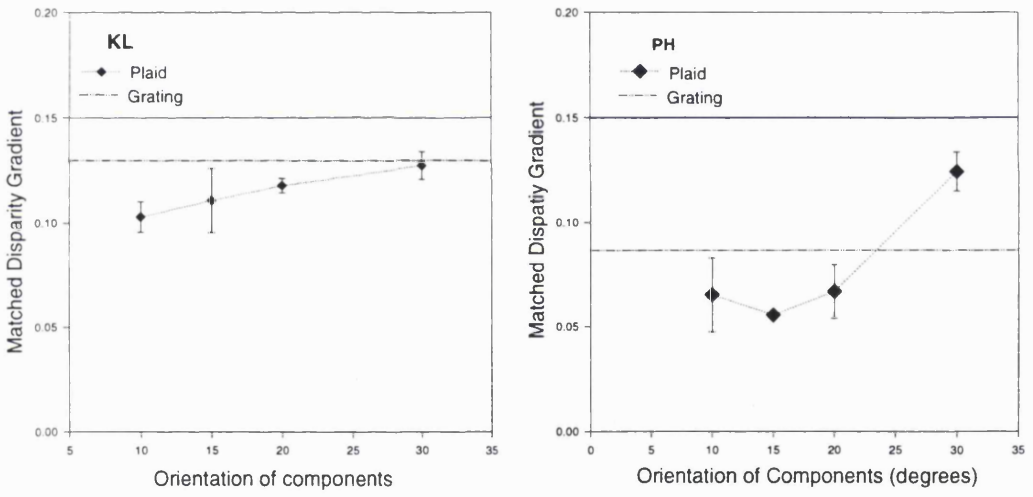
The second experiment investigated the perceived slant of two-component plaid stimuli. It is related to the final experiments of chapter 2, which explored the effect of plaid component orientation on slant thresholds. An important goal of this experiment was to explore the possibility that a non-Fourier mechanism plays a role in stereoscopic slant perception for suprathreshold slants. As reasoned in the introduction, if a non-Fourier mechanism does play a role in stereopsis, it may be more evident for suprathreshold than for threshold stimuli, due to the higher disparity thresholds reported for contrast disparities than for luminance disparities (Liu et al., 1992; Hess and Wilcox, 1994, Lin and Wilson, 1995). If this is the case, perceived slant may not be expected to be related to the orientation of the plaid's components. Conversely, if slant is encoded on the basis of disparities in the plaid's components, perceived slant might be expected to depend on their orientation.

Test stimuli were plaids produced from the superposition of two component sinusoidal gratings. These components had a spatial frequency of 2.5 cycles/degree, and were oriented symmetrically about horizontal. On independent blocks of trials, the components had orientations of $\pm 5°$, $\pm 10°$, $\pm 15°$, or $\pm 20°$, for stimuli slanted about a horizontal axis. For slant about a vertical axis, component orientations were $\pm 10°$, $\pm 15°$, $\pm 20°$ or $\pm 30°$. Figure 3.4 shows examples of the stimuli, with a vertical disparity gradient, and components oriented at $\pm 10°$ and $\pm 30°$; figure 3.2B shows the Fourier transform of the stimulus. Stimuli were again presented with both horizontal and vertical gradients of disparity. In all cases, this gradient had a magnitude of 0.15. The probe stimulus and matching procedure were identical to those used in the first experiment.

3.2.1 Results



(A)



(B)

Figure 3.5: *50% matching points for plaids, as a function of the orientation of plaid components.* **(A)** *Vertical gradient.* **(B)** *Horizontal gradient. The solid horizontal line in each figure represents the results expected if the plaids had appeared to slant equally to a probe stimulus with the same disparity gradient. The dotted line in each figure represents the apparent slant of an equivalent grating stimulus (replotted from figure 3.3).*

	R^2	Intercept	Slope
PH (Shear)	0.7699	0.07159	0.00195
PH (expansion)	0.6476	0.01785	0.00321
KL (Shear)	0.9905	0.08153	0.00282
KL (expansion)	0.97512	0.09183	0.001224

Table 3.2: *Results of linear regressions on the data shown in figure 3.5. R^2 values, and the intercepts and slopes of the regression fits are shown. In all cases, the intercept was significantly less than 0.15 ($p < 0.05$). The slope values, representing the effect of component orientation, were signifcantly different to 0 for subject KL ($p < 0.05$), but not for subject PH.*

Figure 3.5 shows the 50% points of fitted psychometric functions, as a function of orientation. Results are presented independently for horizontal and vertical axes of slant. The thick line in figure 3.5 represents the slant of the test stimuli. If test stimuli had appeared equally slanted as probe stimuli with an equal disparity gradient, the results would lie on this line. Plaid stimuli appeared to slant less than probe stimuli in all cases. Perceived slant decreased as the component gratings tended towards horizontal. Again, linear regressions were performed on the data. While the intercept was found to be significant ($p < 0.05$) in all cases, the effect of component orientation was only found to be significant for subject KL ($p < 0.05$) (table 3.2).

The dotted lines in figure 3.5 represent the perceived slant of single vertical gratings with equal disparity gradients to the plaids. These are drawn on the basis of the results of the first experiment.

3.2.2 Discussion

In all cases, two-component plaid stimuli had less perceived slant than did the three-component probe stimulus. As the orientation of the components approached horizontal, perceived slant decreased. This may be observed in figure 3.4. The stimuli in figures 3.4A and 3.4B have equal vertical disparity gradients. However, 3.4A, which has components oriented at $\pm 5°$, appears to slant less than 3.4B, which has components oriented at $\pm 20°$.

If perceived slant depended on disparities in the plaid components alone, it would be expected to decrease as component orientation tended to horizontal. This effect was apparent in the data, and may be observed in figure 3.4. The results may also be related to the notion that a non-Fourier mechanism encodes slant on the basis of disparities in the vertically oriented contrast envelope. This possibility may be assumed equivalent to the addition of a vertically oriented component to the image. This analysis does not assume that luminance and contrast disparities are processed by a single mechanism. Altering the orientation of the components affected the spatial frequency of the contrast envelope, but not its orientation. Perceived slant is determined mainly by the orientation of image contours (Mitchison and McKee, 1990; Gillam and Ryan, 1992). The effects found here could have resulted from the change in component orientation directly.

It is possible that the effects of component orientation were due to orientation disparities, or to perspective conflict. The latter would predict that, for horizontal gradients of disparity, plaids with components oriented closer to horizontal would appear less slanted. For vertical gradients of disparity, however, smaller effects in the opposite direction would be expected. The orientation disparity content of images would predict a decrease in perceived slant for components oriented closer to horizontal for both axes of slant. Again, while perspective conflict may affect perceived

slant, it cannot explain the results found in all conditions reported here. Decreasing the angular separation of the components would also increase the perceived contrast of the plaid (Georgeson and Shackleton, 1994). In motion, perceived speed has been found to decrease for low contrast stimuli (Thompson, 1982). As perceived slant decreased for decreasing component separation, it may be concluded that the results found do not reflect the effect of a similar contrast dependence in stereopsis.

4 Conclusions

Grating and two-component plaid stimuli appeared less slanted than a three component plaid probe stimulus, which was chosen so as to avoid the stereoscopic aperture problem, and to allow for effective encoding of slant. For two-component plaid stimuli, apparent slant was dependent on component orientation. If contrast envelope disparities contribute to the perception of slant, their contribution is minor in comparison to that of the plaid's Fourier components.

Taken as a whole, the results of these two experiments may be accounted for on the basis of component orientation disparities. It is not necessary to suppose that envelope disparities are used in addition.

While it may be expected that envelope disparities would be more evident for suprathreshold slant, no evidence for this was found. As was reported for slant thresholds, perceived slant was found to be influenced by the orientation of a plaid's components. Again, related results have been reported in motion. Welch (1989) found Weber fractions for the velocity of plaid motion to be predictable from plaid component velocities, rather than the plaid velocity itself, even for relatively high baseline velocities. The perceived speed of moving sinusoidal gratings declines for both high and low spatial frequencies, exhibiting a peak for intermediate frequencies of around

4 cycles/degree (Campbell and Maffei, 1981; Smith and Edgar, 1990). Smith and Edgar reported that the underestimation of the speed of a high frequency grating relative to a lower frequency grating increased as the velocity of the gratings increased. They reported similar results for plaid stimuli. For the experiments reported in this chapter, the underestimation of the slant of a grating relative to the probe stimulus also increased as the slant of the grating increased.

In the study of plaid motion, a distinction has been drawn between Type I plaids, whose component gratings lie on either side of the Intersection of Constraints velocity, and Type II plaids, whose component gratings lie on one side of the IOC direction (Ferrera and Wilson, 1987,1990,1991). Ferrera and Wilson (1991) reported that the perceived speed of Type I, symmetrical plaids (similar to those used here) is underestimated relative to a single grating moving in the IOC direction of the plaid. Further, this underestimation depended on the angular separation of the components of the plaid, such that the degree of underestimation increased as the IOC speed of the plaid increased relative to the speeds of the carriers. These results are consistent with the findings reported here.

Ferrera and Wilson (1990) also studied the perceived direction of motion of plaid stimuli. Whereas the direction of motion of Type I plaids was perceived veridically, they found marked misperception of the direction of motion of Type II plaids. Similar results were reported by Burke and Wenderoth (1993), who found in addition that the degree of this misperception depended on the angular separation of the plaid components. These biases may be accounted for by a model which does not require an explicit, non-Fourier channel (Langley and Fleet, 1995). It might be predicted that similar biases in the perceived tilt of stereoscopically presented plaid stimuli may be observed.

The results of the experiments presented here, and the results of related motion

experiments, provide little evidence for non-Fourier channels in the encoding of properties of plaid stimuli. This is not to say however that such channels do not exist. Rather, they do not appear to play a role in the perception of motion and slant for simple plaid stimuli of the type used here. In the next two chapters, experiments are presented which aim to investigate situations in which non-Fourier mechanisms may be evident in stereopsis.

4. Luminance and contrast disparities in stereoscopic transparency

1 Introduction

The previous two chapters investigated the perception of slant in stereoscopic plaids. For both threshold and suprathreshold stimuli, slant perception showed a marked dependence on the orientation of the component gratings of the plaid. These experiments found no evidence for a stereoscopic non-Fourier channel making direct use of contrast envelope disparities. Similar results have been reported in the study of plaid motion (Welch, 1989).

Other researchers have suggested that stereopsis is supported by disparities in image contrast envelopes. Stereoscopic depth may be influenced by contrast envelope

disparities in both Gabor stimuli (Liu, Schor and Ramachandran, 1992; Hess and Wilcox, 1994; Wilcox and Hess, 1995, 1996) and contrast modulated, random dot stereograms (Sato and Nishida, 1993). These results have been accounted for by proposing that stereopsis has access to independent linear (Fourier) and nonlinear (non-Fourier) channels (Hess and Wilcox, 1994). However, it was argued in the introductory chapter that it may be possible to account for these data using a single channel model, such as that proposed by Fleet, Wagner and Heeger (1996).

For non-Fourier motion stimuli, motion is seen with a velocity for which no Fourier components exist. Fleet and Langley (1994a) demonstrated that many classes of these stimuli have a relatively simple characterisation in frequency space, related to the notions of phase and group velocity. The simplest of these stimuli consist of the motion of a contrast envelope over an underlying carrier. An example of this is a contrast modulated grating, involving the motion of a contrast envelope over an underlying carrier grating (figure 4.1). This stimulus is perceived as transparent, with the contrast envelope and carrier moving with different velocities (Derrington and Badcock, 1985).

Motion transparency may also be observed from additively combined signals. The simplest example of this is transparent motion in plaid stimuli. A one dimensional signal viewed through a circular window will be seen to move in the direction normal to its orientation (Wallach, 1935). If two moving one dimensional structures, with different orientations, are added together, there exists a unique velocity consistent with both normal velocities. This is known as the Intersection of Constraints (or "IOC") velocity (Adelson and Movshon, 1982). A plaid composed of gratings similar in contrast, orientation, and spatial and temporal frequency will be seen to move with this IOC velocity (Adelson and Movshon, 1982; Movshon, Adelson, Gizzi and Newsome, 1985; Nakayama and Silverman, 1985; Farid and Simoncelli, 1994). For other superpositions, the gratings will tend to be seen to move transparently over

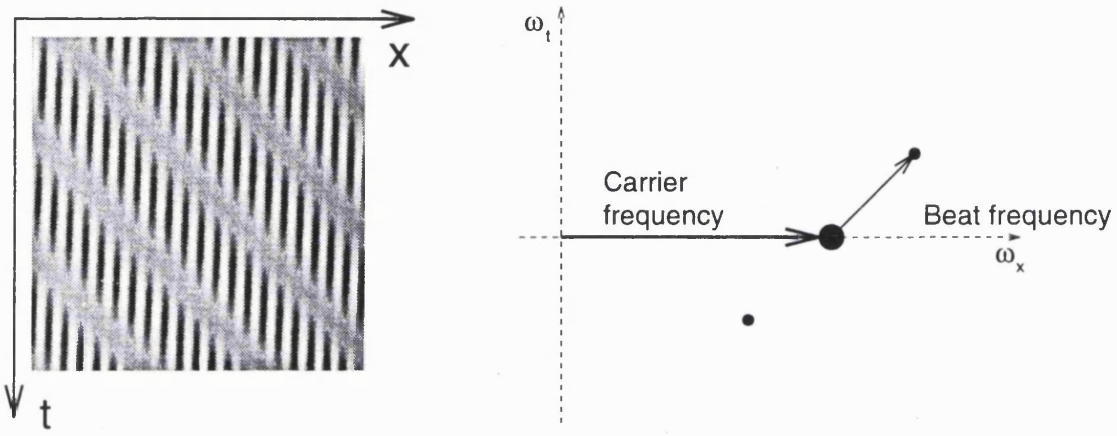


Figure 4.1: **(A)** *Space-time diagram of contrast modulated grating motion. The carrier grating is stationary. As such, it appears as vertical in the space-time plot. The contrast modulating grating moves to the right, and can be seen as a contrast modulation oriented at* 135°. **(B)** *Fourier transform of* **(A)**. *The centroid of power represents the stationary carrier grating. This is signified by the thick horizontal vector, the magnitude of which gives the grating's spatial frequency. The side bands of power are introduced by the contrast modulation; the orientation of the vector from the centroid of power to the sidebands gives the velocity of the contrast envelope.*

one another. Here, the transparency may be decomposed into the velocities of two additively combined components.

Kersten (1991) analysed several physical causes of perceptual transparency. Common to the examples described is the idea of a distant object viewed through a transparent medium. The transparent medium modifies light passing through it from more distant objects. This attenuated signal is combined with luminance arising directly from the transparent medium itself. The modification of the luminance pattern takes the form of a contrast reduction, the transparent medium transmitting only a fraction of incident light. This contrast reduction may be described in terms of the transmittance of the transparent medium, $\tau(x, y)$, which may vary spatially. $\tau(x, y)$ may take values between 0 and 1, a value of 0 corresponding to complete occlusion, a value of 1 to complete transparency. If a luminance pattern $I_1(x, y)$ passes through a transparent object with transmittance $\tau(x, y)$, and is combined with a luminance pattern $I_2(x, y)$ arising from the transparent object, the resulting image $I(x, y)$ is given by:

$$I(x, y) = \tau(x, y)I_1(x, y) + I_2(x, y) \tag{1}$$

Equation (1) involves both multiplicative and additive combinations of underlying signals. Figure 4.2 illustrates how multiplicative and additive signal combinations arise in perceptual transparency. Equation (1) may be constrained by the fact that neither luminance nor transmittance can take negative values. A non-negative signal may be written as the sum of a constant term and a mean zero term:

$$I(x, y) = \gamma(1 - \alpha f(x, y)) \tag{2}$$

where $0 \leq \alpha \leq 1$, $\gamma \geq 0$ and $-1 \leq f(x, y) \leq 1$. Considering only the first term on

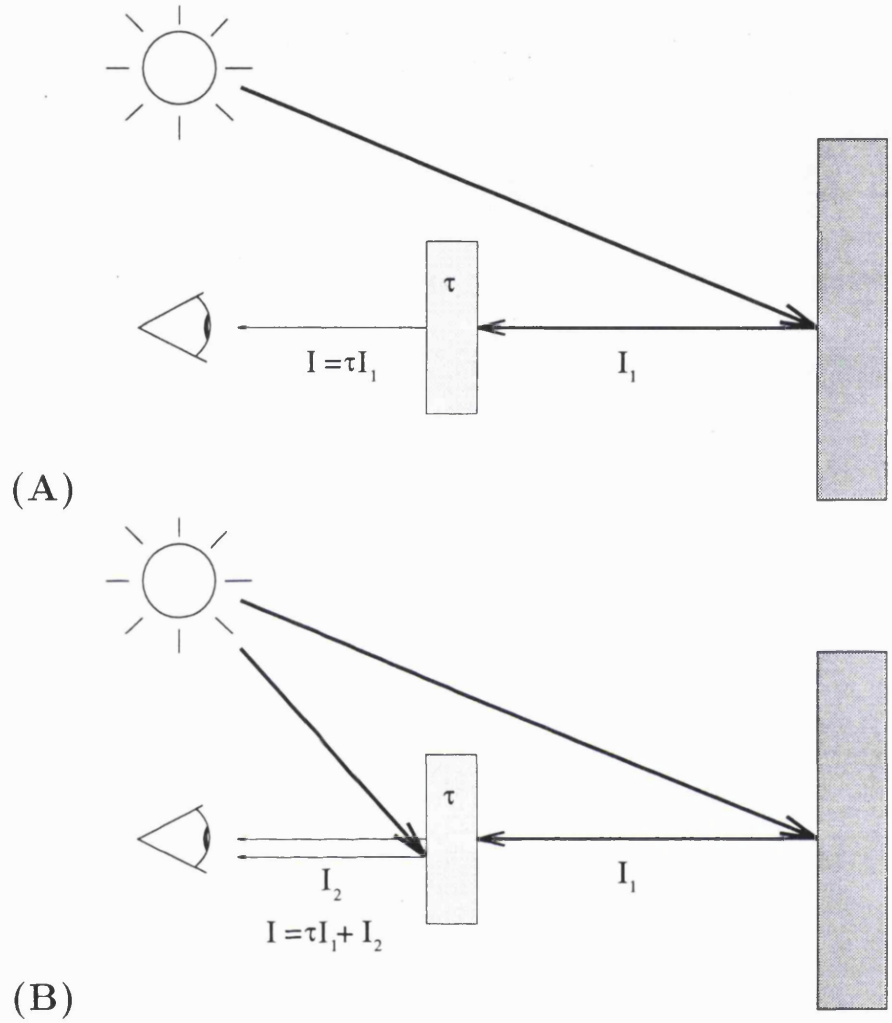


Figure 4.2: **(A)** *Light, with a luminance I_1, is reflected from a distant object is attenuated by passing through a transparent object with transmittance τ. This is an example of a multiplicative transparency.* **(B)**. *The attenuated light may be added to light reflected from the transparent object, I_2, to form an additive component of transparency.*

the right hand side of equation (1), it may be rewritten:

$$I(x,y) = \gamma[1 - \alpha f(x,y) - \beta g(x,y) + \alpha\beta f(x)g(x)] \tag{3}$$

where $0 \leq \alpha, \beta \leq 1$. Equation (3) includes both additive and multiplicative signal combinations. This may be related to the psychophysical findings that both additive and multiplicative signals are able to generate transparency. Perceptual transparency involves the simultaneous representation of surfaces at different depths. A contrast modulated grating, for example, may appear as a contrast modulation moving *in front of* a grating. Transparent plaid stimuli also appear as one grating moving in front of another. Which of the gratings appears as closer may be influenced by the addition of stereoscopic disparity to one of the gratings (Adelson and Movshon, 1984; von Grünau, Dabé and Kwas, 1993). Transparency may also be observed purely on the basis of disparities. If random dot stereograms with differing disparity content are superimposed, two or more depth planes may be observed (Akerstrom and Todd, 1988; Weinshall, 1990; Parker, Johnston, Mansfield and Young, 1991; Langley, Fleet and Hibbard, 1995). These planes, and hence the transparency, are purely cyclopean.

In this chapter, stereoscopic equivalents of several motion transparency stimuli are explored. The principle goal was to assess whether stereoscopic transparency is possible given additive and multiplicative signal combinations. These experiments will provide further evidence relating to the question of whether stereopsis makes use of independent Fourier and non-Fourier channels.

2 Methods

2.1 Procedure

For each trial, subjects were presented with a stimulus, and asked to decide whether they observed a transparency. They were asked to make this decision on the basis of whether or not they could perceive two distinct surfaces, clearly separated in depth. This criterion was used to eliminate the possibility that monocularly visible cues to transparency were being used in isolation, and is the same criterion as is used in motion transparency experiments. All stimuli were consistent with the perception of a single surface, but were constructed such that disparity information should also allow the perception of transparency. In all cases, stimuli were presented in a hard circular window, with a diameter of 7.9 degrees.

In all experiments, one independent variable was varied between trials using the APE adaptive probit analysis algorithm (Watt and Andrews, 1981). Other independent variables were varied between blocks of trials, which were randomly interleaved. For each block of trials, a psychometric function was fit to the data using the APE algorithm. The fitted curve was used to ascertain the point when transparency was reported on 50% of occasions. This fitted point was taken to represent the threshold for perceived transparency. Presented results show the mean and standard error of the 50% points, based on three independently measured psychometric functions. Each function is based on 64 trials.

3 Experiments

3.1 Additive Transparency

This experiment aimed to assess whether depth transparency is possible given an additive combination of signals. In motion, transparency may arise from the summation of moving one dimensional gratings. Transparency and coherence are possible in plaid motion due to the existence of independent normal velocities for the two components, and the IOC direction for the pattern as a whole. To investigate similar transparency in stereoscopic plaid stimuli, gratings must be presented which, if viewed independently, would appear to lie in different depth planes. Stimuli consisted of the sum of a horizontal and vertical grating. A binocularly viewed horizontal grating carries no horizontal disparity information. A vertical grating on the other hand can be given binocular disparity, and will be seen to lie in the plane determined by this disparity. If a horizontal and a vertical grating are added to form a plaid, then the depth plane in which the plaid lies will be determined by the disparities in the vertical grating. Binocular images were created by horizontally shearing a plaid formed from the addition of a horizontal and a vertical grating. This transformation produced a vertical gradient of disparity. A vertical grating with a vertical gradient of disparity will appear to slant about a horizontal axis. In preliminary investigations, it was found that this stimulus under some circumstances appeared transparent, with the horizontal grating appearing in the plane of fixation. This may have resulted from the hard boundary used. If transparency was not observed, the whole stimulus would be expected to appear to slant about a horizontal axis.

Plaid stimuli were created so as to make perceptual transparency likely. In motion, transparency is observed for gratings differing in orientation, spatial and tempo-

ral frequency, and contrast. (Adelson and Movshon, 1982; Movshon et al., 1985; Nakayama and Silverman, 1985; Farid and Simoncelli, 1994). This experiment used plaids formed from orthogonal gratings with differing spatial frequencies. Between trials, the contrast of one of the gratings was altered as an independent variable. Stimuli consisted of a horizontal grating with a spatial frequency of 2.0 cycles/degree, superimposed on a vertical grating with a spatial frequency varying between blocks between 3.5 cycles/degree and 5.0 cycles/degree. The Michelson contrast of the horizontal grating was fixed at 0.70; the contrast of the vertical grating was varied between trials. The vertical grating was subjected to a horizontal binocular shear with a magnitude of 0.1; the sign of the resulting vertical disparity gradient was varied randomly between trials. In all cases, the top of the stimulus had zero disparity, so that the transformation took the form of a gradient of (either crossed or uncrossed) disparity, increasing towards the bottom of the stimulus.

3.1.1 Results

Figure 4.4 shows a typical response function. Transparency was observed when the contrast of the vertical grating was low, but not when it approached the contrast of the horizontal grating. Also shown are the 50% points as a function of spatial frequency. As the spatial frequencies of the component gratings were made more similar, a greater contrast difference was required to observe transparency.

3.1.2 Discussion

In this experiment, transparency was observed from the linear addition of image signals. The case of stereopsis appears equivalent to that of two dimensional motion. The stimuli used here incorporated either crossed or uncrossed disparities in the vertical gratings. Both signs of disparity supported the perception of transparency.

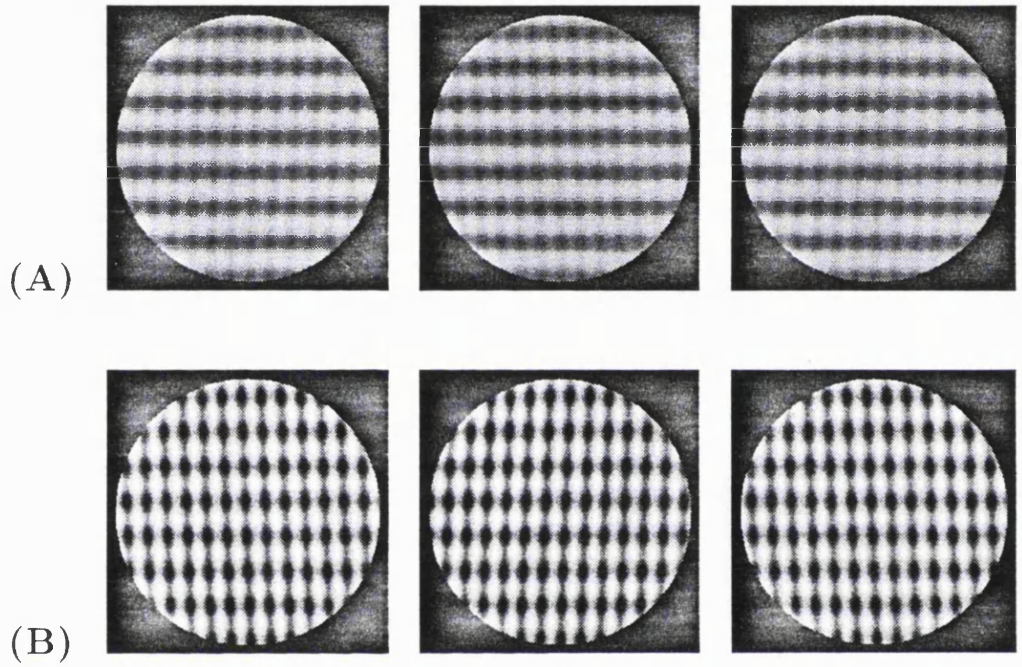
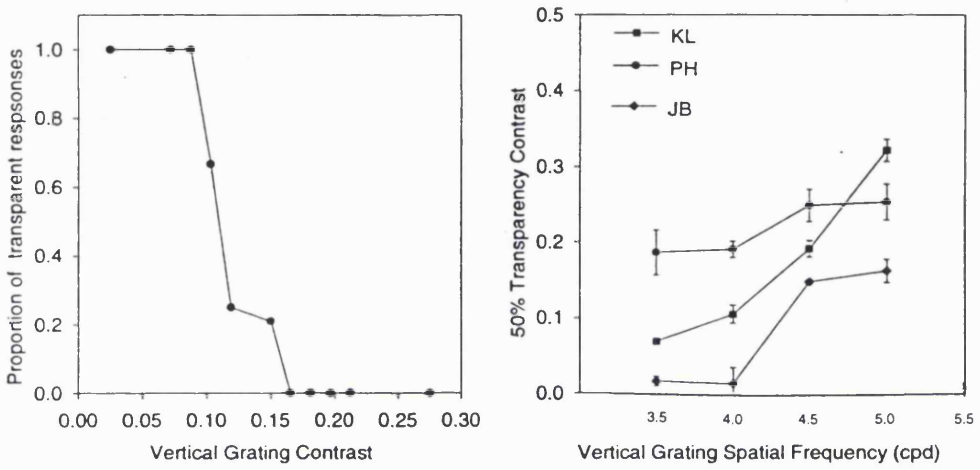


Figure 4.3: *Examples of the stimuli used in the first experiment.* **(A)** *With the vertical grating at low contrast, crossed-eyed fusion reveals a transparency, with the vertical grating appearing to slant in front of* (left)*, or behind* (right) *the horizontal grating.* **(B)** *With both gratings at high contrast, a single slanted surface is perceived.*



(A) **(B)**

Figure 4.4: **(A)** *A typical response function (for subject JB, with a vertical spatial frequency of 4.0 cycles/degree). Transparency was observed for low contrasts of the vertical grating, but not for higher contrasts.* **(B)** *50% points for all subjects plotted against the spatial frequency of the vertical grating. Transparency was perceived over a wider range of contrasts as the difference in frequency between the component gratings was increased.*

Subjects were able to see the vertical grating lying both in front of and behind the horizontal grating.

The stimuli used in this experiment were transformed using an interocular horizontal shear, such that the vertical grating appeared to lie on an inclined plane. A horizontal shear was used partly so that the stimuli could be compared to those used in chapters 2 and 3, where slant was perceived, but not transparency. However, it would expected that similar results would be obtained with simple horizontal disparities, in which case the vertical grating would be expected to lie on a frontoparallel surface in front of or behind the horizontal grating.

The results of this experiment may be contrasted with those of chapters 2 and 3, in which a binocularly transformed plaid was seen as a single, slanted surface. Plaids used in the previous chapter were formed from two components with identical spatial frequencies and contrasts, and similar orientations. Again, these results reflect those reported in motion, for which transparency occurs for stimuli differing in terms of contrast, orientation, and spatial and temporal frequencies. These results may be considered in terms of the bandwidths of orientation- and spatial frequency-tuned mechanisms. Figure 4.5 shows two plaids, with their Fourier spectra. The plaid in figure 4.5A has components with similar orientations and spatial frequencies. In figure 4.5C, the plaid components are orthogonal, and have markedly different spatial frequencies. Figure 4.5C may be compared to the stimuli used in the current experiment, whereas figure 4.5A is similar to those used in chapters 2 and 3. Figures 4.5B and 4.5D show the Fourier spectra of the two plaids. Whereas the components of figure 4.5A may be expected to fall within the bandwidth of a single orientation- and frequency-tuned mechanism, this is unlikely for figure 4.5C. This may explain why transparency was observed for plaids similar to figure 4.5C, but not those similar to 4.5A.

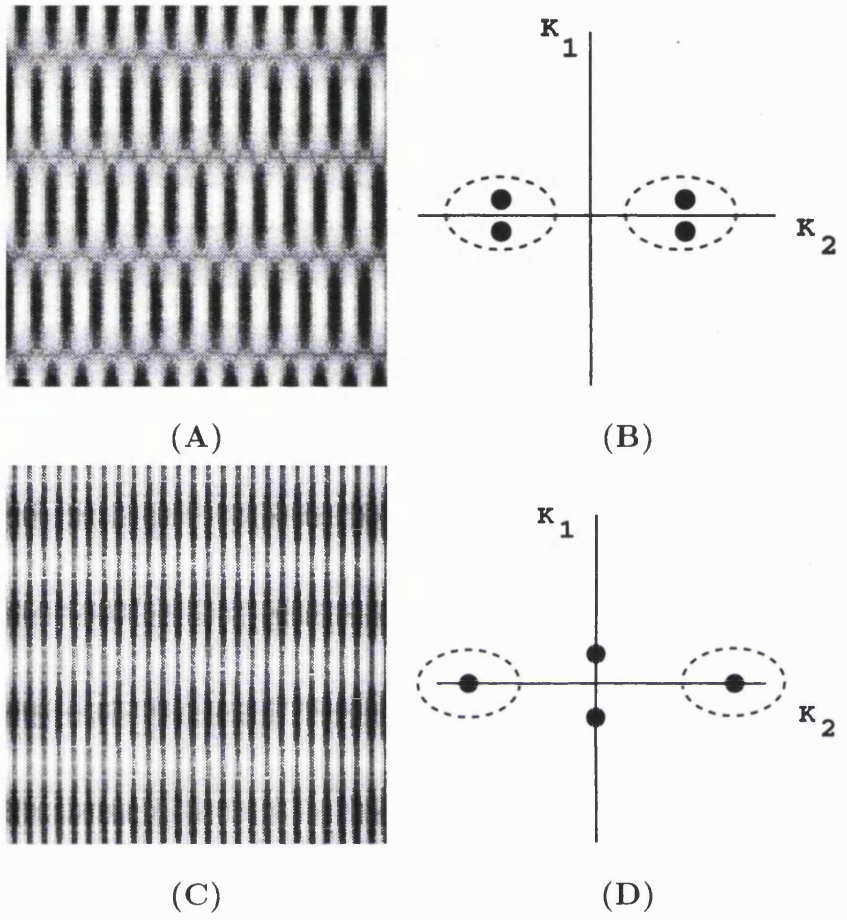


Figure 4.5: **(A)** *A plaid formed from two gratings similar in orientation and frequency.* **(B)** *Fourier transform of* **(A)**. *The two gratings lie within the passband of a single filter.* **(C)** *A plaid formed from two gratings, differing markedly in orientation and frequency.* **(D)** *Fourier transform of* **(C)**. *Here, the two gratings are widely separated in frequency space, and would be expected to be detected by independent bandpass filters.*

3.2 Multiplicative Transparency

Many non-Fourier motion stimuli consist of the motion of a contrast envelope over a carrier structure (e.g. Derrington and Badcock, 1985; Fleet and Langley, 1994a). In the same way, stereopsis may be supported by disparities in image contrast variations, (Liu et al., 1992; Fleet and Langley, 1994b; Hess and Wilcox, 1994; Sato and Nishida, 1993, 1994; Wilcox and Hess, 1995, 1996). This experiment investigated the relationship between depth from contrast disparities, and the perception of stereoscopic transparency.

A very simple stimulus was used, consisting of a high frequency horizontal sinusoidal grating modulated by another, lower frequency vertical grating. For the vertical grating, an approximation to a square wave was formed from the fundamental plus its third and fifth harmonics:

$$I(x,y) = \sin(fy)\left[1 + \lambda\left(\sin(2\pi fx) + \frac{1}{3}\sin(6\pi fx) + \frac{1}{5}\sin(10\pi fx)\right)\right] \qquad (4)$$

This stimulus was chosen as a compromise between a stimulus with sharp edges to the contrast envelope, and one well localised in Fourier frequency space. However, is would be expected that transparency would also be observed with other envelopes (e.g. a true square wave, or just the fundamental). The horizontal carrier had a spatial frequency of 4 cycles/degree. The spatial frequency of the vertical modulation was varied between trials between 0.2 cycles/degree and 0.95 cycles/degree. A horizontal carrier grating was used so that both a transparent and a coherent interpretation of the stimulus were possible. The binocular transformation used took the form of a uniform translation, or a horizontal shear.

Both crossed and uncrossed disparities were used, although no stimulus contained both signs of disparity. For sheared stimuli, points at the top of the stimulus had

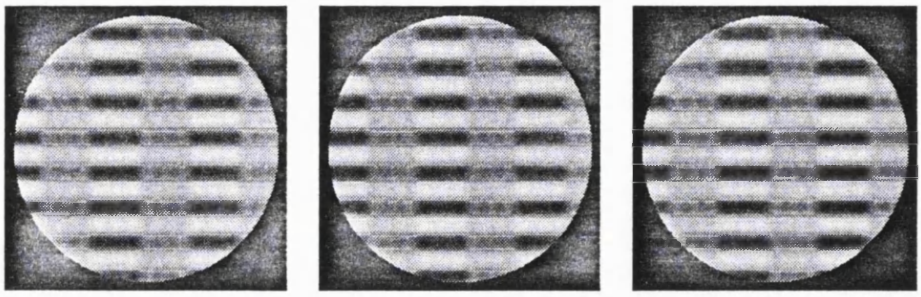
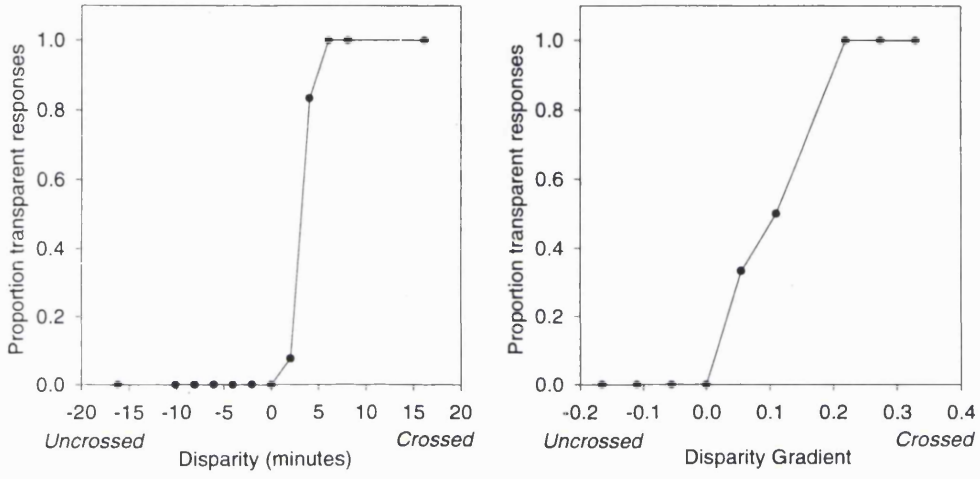


Figure 4.6: *Examples of the stimuli used in the second experiment. Crossed fusion of the left two images should show the contrast envelope hovering transparently in front of the carrier grating. When the right two images are fused, the whole pattern appears to lie on a single surface in depth.*

zero disparity, so that only crossed or only uncrossed disparities were present. For each block of trials, subjects were informed whether they would be presented with slanted or frontoparallel surfaces, and they were asked whether they perceived a single surface, or two surfaces separated in depth. Examples of the stimuli are given in figure 4.6.

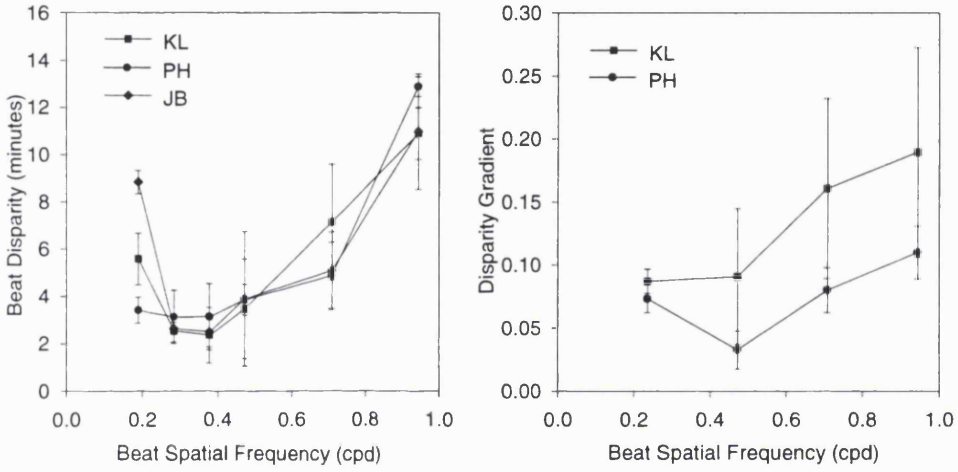
3.2.1 Results

Typical psychometric functions for the two tasks are presented in figure 4.7. For large crossed disparities, transparency was clearly observed, both for frontoparallel and for slanted surfaces. For uncrossed disparities, a single surface was always reported, i.e. the horizontal carrier grating appeared to lie in the same plane as the vertical modulation. For small crossed disparities, transparency was not reported. This may be explained in terms of the task as described to the subjects, who were asked to give a "transparent" response *only* if two surfaces were seen clearly separated in depth. Although these stimuli can be perceived to represent a transparency



(A) **(B)**

Figure 4.7: *Typical response functions* **(A)** *frontoparallel surfaces* **(B)** *slanted surfaces. Results in (B) represent a gradient of uncrossed, or of crossed disparities. In each case, transparency was perceived only for crossed disparities (or gradients of crossed disparities) above a minimum value. Both functions are for subject PH, with an envelope spatial frequency of 0.7 cycles/degree.*



(A) **(B)**

Figure 4.8: *50% points plotted for different spatial frequencies* **(A)** *frontoparallel surfaces* **(B)** *slanted surfaces. As the spatial frequency of the contrast modulation was decreased, the minimum disparity for which transparency was reported also decreased, reaching a minimum for spatial frequencies around 0.4 cycles/degree.*

when no disparity or motion cues are present, the task presented here requires that disparity information is used. Disparities in the contrast envelope must therefore be detectable for transparency to be observed. Figure 4.8 presents 50% points for psychometric functions as a function of the contrast envelope spatial frequency. The minimum crossed disparity or disparity gradient required for transparency to be observed decreased with decreasing contrast modulation spatial frequency, showing a minimum for frequencies around 0.4 cycles/degree.

3.2.2 Discussion

The stimulus used in this experiment was non-Fourier in that the the two independent depth planes perceived cannot be related directly to its additive components. The stimulus appeared as a horizontal sinusoidal grating and a vertical contrast modulation, separated in depth. One possible explanation of these results is that luminance and contrast disparities are processed independently. As an alternative, it is possible that an early nonlinear transformation of the image luminance may introduce Fourier components at the frequency of the contrast modulation. Transparency could then arise as a result of the independent processing of the horizontal and vertical components of the transformed image. However, transparency was only perceived when the contrast modulation had crossed disparity, and was seen to lie in front of the carrier. This is consistent with the physics of transparency (a contrast reducing medium must lie in front of the source of luminance which is modulated). It is also suggests that independent mechanisms process transparency from additive and multiplicative signal combinations. Although not explicitly tested, additive transparency was observed for both crossed and uncrossed disparities. The asymmetry was thus observed only for multiplicative transparency. The transparency reported here does not appear to have resulted from an early nonlinear transformation followed by independent processing of horizontal and vertical gratings, as

was evident in the previous experiment. Other studies of stereoscopic transparency have reported a similar depth asymmetry to that found here. The nature of multiplicative transparency requires that the transparent object (the contrast variation) lies nearer to the observer than do other objects. If binocular depth information inconsistent with this notion is included in such stimuli, the percept of transparency is often destroyed (Nakayama, Shimojo and Ramachandran, 1990; Kersten, 1991). Similar findings have been reported for occlusion (Pastore, 1974).

Finally, it is interesting to note that transparency was most readily perceived for contrast envelopes with a spatial frequency of around 0.4 cycles/degree. Contrast thresholds for the discrimination of the orientation of contrast envelopes also show a minimum for envelopes with this spatial frequency (Langley, Fleet and Hibbard, 1996). This result was interpreted as evidence for the existence of a contrast processing mechanism selective for spatial frequencies around 0.4 cycles/degree. The results presented here are consistent with transparency arising from disparity processing performed on the output of a similar mechanism.

3.3 Transparency in Squarewave Plaid Patterns

The first two experiments showed that both additive and multiplicative signal combinations may give rise to stereoscopic transparency. The final experiment investigated how transparency may arise when signals are combined both additively and multiplicatively. Stoner, Albright and Ramachandran (1990) performed an experiment in which plaids were generated from two square wave gratings. These patterns contained areas with three different intensities, corresponding to the background, the individual bars of the squarewaves, and their intersections. This stimulus allows the contributions of additive and multiplicative signals to be readily altered, by adjusting these three luminances. Stoner et al. (1990) varied the luminance of

the intersections, while holding the other two luminances constant. Increasing the intersection luminance is equivalent to increasing the contribution of the product of the two gratings. They found that, for relatively high and low intersection luminances, coherent motion was observed in the IOC direction. Within a range of luminances, however, the gratings were seen to move transparently in their normal directions. This range coincided with the range of intersection luminances for which multiplicative transparency was a viable interpretation of the image. This experiment aimed to replicate these results, using binocular disparity rather than motion as a cue to transparency.

Horizontal and vertical rectangular waveforms were used, and disparities consisted of a binocular horizontal shear. The stimulus prior to binocular transformation may be described as follows:

$$I(x,y) = [\alpha - \beta H_\lambda(y) - \beta H_\lambda(x) + \gamma H_\lambda(y) H_\lambda(x)] \tag{5}$$

where $H_\lambda(x)$ refers to a rectangular wave grating with a frequency defined by the subscript λ. Stoner et. al. (1990) used a duty cycle (defined as the ratio of narrow bar width: narrow bar width + wide bar width) of 0.286, such that the stimulus appeared as a series of dark bars against a light background. In equation (5), α determines the luminance of the background. The bar and intersection luminances are then given by $\alpha-\beta$ and $\alpha-2\beta+\gamma$ respectively. Gratings with a spatial frequency of 2.0 cycles/degree and a duty cycle of 0.286 were used. The background luminance and the bar luminance were fixed at 70cdm^{-2} and 35cdm^{-2} respectively.

Between experimental trials, the luminance of the bar intersections was varied. Subjects were asked to report whether a single surface, or two surfaces seen transparently in depth were observed. Between blocks of trials, the magnitude of the binocular shear was varied. Example stimuli are shown in figure 4.9.

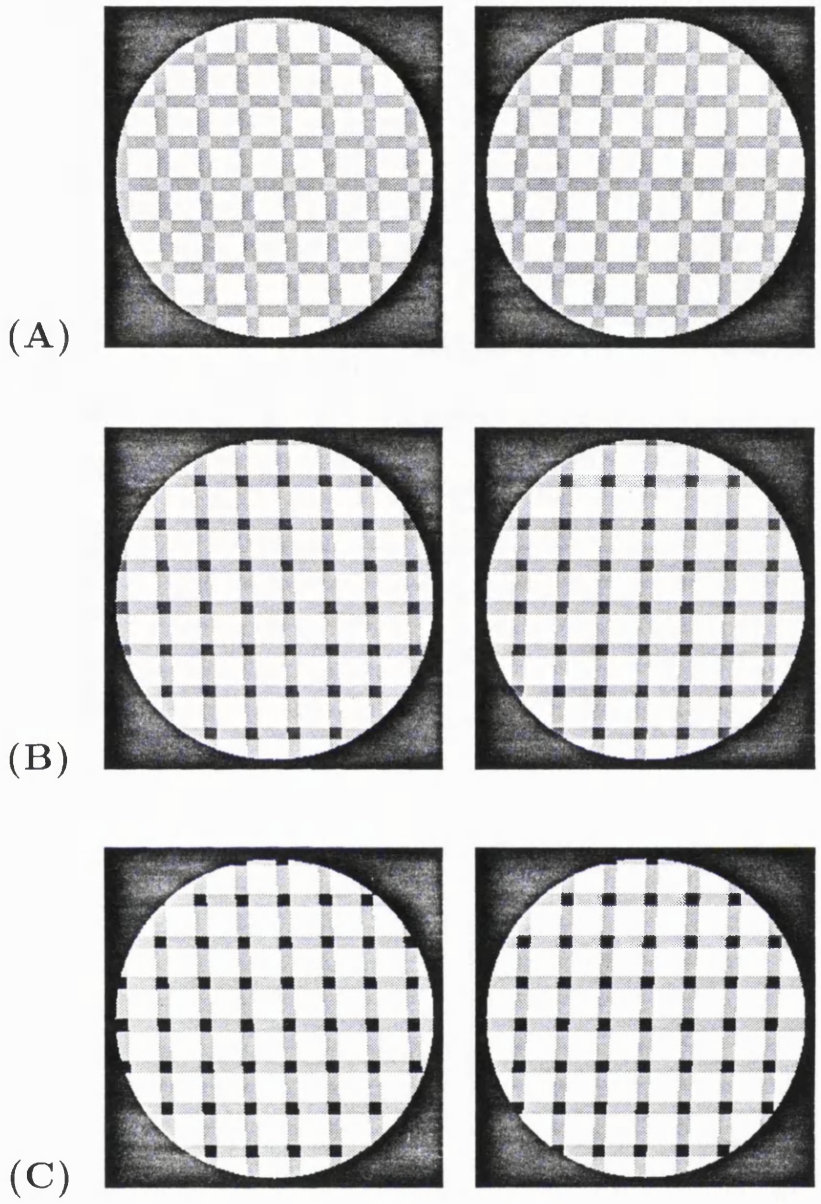


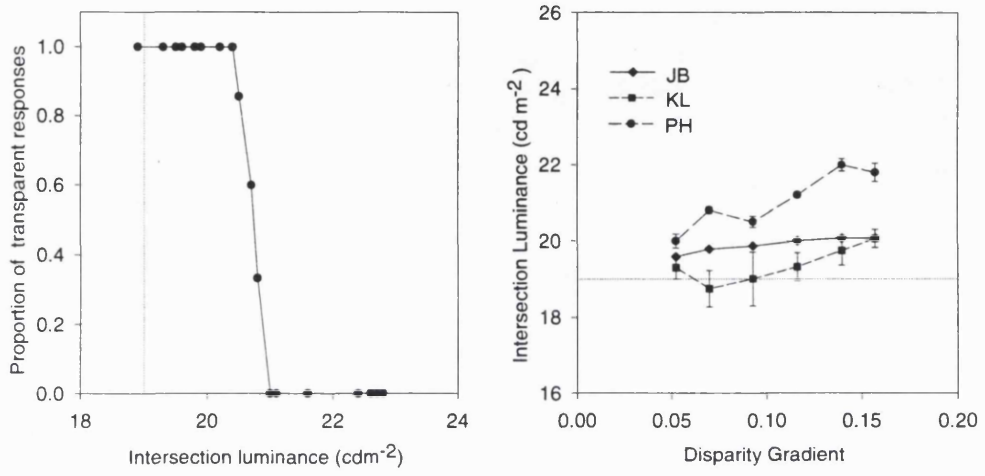
Figure 4.9: *Examples of the stimuli used in the third experiment.* **(A)** *With bright intersections, transparency was not observed.* **(B)** *With intersections ~~below~~ the luminance of the bars, transparency was observed, as reported by Stoner et al. (1990) in motion.* **(C)** *With vary dark intersections, transparency was again reported. This time however it took the form of wide, light vertical bars seen transparently in front of thin, dark, horizontal bars (see Appendix B for discussion).*

3.3.1 Results

The results of Stoner et al. (1990) suggested that transparency would only be observed over an intermediate range of intersection luminances. However, pilot studies revealed that transparency was also perceived with very dark intersections. Only for relatively high intersection luminances was a single surface reported. The APE algorithm was again used to measure the intersection luminance for which transparency was reported on 50% of trials. A typical response function, and 50% points measured for different magnitudes of surface slant, are given in figure 4.10. As the gradient of disparity was increased, transparency was observed over a wider range of intersection luminances.

3.3.2 Discussion

Above the intersection luminance at which the stimuli are no longer consistent with a film transparency, a single surface was reported. For luminances below this value, transparency was reported. Transparency was also reported, however, for very dark intersection luminances, where Stoner et al. (1990) reported coherent plaid motion for their stimuli. While this appears at first glance to indicate a difference between the domains of motion and stereopsis, there exists another, possibly important difference between the stimuli used here, and those used by Stoner et al. (1990). Stoner and Albright (1992) emphasised the importance of figure/ground assignment in the perception of transparency, and indeed the small duty used was chosen to bias the figure/ground percept to one of dark bars on a lighter background. In this experiment, some of the percepts of transparency were accompanied by a shift in the figural interpretation of the stimuli. Whereas most of the stimuli were interpreted as expected, those with the darkest intersections appeared as a series of thin dark bars on a lighter background, seen thrcugh thick, *light* transparent bars. The



(A) **(B)**

Figure 4.10: **(A)** *A typical response function (subject JB, with a disparity gradient of 0.1). With bright intersections, transparency was not reported. With darker intersections, transparency became apparent; there appeared to be no lower bound on intersection luminance below which transparency ceased to be reported.* **(B)** *50% points for different magnitudes of surface slant. As slant was increased, transparency was observed over a wider range of intersection luminances. The dotted vertical line in the left-hand graph, and the dotted horizontal line in the right-hand graph, represent the intersection luminance associated with a purely multiplicative transparency.*

very dark and very light areas of the image appeared as figure, while the midrange luminances appeared as both figure and ground. This may be seen in the stimulus in the bottom row of figure 4.9.

This interpretation of the stimuli is consistent with a transparency in which the thick, transparent bars have a transmittance of less than one, but contribute additional brightness to the image. This percept may not have been reported by Stoner et al. (1990) because of the luminances chosen for their background and bars, which did not permit this interpretation of the image (see Appendix B for details).

4 Conclusions

Binocular transparency was observed on the basis of both additive and multiplicative signal combinations. These results are similar to previously reported experiments in motion (Adelson and Movshon, 1982; Stoner et al., 1990; Farid and Simoncelli, 1994; Derrington and Badcock, 1985). These findings suggest that the domains of motion and stereopsis apply similar underlying assumptions in the interpretation of transparency. These conclusions support the notion of a non-Fourier channel in stereopsis.

The use of stereoscopic disparity revealed a distinction between multiplicative and additive transparency. Transparent motion seen in contrast modulated stimuli involves the motion of a contrast modulation over a carrier (Derrington and Badcock, 1985; Fleet and Langley, 1994a). Using stereoscopic disparity, it was possible to make either the carrier or contrast modulation appear closer to the observer. Only when the contrast modulation appeared in front of the carrier was transparency observed. No such asymmetry was reported for additive combinations. This difference suggests that two distinct models of transparency are employed by the visual

system.

Given that real-world perceptual transparency will always involve a multiplicative component, one might question why the visual system employs an additive transparency model at all. The answer to this may lie in equation (3). Even when a transparency is produced purely from the product of a luminance pattern and a transmittance pattern, the signal is dominated by the additive terms. This is a consequence of the non-negative nature of luminance and transmittance values. Although examples of transparency will be accompanied by a product component, its magnitude will generally be smaller than the additive components. An additive approximation thus suffices to interpret the transparency.

Plaid motion studies have revealed situations in which coherent or transparent motion are likely to occur. Transparency is less likely when power is localised in frequency space, as when component gratings are similar in orientation and spatial frequency. Similar findings are reported here for stereopsis. In the first experiment, transparency was more likely to be observed from added sinusoidal gratings when the gratings had widely different spatial frequencies. However, in the second experiment, transparency was more readily perceived for contrast envelopes with a spatial frequency of 0.4 cycles/degree. This value may be taken to represent an optimum frequency for the processing of disparity on the basis of a contrast detecting mechanism operating locally in frequency space, subsequent to a stage of linear filtering (Fleet and Langley, 1994a; Langley et al., 1996).

These results taken together suggest that additive and multiplicative transparency occur as the result of different processes. Additive transparency may be perceived when image components occupy markedly different regions of frequency space, and are treated independently. Multiplicative transparency may arise due to processing of the contrast envelope by some non-Fourier process. These conclusions are

in agreement with computational strategies suggested by Wilson, Ferrera and Yo (1992), and support a distinction between Fourier and non-Fourier processes in stereopsis.

5. Asymmetry in the perception of transparency from contrast disparities

1 Introduction

The previous chapter revealed a marked asymmetry in binocular transparency. Contrast modulated horizontal grating stimuli were seen as transparent if the contrast modulations had crossed disparities relative to the carrier grating. The contrast modulations were seen to float transparently in front of the carrier. Transparency was not however observed if the contrast modulations had uncrossed disparities relative to the carrier grating. Rather, the whole pattern was seen to lie behind the fixation plane, in the plane defined by the disparity present in the contrast envelope. This asymmetry was not evident for plaids formed from the additive superposition of a horizontal and a vertical grating. Transparency was observed when the vertical

grating had both crossed and uncrossed disparities, and was seen either in front of or behind the horizontal grating. This distinction may be observed in figures 5.1 and 5.2. These results were taken as evidence that the perception of transparency in contrast modulated stimuli was not the result of distortion products, introduced by an early nonlinearity, activating normal disparity mechanisms (Burton, 1973; Henning, Hertz and Broadbent, 1975). If this were the case, there would be no difference between transparency observed from additive plaid stimuli and contrast modulated grating stimuli, and both would be apparent for disparities of either sign. Rather, it appears that transparencies perceived on the basis of additive and multiplicative signals are the results of distinct stereoscopic processing.

This asymmetry in the perception of transparency appears to be related to disparities defined by nonlinear image properties. Similar asymmetries have been found in other stereoscopic tasks. A sense of depth may be created if a vertical bar is drawn so as to occlude a horizontal bar. Pastore (1974) demonstrated that the sense of occlusion, and the inferred depth ordering, is destroyed if the vertical bar is given uncrossed binocular disparity. Here, occlusion is observed for zero or crossed disparities, but not for uncrossed disparities. Kersten (1991) showed that stimuli in which transparency is observed with an unambiguous depth ordering of surfaces may appear as rivalrous if inconsistent disparities are introduced. Nakayama, Shimojo and Ramachanrdan (1990) reported similar results in relation to neon colour spreading, a phenomenon in which luminance or chromaticity may spread from one image region to another so as to appear as an illusory transparency. Neon colour spreading may be suppressed by binocular disparity in conflict with this transparency. Stereoscopic capture by subjective contours, which are perceived as belonging to an occluding figure, is also only observed for crossed disparities (Ramachandran and Cavanagh, 1985). In these examples, transparency or occlusion act as a qualitative depth cue. When disparity consistent with this cue is introduced, perceived depth relationships are enhanced. When however disparity information conflicts with the transparency

or occlusion, a new solution consistent with all depth cues may be reached (Pastore, 1974; Nakayama, Shimojo and Ramachandran, 1990). Alternatively, the result may be diplopia and binocular rivalry, and the loss of the sensation of depth (Ramachandran and Cavanagh, 1985; Kersten, 1991). In all cases, a percept of depth that is evident with crossed disparities, or when stereoscopic information is absent, is destroyed or altered when uncrossed disparities provide information inconsistent with the existing depth cues.

In this chapter, the asymmetry of transparency resulting from disparity in multiplicatively defined image structures was explored further. Stimuli were related to those used in the previous chapter, and consisted of contrast modulated sinusoidal gratings. Contrast modulations in all cases had uncrossed disparities. As such, transparency would not be expected to be apparent. Figure 5.1 shows examples of the stimuli, with both crossed and uncrossed disparities. Details of all stimuli are given below. Transparency may be observed with crossed, but not with uncrossed disparities. It was observed that if the contrast modulation was replaced by an added luminance pattern, transparency was observed for both crossed and uncrossed disparities. This distinction between additive and multiplicative transparency is predicted by the results of chapter 4. It was also observed that transparency could become apparent if luminance patterns were added to the contrast modulated stimuli. In figure 5.2, luminance patterns have been added to the stimuli of figure 5.1. Transparency may now be observed with both crossed and uncrossed disparity.

The experiments presented here investigated the contrast of luminance patterns it was necessary to add to the contrast modulated stimuli in order to observe transparency. These experiments were used to investigate the extent to which Fourier and non-Fourier stereoscopic channels may be viewed as independent.

2 Methods

2.1 Stimuli

All stimuli consisted of a contrast modulated sinusoidal grating. The carrier grating was either horizontal or vertical, and had a spatial frequency of 4.5 cycles/degree, and a fixed contrast of 0.30. The contrast modulation took one of three forms. The first was an approximation to a vertical square grating, formed from the fundamental and its third and fifth harmonics, given by:

$$S(x,y) = \sin(2\pi f_s x) + \frac{1}{3}\sin(6\pi f_s x) + \frac{1}{5}\sin(10\pi f_s x) \quad (1)$$

where f_s represents the frequency of the square wave. The frequency used was $f_s = 0.5$ cycles/degree. This stimulus is directly related to that used in the previous experiments.

The second stimulus was a vertically oriented Gabor patch, given by:

$$G(x,y) = \exp\left(-\frac{(x^2+y^2)}{2\sigma^2}\right)\sin(2\pi f_g x) \quad (2)$$

where f_g represents the spatial frequency of the sinusoidal modulation, and σ^2 the standard deviation of the Gaussian envelope. A spatial frequency of $f_g = 0.5$ cycles/degree and a standard deviation of $\sigma^2 = 5$ degrees were used.

The final stimulus used was a central square patch, which had a width of 4.6 degrees.

The contrast modulation had binocular disparity taking the form either of a full field binocular disparity of 20 minutes of arc, or of a vertical gradient of disparity

with a magnitude of 0.14. In both cases, horizontal disparities were all uncrossed, so as to be consistent either with a frontoparallel plane behind the plane of fixation, or a plane with a horizontal axis of slant, again slanting behind the plane of fixation. When contrast disparities represented a frontoparallel plane, a horizontal carrier grating was used. When disparities represented a slanting surface, a vertical carrier grating was used. For the frontoparallel surface condition, the square wave and square modulations were employed. For the slanting surface condition, all three contrast modulation patterns were used. Stimuli were presented through a hard circular window, with a diameter of 8.6 degrees.

The stimuli are related to those in the multiplicative transparency experiments in the previous chapter which failed to exhibit transparency. In this experiment, luminance patterns were added to the stimuli in an attempt to observe transparency. Luminance patterns took the same form as the contrast modulation, and in all cases were added in phase with the modulation: luminance was subtracted from points where contrast was reduced, and added to the high contrast image regions. Between trials, the contrast of the added luminance components was varied. The mean level of illumination was held constant at 37.8cdm^{-2} across all experimental conditions. Between blocks of trials, the contrast modulation depth was varied in the range 0.1-0.9, for all combinations of contrast modulations and binocular transformations.

The stimuli used may be described by the following equations:

$$I_l(x,y) = \alpha C(x,y)\,[1 + \beta M_l(x,y)] + \gamma M_l(x,y) \tag{3}$$

$$I_r(x,y) = \alpha C(x,y)\,[1 + \beta M_r(x,y)] + \gamma M_r(x,y) \tag{4}$$

where $C(x,y)$ represents the sinusoidal carrier grating, and $M_{l,r}(x,y)$ the contrast modulation. α, β and γ give the contrast of the carrier, the contrast modulation

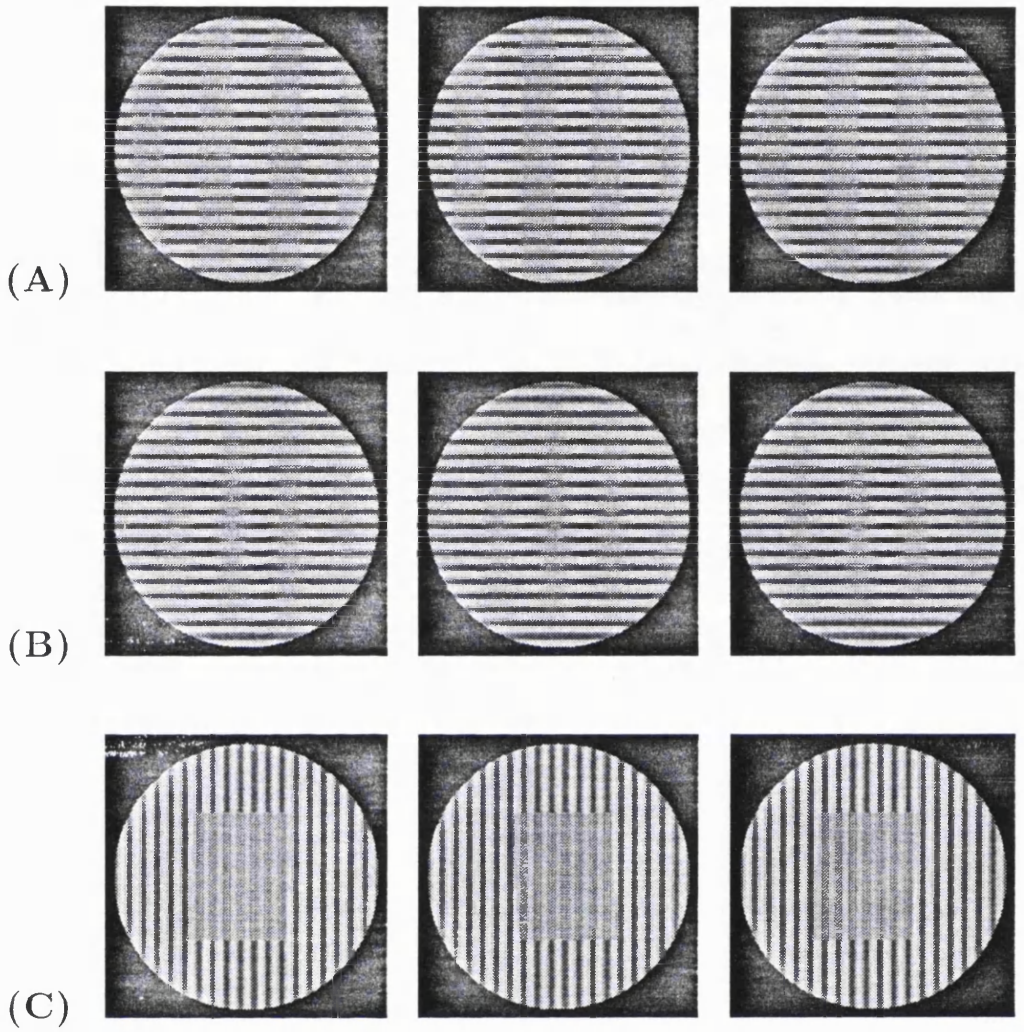


Figure 5.1: *Examples of the stimuli used. Shown here are contrast modulated grating stimuli, with no additional Fourier energy. The three stimuli show the three contrast modulation patterns used.* **(A):** *Square-wave modulation.* **(B):** *Gabor modulation.* **(C):** *Square modulation. In all cases, cross-eyed fusion of the left two images should reveal the contrast modulation floating transparently in front of the carrier grating. In* **(A)** *and* **(B)**, *cross-eyed fusion of the right two images should result in the perception of a single surface seen behind the plane of the paper. In* **(C)**, *a vertical carrier is modulated by a binocularly sheared square, and appears as rivalrous, with no sensation of depth. See text for discussion.*

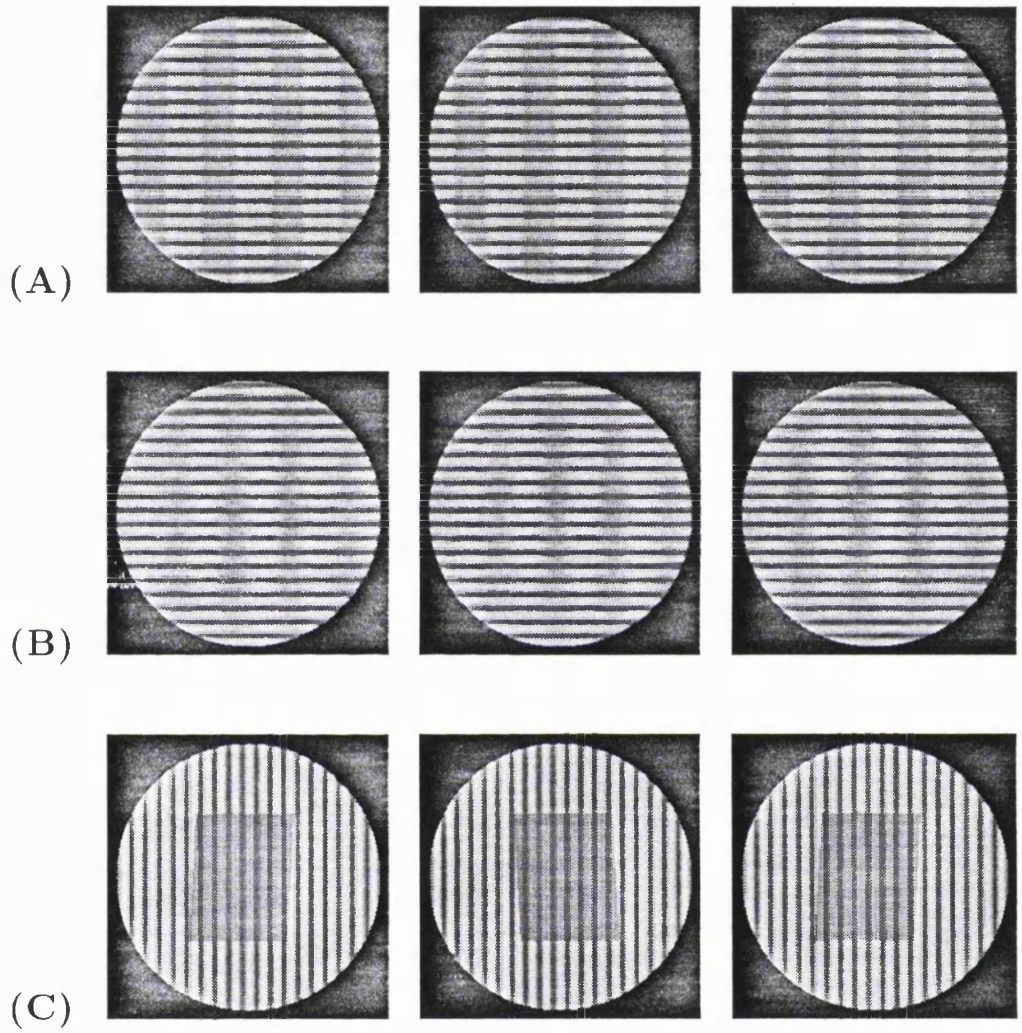


Figure 5.2: *These stimuli are identical to those shown in figure 5.1, except that luminance has been added in phase with the contrast modulation. Cross-eyed fusion of either the left- or right- image pair should appear as a transparency, with the contrast modulation pattern appearing in front of or behind the carrier grating respectively.*

depth, and the contrast of the added luminance pattern, respectively. $I_l(x,y)$ and $I_r(x,y)$ refer to the left and right binocular images, formed from the contrast modulation patterns $M_l(x,y)$ and $M_r(x,y)$. $M_l(x,y)$ and $M_r(x,y)$ were in all cases related by the binocular translation or horizontal shear described above. Examples of these stimuli, both contrast modulated and with additional luminance components, are presented in figures 5.1 and 5.2.

2.2 Procedure

For each trial, subjects were presented with a stimulus, which remained visible until a response was made. Subjects were asked to decide whether a transparency was observed, on the grounds of whether two surfaces were seen, transparently and clearly separated in depth.

3 Results

Figure 5.3 shows a typical psychometric function. For low contrasts of added luminance patterns, transparency was not reported, as would be expected from the results of chapter 4. For higher contrasts, transparency was reported. These results may be interpreted as an additive form of transparency. Figures 5.4 and 5.5 present the 50% points of measured psychometric functions for all conditions and subjects, as a function of contrast modulation depth. The contrast of the added luminance pattern required to evoke a percept of transparency ranged between 0.10 and 0.30, but appeared independent of the depth of modulation. Again, this result is suggestive of the transparency arising from an additive signal combination, occurring independently of the multiplicative contrast modulation depth.

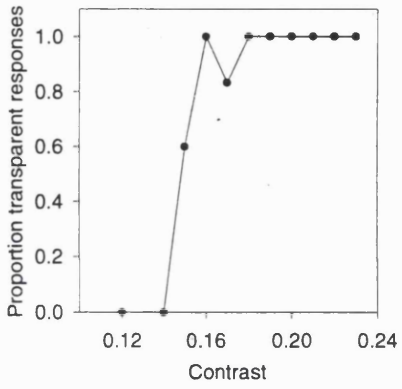


Figure 5.3: *A typical response function. For low contrasts of added luminance patterns, transparency was not reported. As the contrast of the added luminance was increased, transparency was observed. This function is for subject JB, with a square grating stimulus, a horizontal carrier, and a contrast modulation depth of 0.6.*

4 Conclusions

The previous chapter reported that transparency was apparent given crossed, but not uncrossed disparities in an image contrast envelope. Here it was demonstrated that, by incorporating an additional luminance pattern, in phase with the contrast modulation, transparency could be evoked even given uncrossed disparities. This additional luminance added power to the image at the orientation and spatial frequency of the contrast modulation. The contrast required for transparency to be observed was found to be independent of the contrast modulation depth. These results suggest that the perception of transparency from additive and multiplicative signal combinations are the result of separate, independent processes. The asymmetry with respect to sign of contrast disparity was not evident for additively produced

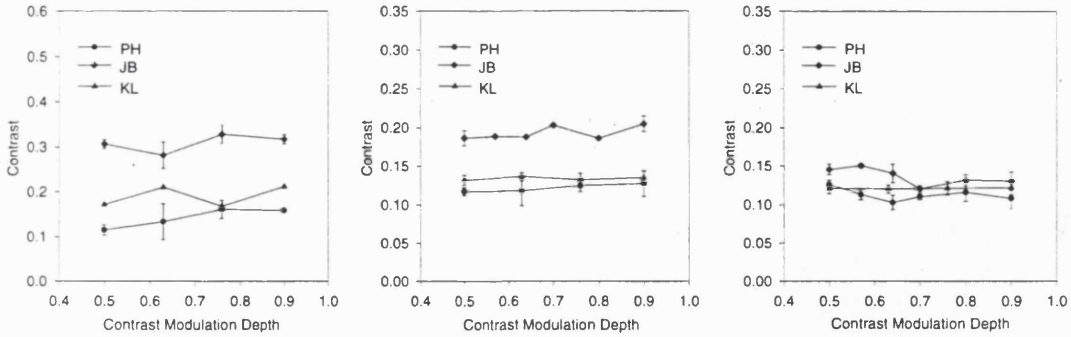


Figure 5.4: *The results presented here represent 50% points of psychometric functions such as that shown in figure 5.3, for the situations in which the transparency had a uniform uncrossed disparity. 50% points are plotted as a function of modulation depth. Results are presented independently for the three different stimulus types.* **(A)**: *Square-wave.* **(B)**: *Gabor.* **(C)**: *Square.*

patterns. Hence, additive and multiplicative transparency are qualitatively different. This would appear to rule out the possibility that transparency, and hence depth, in the non-Fourier stimuli resulted from an early non-linearity introducing additional Fourier components, which were then processed by normal disparity detecting mechanisms. This is not to say that such early non-linearities, introducing additional Fourier components (distortion products), do not exist. Adaptation and masking studies have suggested that distortion products are introduced by early non-linearities (Burton, 1973; Henning, Hertz and Broadbent, 1975). The contrast of distortion products introduced by nonlinearity has been quantified at around a few percent (Henning et al., 1975). It would appear that this contrast is not sufficient to support transparency in the stimuli used here. The contrast required to evoke transparency in the current experiment was around 0.20, which is significantly

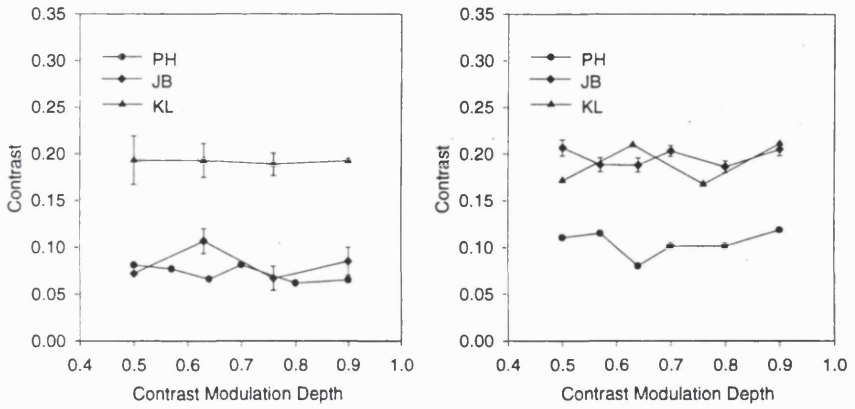


Figure 5.5: *Results presented here correspond to the cases in which the transparency was slanted behind the fixation plane.* **Left:** *Square-wave.* **Right:** *Square.*

greater than the predicted contrast of distortion products. Further, the contrast of the added luminance pattern required to evoke transparency was independent of the contrast modulation depth of the stimuli. The contrast of any distortion products will depend on the depth of contrast modulation, and would be expected to sum linearly with the experimentally introduced luminance patterns (Henning et. al., 1975). That contrast modulation depth did not appear to affect the added contrast required to observe transparency suggests that the contrast of any distortion products was relatively small. Luminance patterns were always added in phase with the contrast modulation. Transparency would also be predicted to be observed if the luminance were added out of phase. This would reverse the contrast of the added distortion products, and would be expected to alter the contrast required to evoke transparency. Henning et al. (1975) found that masking effects were greater for a grating out of phase with the expected distortion product (i.e. greater masking wa

observed when the grating was added such that the highest luminance values of the grating coincided with the regions of highest contrast in the image). This would lead us to predict that the added contrast required would be less if the luminance were added out of phase than if it were added in phase. However, as the modulation depth did not affect the required contrast, the difference would be expected to be negligible.

The contrast at which transparency was observed here was significantly greater than that found to support transparency in the additive transparency experiments of chapter 4, where it was found that increasing the contrast of the vertical grating diminished the tendency to observe transparency. In the current experiments, the smallest difference in spatial frequency between the carrier grating and the added pattern, occuring in the vertical squarewave condition, involved a difference in frequency of more than 3 octaves. Given this separation in frequency, one would expect very little of the interaction between channels necessary for the integration of disparity signals.

These results suggest that contrast disparities are processed independently to luminance disparities. As such, they support a distinction between Fourier and non-Fourier channels in stereopsis. Central to models of non-Fourier processing is the notion of a nonlinearity, making explicit the contrast envelope of a signal. An important question is the stage at which this nonlinearity occurs. Chubb and Sperling (1988) suggested that a full-wave rectification occurs prior to bandpass filtering. Models have also been suggested in which nonlinearities occur after a stage of oriented, bandpass filtering (Wilson, Ferrera and Yo, 1992; Zhou and Baker, 1993; Fleet and Langley, 1994a). This distinction may be related to whether the nonlinearity occurs before or after orientation and spatial frequency selective filtering in the primary visual cortex. (Movshon, Thompson and Tolhurst, 1978). The results presented in this and the previous chapter are consistent with models involving

either a late non-linearity, or an early nonlinearity occuring in an independent non-Fourier channel. An alternative explanation would be that the perception of contrast variations relies on nonlinearities occuring as early as the LGN, as have been found in the responses of cat X-cells (Derrington, 1987). This position is not supported by the results presented here.

Finally, it is interesting to describe the appearance of the stimuli when transparency was not observed. When disparity took the form of a uniform binocular translation, capture of the carrier grating was observed, and the whole pattern was seen to lie in the depth plane defined by the contrast disparity. This effect may be observed in the stimuli in the top two rows of figure 5.1. When the disparity defined a slanting surface, and the carrier grating was horizontal, capture of the entire plane was again observed. This time, the whole pattern was seen to slant about a horizontal axis. When the carrier was vertical, however, rivalry and diplopia were reported, with neither transparency nor slant apparent. The contrast modulation appeared as rivalrous, as may be observed in the bottom row of figure 5.1. Classical stereoscopic capture effects involving subjective contours are reported only for crossed disparities, uncrossed disparities resulting in rivalry and an absence of depth from disparity. Additionally, capture of lines or texture elements lying outside the capturing contours is not observed (Ramachandran and Cavanagh, 1985). Vallortigara and Bressan (1994) interpreted stereoscopic capture as a solution to conflicting depth cues. Subjective contours give the impression of occlusion, which is not consistent with the contours having uncrossed disparities relative to the captured texture. By manipulating the depth relations suggested by occlusion cues, Vallortigara and Bressan (1994) demonstrated that capture of elements lying outside of the subjective figures can occur. In the current experiments, transparency suggested the opposite depth relationships between the carrier grating and the transparent figure—the carrier grating must lie behind the plane defined by the transparency. No additional occlusion cues were in conflict with this interpretation of the surface relationships.

The capture effects described here support the arguments of Vallortigara and Bressan (1994) that capture reflects a solution to potentially conflicting depth cues, and as such is intimately related to transparency and occlusion. The only situations in which neither transparency nor capture were observed involved a vertical grating and a slanted transparent figure. Here there is conflict between the slants determined by the transparent figure and the vertical grating, and it is not possible for any capture solution to integrate these cues. In this case, rivalry and diplopia occurred as a result of a failure to integrate inconsistent depth cues (Ramachandran and Cavanagh, 1985).

6. The site of nonlinearity in contrast envelope processing

1 Introduction

The results of chapters 4 and 5 demonstrated that depth may be observed from disparities in the contrast envelope of an image. Burton (1973) proposed that contrast beats in plaid stimuli are perceived from distortion products, introduced by nonlinearities acting prior to orientation and frequency selective processing. These distortion products would be detected by the same mechanism that detects luminance gratings. However, Derrington and Badcock (1985) provided evidence that image contrast variations are *not* perceived as the result of an early nonlinearity.

Hess and Wilcox (1994; Wilcox and Hess, 1996) suggested that contrast disparities are processed by an independent, nonlinear channel in stereopsis. This notion is related to models of motion processing incorporating separate Fourier and non-Fourier channels (Chubb and Sperling, 1988; Wilson, Ferrera and Yo, 1992; Zhou

and Baker, 1993; Fleet and Langley, 1994a). The two channel model proposed by Chubb and Sperling (1988) involves one channel processing luminance information, and another, independent channel detecting motion after broadband filtering and fullwave rectification. It is the second channel that processes contrast modulations. Wilson et al. (1992), Zhou and Baker (1993) and Fleet and Langley (1994a) proposed models incorporating a late nonlinearity, occurring after an initial stage of orientation- and spatial-frequency specific filtering. In the models of Wilson et al. (1992) and Zhou and Baker (1993), rectification of the outputs of bandpass filters is followed by a further stage of orientation- and frequency- selective filtering. As an alternative, Fleet and Langley (1994a) proposed that contrast beats can be detected via the spatial gradient of energy. This last model does not therefore involve a second stage of oriented filtering.

These models may be classified in terms of the stage of significant nonlinearities. All the above models involve a nonlinearity preceded by a stage of filtering. This initial filtering stage is broadband in the model of Chubb and Sperling (1988), but bandpass in the models of Wilson et al. (1992), Zhou and Baker (1993) and Fleet and Langley (1994a). Whether nonlinearities occur before or after orientation and spatial frequency specific filtering is an important issue, since orientation and spatial frequency specificity are first evident in the striate cortex (Movshon, Thompson and Tolhurst, 1978). The various models may thus be distinguished on the grounds of whether significant nonlinearities occur cortically or precortically. In addition, Wilcox and Hess (1996) have presented evidence that nonlinearities occur before the binocular integration of information.

The results of chapter 5 provide support for the notion that luminance and contrast disparities are processed by distinct mechanisms. Transparency was found to be asymmetric with respect to disparity for contrast modulated, but not for additive stimuli, indicating that the processing of contrast disparities does not rely on dis-

tortion products. Rather, it would appear that the two forms of transparency result independently from distinct Fourier and non-Fourier processing.

This chapter addresses the question of the site of significant nonlinearities in disparity processing, using an adaptation paradigm. Blakemore and Campbell (1969) showed that prolonged presentation of a high-contrast sinewave grating increases the minimum contrast required to detect gratings of a similar orientation and spatial frequency. The elevation in contrast thresholds was found to decrease as orientation and spatial frequency differences between adapting and test stimuli increased. In addition to increasing contrast detection thresholds, adaptation also reduces the perceived contrast of suprathreshold stimuli (Greenlee and Heitger, 1988). As stereoacuity is markedly contrast dependent (Legge and Gu, 1980), it may be expected that adaptation to a grating at the correct disparity should affect disparity detectability. Specifically, adaptation would be expected to increase the minimum contrast at which a given disparity could be discriminated. Masking studies have shown stereopsis to be tuned to both spatial frequency (Julesz and Miller, 1975; Yang and Blake, 1991) and orientation (Mansfield and Parker, 1994). Any adaptation effects occurring in stereopsis would be expected to be similarly orientation and spatial frequency specific.

The experiments presented here were designed to test whether the detection of disparities in the contrast envelope of a grating stimulus is preceded by a stage of orientation and spatial frequency specific filtering. Further, the experiments were intended to assess the orientation and frequency tuning of any such filtering stage.

It is anticipated that, if any such tuning is demonstrated, it will reflect the nature of the early stages of contrast envelope disparity processing. processing. If such processing occurs after a stage of oriented bandpass filtering, adaptation to a grating with the orientation and spatial frequency of the carrier of a contrast modulation

should affect the contrast at which disparity in the contrast modulation can be detected. For a contrast modulated sinusoidal grating, it might also be predicted that the greatest contrast threshold elevation might be obtained by adaptation to a grating with the orientation and spatial frequency of Fourier components of the stimulus other than the carrier. With a horizontal grating, for example, disparity information is carried by components other than the carrier. Adaptation to gratings with orientations and spatial frequencies lying in the sidebands of power this stimulus might be expected to have the greatest increase in the contrast required to discriminate depth. If, however, disparities in contrast envelopes are perceived on the basis of a nonlinearity introducing Fourier power prior to a stage of bandpass filtering, adaptation to a grating with the orientation and spatial frequency of either the carrier, or lying in the sidebands of power, would have relatively little effect on ability to detect contrast envelope disparities. Rather, adaptation to a grating with the orientation and spatial frequency of the contrast envelope itself might be expected to have the greatest effect. This chapter assessed the effects of adaptation to gratings of varying spatial frequencies and orientations on a stereoscopic task involving the detection of disparity in a contrast envelope. The orientation and spatial frequency of the adapting grating were varied relative to both the carrier and modulation gratings. The results were analysed in terms of whether significant nonlinearities occur prior or subsequent to orientation- and spatial frequency-tuned filtering.

2 Methods

2.1 Stimuli

Test stimuli consisted of contrast modulated horizontal gratings. The contrast modulation used was an approximation to a vertical grating, formed from a fundamental plus its third and fifth harmonics, as was used in chapters 4 and 5. Figure 6.1 demonstrates the Fourier transform of the image. Two different carrier gratings were used, with spatial frequencies of 2.0 cycles/degree and 4.0 cycles/degree. The beat had a spatial frequency of 0.45 cycles/degree. The modulation depth was 1.0; the contrast of the carrier grating was varied between trials. The contrast beat had a binocular disparity of 20 minutes; this was randomly crossed or uncrossed, and the subjects' task was to discriminate the sign of the disparity.

In the adaptation task, subjects were presented with an adapting sinusoidal grating prior to the test stimuli, and again between presentations of the test stimuli. In some conditions, the adapting grating was horizontal, and thus carried no disparity information. In all other conditions, the adapting grating was presented with zero disparity. The adapting grating had a contrast of 98%, and was counterphase flickered at a rate of 4Hz to avoid phase dependent after effects (Georgeson, 1987). Both the adapting grating, and the test stimuli, were presented in a hard circular window, with a diameter of 7.9 degrees. The experiment consisted of three sets of sessions:

- In the first set of sessions, the frequency of the adapting grating was equal to the carrier frequency of the test stimulus, and the angle between the adapting grating and carrier grating was varied. This allowed for the examination of the importance of the orientation of the adapting grating, for frequencies close

to the carrier. Angles of 0, 10, 25, 45, and 90 deg were used. The fundamental frequency of the contrast envelope was fixed at 0.45 cycles/degree, while the frequencies of the carrier and adapting grating were fixed at 4.0 cycles/ degree.

- In the second set of sessions, the orientations of the carrier and adapting gratings were identical, and the spatial frequency of the adapting grating was varied. Two different carrier frequencies (2.0 cycles/degree and 4.0 cycles/degree) were used, to assess the spatial frequency tuning of the adaptation. The spatial frequency of the adapting grating differed from that of the carrier by factors of 0.5, 1.0, 1.414 and 2.0. As a fifth point, the frequency of the adapting grating was equal to the fundamental frequency of the envelope.

- In the third set of sessions, the frequency of the adapting grating was equal to the fundamental frequency of the envelope, while the angle between the adapting grating and the carrier was varied. Angles of 0, 45 and 90 deg were used.

2.2 Procedure

2.2.1 Baseline task

The baseline task, without adaptation, required subjects to discriminate the disparity of a contrast modulated pattern. On each trial, subjects were presented with a test stimulus for 500ms. The subjects' task was to decide whether the contrast beats appeared in front of or behind the monitor. In the multiplicative transparency experiments of chapter 4, it was found that a contrast modulation with a crossed disparity relative to its carrier would appear to lie transparently in front of the carrier. With uncrossed disparities, however, the whole stimulus was seen to lie in the plane defined by the disparity of the contrast envelope. It would therefore

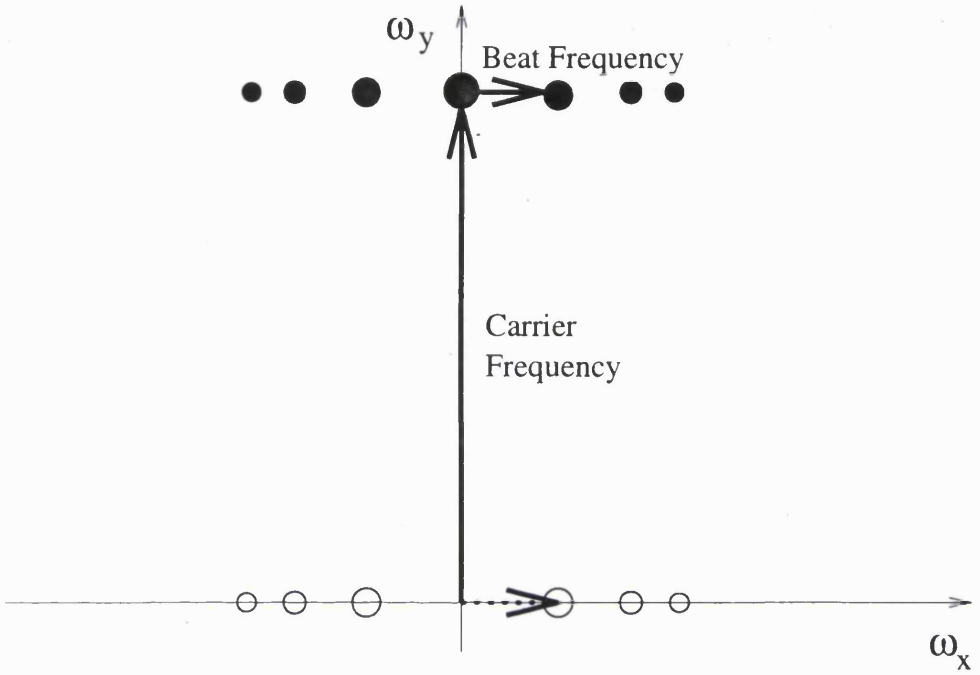


Figure 6.1: *Representation of the Fourier spectrum of the stimulus. The stimulus has four non-zero Fourier components, denoted by the black circles. These are determined by the carrier frequency, the beat frequency, and the two harmonics of the beat. The horizontal carrier is located along the ω_y -axis, as denoted by the solid vector passing through the origin. The length and direction of the vector give the spatial frequency and orientation of the carrier, respectively. The beat spatial frequency and orientation are given in a similar way by the horizontal vector from the carrier to the component corresponding to the fundamental frequency of the beat. The empty circles, and the dotted vector, show the locations of power that would be introduced by an early nonlinearity.*

have been possible for subjects to perform the experimental task described above on the grounds of whether of not a transparency was observed. However, as the depth discrimination and transparency detection tasks rely on the same disparity information, this potential alternative strategy would not be expected to influence the results obtained. Between experimental trials, the contrast of the stimulus was varied. 10 fixed contrast levels were used. Each was presented 16 times, in random order. Two sessions were run, each with 8 repetitions of each contrast level.

2.2.2 Adaptation Task

For the adaptation task, each block of trials was preceded by presentation of an adapting grating for 2 minutes. Subjects were then presented with a test pattern, identical to those used in the baseline task, and again asked to determine whether the contrast beats appeared in front of or behind the monitor. Subsequent trials were preceded by a "top-up" adaptation period of 6s. Subjects were required to respond during this top-up period. Again, 10 levels of contrast of the test pattern were used, and each was presented 16 times.

2.3 Subjects

The two experimenters plus one other volunteer acted as subjects in this experiment. All subjects had normal vision; the experimenters had acted as subjects in experiments presented in previous chapters. The third subject had not participated in other experiments, and was naive to the purposes and aims of the experiment.

3 Results

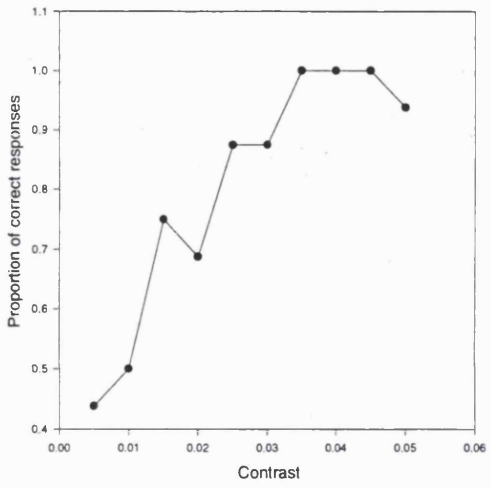
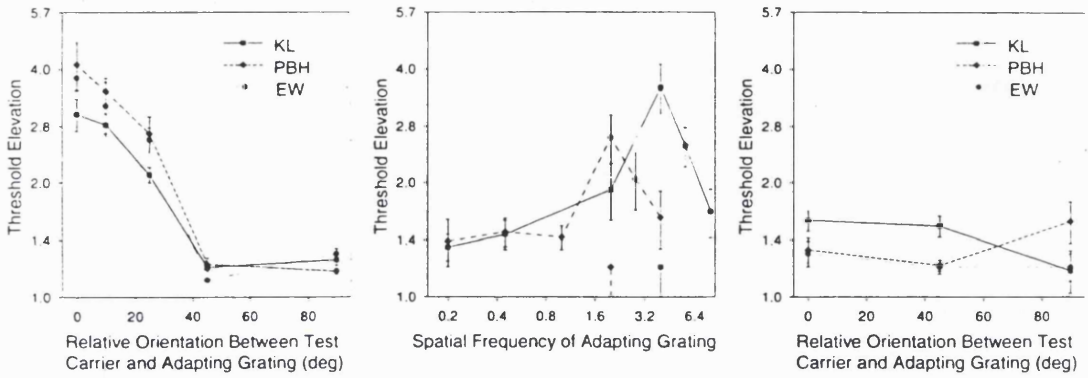


Figure 6.2: *Response function, for subject KL, for the baseline task. Subjects' ability to perform the disparity detection task was affected by stimulus contrast.*

Figure 6.2 shows an example response function, derived from the baseline task. Subjects were consistently able to determine whether the beat was in front of or behind the monitor at high contrasts; performance fell to chance at low contrasts. For each session, a logistic function ranging between 50% and 100% was fit to the data. This fitted curve was used to estimate the contrast at which subjects responded correctly on 75% of trials. This procedure was used to determine the baseline contrast threshold for each experimental session, in the absence of any adaptation. For each set of data collected in the adaptation condition, the estimated contrast threshold was divided by this baseline threshold, so as to determine the threshold elevation resulting from adaptation. Figure 6.3 shows threshold elevations as a function of the relative orientation and spatial frequency of the adapting grating. Also shown are results for the condition in which the adapting grating had the same spatial frequency as the beat. Threshold elevations were greatest when both the orientation and spatial frequency of the adapting grating and test carrier were equal. As the spatial



(A) (B) (C)

*Figure 6.3: (**A**) Mean threshold elevations, on log-log axes, as a function of the angle between the carrier and the adapting grating. Threshold elevations were maximal when the orientations were identical. (**B**): Means of the threshold elevations for the three subjects, as a function of the spatial frequency of the adapting grating. The two curves represent carrier frequencies of 2.0 and 4.0 cycles/degree, which are marked on the horiztonal axis. Threshold elevations were maximal when the frequency of the adapting grating matched the carrier. (**C**): Mean threshold elevations when subjects adapted to a grating with the frequency of the beat. Elevation is markedly lowe than when subjects adapted to a grating with the frequency of the carrier. Results are plotted against the angle between the adapting grating and the carrier.*

frequency or orientation of the adapting grating was varied relative to the carrier, threshold elevation reduced. Relatively low levels of threshold elevation were found when subjects were adapted to the beat frequency. It has been argued that there are no adaptable channels tuned to below 1.5 cycles/degree, and that adaptation to frequencies below this will produce maximum threshold elevations for gratings with a frequency of 1.5 cycles/degree (e.g. Blakemore and Campbell, 1969). This could explain the failure to find significant threshold elevation subsequent to adaptation to the beat frequency here. However, Stromeyer, Klein, Dawson and Spillman (1982) found adaptation and masking effects which were centered on spatial frequencies as low as 0.2 cycles/degree, so adaptation of channels tuned to 0.45 cycles/degree could have been possible. Another reason why strong contrast elevation was not evident could be that the adapting grating was always presented with zero disparity, whereas the test stimulus had a (crossed or uncrossed) disparity of 20 minutes. As the effects of adaptation are disparity specific (Blakemore and Hague, 1972), this difference in disparity may have reduced any potential adaptation.

4 Conclusions

The adaptation effects reported here showed marked orientation- and spatial frequency specific tuning. Threshold elevation was greatest when the adapting grating had the same orientation and spatial frequency as the test carrier grating. Strong threshold elevation was not obtained subsequent to adpatation to a grating with the spatial frequency and orientation of the contrast envelope. This may be related to the low spatial frequency of the envelope, or the difference in disparity between the adapting and test stimuli. However, the threshold elevation produced by adapting to the carrier of the stimulus is inconsistent with the notion that the visual system relies upon early nonlinearities to detect disparities in the contrast beat. The data support the

idea that disparity sensitive processing of contrast beats occurs after orientation- and spatial frequency- selective filtering, with significant nonlinearities occurring in the visual cortex.

In studies of stereoscopic masking, masking effects have been found to be greatest when test and mask have the same orientation (Mansfield and Parker, 1994). The orientation bandwidth of these masking effects, which is around 25°, is similar to that found for orientation discrimination (Campbell and Robson, 1966). While masking is also spatial frequency tuned, the greatest effects may occur for spatial frequencies other than that of the test stimulus (Yang and Blake, 1991). Masking effects may be characterised by two fairly broad channels, tuned to 3 and 5 cycles/degree. These channels are most likely to be the result of pooling across more narrowly tuned disparity detectors, of the type found in physiological studies (Ohzawa and Freeman, 1986). The results reported here are likely to be the result of adaptation prior to the stage of pooling of disparity detector responses.

The half-width at half-height of the adaptation effects was found to be 25° in orientation, and 1 octave in spatial frequency. These findings are similar to results reported for the detection of sinusoidal gratings subsequent to adaptation, for gratings presented both monocularly, and with binocular disparity. (Blakemore and Campbell, 1969; Blakemore and Hague, 1972). Similar results were also found in a related experiment, involving a two dimensional orientation discrimination task (Langley, Fleet and Hibbard, 1996). One marked difference between the results reported for the two experiments is that threshold elevations were approximately a factor of 10 higher for the spatial orientation discrimination task. Again, this may be related to the findings reported by Blakemore and Hague (1972), that adaptation effects were disparity specific. It was not possible to present the adapting grating at the same disparity as the test stimuli, since the subjects' task was to discriminate the test stimulus disparity. Consequently, the adapting grating was always presented at zero

disparity, and as such had a disparity relative to all test stimuli of 20 minutes. This disparity would be expected to produce reduced adaptation effects in comparison to those found for the two dimensional orientation discrimination task. The reduced contrast threshold elevations reported here may also be influenced by the relatively high baseline contrast thresholds. As a result, the difference in contrast between the adapting grating and test stimuli used here was considerably lower than it would have been in the two-dimensional discrimination task. This reduction in the difference in contrast between adaptation and test stimuli would be expected to produce correspondingly smaller adaptation effects (Greenlee and Heitger, 1988).

In conclusion, these data support the notion that early nonlinearities do not contribute significantly to binocular depth from disparities in image contrast envelopes, at least of the type used in this study. Rather, it would appear that significant nonlinearities are apparent subsequent to a stage of orientation- and spatial frequency-selective filtering. For the stimuli used here, it would appear that processing of contrast disparities is preceded by a stage of linear filtering tuned to horizontally oriented stimuli. These results therefore support the notion that stereoscopic processing may make proceed on the basis of the responses of horizontally tuned mechanisms (Fleet and Langley, 1994b). These results are similar to those reported in related experiments involving the discrimination of the two dimensional orientation of contrast beats (Langley et al., 1996). As such, they suggest that similar mechanisms are employed in the processing of spatial and stereoscopic information manifest in image contrast variations.

7. A model of slant perception based on differences in instantaneous frequency

1 Introduction

Computational models of stereopsis have been suggested to explain how the correspondence problem is solved, and how binocular disparities in images may be represented (e.g. Marr and Poggio, 1979; Grimson, 1980; Mayhew and Frisby, 1980). While some models have considered both horizontal and vertical disparities (e.g. Langley, Atherton, Wilson and Larcombe, 1990), most have considered only horizontal disparities. This is a result of the preponderance of horizontal disparities in stereoscopic viewing, and their importance as a source of depth information. By estimating horizontal disparity for all points in an image, a disparity map may be built up, which can provide important depth information. However, properties of

object and surface shape, such as slant and curvature, are apparent only implicitly in this representation. In this chapter, a computational model is presented which demonstrates how such a direct estimation of surface orientation may be achieved.

The notion that stereopsis in concerned primarily with the representation of horizontal disparities has led physiological researchers to consider the responses of binocular neurons only to horizontal disparities (e.g. Poggio and Fischer, 1977). However, Freeman and Ohzawa (1990) have criticised these studies for assuming that disparity sensitivity plays a functional role in the analysis of depth from horizontal disparities. Other studies (Ohzawa, DeAngelis and Freeman, 1990; DeAngelis, Ohzawa and Freeman, 1991, 1995) have addressed this issue by independently considering the two dimensional monocular receptive fields of binocular neurons, and relating disparity tuning to differences in the position and shape of left and right monocular receptive fields.

Physiological studies have also considered the monocular orientation tuning of binocular neurons. Specifically, it has been reported that left and right monocular receptive fields of binocular neurons may show differences in preferred orientation (Blakemore, Fiorentini and Maffei, 1972; Nelson, Kato and Bishop, 1977). Cells falling into this category have been shown to respond to orientation disparity (Nelson et al., 1977; Hänny, von der Heydt and Poggio, 1980). It has been suggested that orientation disparities may play a role in the representation of surface orientation (e.g. Koenderink and van Doorn, 1976; von der Heydt, Hänny and Dürsteller, 1981; Rogers and Graham, 1983). The orientation disparity tuning found in physiological studies may support this computational strategy.

Computational models of stereopsis have been proposed which make use of orientation disparities. Wildes (1991) presented a model in which surface orientation is estimated from angular disparities in pairs of image edges. This model is related

to the geometrical analysis of Koenderink and van Doorn (1976). However, the model employed a previously obtained edge orientation map, so did not address the problem of how orientation may be represented. Jones and Malik (1992) proposed a model in which surface orientation is estimated from orientation and spatial frequency differences in the outputs of bandpass filters. This model is able to infer inclination (slant around a horizontal axis) from stimuli which do not contain systematic positional correspondences between image points (such as those used by von der Heydt et al., (1981)).

The disparity field associated with a planar surface may be approximated as an affine transformation between binocular image pairs. The model presented in this chapter estimates this transformation, using locally obtained image measurements. This estimation is achieved without first representing a positional disparity map, and is intended as a first stage in the estimation of surface orientation. This model is related to models of stereoscopic processing based on the analysis of phase disparities (e.g. Sanger, 1988; Langley et al., 1990).

In section 2, *The affine model of disparity*, the binocular disparity field is related to image spatial gradients. Section 3, *Responses of bandpass filters*, describes how this relationship may be exploited to provide estimates of the affine parameters from the responses of bandpass filters. The derived model is implemented in section 4, *Simulations*. Finally, section 5, *Conclusions*, discusses the role of this model in furthering an understanding of the representation of stereoscopic slant and transparency, and suggests how it may be extended to relate more directly to both psychophysical and physiological evidence.

2 The affine model of the disparity field

With the exception of disparity discontinuities occurring at surface boundaries, both horizontal and vertical disparities will typically vary smoothly with space, and may be represented by the continuous functions $d_h(x,y)$ and $d_v(x,y)$, respectively. These functions may be represented in the neighbourhood of the point (x_0, y_0) using the Taylor series expansion:

$$d_h(x,y) = d_h(x_0,y_0) + \frac{\partial}{\partial x}(d_h(x_0,y_0))x + \frac{\partial}{\partial y}(d_h(x_0,y_0))y + \ldots \quad (1)$$

$$d_v(x,y) = d_v(x_0,y_0) + \frac{\partial}{\partial x}(d_v(x_0,y_0))x + \frac{\partial}{\partial y}(d_v(x_0,y_0))y + \ldots \quad (2)$$

which may be written as a first order approximation:

$$\begin{bmatrix} d_h \\ d_v \end{bmatrix} \approx \begin{bmatrix} h_0 \\ v_0 \end{bmatrix} + \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (3)$$

where $h_0 = d_h(x_0,y_0)$, $v_0 = d_v(x_0,y_0)$, $a = \frac{\partial}{\partial x}d_h$, $b = \frac{\partial}{\partial y}d_h$, $c = \frac{\partial}{\partial x}d_v$, and $d = \frac{\partial}{\partial y}d_v$. Equation (3) represents an affine transformation between left and right images, and gives a close approximation to the disparity field for planar surfaces, for which second and higher order terms in d_h and d_v are small.

The model presented in this chapter estimates this affine transformation for stereoscopic image pairs. Let the luminance of a point (x,y) in the left- and right-images be $E_l(x,y)$ and $E_r(x,y)$, respectively. $E_l(x,y)$ and $E_r(x,y)$ are related by the horizontal and vertical disparities, $d_h(x,y)$ and $d_v(x,y)$, such that:

$$E_r(x,y) = E_l(x + d_h(x,y), y + d_v(x,y)) \quad (4)$$

Expanding the right hand side of equation (4),

about the point (x,y), gives:

$$E_r(x,y) = E_l(x,y) + d_h(x,y)\frac{\partial E_l}{\partial x} + d_v(x,y)\frac{\partial E_l}{\partial y} + \ldots \tag{5}$$

Subtracting $E_r(x,y)$ from both sides, and ignoring second and higher order terms in d_h and d_y, gives:

$$E_x d_h + E_y d_y + (E_l - E_r) = 0 \tag{6}$$

where, $E_x = \frac{\partial E_l}{\partial x}$ and $E_y = \frac{\partial E_l}{\partial y}$ give the partial derivatives of the image. An affine model may be derived from (6) using the two-dimensional, spatial gradient operator, $\nabla(\psi) = (\frac{\partial \psi}{\partial x}\hat{\imath} + \frac{\partial \psi}{\partial y}\hat{\jmath})$:

$$\begin{aligned} \left(E_{xx}d_h + E_x\frac{\partial d_h}{\partial x} + E_{xy}d_v + E_y\frac{\partial d_v}{\partial x} + \Delta E_x\right)\hat{\imath} \; + \\ \left(E_{xy}d_h + E_x\frac{\partial d_h}{\partial y} + E_{yy}d_v + E_y\frac{\partial d_v}{\partial y} + \Delta E_y\right)\hat{\jmath} \;\; = \;\; \mathbf{0} \end{aligned} \tag{7}$$

$$\text{i.e.} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} E_x \\ E_y \end{bmatrix} + \begin{bmatrix} E_{xx} & E_{xy} \\ E_{xy} & E_{yy} \end{bmatrix} \begin{bmatrix} h_0 \\ v_0 \end{bmatrix} + \begin{bmatrix} \Delta E_x \\ \Delta E_y \end{bmatrix} = \mathbf{0} \tag{8}$$

where h_0, v_0, a, b, c and d are defined as before, and ΔE_x and ΔE_y represent the interocular differences in the partial derivatives, given by:

$$\Delta E_x(x,y) = [E_x(x,y)]_l - [E_x(x,y)]_r \tag{9}$$

$$\Delta E_y(x,y) = [E_y(x,y)]_l - [E_y(x,y)]_r \tag{10}$$

Equation (8) may be written in terms of the differential invariants of translation, dilation, rotation and deformation:

$$\begin{bmatrix} E_{xx} & E_{xy} \\ E_{xy} & E_{yy} \end{bmatrix} \mathbf{T} + \{\mathbf{D} + \mathbf{R} + \mathbf{S}\} \begin{bmatrix} E_x \\ E_y \end{bmatrix} + \begin{bmatrix} \Delta E_x \\ \Delta E_y \end{bmatrix} = \mathbf{0} \tag{11}$$

where $\mathbf{T}$ represents translation, $\mathbf{D}$ dilation, $\mathbf{R}$ rotation, and $\mathbf{S}$ deformation, given by:

$$\begin{aligned} \mathbf{T} &= \begin{bmatrix} h_0 \\ v_0 \end{bmatrix}, \qquad \mathbf{D} = \begin{bmatrix} e & 0 \\ 0 & e \end{bmatrix}, \\ \mathbf{R} &= \begin{bmatrix} 0 & -r \\ r & 0 \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} -p & q \\ q & p \end{bmatrix}. \end{aligned} \tag{12}$$

Here, e and r represent the magnitude of dilation and rotation, respectively, and (p, q) represents a deformation of magnitude $\sqrt{p^2 + q^2}$.

Equation (11) relates the 6 parameters of the disparity field, $\{h_0, v_0, e, r, p, q\}$, to the image derivatives $\{E_x, E_y, E_{xx}, E_{xy}, E_{yy}, E_{xt}, E_{yt}\}$. For two dimensional motion, only the component of velocity normal to image contours may be measured. This limitation is known as the aperture problem (Wallach, 1935), and is an example of degeneracy. Yamamoto (1989) described the general aperture problem for the estimation of the three-dimensional motion parameters. He showed how certain image configurations mean that some of the motion parameters will remain undetermined. Two dimensional motion is uniquely determined if there exist two or more image contours with different orientations. To estimate the six parameters of the affine disparity field would require six linearly independent sets of measures of image derivatives.

The model presented in this chapter is local, with estimates based on measurements obtained from a single image location. As such, the degree of information required to obtain a unique solution to equation (11) is unlikely to be available. This problem may be alleviated by reducing the number of parameters to be estimated. Here, this is achieved in two ways. First, it is assumed that an estimate of the translation component, $\mathbf{T}' = [h'_0, v'_0]^T$, may be obtained independently. Image derivatives at the point (x, y) in the left image may then be compared with those at the point $(x + h'_0, y + v'_0)$ in the right image, giving:

$$\{\mathbf{D} + \mathbf{R} + \mathbf{S}\} \begin{bmatrix} E_x \\ E_y \end{bmatrix} + \begin{bmatrix} \Delta E'_x \\ \Delta E'_y \end{bmatrix} = \mathbf{0} \tag{13}$$

where:

$$\Delta E'_x(x, y) = [E_x(x, y)]_l - [E_x(x + h'_0, y + v'_0)]_r \tag{14}$$

$$\Delta E'_y(x, y) = [E_y(x, y)]_l - [E_y(x + h'_0, y + v'_0)]_r \tag{15}$$

Second, the dilation component may be removed from (13) by premultiplying all terms by $[-E_y, E_x]$ giving:

$$(E_x^2 + E_y^2)r + (E_y^2 - E_x^2)p + (2E_xE_y)q + \left(E_x(\Delta E'_y) - E_y(\Delta E'_x)\right) = 0 \tag{16}$$

Equation (16) contains 3 unknowns, representing the rotation and deformation components of the affine transformation, respectively. This solution is attractive, as it requires the estimation of a reduced number of parameters. Further, these parameters will account for the orientation disparities introduced by the transformation,

and have been suggested on theoretical grounds as a representation which may be used to recover surface orientation (Koenderink and van Doorn, 1976).

3 Responses of bandpass filters

The model presented here employs quadrature pairs of bandpass filters as a first stage in processing. Simple cells in primary visual cortex are typically modelled as linear, bandpass neurons, whose responses are tuned for stimulus position, orientation and spatial frequency (Hubel and Wiesel, 1962; Campbell, Cleland, Cooper and Enroth-Cugell, 1968; Campbell, Cooper and Enroth-Cugell; 1969; Movshon, Thompson and Tolhurst, 1978; Ohzawa and Freeman, 1986; Hamilton, Albrecht and Geisler, 1989). These properties have been incorporated into models of early vision (e.g. Wilson, Levi, Maffei, Rovamo, DeValois, 1990; Heeger, 1992; Wilson, Ferrera and Yo, 1992; Fleet and Langley, 1994a). In addition, it has been suggested that simple cells form quadrature pairs, with similar amplitude spectra, but exhibitng a 90° phase shift (e.g. Marcelja, 1980; Daugman, 1985). This characterisation is motivated partially by the finding that adjacent simple cells tend to exhibit 90° or 180° phase relationships (Palmer and Davis, 1981; Pollen and Ronner, 1981; Foster, Gaska, Marcelja and Pollen, 1983; Liu, Gaska, Jacobson and Pollen, 1992). However, phase and amplitude may also be derived from a population of cells with a range of phases of receptive fields (Fleet, Wagner and Heeger, 1996).

Complex cells have subsequently been modelled as energy neurons, summing the squared responses of quadrature pairs of simple cells (Pollen and Ronner, 1983; Adelson and Bergen, 1985; Emerson, Bergen and Adelson, 1992; Heeger, 1992). Energy responses of quadrature filters have been used in models of both motion processing (Adelson and Bergen, 1985; Emerson et al., 1992) and stereopsis (Fleet et al, 1996). Quadrature filters have also been employed in phase based models of

visual processing (e.g. Sanger, 1988; Fleet and Jepson, 1990; Langley et al., 1990; Fleet, Jepson and Jenkin, 1991; Jenkin and Jepson, 1991; Sanger, 1988; Fleet and Langley, 1994a). These models provide estimates of disparity or velocity. Here, this notion is extended so that equation (11) may be solved for the parameters of rotation and deformation, using the phase responses of bandpass filters.

3.1 Phase and Amplitude responses

A quadrature filter may be expressed conveniently as a single complex valued filter. An example of such a filter is a two-dimensional Gabor function[1] (Gabor, 1946):

$$G(x,y;k_1,k_2,\sigma) = \frac{1}{2\pi\sigma^2}\exp\left[-\frac{(x^2+y^2)}{\sigma^2}\right][\cos(k_1x+k_2y)+i\sin(k_2x+k_2y)] \quad (17)$$

where $i^2 = -1$. Here, $\theta = \tan^{-1}\frac{k_1}{k_2}$ and $f = \sqrt{k_1^2+k_2^2}$ represent the orientation and frequency tuning of the filter, respectively. For a given frequency tuning, the bandwidth of the filter is determined by σ^2, the standard deviation of the Gaussian envelope[2]. The response of $G(x,y;k_1,k_2,\sigma)$ to a two dimensional input image I(x,y) is given by:

$$R(x,y) = G(x,y;k_1,k_2,\sigma) * I(x,y) \quad (18)$$

where $*$ represents the convolution operator (Bracewell, 1986). As $R(x,y)$ represents the convolution of a real-valued image signal $I(x,y)$ and a complex-valued filter, it

[1]while Gabor filters are not quadrature pairs, they provide a reasonable approximation for sufficiently small bandwidths, of around 1 octave or less.

[2]The bandwidth of the Gabor filter in equation (17) is given by $\log_2\left[\frac{f+\sigma}{f-\sigma}\right]$ (Fleet and Jepson; 1990).

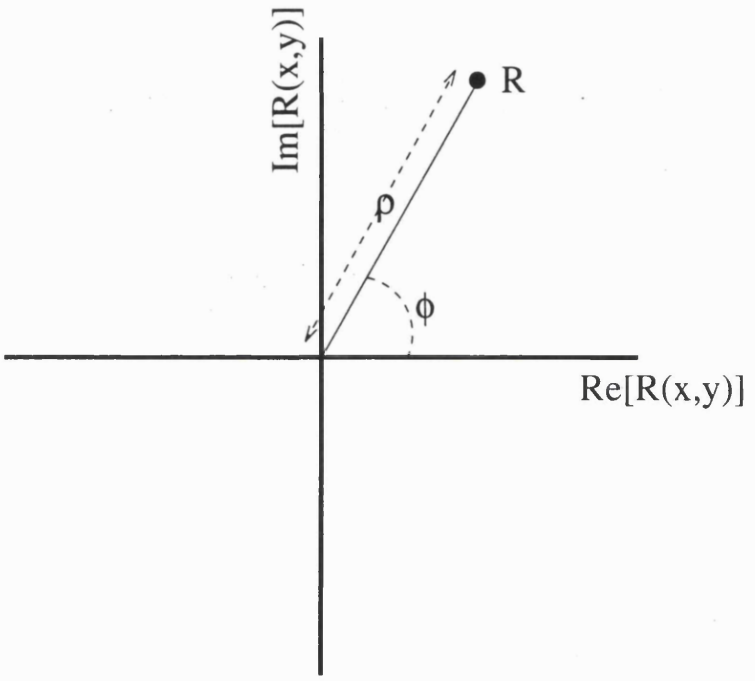


Figure 7.1: *The complex response $R(x, y)$ is shown here in the complex plane. Phase and amplitude form a polar representation of the response.*

too is a complex valued function:

$$R(x, y) = \mathrm{Re}[R(x, y)] + i\mathrm{Im}[R(x, y)] \tag{19}$$

which may be represented using polar coordinates in the complex plane (Figure 7.1):

$$\begin{aligned} R(x, y) &= \rho(x, y)e^{i\phi(x,y)} \\ &= \rho(x, y)\left[\cos\phi(x, y) + i\sin\phi(x, y)\right] \end{aligned} \tag{20}$$

where $\rho(x, y)$ and $\phi(x, y)$, which represent the amplitude and phase components of

the signal, respectively, are given by:

$$\begin{aligned} \rho(x,y) &= |R(x,y)| \\ &\equiv \sqrt{\mathrm{Re}[R(x,y)]^2 + \mathrm{Im}[R(x,y)]^2} \qquad (21) \\ \phi(x,y) &= \arg[R(x,y)] \\ &\equiv \arctan\left(\frac{\mathrm{Im}[R(x,y)]}{\mathrm{Re}[R(x,y)]}\right) \qquad (22) \end{aligned}$$

Given the local nature of $\phi(x,y)$ and $\rho(x,y)$, they are referred to as *instantaneous phase*, and *instantaneous amplitude*, respectively.

The local behaviour of the response of a bandpass filter is approximately sinusoidal, with a frequency close to the central tuning frequency of the filter, and a slowly varying amplitude. As a consequence, the phase response within a local spatial neighbourhood will be approximately linear (Fleet and Jepson, 1993), and the amplitude response lowpass (Knutsson, 1982). As an example, a pure sinusoid has linear phase, and a constant amplitude. Fleet and Jepson (1993) demonstrated the robustness of local phase information for measuring image velocity and binocular disparity. The disparity model derived in section 2 requires measurements of image derivatives. Given the linearity of phase in a local neighbourhood, and the relatively slowly varying nature of amplitude, measurements of spatial derivatives of phase would again be expected to exhibit greater stability than would spatial derivatives of amplitude.

3.2 Instantaneous frequency

Related to the concept of instantaneous phase is that of instantaneous frequency, which is defined here as the spatial derivative of instantaneous phase:

$$\tilde{k_1}(x,y) = \frac{\partial\phi(x,y)}{\partial x} = \phi_x, \quad \tilde{k_2}(x,y) = \frac{\partial\phi(x,y)}{\partial y} = \phi_y \tag{23}$$

Instantaneous frequency represents a local approximation to the frequency and orientation of a signal, which are given by:

$$\tilde{f}_\phi(x,y) = \sqrt{\phi_x^2 + \phi_y^2} \tag{24}$$

$$\tilde{\theta}_\phi(x,y) = \arg[\phi_x, \phi_y] \tag{25}$$

respectively. Instantaneous frequency may be estimated by taking spatial derivatives of the phase response in a local spatial neighbourhood. As an alternative, it is possible to obtain instantaneous frequency estimates by convolving images with both a quadrature filter $G(x,y;k_1,k_2,\sigma)$, and its spatial derivatives $G_x(x,y;k_1,k_2,\sigma) = \frac{\partial G}{\partial x}$ and $G_y(x,y;k_1,k_2,\sigma) = \frac{\partial G}{\partial y}$, (Fleet, 1990), giving:

$$R(x,y) = G(x,y;k_1,k_2,\sigma) * I(x,y) \tag{26}$$

$$R_x(x,y) = G_x(x,y;k_1,k_2,\sigma) * I(x,y) \tag{27}$$

$$R_y(x,y) = G_y(x,y;k_1,k_2,\sigma) * I(x,y) \tag{28}$$

Instantaneous frequency is then given by:

$$\phi_x = \frac{\mathrm{Im}[R_x]\mathrm{Re}[R] - \mathrm{Re}[R_x]\mathrm{Im}[R]}{\mathrm{Re}[R]^2 + \mathrm{Im}[R]^2} \tag{29}$$

$$\phi_y = \frac{\mathrm{Im}[R_y]\mathrm{Re}[R] - \mathrm{Re}[R_y]\mathrm{Im}[R]}{\mathrm{Re}[R]^2 + \mathrm{Im}[R]^2} \tag{30}$$

In the simulations presented here, Gabor filters with an isotropic envelope, and a bandwidth of 1 octave were used. Figure 7.2 shows the response of a horizontally

tuned filter to a plaid stimulus similar to those used in chapters 2 and 3. The stimulus used here consisted of two sinusoidal component gratings, with equal magnitudes of spatial frequency and contrast, oriented at $\pm 15°$ from horizontal. The stimulus, the response of the real part of the filter kernel, and the phase and amplitude responses are shown. It can be seen that the amplitude response varies more slowly than does the phase response. Further, the structure of the carrier grating is captured in the phase response, whereas the contrast envelope is reflected in the amplitude response.

4 Simulations

Equation (16) may be used to provide estimates of rotation and deformation on the basis of instantaneous frequency measurements. A complete solution to (16) requires three linearly independent sets of instantaneous frequency measurements, which may be obtained for phase contours with different orientations. Estimates may be obtained by combining measurements obtained from filters at a single position, with different orientation tuning:

$$\begin{bmatrix} (\phi_{x_1}^2 + \phi_{y_1}^2) & (\phi_{y_1}^2 - \phi_{x_1}^2) & (2\phi_{x_1}\phi_{y_1}) \\ \vdots & \vdots & \vdots \\ (\phi_{x_n}^2 + \phi_{y_n}^2) & (\phi_{y_n}^2 - \phi_{x_n}^2) & (2\phi_{x_n}\phi_{y_n}) \end{bmatrix} \begin{bmatrix} r \\ p \\ q \end{bmatrix} = \begin{bmatrix} \phi_{x_1}\Delta\phi_{y_1} - \phi_{y_1}\Delta\phi_{x_1} \\ \vdots \\ \phi_{x_n}\Delta\phi_{y_n} - \phi_{y_n}\Delta\phi_{x_n} \end{bmatrix} \quad (31)$$

where $\phi_{x_1}, \ldots, \phi_{x_n}$ and $\phi_{y_1}, \ldots, \phi_{y_n}$ represent instantaneous frequency measurements obtained from filters tuned to n orientations. Equation (31) may be written as:

$$\mathbf{Ax} = \mathbf{b} \quad (32)$$

Equation (32) may be solved using the method of least squares, to provide a max-

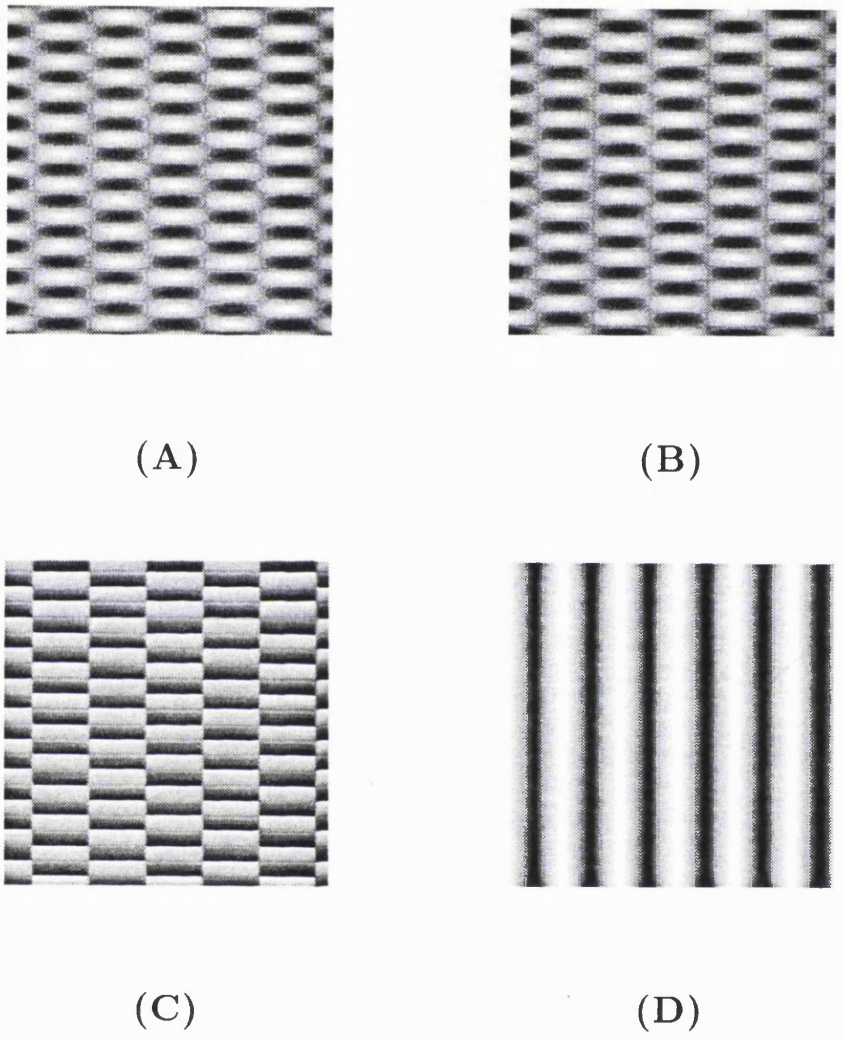


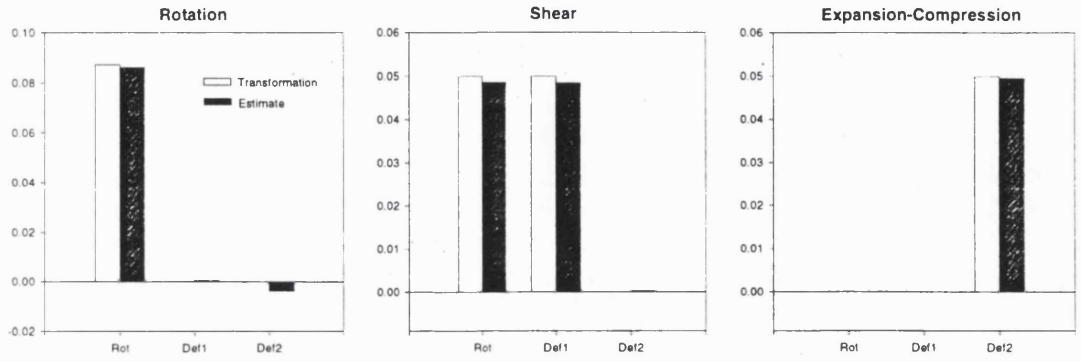
Figure 7.2: **(A):** *Plaid Stimulus.* **(B):** *Response of real part of filter.* **(C):** *Phase response.* **(D):** *Amplitude response.*

imum likelihood estimator of $\mathbf{x}$ (Press, Flannery, Teukolsky and Vetterlin, 1992). Figure (7.3) shows estimates for rotation, and horizontal shear and expansion-compression transformations. The stimulus used in these simulations was a plaid formed from four components, with orientations of 0°, 45°, 90° and 135°. This is an idealised stimulus, containing sinusoidal components with a spatial frequency equal to the tuning frequency of the filters used. By including components at four distinct orientations, it was ensured that filters tuned to different orientations would respond to different components, and thus provide independent estimates of instantaneous frequency. The results were obtained by convolving the stereoscopic image pairs with Gabor filters similar to those used in section 3; filters at 12 different orientation were used. Instantaneous frequency estimates were obtained by also convolving with the spatial derivatives of these filters. On the basis of these estimates, a solution to equation (31) was obtained using the generalised (Moore-Penrose) inverse (Rao, 1971).

The results demonstrate that, provided sufficient independent estimates of instantaneous frequency may be obtained, the affine parameters of rotation and deformation may be successfully estimated.

If sufficient information in not available locally, the matrix $\mathbf{A}$ in equation (32) will be ill-conditioned (close to singular), and the vector $\mathbf{x}$ cannot be uniquely determined. This situation leads to unreliable parameter estimates. The problem of ill-conditioning may be alleviated by constraining the solution to equation (31). Constrained minimisation may be achieved using the method of Lagrange multipliers (Horn, 1986). The problem then is to minimise the sum of the errors in equation (16), E_d, given by:

$$E_d = (E_x^2 + E_y^2)r + (E_y^2 - E_x^2)p + (2E_xE_y)q + \left(E_x(\Delta E'_y) - E_y(\Delta E'_x)\right) \qquad (33)$$



(A) (B) (C)

Figure 7.3: *Results for four component plaid stimuli* (**A**): *Rotation* (**B**): *Shear.* (C): *Expansion-Compression. White bars depict the actual transformation, while dark bars give the model estimates.*

and errors associated with an additional constraint, E_c. Let the total error to be minimised be ξ, which is given by:

$$\xi = E_d^2 + \lambda E_c^2 \tag{34}$$

where λ is a constant which weighs errors in the least squares equation relative to departures from the constraint equation. Minimisation was achieved subject to two independent constraints. The first assumed that viewed surfaces were frontoparallel, i.e. $\mathbf{x} = \mathbf{0}$. This method is equivalent to Bayesian estimation, and results in biases towards a frontoparallel surface. These biases will lead to the underestimation of surface slant. Such underestimation has been reported psychophysically (Gillam, Flagg, and Finlay, 1984; Mitchison and McKee, 1990; van Ee and Erkelens, 1996).

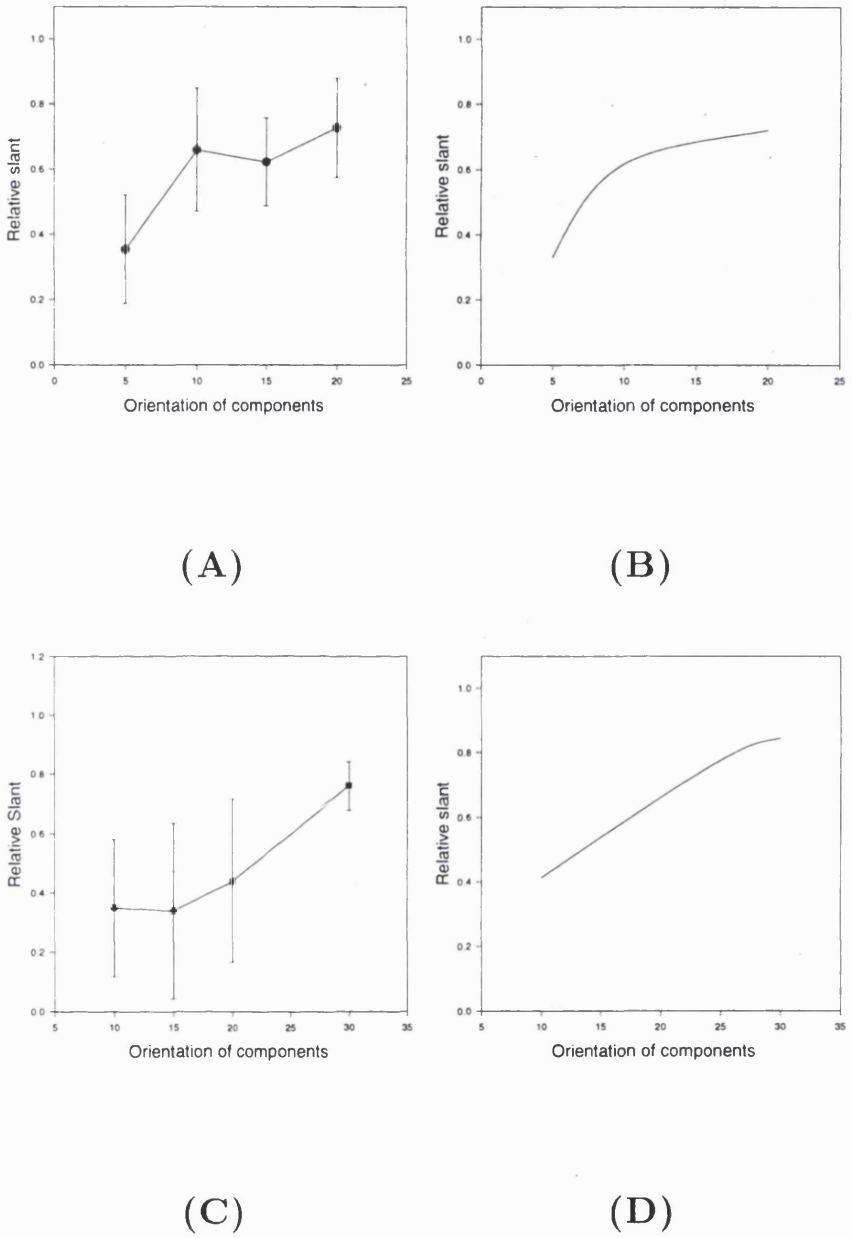


Figure 7.4: *Slant estimation for plaid stimuli.* **(A)** *Subject means for the psychophysical results from chapter 3, for horizontal shear.* **(B)** *Model estimates horizontal shear.* **(C)** *Subject means for the expansion-compression condition.* **(D)** *Model estimates for the expansion-compression. Results are plotted perceived or measured slant of the probe stimulus in all cases.*

The underestimation will be particularly severe when image measurements provide only limited information to the actual slant.

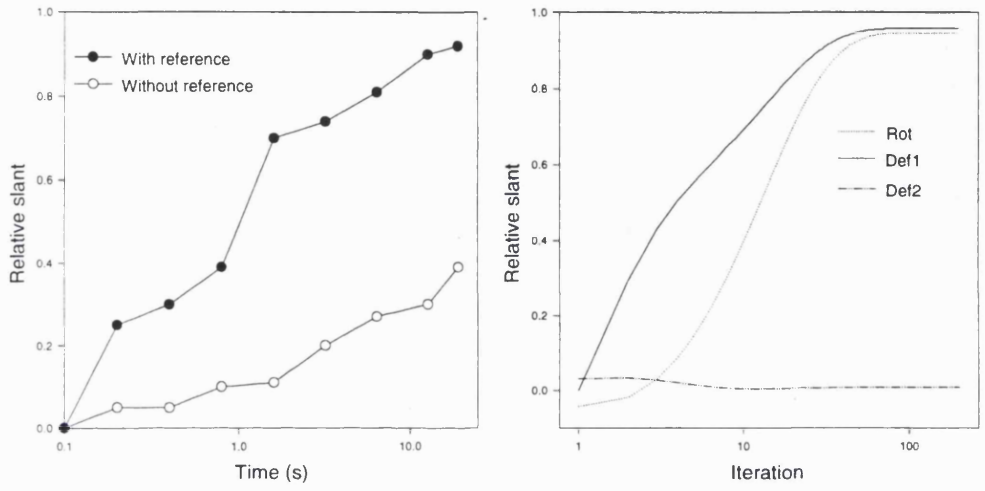
In chapter 3, it was shown that the slant of a two component plaid stimulus, with components close to horizontal, was underestimated relative to a stimulus with three components, equally spaced through orientation. This underestimation is predicted by the constraint used. Figure 7.4 shows the measured magnitude of the deformation component p, for the stimuli used in chapter 3. These results were obtained with a prior that was 10% of the maximum magnitude of components of **A**. Also shown are the means of the psychophysical results. The model predicts the decrease in perceived slant as the component orientation separation is decreased. The model cannot, however, account for the threshold results obtained in chapter 2, as the issue of slant discriminability is not addressed.

In the second simulation, a smoothness constraint was applied to the solution. This constraint assumes that the parameters of rotation and deformation vary slowly as a function of spatial position. Departures from smoothness are defined in terms of the partial spatial derivatives of the parameter estimates. In this case, the function E_c to be minimised is given by:

$$E_c = \left(\frac{\partial r}{\partial x}\right)^2 + \left(\frac{\partial r}{\partial y}\right)^2 + \left(\frac{\partial p}{\partial x}\right)^2 + \left(\frac{\partial p}{\partial y}\right)^2 + \left(\frac{\partial q}{\partial x}\right)^2 + \left(\frac{\partial q}{\partial y}\right)^2 \qquad (35)$$

Horn and Schunck (1981) employed a similar smoothness constraint to the problem of estimating optical flow. Following Horn and Schunck, equation (34) may be solved iteratively. Given an estimate of the parameters (r^n, p^n, q^n), we may obtain the new estimate:

$$r^{n+1} = \overline{r}^n + \frac{R[R\overline{r}^n + P\overline{p}^n + Q\overline{q}^n + S]}{\lambda + P^2 + Q^2 + R^2} \qquad (36)$$



(A) **(B)**

Figure 7.5: **(A)** *Perceived horizontal axis slant, as a function of presentation time, relative to geometrically predicted slant. (Replotted from Van Ee and Erkelens (1996), for subject OS, for the shear condition). Slant increases over time, for stimuli both with and without a zero disparity reference.* **(B)** *Model estimates of the parameters of rotation and deformation, for a horizontally sheared plaid stimuli, relative to the magnitude of transformation. Parameter estimates increase with successive iterations.*

$$p^{n+1} = \overline{p}^n + \frac{P[R\overline{r}^n + P\overline{p}^n + Q\overline{q}^n + S]}{\lambda + P^2 + Q^2 + R^2} \quad (37)$$

$$q^{n+1} = \overline{q}^n + \frac{Q[R\overline{r}^n + P\overline{p}^n + Q\overline{q}^n + S]}{\lambda + P^2 + Q^2 + R^2} \quad (38)$$

where:

$$\begin{aligned} R &= E_x^2 + E_y^2 \quad P = E_y^2 - E_x^2 \\ Q &= 2E_xE_y \quad\quad S = E_x(\Delta E_y') - E_y(\Delta E_x') \end{aligned} \quad (39)$$

and $(\overline{r}^n, \overline{p}^n, \overline{q}^n)$ gives the spatial mean of the previous estimates:

$$\begin{aligned} \overline{r}^n &= \frac{1}{6}\left(r^n_{i-1,j} + r^n_{i+1,j} + r^n_{i,j-1} + r^n_{i,j+1}\right) \quad (40) \\ &+ \frac{1}{12}\left(r^n_{i-1,j-1} + r^n_{i-1,j+1} + r^n_{i+1,j-1} + r^n_{i+1,j+1}\right) \\ \overline{p}^n &= \frac{1}{6}\left(p^n_{i-1,j} + p^n_{i+1,j} + p^n_{i,j-1} + p^n_{i,j+1}\right) \quad (41) \\ &+ \frac{1}{12}\left(p^n_{i-1,j-1} + p^n_{i-1,j+1} + p^n_{i+1,j-1} + p^n_{i+1,j+1}\right) \\ \overline{q}^n &= \frac{1}{6}\left(q^n_{i-1,j} + q^n_{i+1,j} + q^n_{i,j-1} + q^n_{i,j+1}\right) \quad (42) \\ &+ \frac{1}{12}\left(q^n_{i-1,j-1} + q^n_{i-1,j+1} + q^n_{i+1,j-1} + q^n_{i+1,j+1}\right) \end{aligned}$$

The regularisation scheme may be related to the time course of slant perception. van Ee and Erkelens (1996) found that perceived inclination increased with increasing presentation times. Data for their subject OS are replotted in figure 7.5A. Figure 7.5B shows parameter estimates for the horizontally sheared four component plaid stimulus that was used in the earlier simulations. These parameter estimates increase with successive iterations. If the iterations of the model may be related to time, and inclination encoding made use of this estimate of deformation, perceived inclination would be expected to increase over time, in a manner similar to that reported by van Ee and Erkelens (1996).

5 Conclusions

The model presented in this chapter describes the disparities in a local spatial neighbourhood as an affine transformation between a binocular pair of images. A reduced model of this transformation was presented, from which estimates of rotation and deformation were obtained. Estimates were derived from local measures of interocular differences in instantaneous frequency, the spatial derivative of the phase response of quadrature, bandpass filters. However, the derivative model presented in section 2 may potentially be implemented using derivatives of other image properties. One possibility that was discussed is the use of derivatives of the amplitude response of bandpass filters. Figure 7.2 shows how energy responses reflect the structure of the contrast envelope of a stimulus. The model presented could in principle be implemented using energy derivatives, and would then produce estimates of the local affine transformation of the image contrast envelope.

Energy derivatives represent an additional potential source of information in describing the affine transformation (Hibbard and Langley, 1994). They may also be useful in the representation of transparency (Hibbard, Langley and Fleet, 1994). Fleet and Langley (1994a) proposed that phase and amplitude responses may be used to represent Fourier and non-Fourier motion, respectively. Similarly, phase and amplitude may be used as a basis for the representation of luminance and contrast disparities.

Estimates obtained from measures of local interocular differences in amplitude gradients may present particular difficulties from a computational viewpoint. Fleet and Jepson (1993) emphasised the stability of phase information, and its suitability for the estimation of image changes. Conversely, amplitude responses were suggested to form a less reliable basis for such estimation. Given the slowly varying nature of amplitude responses, estimates of amplitude gradients would be expected to be less reliable than would estimates of phase gradients.

The model presented is able to account for two important properties of slant perception. First, the model estimates of deformation would lead to an underestimation of slant. This is a consequence of the Lagrangian multiplier incorporated into the model which introduces a bias towards small estimates. Further, for plaid stimuli, this underestimation was found to increase as the plaid's components approached horizontal. This results from the increased importance of the *a priori* distribution when presented with more limited information. Similar constraints have also been used in models of motion (Simoncelli and Heeger, 1992), and may account for perceptual biases towards slow speeds (e.g. Thompson, 1982; Smith and Edgar, 1990; Farid and Simoncelli, 1994).

Second, the estimates of rotation and deformation provided increase with successive iterations; this property of the model is a consequence of the smoothness constraint incorporated. This aspect of the model's behaviour is the result of the initialisation of the estimates to zero. Perceived slant has similarly been found to build up over a time course of several seconds (van Ee and Erkelens, 1996).

Although the model presented is based on physiologically plausible filters, it is not intended as a model of the neural processes underlying the perception of slant. Rather, the model demonstrates how information contained in the responses of such bandpass filters may yield estimates of useful parameters. However, it may be possible to implement the model in a similar manner to the energy based model of Fleet et al. (1996), which incorporated disparity detectors with both interocular position and phase shifts. Physiological support exists for binocular neurons tuned to difference orientations in left and right eyes (Blakemore et al, 1972; Nelson et al., 1977, Hänny et al., 1980). The model presented here may be implemented using disparity detectors, to encode instantaneous frequency difference from interocular differences in orientation and spatial frequency tuning.

8. Conclusions

1 A model of surface representation from stereopsis

This thesis investigated the roles of luminance and contrast disparities in the representation of surfaces in three dimensional space. The investigations focused on the processing of slant, and on the representation of transparency. The conclusions drawn from the psychophysical evidence obtained are summarised in the model presented in figure 8.1. This model outlines the computational strategy which, it is proposed, allows the representation of surfaces in depth on the basis of luminance and contrast disparity cues. This concluding chapter discusses the psychophysical data, and outlines further questions raised by the research.

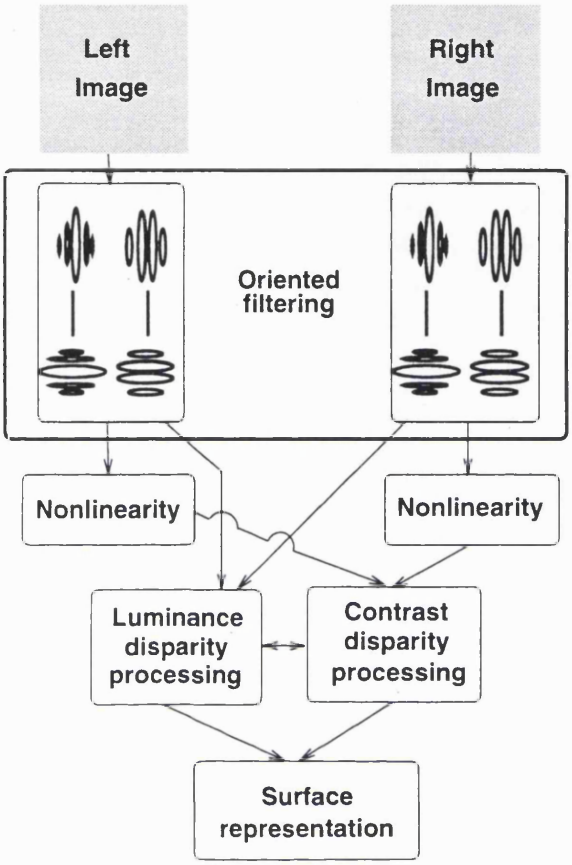


Figure 8.1: *A model of luminance and contrast disparity processing.*

1.1 Linear filtering

Disparity detectors have been characterised as linear neurons, tuned to stimulus orientation and spatial frequency (e.g. Ohzawa, DeAngelis and Freeman, 1990; DeAngelis, Ohzawa and Freeman, 1991, 1995). Psychophysical studies have demonstrated that human stereopsis is similarly tuned to both orientation and frequency (Mansfield and Parker, 1994; Yang and Blake, 1991). It is proposed that the processing of luminance and contrast disparities share a common, linear filtering stage. For contrast disparities, this stage will exhibit tuning to the orientation and frequency

of the carrier of the stimulus. In chapter 6, it was demonstrated that adaptation to a sinusoidal grating with the orientation and frequency of the carrier of a contrast modulated stimulus increased the contrast at which disparity in the contrast envelope of the stimulus could be detected. Conversely, adaptation to a grating with the orientation and frequency of the contrast beats of such a stimulus had relatively little effect on the contrast at which disparity could be detected. Similar results were reported in a two dimensional orientation discrimination task (Langley, Fleet and Hibbard, 1996). It has been proposed that nonlinearities occur early in vision, prior to a stage of bandpass filtering in primary visual cortex (e.g. Burton, 1973; Henning, Hertz and Broadbent, 1975; Smallman and Harris, 1996). However, it is argued here that these nonlinearities do not play a significant functional role in the processing of contrast disparities.

1.2 Post-filtering nonlinearity

The detection of image contrast variations requires a nonlinear stage in processing (Chubb and Sperling, 1988; Fleet and Langley, 1994a). Langley et al. (1996) suggested that this nonlinearity occured after a stage of linear, bandpass filtering, which is sensitive to stimulus orientation and spatial frequency. In addition, the stage of nonlinear contrast processing was proposed to exhibit tuning to the spatial frequency of contrast modulations. It was found that contrast sensitivity for contrast modulations showed a peak at 0.4 cycles/degree. Again, this finding was taken as evidence that contrast modulations are *not* processed on the basis of nonlinear processing occurring prior to a stage of bandpass filtering; such processing exhibits peak sensitivity for frequencies of 3-5 cycles/degree (e.g. DeValois, Morgan and Snodderly, 1974). In chapter 4, it was found that, for contrast modulated, sinusoidal gratings, transparency was most readily perceived for contrast modulations with a spatial frequency of around 0.4 cycles/degree. Although these experiments studied

transparency, as a function of disparity, rather that contrast sensitivity, they do provide evidence for spatial frequency tuning in the processing of contrast envelope disparities. Again, this tuning is distinct to that found for luminance disparities, stereoacuity for which peaks for stimuli with a spatial frequency of 2.5 cycles/degree (Schor and Wood, 1983; Kontsevich and Tyler, 1994; Hess and Wilcox, 1994).

1.3 Disparity processing

Binocular neurons have been found which are sensitive to position disparities (e.g. Poggio and Fischer, 1977; Ohzawa et al., 1990; DeAngelis et al., 1991, 1995) and orientation disparities (Blakemore, Fiorentini and Maffei, 1972; Nelson, Kato and Bishop, 1977; Hänny, von der Heydt and Poggio, 1980). It has been argued that the representation of slant relies not on position disparities, but on orientation disparities. (e.g. von der Heydt, Hänny and Dürsteller, 1981; Rogers and Graham, 1983). Additionally, it has been suggested that slant may be encoded from interocular spatial frequency differences (Blakemore, 1970; Tyler and Sutter, 1979; Tyler, 1990).

In chapter 2, slant thresholds were measured for grating stimuli, with interocular differences in orientation and/or spatial frequency. It was found that thresholds depended both on the cyclopean orientation of the grating, and on the nature of the binocular transformation. It was found that thresholds could be predicted primarily on the basis of the magnitude of the orientation disparity between the left- and right-eye gratings. However, slant was also perceived in stimuli containing no orientation disparities. For example, a vertical grating with an interocular spatial frequency difference was perceived to slant about a vertical axis. Slant must therefore be perceived from other cues in addition to orientation disparities. It was found that the results could be modelled by assuming that either gradients of phase dis-

parities, or frequency differences, were combined with orientation disparities in the representation of slant.

Additionally, it was found that slant thresholds for plaid stimuli could be predicted from disparities in their Fourier components. It was found that, for plaids with components oriented between horizontal and $\pm 45°$, slant thresholds were lower for surfaces with a vertical slant axis than for surfaces with a horizontal slant axis. This result is a reversal of the usual anisotropy in slant (e.g. Wallach and Bacon, 1976; Rogers and Graham, 1983; Cagenello and Rogers, 1993), but is predicted from a consideration of orientation disparities, and is consistent with the finding that the anisotropy is not evident for stimuli containing lines oriented at $\pm 45°$ (Cagenello and Rogers, 1993). These results lend further support to the notion that slant perception relies primarily on the encoding of orientation disparities.

For suprathreshold magnitudes of slant, perceived slant for plaid stimuli was dependent on the orientation of the plaid's component gratings. Perceived magnitude was less for plaids with components close to horizontal. Again, this result is consistent with the magnitude of orientation disparities contained in the stimuli.

Although the perception of slant in plaid stimuli may be characterised in terms of disparities in their component gratings, contrast disparities were found to play a role in the perception of transparency. This result demonstrates that depth may be observed from disparities in contrast modulation. The smallest disparity for which transparency was observed occured for stimuli with a contrast envelope whose spatial frequency was around 0.4 cycles/degree. A similar result has been reported for contrast sensitivity for contrast modulations (Langley, Fleet and Hibbard, 1996). This result suggests first that contrast modulations are processed by a spatial frequency tuned mechanism, with peak sensitivity at 0.4 cycles/degree, and second that contrast modulation disparities are processed by mechanisms similar to those

responsible for the detection of the modulation.

Transparency was only observed for contrast envelopes with crossed disparities, stimuli for which the envelope appeared to lie in front of its carrier. This asymmetry was not observed for additive transparency, suggesting that a distinction between contrast and luminance defined disparities is preserved in stereoscopic processing. This phenomenon was investigated further, by presenting stimuli containing both contrast modulation and additional luminance, taking the same form as the modulation. Transparency was not reported for stimuli containing only the contrast modulation, if this modulation had uncrossed disparities relative to the carrier. Transparency was however observed when luminance was subtracted from the stimuli. The contrast of the luminance required was not influenced by the depth of contrast modulation. These results suggest that luminance and contrast disparities are treated as separate sources of information, which are processed independently in stereopsis.

1.4 Surface representation

Surface orientation may be inferred from the local spatial variation in binocular disparities. Koenderink and van Doorn (1976) suggested that an analysis of the deformation component of this local variation may be used to provide reliable estimates of slant. Howard and Kaneko (1994; Kaneko and Howard, 1994) proposed that surface slant is represented on the basis of the difference between horizontal and vertical shear, and between horizontal and vertical expansion-compression. Either of these strategies may make use of orientation, positional, or frequency disparities.

In chapter 7, a model of slant was presented in which the affine parameters of deformation and rotation are estimated from interocular differences in instantaneous frequency. These estimates may form the basis of a representation of surface slant.

The model proposed requires the estimation of three parameters, to determine the magnitude of rotation, and the magnitude and direction of deformation. For single sinusoidal gratings, and for plaids formed from two sinusoidal components, a unique solution is not possible. In two dimensional motion, only the component of velocity normal to a contour may be estimated (Wallach, 1935). Similarly, a slanted sinusoidal grating may be seen with a tilt that is normal to its orientation (Wilson, Blake and Halpern, 1996). When a unique solution is not available, the problem of estimating the affine parameters degenerates, and parameter estimates may become unreliable. This problem was addressed by constraining the solution obtained, by incorporating an *a priori* bias toward frontoparallel surfaces. This bias leads to underestimation of slant, and may account for the finding that the perceived magnitude of slant for plaid stimuli decreased as the orientation of the plaid's component gratings approached horizontal. Further, a smoothness constraint was applied, so that a solution was obtained which minimised the differences in slant estimates obtained from adjacent image locations. This constraint led to a slow build up of parameter estimates over succesive iterations. Similarly, perceived slant and inclination have been found to increas over time. (van Ee and Erkelens, 1996).

The model was implemented using estimates of instantaneous frequency, the spatial derivative of phase. However, it was suggested that the model may also be implemented using derivatives of the amplitude responses of bandpass filters. The amplitude response captures slow variations in the contrast of stimuli, and may therefore be used to represent disparities in an image contrast envelope. Phase and amplitude responses may be used to characterise luminance and contrast disparities, from which surfaces may be represented independently. This would allow for the representation of transparency.

Transparency may be observed in stimuli for which different elements have different disparities. Transparency may be evident, for example, from a superposition of ran-

dom dot stereograms (Todd and Akerstrom, 1988; Weinshall, 1990; Parker et al., 1991; Langley, Fleet and Hibbard, 1995). For these stimuli, there is no single interpretation consistent with all the disparity cues present. Under situations in which transparency is not observed in these stimuli, depth averaging between surfaces occurs. The experiments presented in this thesis concerned transparency in stimuli for which an interpretation of the disparities as resulting from the viewing of a single surface was possible. These stimuli may be considered analogous to transparent motion from plaid stimuli (for additive transparency) or in contrast modulated stimuli (for multiplicative transparency).

For plaid stimuli, transparent motion has been taken to indicate the activation of independent motion detectors (Adelson and Movshon, 1982). Experiments which have studied conditions under which either transparency or coherent motion have been observed have been used to characterise the tuning properties of motion detection. Transparency is observed in plaids with components differing in orientation, spatial and temporal frequency, and contrast. (Adelson and Movshon, 1982; Movshon, Adelson, Gizzi and Newsome, 1985). On the basis of these results, it has been proposed that motion detectors are tuned to orientation, and spatial and temporal frequency. Similarly, the results of the additive transparency experiments presented in chapter 4 suggest that stereoscopic channels are tuned to spatial frequency, and that interactions between channels are dependent on both spatial frequency and contrast differences in stimulus components. Both spatial frequency tuning (Schor and Wood, 1983) and interactions across scale (e.g. Wilson, Blake and Halpern, 1991; Smallman, 1995) have been reported in stereopsis. Wilson et al. showed that disambiguation of fine spatial scales only occured for coarser scale information if the difference in scale was not too great. They found interactions for a difference of 2 octaves, but not for a difference of 4 octaves. In chapter 4, the difference in scale was close to 1 octave, yet transparency was nevertheless observed, indicating a lack of integration between differently tuned mechanisms. However, the stimulus com-

ponents also differed in orientation (being orthogonal) and contrast. It was found that when the difference in contrast was sufficiently small, integration did occur, and transparency was not observed.

The asymmetry reported for multiplicative transparency reflects inherent knowledge of physical situations that will give rise to perceptual transparency. Multiplicative transparency is associated with contrast reducing media which, to be perceived as transparent, must lie in front of another viewed object. An example is a contrast reducing film (or neutral density filter) which will only be perceived as transparent if it is placed in front of an opaque background. Altering the configuration of surfaces, to remove cues to the contrast reducing nature of the filter, will destroy the perception of transparency (Metelli, 1974). Although such phenomena may be rare in nature, the detection of contrast modulations may be useful in the perception of shadows and lighting variations (Kersten 1991). A similar asymmetry has been reported for occlusion. This asymmetry may reflect a general strategy which is applied both to contrast reducing, and occluding figures.

It may be argued that the asymmetry results not from constraints on transparency applied with respect to the sign of contrast disparities directly, but from the integration of stereoscopic cues with monocular depth cues. For a contrast modulated stimulus that is perceived as a multiplicative transparency, there exists a monocular cue that the contrast reducing medium is in front of the carrier of the stimulus. The asymmetry reported might therefore be explained in terms of competition between stereoscopic and monocular depth cues. However, when additional luminance disparities were incorporated into the stimulus, transparency was observed for both crossed and uncrossed disparities. This stimulus appears to have the same transparency relationship, when viewed monocularly, as when only the contrast modulation is present in the stimulus. The asymmetry would therefore appear to relate to the presence of contrast disparities, rather than luminance disparities in the stimuli.

While this issue requires further investigation, these results would appear to support the notion that contrast and luminance disparities represent independent sources of information in stereopsis.

2 Further Questions

Mitchison and McKee (1990) demonstrated that, for a given magnitude of disparity gradient, perceived slant depends on the direction of the gradient. In chapter 3, it was demonstrated that perceived slant is influenced by the orientation of stimulus contours. These results might lead one to predict biases in the perceived tilt of stimuli. For surfaces with a tilt other than 0° or 90°, slant may be considered in terms of components about horizontal and vertical axes. If the component of slant about a horizontal axis is encoded with proportionally more slant than the component of slant about a vertical axis, one would predict a bias in the perceived tilt of a surface toward 0°.

In chapters 4 and 5, it was demonstrated that stereoscopic transparency may be observed for contrast modulations with crossed disparities. These experiments did not, however, distinguish between the disparity of the contrast modulation, and the sign of its disparity relative to that of the carrier of the stimulus. By using stimuli with a carrier with a non-zero disparity, it would be possible to determine whether it is the sign of the contrast modulation disparities relative to fixation, or the inferred depth relationships between the carrier and the modulation, which determine the perception of transparency. While the results of this experiment would not affect the previous discussion, they may shed light on the site in processing at which the asymmetry in the perception of transparency arises. If the sign of the disparities relative to fixation, rather than relative to the carrier, was found to be crucial, this could taken as evidence that the constraint is applied early in processing, and

would provide further evidence for a distinction between contrast and luminance disparities.

The results of chapter 6 suggest that contrast disparities are processed after a stage of orientation and frequency specific filtering, and that the processing of the contrast envelope is itself spatial frequency tuned. Contrast adaptation, and disparity sensitivity, have been suggested to occur in cortical simple cells. It would appear physiologically plausible therefore if the processing of contrast disparities occured via a separate mechanism to the processing of luminance disparities. The asymmetry results presented in chapter 5 again suggested that contrast and luminance disparities are treated independently, and are subject to different constraints in the interpretation of depth. However, as argued above, it is possible that this asymmetry reflects assumptions imparted at higher levels of processing, at which stereoscopic information is integrated with other depth cues. If this were the case, it might be argued that disparities in contrast and luminance contours are processed by identical mechanisms. One way in which this may be tested would be to present luminance contours to one eye, and contrast contours to the other eye. Lin and Wilson (1995) provided evidence that depth may be observed on the basis of disparities between a luminance contour and a contrast contour. However, this finding does not rule out the possibility that luminance and contrast contours are represented independently. Conversely, if a failure to integrate luminance and contrast contours in stereopsis could be demonstrated for some stimuli, this would provide evidence that they are indeed processed differently. It has been demonstrated that interleaved frames of moving Fourier and non-Fourier stimuli are not integrated to produce the perception of motion (Scott-Samuel and Georgeson, 1995). Similar results in stereopsis would provide further evidence for the existence of independent, Fourier and non-Fourier mechanisms.

It was observed in chapter 4 that a transparent, slanted surface could be perceived

from a contrast envelope. This opens the possibility that contrast disparity processing may be directly sensitive to surface orientation, in the same way that it has been argued that surface orientation for luminance disparity defined surfaces is not encoded on the basis of position disparities. If this were the case, one might predict similar anisotropies in the perception of slant from contrast modulations, as are generally found in stereopsis and shape from motion studies (Wallach and Bacon, 1976; Rogers and Graham 1983; Gillam, Flagg and Finlay, 1984; Gillam, Chambers and Russon, 1988; Michison and McKee, 1990; Mitchison and Westheimer, 1990; Gillam and Ryan, 1992; Cagenello and Rogers, 1993). Related to this possibility is the notion that slant may be encoded on the basis of orientation disparities in contrast envelopes (e.g. Rogers and Graham, 1983).

Langley et al. (1996) demonstrated that contrast sensitivity for image contrast modulations showed a peak for contrast modulations with a spatial frequency of 0.4 cycles/degree. Similarly, transparency was most readily observed in contrast modulated stimuli with contrast modulations around this frequency. This finding may be extended, to provide a full contrast sensitivity function for contrast modulations. Further, masking studies could be performed to assess the number, and orientation and spatial frequency bandwidths of contrast detecting mechanisms. These experiments could be performed for both two dimensional, and stereoscopic tasks. By comparing orientation and frequency tuning in the two domains, it would be possible to assess the extent to which the tasks appeared to share underlying contrast sensitive mechanisms.

References

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

Adelson, E.H. and Movshon, J.A. (1982) Phenomenal coherence of moving visual patterns. *Nature*, 300:523-525.

Adelson, E.H. and Movshon, J.A. (1984) Binocular disparity and the computation of of two-dimensional motion *Journal of the Optical and Optical Society of America* 1A:1266.

Akerstrom, R.A. and Todd, T.J.T. (1988) The Perception of Stereoscopic Transparency *Perception and Psychophysics*, 44:421-432.

Albrecht, D.G. and DeValois, R.L. (1981) Striate cortex responses to periodic patterns with and without the fundamental harmonics. *Journal of Physiology (London)*, 319:497-514.

Bergman, R. and Gibson, J.J. (1959) The negative after-effect of a surface slanted in the third dimension. *American Journal of Psychology*, 72:364-374.

Blakemore, C.(1970) A new kind of stereoscopic vision. *Vision Research*, 10:1181-1200.

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

Blakemore, C., Fiorentini, A., and Maffei, L. (1972) A second neural mechanism of binocular depth discrimination. *Journal of Physiology*, 226:725-749, 1972.

Blakemore, C. and Hague, B. (1972) Evidence for disparity detecting neurones in the human visual system. *Journal of Physiology (London)*, 225:437-455.

Blakemore, C. and Nachmias, J. (1971) The orientation specificity of two visual after-effects. *Journal of Physiology (London)*, 213:157-174.

Bowne, S.F. (1990) Contrast discrimination cannot explain spatial frequency, orientation or temporal frequency discrimination. *Vision Research*, 30(3):449-461.

Bracewell, R.N. (1986) The Fourier transform and its applications. *McGraw-Hill Book CO., London.*

Bradley, A. and Skottun, B.C. (1984) The effects of large orientation and spatial frequency shifts on spatial discriminations. *Vision Research*, 24:1889-1896.

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

Burbeck, C.A. and Regan, D. (1983) Independence of orientation and size in spatial frequency discriminations. *The Journal of the Optical Society of America* A,72(12):1691-1694.

Burke, D. and Wenderoth, P. (1993) The effect of interactions between one-dimensional component gratings on two-dimensional motion perception. *Vision Research*, 33:343-350.

Burr, D.C. and Wijesundra, S.A. (1991) Orientation discrimination depends on spatial-frequency. *Vision Research*,31:1449-1452.

Burt, P. and Julesz, B. (1980) Modifications of the classical notion of Panum's fusional

areas. *Perception*, 9:671-682.

Burton, G.J. (1973) Evidence for non-linear response process in the visual system from measurements on the thresholds of spatial beat frequencies. *Vision Research*, 13:1211-55.

Cagenello, R. and Rogers, B.J. (1988) Local orientation differences affect the perceived slant of stereoscopic surfaces. *Investigative Ophthalmology and Visual Science*, 29 (Suppl):399.

Cagenello, R. and Rogers, B.J. (1993) Anisotropies in the perception of stereoscopic surfaces: the role of orientation disparity. *Vision Research*, 33:2189-2201.

Campbell, F.W., Cleland, B.G., Cooper, G.F. and Enroth-Cugell, C. (1968) The angular selectivity of visual cortical cells to moving gratings. *Journal of Physiology (London)*, 198:237-250.

Campbell, F.W., Cooper, G.F. and Enroth-Cugell, C. (1969) The spatial selectivity of visual cells of the cat. *Journal of Physiology (London)*, 203:223-235.

Campbell, F.W. and Maffei, L. (1981) The influence of spatial frequency and contrast on the perception of moving patterns. *Vision Research*, 21:713-721.

Campbell, F.W. and Robson, J.G. (1966) Application of Fourier analysis to the visibility of gratings. *Journal of Physiology (London)*, 197:551-566.

Chen, B., Makous, W. and Williams, D.R. (1993) Serial spatial filters in vision *Vision Research*, 33:413-427.

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

Daugman, J.G. (1985) Uncertainty relation for resolution in space, spatial frequency and orientation optimised by two-dimensional visual cortical filters. *Journal of the Optical Society of America* A, 2:1160-1168.

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

DeAngelis, G.C., Ohzawa, I. and Freeman, R.D. (1995) Neuronal mechanisms underlying stereopsis: How do simple cells in the visual cortex encode binocular disparity? *Perception*, 24:3-32.

Derrington, A. M. (1987) Distortion products in geniculate X-cells: a physiological basis for masking by spatially modulated gratings. *Vision Research*, 27:1377-86.

Derrington, A.M. (1990) Mechanisms for coding luminance patterns: are they really linear? In *Vision: Coding and Efficiency*, Blakemore, C. (ed), Cambridge University Press, Cambridge.

Derrington, A.M. and Badcock, D.R. (1985) Separate detectors for simple and complex grating patterns *Vision Research*, 25:1869-1878.

Derrington, A.M. and Badcock, D.R. (1986) Detecting spatial beats: non-linearity or contrast increment detection? *Vision Research*, 27:343-8.

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

DeValois, R.R. and DeValois, K.K. (1990) *Spatial Vision* Oxford University Press, New York.

DeValois, R.L., Morgan, H., and Snodderly, D.M. (1974) Psychophysical studies of monkey vision III. Spatial luminance contrast sensitivity tests of macaque and human observers. *Vision Research*, 14:75-81.

DeValois, R.L. Yund, E.W. and Hepler, N. (1982) The orientation and direction selectivity of cells in macaque visual *Vision Research*, 22:531-544.

Emerson, R.C., Bergen, J.R. and Adelson, E.H. (1992) Directionally selective complex cells and the computation of motion energy in cat visual cortex. *Vision Research*, 32:203-218.

Farid, H. and Simoncelli, E.P. (1994) The perception of transparency in moving quare-wave plaids. *Investigative Ophthalmology and Visual Science*, 35 (Suppl): 1271.

Felton, T.B., Richards, W. and Smith, R.A. (1972) Disparity processing of spatial frequencies in man. *Journal of Physiology (London)*, 225:349-362.

Ferrera, V. and Wilson, H.R. (1987) Direction specific masking and the analysis of motion in two-dimensions. *Vision Research*, 27:1783-1796.

Ferrera, V. and Wilson, H.R. (1990) Perceived direction of moving two-dimensional patterns. , *Vision Research*, 30:273-287.

Ferrera, V. and Wilson, H.R. (1991) Perceived speed of moving two-dimensional patterns. *Vision Research*, 31:877-893.

Fleet, D.J. (1990) *Measurement of image velocity*, PhD Thesis, University of Toronto.

Fleet, D. and Jepson, A.D. (1989) Computation of normal velocity from local phase information. *IEEE CVPR 1989,* 379-386.

Fleet, D.J. and Jepson, A.D. (1990) Computation of component image velocity from local phase information *International Journal of Computer Vision*, 5:77-104.

Fleet, D.J. and Jepson, A.D. (1993) Stability of phase information. *IEEE transactions on pattern analysis and machine intelligence*, 15:1253-1268.

Fleet, D.J., Jepson, A.D. and Jenkin, M.R.M. (1991) Phase-based disparity measurement *Computer Vision Graphics and Image Processing-Image Understanding*, 53(2):198-210

Fleet, D.J. and Langley, K. (1994a) Computational analysis of non-Fourier motion. *Vision Research*, 34: 3057-3079.

Fleet, D.J. and Langley, K. (1994b) Non-Fourier channels in stereopsis and motion. *Perception* , 23 (suppl), 35.

Fleet, D.J., Wagner, H. and Heeger, D.J. (1996) Neural encoding of binocular disparity: energy models, position shifts and phase shifts. *Vision Research*, 36:1839-1857.

Foster, K.H. Gaska, J.P. and Pollen, D.A. (1983) Spatial and temporal frequency selectivity of v1 neurons in the macaque monkey. *Investigative Ophthalmology and Visual Science (Suppl)*, 22:228.

Foster, K.H., Gaska, J.P., Marcelja, S. and Pollen, D.A. (1983) Phase relationships between adjacent simple cells in the feline visual cortex. *Journal of Physiology*, 345:22.

Gabor, D. (1946) Theory of communication. *Journal of IEE*, 93:429-457.

Gårding, J., Porrill, J, Mayhew, J.E.W. and Frisby, J.P. (1996) Stereopsis, vertical

disparity and relief transformations. *Vision Research*, 35:703-722.

Georgeson, M.A. (1987) Temporal properties of spatial contrast vision. Vision Research, 27:765-780.

Georgeson, M.A. and Shackleton, T.M. (1994) Perceived Contrast of Gratings and Plaids: Non-Linear Summation Across Oriented Filters. *Vision Research*, 34:1061-1075.

Gillam, B. (1968) Perception of slant when perspective and stereopsis conflict: experiments with aniseikonic lenses. *Journal of Experimental Psychology*, 79:299-305.

Gillam, B., Chambers, D. and Russo, T. (1988) Postfusional Latency in Stereoscopic Slant Perception and the Primitives of Stereopsis. *Journal of Experimental Psychology: Human Perception and Performance*, 14(2):163-175.

Gillam, B. Flagg, T. and Finlay, D. (1984) Evidence for disparity change as the primary stimulus for stereoscopic processing. *Perception and Psychophysics*, 36, 559-564.

Gillam, B. and Rogers, B. (1991) Orientation disparity, deformation, and stereoscopic slant perception. *Perception*, 20:441-448.

Gillam,B. and Ryan,C. (1992) Perspective, orientation disparity, and anisotropy in stereoscopic slant perception. *Perception*, 21:427-439.

Grimson, W.E.L. (1980) A computer implementation of a theory of human stereo vision. *(A.I. Memo No 565), Cambridge, MA: M.I.T.*: Artificial Intelligence Library.

Greenlee, M.W. and Heitger, F. (1988) Functional role of contrast adaptation. *Vision Research,* 15:887-897.

Halpern, D.L. and Blake, R.R. (1988) How contrast affects stereoacuity. *Perception*, 17:483-495.

Halpern, D.L., Patterson, R. and Blake, R. (1987) What causes stereoscopic tilt from spatial frequency disparity? *Vision Research*, 27:1619-1629.

Halpern, D.L., Wilson, H.R, and Blake, R. (1996) Stereopsis from interocular spatial-frequency differences is not robust. *Vision Research*, 36:2263-2270.

Hamilton, D.B., Albrecht, D.G. and Geisler, W.S. (1989) Visual cortical receptive field in monkey and cat: Spatial and temporal phase transfer function. *Vision Research*, 29:1285-1308.

Hänny, P., von der Heydt, R. and Poggio, G.F. (1980) Binocular neuron responses to tilt in depth in the monkey visual cortex. Evidence for orientation disparity processing. *Experimental Brain Research*, 41:A26.

Heeger, D.J. (1992) Normalization of cell responses in cat striate cortex. *Visual Neuroscience*, 9:181-198.

Heeley, D.W. and Buchanan-Smith, H.M. (1994) Evidence for Separate, Task Dependent Noise Processes in Orientation and Size Perception. *Vision Research*, 34(16),2059-2069.

Heeley, D.W. and Timney, B. (1988) Meridional anisotropies of orientation discrimination for sine wave gratings. *Vision Research*, 28(2),337-345.

Heeley, D.W. and Timney, B. (1989) Spatial frequency discrimination at different orientations. *Vision Research*, 29,1221-1228.

Helmholtz, H. von (1909) *Physiological Optics*, Republished by the Optical Society 1924 edition; translated from the third German edition (New York: Dover).

Henning, G.B., Hertz, B.G. and Broadbent, D.E. (1975) Some experiments bearing on the hypothesis that the visual system analyses spatial patterns in independent bands of spatial frequency. *Vision Research*, 15:887-898.

Hess, R.F. and Wilcox, L.M. (1994) Linear and non-linear filtering in stereopsis. *Vision Research*, 34(18):2431-2438.

Hibbard, P.B. and Langley, K. (1994) Using local phase information to detect image transformations. *Ophthalmic and Physiological Optics*, 14:438.

Hibbard, P.B., Langley, K and Fleet, D.J. (1994) A computational model for stereoscopic slant based upon orientational differences of Fourier and non-Fourier mechanisms. *Perception*, 23 (suppl):35.

Hirsch, J. and Hylton, R. (1982) Limits of spatial frequency discrimination as evidence of neural interpolation. *Journal of the Optical Society of America*, 72:1367-1374.

Horn, B.K.P. (1986) *Robot Vision*, London, MIT Press.

Horn, B.K.P. and Schunk, B.G. (1981) Determining Optic Flow. *Artificial Intelligence*, 17:185-203.

Howard, I.P. and Kaneko, H. (1994) Relative Shear Disparities and the Perception of Surface Inclination. *Vision Research*, 34(9):2505-2517.

Howard, I.P., Ohmi, M. and Sun., L. (1993) Cyclovergence: a comparison of objective and psychophysical measurements. *Experimental Brain Research*, 97:349-355.

Howard, I.P. and Rogers, B.J. (1995) *Bincocular vision and stereopsis.* Oxford Psychology Series, No 29, Oxford University Press.

Hubel, D.H. and Wiesel, T.N. (1962) Receptive fields, binocular interaction and fundamental architecture in the cat's visual cortex. *Journal of Physiology (London)*, 150:106-154.

Jenkin, M.R.M. and Jepson, A.D. (1991) Techniques for disparity measurement. *CVGIP: Image Understanding*, 43:14-30.

Jones, D.G. and Malik, J. (1992) Determining 3-dimensional shape from orientation and spatial-frequency disparities. *Lecture Notes in Computer Science*, 588:661-669.

Jones, J.P. and Palmer, L.A. (1987) The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *Journal of Physiology (London)*, 58:1187-1211.

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

Julesz, B. (1971) *Foundations of cyclopean perception.* Chicago, University of Chicago Press.

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

Kaneko, H. and Howard, I. (1994) Compression magnification disparities and the perception of slant. *Perception*, 23(Suppl):33.

Kersten, D. (1991) Transparency and cooperative computation of scene attributes. In *Computational Models of Visual Processing.* M. Landy and J.A. Movshon. (eds) MIT Press. London.

Knutsson, H. (1982) *Filtering and reconstruction in image processing.* PhD dissertation, Department of Electrical Engineering, Linkoping University.

Koenderink, J.J. and Van Doorn, A.J. (1975) Invariant properties of the motion parallax field due to the movement of rigid bodies relative to the observer. *Optica Acta*, 22:773-791.

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

Kohler, W. and Emery, D.A. (1947) Figural after-effects in the third dimension of visual space. *American Journal of Psychology*, 60:159-201.

Kontsevich. L.L. and Tyler, C.W. (1994) Analysis of Stereothresholds for Stimuli Below 2.5c/deg. *Vision Research*, 34:2317-2329.

Langley, K., Atherton, T.J., Wilson, R.G. and Larcombe, M.H.E. (1990) Vertical and horizontal disparities from phase. *Proceedings ECCV*, Faugeras, O. (ed), Springer-Verlag.

Langley, K. and Fleet, D.J. (1995) A model for for coherent and multiplicatively transparent plaids. *Investigative Ophthalmology and Physiological Optics*, 36(Suppl): S1048.

Langley, K., Fleet, D.J. and Hibbard, P.B. (1995) Scale dependence of transparent random-dot stereograms. *Perception*, 24 (Suppl): 137.

Langley, K., Fleet, D.J. and Hibbard, P.B. (1996) Linear filtering precedes nonlinear processing in early vision. *Current Biology*, 6:891-896.

Legge, G.E. (1981) A power law for contrast discrimination. *Vision Research*, 21:457-467.

Legge, G.E. and Gu, Y. (1987) Stereopsis and contrast. *Vision Research*, 8:989-1004.

Lin, L. and Wilson, H.R. (1995) Stereoscopic integration of Fourier and non-Fourier patterns. *Investigative Ophthalmology and Visual Science* (Suppl),36:S364.

Liu, A., Gaska, J.P., Jacobson, L.D. and Pollen, D.A. (1992) Interneuronal interaction between members of quadrature phase and anti-phase pairs in the cat's visual cortex. *Vision Research*, 32:1193-1198.

Liu, L. Schor, C.W and Ramachandran, V.S. (1992) Positional disparity is more efficient in encoding depth than phase disparity. *Investigative Ophthalmology and Visual Science* (Suppl), 33(1373).

Longuet-Higgins, H.C. and Prazdny, K. (1981) The interpretation of a moving retinal image. *Proceedings of the Royal Society of London* ,B, 208:385-387.

Longuet-Higgins, H.C. (1982) The role of the vertical dimension in stereoscopic vision. *Perception*, 11:377-386.

Mansfield, J.S. and Parker, A.J. (1994) An orientation-tuned component in the contrast masking of stereopsis. *Vision Research*, 33:1535-1544.

Marcelja, S. (1980) Mathematical description of the responses of simple cortical cells. *Journal of the Optical Society of America* A, 70:1297-1300.

Marr, D. and Poggio, T. (1979) A Computational Theory of Human Stereo Vision. *Proceedings of the Royal Society of London*, B, 204:301-328.

Mayer, M.J. and Kim, C.B.Y. (1986) Smooth frequency discrimination functions for foveal, high contrast, mid spatial frequencies. *Journal of the Optical Society of America* A, 3:1957-1969.

Mayhew, J.E. and Frisby, J.P. (1980) The computation of binocular edges. *Perception*, 9:69-86.

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

McKee, S.P. (1983) The spatial requirement for fine stereoacuity. *Vision Research*, 23:191-198.

Metelli, F. (1974) The perception of transparency. *Scientific American*, 230:90-98.

Mitchison, G.J. and Mckee, S.P. (1990) Mechanisms underlying the anisotropy of stereoscopic tilt perception. *Vision Research*, 30:1781-1791.

Mitchison, G.J. and Westheimer, G. (1990) Viewing geometry and gradients of orizontal disparity. In *Vision: coding and efficiency*, (ed C. Blakemor) pp302-309. Cambridge University Press.

Morgan, M.J. and Castet, E. (1995) Stereoacuity for oblique gratings predicted by phase shifts, not disparities. *Investigative Ophthalmology and Visual Science* (Suppl), 36:S231.

Movshon, J.A and Blakemore, C. (1973) Orientation specificity and spatial selectivity in human vision. *Perception*, 2.

Movshon, J.A. Thompson, I.D. and Tolhurst, D.J. (1978) Spatial and temporal contrast sensitivity of neurons in areas 17 and 18 of the cat's visual cortex. *Journal of Physiology (London)*, 283:101-120.

Movshon, J.A., Adelson, E.H., Gizzi, M.S., and Newsome, W.T. (1985) The analysis of moving visual patterns. In *Pattern recognition mechanisms,* (ed. C. Chigas, R. Gattas, and C. Gross), pp. 117-151. Springer-Verlag, New York.

Nakayama, K. and Silverman, G.H. (1985)Detection and discrimination of sinusoidal

grating displacements *Journal of the Optical Society of America* A(2) 267-274.

Nakayama, K., Shimojo, S. and Ramachandran, V.S. (1990) Transparency: relation to depth, subjective contours, luminance, and neon color spreading. *Perception*, 19:497-513.

Nelson, J.I. (1975) Globality and stereoscopic fusion in binocular vision. *Journal of Theoretical Biology*, 49:1-80.

Nelson, J.I., Kato, H. and Bishop, P.O. (1977) Discrimination of orientation and position disparities by binocularly activated neurons in cat striate cortex. *Journal of Neurophysiology*, 40:260-283.

Ninio, J. (1985) Orientational versus horizontal disparity in the stereoscopic appreciation of slant. *Perception*, 14:305-314.

Ogle, K.N. (1938) Induced size effect. I. A new phenomenon in binocular space-perception associated with the relative sizes of the images of the two eyes. *AMA Archives of Ophthalmology*, 20:604-623.

Ogle, K.N. (1950) *Researches in binocular vision*. Philadelphia: W B Saunders.

Ohzawa, I., DeAngelis, G. and Freeman, R.D. (1990) Stereoscopic depth discrimination in the visual cortex: Neurons ideally suited as disparity detectors. *Science*, 249:1037-1041.

Ohzawa, I. and Freeman, R.D. (1986) The binocular organization of simple cells in the cat's visual cortex. *Journal of Neurophysiology*, 56:243-259.

Orban, G. VandenBussche, E and Vogels, R. (1984) Human orientation discrimination tested with long stimuli. *Vision Research*, 24:121-128.

Palmer, L.A. and Davis, T.L. (1981) Receptive field structure in cat striate cortex. *Journal of Neurophysiology*, 46:260-276.

Parker, A.J., Johnston, E.B., Mansfield, J.S. and Yang, Y (1991) Stereo, surfaces and shape. In *Computational Models of Visual Processing*. M. Landy and J.A. Movshon. (eds) MIT Press. London.

Pastore, N. (1974) Binocular depth perception. *American Scientist*, 62:262.

Piggins, D.(1978) Moirés maintained internally by binocular fusion. *Perception*, 7:679-681.

Poggio, G.F., and Fischer, B. (1977). Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. *Journal of Neurophysiology*, 40:1393-1405.

Pollen, D.A. and Ronner, S. (1981) Phase relationships between adjacent simple cells in the visual cortex. *Science*, 212:1409-1411.

Pollen, D.A. and Ronner, S. (1983) Visual cortical neurons as localized spatial frequency filters. *IEE transactions of systems, man and cybernetics*, 13:907-916.

Press, W.H., Flannery, B.P., Teukolsky, S.A., and Vetterlin, W.T. (1992) *Numerical Recipes in C: the art of scientific computing*, Cambridge: Cambridge University Press.

Ramachandran, V.S. and Cavanagh, P. (1985) Subjective contours capture stereopsis. *Nature*, 317:527-530.

Rao, C.R. (1971) *Generalized inverse of matrices and its applications.* New York, Wiley.

Rogers, B.J. (1992) The perception and representation of depth and slant in stereoscopic surfaces. In *Artifical and biological vision systems*, (ed. G.A. Orban and H.H Nagel),pp. 241-266. Springer-Verlag, Berlin.

Rogers, B.J. and Bradshaw, M.F. (1994) Is dif-frequency a stimulus for stereoscopic slant? *Investigative Ophthalmology and Visual Science* (Suppl), 35:1316.
Rogers, B.J. and Bradshaw, M.F. (1996) Disparity scaling and the perception of frontoparallel surfaces. *Perception*, 24:155-179.

Rogers, B.J and Cagenello, R. (1989) Disparity curvature and the perception of three-dimensional surfaces. *Nature*, 339:135-137.

Rogers, B.J. and Graham, M. (1983) Anisotropies in the perception of three-dimensional surfaces. *Science*, 221:1409-1411.

Ryan, C. and Gillam, B. (1993) A proximity-contingent stereoscopic depth aftereffect: evidence for adaptation to disparity gradients. *Perception*, 22:403-418.

Ryan, C. and Gillam, B. (1994) Cue conflict and stereoscopic slant about horizontal and vertical axes. *Perception,* 23:645-658.

Sanger, T.D. (1988) Stereo disparity computation using Gabor filters. *Biological Cybernetics*, 59:405-418.

Sato, T. and Nishida, S. (1993) Second order depth perception with texture-defined random check stereograms. *Investigative Ophthalmology and Visual Science* (Suppl), 34:1438.

Sato, T. and Nishida, S. (1994) Does an envelope detecting mechanism mediate stereopsis for bandlimited stimuli? *Investigative Ophthalmology and Visual Science*, 35 (Suppl):1916.

Schor, C.M and Wood, I. (1983) Disparity Range for Local Stereopsis as a Function of Luminance Spatial Frequency. *Vision Research,* 23:1649-1654.

Scott-Samuel, N.E. and Georgeson, M.A. (1995) Does early nonlinearity account for second-order motion. *Perception* 24(suppl):104.

Sherrington, C.S. (1906) *Integrative action of the nervous system,* New Haven: Yale University Press.

Simoncelli, E. and Heeger, D. (1992)A computational model for perception of 2-dimensional pattern velocities *Investigative Ophthalmology and Visual Science,* 33(4):954.

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

Smallman, H.S. and Harris, J.M. (1996) Nonlinear visual distortion: an effective nonlinearity from asymmetry in on and off pathways. *Investigative Ophthalmology and Visual Science,* 27 (suppl):S232.

Smith, A.T. and Edgar, G.K. (1990) The influence of spatial frequency on perceived temporal frequency and perceived speed. *Vision Research,* 30:1467-1474.

Smith, A.T. and Edgar, G.K. (1991) Perceived speed and direction of complex gratings and plaids. *Journal of the Optical Society of America* A, 8:1161-1171.

Spitzer, H. and Hochstein, S. (1985) A complex-cell receptive-field model *Journal of neurophysiology,* 53:1266-1286.

Stevens, K.A. (1983) Surface tilt (the direction of slant): a neglected psychophysical variable. *Perception and Psychophysics,* 33:241-250.

Stevens, K.A and Brookes, A. (1987) Depth reconstruction in stereopsis. In *Proceedings of the first IEEE International Conference in Computer Vision*, 682-686. IEEE Computer Society.

Stoner, G.R., and Albright, T.D. (1994) Visual motion integration: a neurophysiological and psychophysical perspective. In *Visual Detection of Motion* Smith, A.T. and Snowden, R.J. (eds), Academic Press, London.

Stoner, G.R., Albright, T.D. and Ramachandran, V.S. (1990) Transparency and coherence in human motion perception. *Nature*, 344:153-155.

Stromeyer, C.F. III, Klein, S., Dawson, D.M. and Spillmann, L. (1982) Low spatial-frequency channels in human vision: adaptation and masking. *Vision Research*, 22:225-233.

Swash, S.A., Rogers, B.J. Bradshaw, M.F. and Cagenello, R. (1995) The role of cyclovergence in the perceived slant of stereoscopic images related by vertical shear, rotation and deformation. *Investigative Ophthalmology and Visual Science* Suppl),36:S369.

Thomas, J.P. (1983) Underlying psychometric function for detecting gratings and identifying spatial frequency. *Journal of the Optical Society of America*, 73:751-758.

Thompson, P. (1982) Perceived rate of movement depends on contrast. *Vision Research*, 22:377-380.

Treisman, M. and Watts, T.R. (1966) Relation between signal detectability theory and the traditional procedures for measuring thresholds: estimating $d^'$ from results given by the method of constant stimuli. *Psychological Bulletin*, 66(6):438-454.

Tyler, C.W. (1971) Stereoscopic depth movement: two eyes less sensitive than one. *Science*, 174:858-961.

Tyler, C.W. (1975) Characteristics of stereomovement suppression. *Perception and Psychophysics*, 17:225-230.

Tyler, C.W. (1990) A stereoscopic view of visual processing streams. *Vision Research*, 30:1877-1895.

Tyler, C.W. and Cavanagh. P. (1989) Purely chromatic stereomotion perception. *Investigative Ophthalmology and Visual Science*, 30 (Suppl):324.

Tyler, CW. and Sutter, E.E. (1979) Depth from spatial frequency differences,: An old type of stereopsis? *Vision Research.*, 19:859-865.

Vallortigara, G. and Bressan, P. (1994) Occlusion, transparency, and stereopsis: a new explanation for stereo capture. *Vision Research,* 34:2891-2896.

Victor, J.D. and Conte, M.M. (1992) Coherence and transparency of moving plaids composed of Fourier and non-Fourier gratings. *Perception and Psychophysics,* 52:403-414.

Volkmann, A.W. (1864) P*hysiologische Untersuchungen in Gebiete der Optik*, Heft 2, Leipzig.

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

von der Heydt, R. Hänny, P. and Dürsteller, M.R. (1981) The role of orientation disparity in stereoscopic perception and the development of binocular correspondence. *Advances in Physiological Science*, 16:461-470.

Von Grünau, M., Dabé, S. and Kwas, M. (1993) The effect of disparity on motion coherence. *Spatial Vision*, 7:227-241.

Wallach, H. (1935) Uber visuell wahrgenomme Bewegungrichtung. P*sychologische Forschung,* 20:325-380.

Wallach, H. and Bacon, J. (1976) Two forms of retinal disparity. *Perception and Psychophysics*, 19:375-382.

Watt, R.J. and Andrews, D.P. (1981) APE: Adaptive Probit Estimation of psychometric functions. *Current Psychological Reviews*, 1:205-214.

Weinshall, D. (1990) Seeing `ghost' planes in stereo vision. *Vision Research,* 31:1731-1749.

Welch, L. (1989) The perception of moving plaids reveals two motion-processing stages. *Nature*, 337:734-736.

Wenderoth, P.M. (1970) A visual spatial aftereffect of surface slant. *American Journal of Psychology*, 83:576-590.

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

Wilcox, L.M. and Hess, R.F. (1993) Linear and non-linear contributions to stereopsis. *Investigative Ophthalmology and Visual Science (Suppl)*, 34:1187.

Wilcox, L.M. and Hess, R.F. (1994) Is the site of non-linear filtering in stereopsis before or after binocular combination *Investigative Ophthalmology and Visual Science* Suppl), 35:1490.

Wilcox, L.M. and Hess, R.F. (1995) Stereoacuity for amplitude modulated stimuli. *Investigative Ophthalmology and Visual Science* (Suppl), 36:S365.

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

Wildes, R.P. (1991) Direct recovery of three-dimensional scene geometry from binocular stereo disparity. *IEEE transactions on Pattern Analysis and Machine Intelligence*, 13:761-774.

Wilson, H.R. (1994) Models of two-dimensional motion perception. In *Visual Detection of Motion Smith*, A.T. and Snowden, R.J. (eds), Academic Press, London.

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

Wilson, H.R. Ferrera, V.P. and Yo, C.(1992) A psychophysically motivated model for two-dimensional motion perception. *Visual Neuroscience*, 9(1):79-97.

Wilson, H.R., Levi, D., Maffei, L., Rovamo, J. and DeValois, R.L. (1990) The perception of form: retina to striate cortex. In *Visual Perception: the europhysiological foundations* Spillman, L., Werner, JS (eds), San Diego, Academic.

Yamamoto, M. (1989) A general aperture problem for direct estimation of 3-d motion parameters. *IEEE transactions on pattern analysis and machine vision*, 11:528-536.

Yang, Y. and Blake, R. (1991) Spatial frequency tuning of human stereopsis. *Vision Research*, 31:1177-1189.

Zhou, Y.X. and Baker, C.L. (1993) Envelope-responsive neurons in area-17 and area-18 of cat *Journal of Neurophysiology*, 72:2134-2150.

A. Disparities in Slanted Surfaces

1 Stimulus generation and analysis of disparities

Stereograms were generated from an initial sinusoidal grating or plaid image. Let (x_p, y_p) be the position of a point in the original image. Let (x_l, y_l) and (x_r, y_r) be the position of this point in the left- and right-eye images respectively. The three binocular transformations used in the experiments were

Shear:

$$x_l = x_p + \tau y_p \qquad (1)$$

$$y_l = y_p \qquad (2)$$

$$x_r = x_p - \tau y \qquad (3)$$

$$y_r = y_p \qquad (4)$$

Expansion-Compression:

$$x_l = x_p + \tau x_p \quad (5)$$

$$y_l = y_p \quad (6)$$

$$x_r = x_p - \tau x_p \quad (7)$$

$$y_r = y_p \quad (8)$$

Rotation:

$$x_l = x_p + \tau y_p \quad (9)$$

$$y_l = y_p - \tau x_p \quad (10)$$

$$x_r = x_p - \tau y_p \quad (11)$$

$$y_r = y_p + \tau x_p \quad (12)$$

We consider the effects of this transformation on an original sinusoidal grating with orientation (relative to horizontal) θ and spatial frequency f:

$$I(x,y) = \sin(2\pi f(x_p \sin\theta + y_p \cos\theta)) \quad (13)$$

Binocular image pairs are then given by

$$I_l(x,y) = \sin(2\pi f(x_l \sin\theta + y_l \cos\theta)) \quad (14)$$

$$I_r(x,y) = \sin(2\pi f(x_r \sin\theta + y_r \cos\theta)) \quad (15)$$

$$(16)$$

with (x_l, y_l) and (x_r, y_r) given by equations (1- 9). All three transformations result in binocular pairs of sinusoidal gratings, with differing orientations and spatial

frequencies. From equations (1-9), and equation (14), we get the orientation and spatial frequency disparities for the three transformations:

Shear:

$$\begin{aligned} \Delta\theta &\approx \frac{1}{2}\tau(1-\cos 2\theta) \\ \frac{\Delta f}{f} &\approx \tau \sin\theta \cos\theta \end{aligned} \tag{17}$$

Expansion-Compression:

$$\begin{aligned} \Delta\theta &\approx \frac{1}{2}\tau \sin 2\theta \\ \frac{\Delta f}{f} &\approx \frac{1}{2}\tau \sin^2\theta \end{aligned} \tag{18}$$

Rotation:

$$\Delta\theta = \tau \tag{19}$$

$$\frac{\Delta f}{f} = 0 \tag{20}$$

We consider also the disparity gradients generated by the transformations. We consider gradients of phase disparity, rather than horizontal disparity. Morgan and Castet (1995) showed that, for one dimensional stimuli, disparity sensitivity is determined by disparities orthogonal to the image contours. Phase disparity is given by

$$d_\phi = \frac{\sqrt{(x_r - x_l)^2 + (y_r - y_l)^2}}{\sin\theta} \tag{21}$$

The gradient of phase disparity for each of the three transformations, $\nabla[d_\phi]$ is then given by:

Shear:

$$\nabla[d_\phi] = \tau f \sin\theta \tag{22}$$

Expansion-Compression:

$$\nabla[d_\phi] = \tau f \sin\theta \tag{23}$$

Rotation:

$$\nabla[d_\phi] \approx \tau f \tag{24}$$

where a first order small angle approximation has been used in the latter case.

2 Predicting slant thresholds

Equations (19-24) give the magnitude of orientation disparity, spatial frequency disparity and disparity gradient produced by a given transformation. They also give the magnitude of transformation that will generate a threshold binocular difference. As such, they may be used to predict slant thresholds, if it is assumed that they are dependent on a single disparity type. We consider a simple model in which slant detection makes use of a combination of different disparity cues. We assume that the detection of orientation disparities, spatial frequency disparities and disparity

gradients are independent, and that errors in the measurements of these parameters are uncorrelated. We adopt a simple model in which sensitivity to a combination of disparity cues is given by the Euclidean summation of the sensitivities to the individual cues:

$$d'_{comb} = \left[{d'_1}^2 + {d'_2}^2 + \ldots + {d'_n}^2 \right]^{\frac{1}{2}} \tag{25}$$

A second assumption we make is that discrimination thresholds are inversely proportional to sensitivity (Treisman and Watts, 1966). This allows us to use equation (25) to predict slant thresholds given a combination of cues on the basis of thresholds for the individual cues:

$$\tau_T = \left[\tau_{T1}^{-2} + \tau_{T2}^{-2} + \ldots + \tau_{Tn}^{-2} \right]^{-\frac{1}{2}} \tag{26}$$

These expressions were used in the curve fits. We assume a model in which orientation disparity is combined with either disparity gradients, or spatial frequency disparities. We assume that both cues contribute to the perception of both horizontal and vertical disparity gradients, and allow for unequal sensitivity to the two cues. The model was fit to the data for both shear and expansion- compression simultaneously. The fitted curves are given by

$$\tau_T = \begin{cases} \left[\left(\frac{1-\cos 2\theta}{\sigma_\theta^2} \right)^2 + \left(\frac{\sin 2\theta}{\sigma_f^2} \right)^2 \right]^{-\frac{1}{2}} & \text{for shear} \\ \left[\left(\frac{\sin 2\theta}{\sigma_\theta^2} \right)^2 + \left(\frac{1-\cos 2\theta}{\sigma_f^2} \right)^2 \right]^{-\frac{1}{2}} & \text{for expansion} - \text{compression} \end{cases} \tag{27}$$

$$\tau_T = \begin{cases} \left[\left(\frac{1-\cos 2\theta}{\sigma^2_\theta}\right)^2 + \left(\frac{\sin\theta}{\sigma^2_{d_\phi}}\right)^2\right]^{-\frac{1}{2}} & \text{for shear} \\ \left[\left(\frac{\sin 2\theta}{\sigma^2_\theta}\right)^2 + \left(\frac{\sin\theta}{\sigma^2_{d_o}}\right)^2\right]^{-\frac{1}{2}} & \text{for expansion} - \text{compression} \end{cases} \quad (28)$$

These curves were fitted to both the grating and plaid data using the Sigma-Plot non-linear curve fitter.

Orientation disparities at threshold for the discrimination of surfaces defined by gratings, for the shear, rotation and expansion conditions.

	Orientation Disparity								
	PH			KL			JB		
Orientation	Shear	Rot	Exp.	Shear	Rot	Exp.	Shear	Rot	Exp.
10	1.98	-		1.28	-	-	4.99	5.84	-
15	3.48	7.0	1.88	2.76	3.78	1.31	7.20	-	2.57
20	5.05	5.6	-	0.71	2.45	-	11.86	-	-
30	2.96	5.6	1.56	2.70	1.04	1.67	3.40	13.5	3.68
45	5.19	9.0	1.15	-	1.36	1.31	-	18.9	2.91
60	-	7.6	1.28	-	1.27	0.78	-	-	3.26
67.5	4.03	-	-	3.37	-	-	-	-	-
75	-	-	1.31	-	1.43	0.53	-	6.16	2.04
90	5.2	5.2	0	0.94	0.94	0	6.15	6.15	0
Mean	3.98	6.67	1.20	1.96	1.75	0.93	6.72	10.11	2.41

Orientation disparities at threshold for the discrimination of surfaces defined by plaids, for the shear and expansion conditions.

	Orientation Disparity					
	PH		KL		JB	
Orientation	Shear	Expansion	Shear	Expansion	Shear	Expansion
5	0.271	1.091	0.134	0.231	0.386	0.341
10	0.863	1.46	0.050	0.285	0.384	0.623
15	0.764	2.18	.076	0.361	0.445	0.814
20	1.146	2.75	0.141	0.311	-	0.877
25	1.708	2.70	0.166	0.429	1.470	0.998
30	1.75	3.35	0.200	0.409	2.057	0.958
35	2.187	3.07	1.28	0.264	-	0.673
40	2.92	2.17	1.63	0.176	2.21	0.730
Mean	1.451	2.346	0.46	0.308	1.159	0.752

B. Transparency in Squarewave Plaids

Stoner et al. (1990) showed that the transparent motion observed in their stimuli was consistent with that of two sets of thin bars, with the characteristics of neutral density filters, in front of a light background. Let I_1, I_2 and I_3 represent the intensities of the background, bars and intersections, respectively. Let L_b and L_l represent the luminances of the background and the bars, respectively, and let τ represent the transmittance of the bars. We can then write:

$$I_1 = L_b \qquad (1)$$

$$I_2 = \tau L_b + L_l \qquad (2)$$

$$I_3 = \tau^2 L_b + \tau L_l + L_l \qquad (3)$$

to desribe the background, bar and intersection luminances in terms of the intensities of the background and bars, and the transmittance of the bars. Another form of

transparency is however possible. Although a small duty cycle was chosen to bias figure-ground assignments, we consider the case in which the light vertical bars represent a transparent medium throught which the grey horizontal bars are viewed. The dark "intersections" now represent these horizontal bars, while the grey vertical bars represent the background. Let the background, horizontal bars and vertical bars have luminances of L_1, L_2 and L_3, and the transparent bars have a transmittance τ. The three intensity regions in the image are then formed as:

$$I_1 = \tau L_1 + L_3 \tag{4}$$

$$I_2 = \tau L_2 + L_3 \tag{5}$$

$$I_2 = L_1 \tag{6}$$

$$I_3 = L_2 \tag{7}$$

Using the constraints $0 \leq \tau \leq 1$, $I_1 > I_2$ and $I_3 \geq 0$, it can be shown that:

$$I_3 \leq I_2 \tag{8}$$

$$I_2 \geq \frac{1}{2} I_1 \tag{9}$$

Stoner et al. (1990) used values of $I_1 = 200\text{cdm}^{-2}$ and $I_2 = 90\text{cdm}^{-2}$, which do not satisfy (9) above. Values of $I_1 = 70\text{cdm}^{-2}$ and $I_2 = 35\text{cdm}^{-2}$ were used here. These values satisfy all the constraints given above, and thus were able to support both transparent interpretations of the stimuli.