Pergamon

PII: S0042-6989(97)00275-7

Vision Res., Vol. 38, No. 8, pp. 1073–1084, 1998

Printed in Great Britain
0042-6989/98 $19.00 + 0.00

Plaid Slant and Inclination Thresholds can be Predicted from Components*

PAUL B. HIBBARD, KEITH LANGLEY†‡

Received 14 June 1996; in revised form 12 May 1997; in final form 30 July 1997

We investigated whether stereoscopic slant and inclination thresholds for surfaces defined by two component plaids could be predicted from the interocular differences in their individual component gratings. Thresholds were measured for binocular images defined by single sinusoidal gratings and two component plaids. In both cases thresholds showed a marked dependence on component orientation. For absolute component orientations greater than 45 deg we found that inclination thresholds were smaller than slant thresholds. However, for absolute component orientations less than 45 deg, we found a reversal: slant thresholds were smaller than inclination thresholds. We considered three models that might account for these data. One assumed that thresholds stemmed from interocular position differences of corresponding image points. The other two assumed a combination of position, orientation and/or spatial-frequency differences. The best fits were obtained from those models that explicitly represented orientation differences. From the model combining orientation and spatial-frequency differences, we estimated the relative cue sensitivity to be 1.7:1, respectively. For plaids, we found that thresholds obtained from the individual components could be used to predict thresholds for plaids, even though an additional disparity cue from the contrast beat was available. © **1998 Elsevier Science Ltd. All rights reserved.**

Orientation disparity Spatial-frequency disparity Slant anisotropy Stereopsis

INTRODUCTION

One remarkable feature of stereopsis is the anisotropy that exists in sensitivity to surfaces for which depth changes as a function of image position (Rogers & Graham, 1983; Gillam, Flagg & Finlay, 1984; Mitchison & McKee, 1990; Cagenello & Rogers, 1993). Thresholds for surfaces rotated in depth about a horizontal axis (referred to as inclined surfaces) are, in general, smaller than those for surfaces rotated in depth about a vertical axis (slanted surfaces). Although there are marked individual differences in the magnitude of the anisotropy, only approximately 10% of individuals are equally sensitive to both slanted and inclined surfaces (Rogers, 1997, pers. comm.). This anisotropy is not a unique feature of binocular processing. Rogers and Graham (1982) have shown a similar anisotropy for surfaces defined by motion parallax. This suggests in part that there are similarities between the mechanisms which detect spatial changes in surface depth from binocular stereopsis and visual motion. Moreover, while Mitchison and McKee (1990) have suggested that this anisotropy may be related to the existence of an internal reference for inclined surfaces, the anisotropy may also reveal properties of the underlying mechanism that processes binocular disparities.

This paper considers stereoscopically defined inclination and slant. Binocular disparities for these surfaces may be described by the transformations of horizontal shear and horizontal compression (Koenderink & van Doorn, 1976). Inclination introduces a vertical gradient of horizontal disparity (horizontal shear) to binocular images. Similarly, slant will introduce a horizontal gradient of horizontal disparity (horizontal compression). These gradients of disparity may be described by an affine transformation, comprising the scalar quantities of rotation and dilation, and the vector quantities of translation and deformation. Equally, they may be described in terms of interocular differences in orientation and spatial frequency of the Fourier components of the image.

Binocular differences in orientation have proved a useful representation to account for the anisotropy. This may be understood with reference to Fig. 1(A). This figure shows how binocular differences in orientation vary with orientation, for both slanted and inclined surfaces, with equal magnitude of disparity gradient. Figure 1(B) shows how binocular differences in spatial frequency vary as a function of orientation. These curves are reversed in comparison with Fig. 1(A). In Fig. 1(C) we show how binocular differences in phase would

*This research was first presented at ARVO 1995.

†To whom all correspondence should be addressed [*Fax*: +44 1714364276; *Email*: k.langley@psychol.ucl.ac.uk].

‡Department of Psychology, University College London, London, U.K.

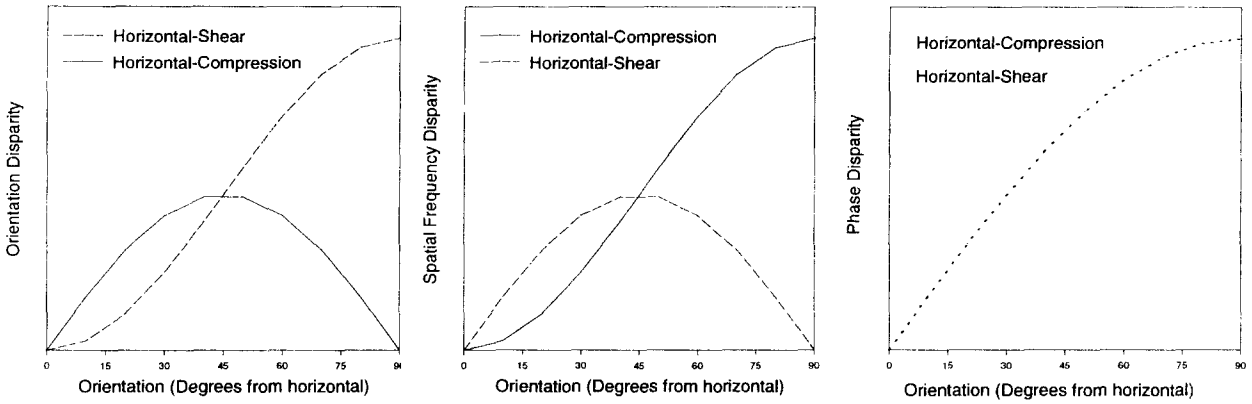

FIGURE 1. The magnitude of orientation, spatial-frequency and phase differences as a function of absolute orientation, for a single sinusoidal grating subject to a constant magnitude of horizontal shear and horizontal compression (see Appendix). (A) Magnitudes of orientation differences. The horizontal shear transformation yields maximum orientation differences for vertically oriented gratings, while for the horizontal compression transformation the maximum occurs at ±45 deg. (B) Magnitudes of spatial-frequency differences are a reversal of those of orientation differences. (C) Magnitudes of phase differences. This model predicts no difference in discrimination thresholds between the two transformations considered.

similarly change as a function of orientation for a constant disparity gradient. The curves for gradients of phase are equal for both transformations. From these figures, it can be seen that binocular orientation differences for inclined surfaces are greater than those for slanted surfaces when the absolute orientation of the image stimuli exceeds 45 deg. For slanted surfaces binocular orientation differences are greatest for orientations of ±45 deg. Also, for orientations less than 45 deg the binocular orientation differences of slanted surfaces exceed those of inclined surfaces. If binocular differences in orientation are exploited by the visual system, then the anisotropy can be explained by the magnitude of orientation differences as a function of absolute orientation. One would predict for instance, smaller thresholds for inclined surfaces than for slanted surfaces for binocular images composed of vertically oriented structure. For images composed of oriented structure at ±45 deg one would predict no difference in thresholds. Cagenello and Rogers (1993) have verified the predictions made from an orientation difference model. They concluded that the anisotropy for slanted and inclined surfaces was consistent with a mechanism that exploited interocular orientation differences. Cagenello and Rogers' insight has support from Blakemore, Fiorentini, and Maffei's (Blakemore, Fiorentini & Maffei, 1972) finding that binocular cells in the cat primary visual cortex receive input from monocular cells with differences in orientation tuning between left- and right eyes (see also Nelson, Kato & Bishop, 1977). However, these data only expose likely candidates that may explain the differences in perceived slant and inclination thresholds. Direct evidence for an explicit mechanism that mediates binocular orientation differences remains elusive (Cagenello & Rogers, 1993).

This paper further explores the anisotropy reported by Rogers and Graham (1983) through three experiments. The first experiment may be viewed as a series of controls to establish the effects of contrast and spatial frequency on slant and inclination thresholds for sinusoidal grating and plaid patterns. In the second experiment we assessed the importance of orientation differences in comparison to other cues. To do this, we measured thresholds for slanted and inclined surfaces defined by single sinusoidal gratings as a function of the grating's cyclopean orientation. By fitting models to these data we were able to assess the relative sensitivity to different disparity cues. Three models were fitted to the data. These models embodied the hypotheses that thresholds for binocular slant and inclination could be explained by position differences, a combination of position and orientation differences or a combination of spatial frequency and orientation differences. Finally, we determined whether thresholds for gratings could be used to predict those for plaid patterns. We used the relative sensitivities to different disparity cues estimated from the model which best fit the data for gratings to find out whether these sensitivities could also explain the thresholds obtained for plaid patterns.

METHODS

Subjects

Three subjects were used. One was naïve to the purpose of the experiment. All subjects had normal or corrected to normal vision.

Stimulus generation and display

Stereoscopic images were presented on a colour monitor with a refresh rate of 76 Hz, using a grey scale

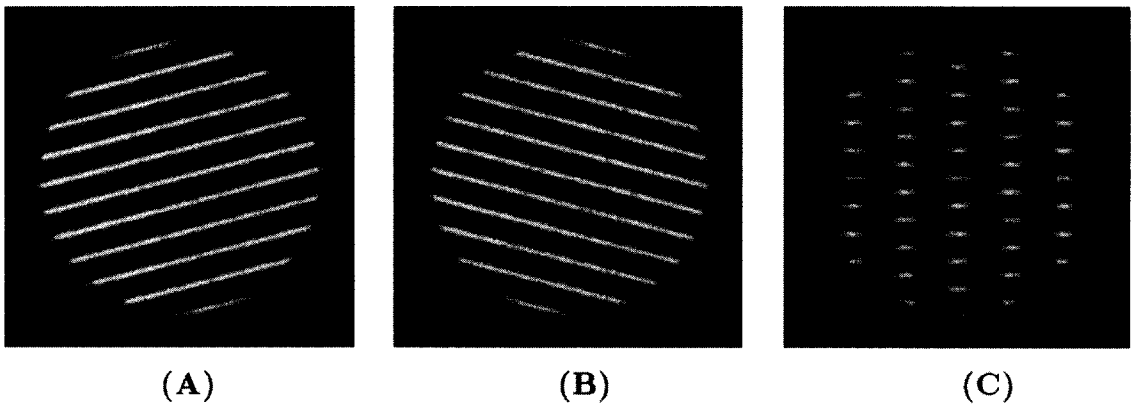

FIGURE 2. (A) and (B) Examples of the sinusoidal gratings used. (C) An example of one of the plaids used. The plaid image was produced from the sum of the two gratings shown in (A) and (B).

with 256 steps. The display had been carefully linearised by taking luminance measurements using a photometer, to which a logarithmic curve was fitted and used to generate a linear look-up table. The departure from linearity was 0.2%. The mean luminance of the images was 37.8 cdm^{-2}. Image pixels were arranged on a square lattice and had a width 2.0 min of visual angle. Stimuli were windowed in software using a circular aperture with a diameter of 4.7 deg of visual angle. The borders of each stimulus were softened with a gaussian window whose standard deviation was 0.94 deg. The viewing distance was fixed at 44 cm. Images were viewed using a modified Wheatstone stereoscope. Images were generated and stored in the RAM of a SUN SPARC 20 Workstation with 32 Mbytes of RAM.

Stimuli

Image stimuli were either single sinusoidal gratings, or plaid patterns generated from the sum of two sinusoidal gratings. For the plaid stimuli, the two component gratings had equal magnitudes of spatial frequency and contrast. The component gratings were always arranged so that, prior to binocular transformation, the plaid beat (or contrasts envelope,) was always vertical, and the plaid carrier (or average spatial frequency) was horizontal. Examples of the plaids used are shown in Fig. 2 and Fig. 3.

Grating and plaid stimuli were transformed using an affine model (see Appendix), so that the binocular images were consistent with either inclined or slanted surfaces. Perspective cues were not introduced into the images, so that thresholds could be related solely to image disparity gradients. As shown in Figs 1 and 9, both slant and inclination will introduce orientation and spatial-frequency disparities. One should also note that the magnitudes of these disparities for a given magnitude of slant or inclination vary as a function of absolute stimulus orientation. This relationship is shown explicitly in Fig. 1.

Procedure

Initially, a gaussian noise image was presented to each eye for 1 sec. This image appeared as frontoparallel, and served as a reference plane for judging the slant or inclination of the test images. An additional function of the noise stimulus was to prevent the build-up of depth

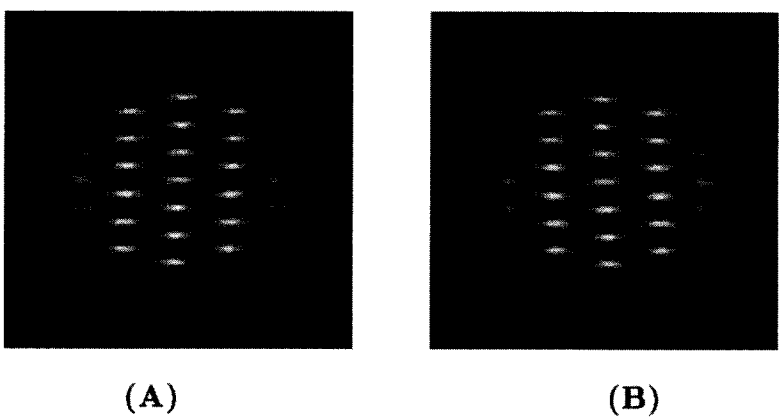

FIGURE 3. An example of the binocular plaid patterns used. The binocular images were generated by introducing a horizontal shear to the image shown in Fig. 2. Cross-eyed fusion of (A) and (B) reveals a surface inclined in depth.

aftereffects throughout each session. The monitor was then cleared to its mean luminance for 500 msec, after which a test stimulus was presented. The test remained visible until the subject responded. The screen was then cleared to the mean luminance value again for a further 500 msec, which completed each trial. Subjects were permitted to examine the stimuli, but instructed to respond quickly. This protocol was used because presentation time is a critical factor in the perception of binocular slant and inclination (Gillam, Chambers & Russo, 1988). This factor could have affected comparisons across different stimuli, and hence was left unconstrained during the experiments. It is, however, possible that an unconstrained viewing time could have introduced cyclovergence, which may affect disparity thresholds. However, Howard and Kaneko (1994) reported that cyclovergence does not significantly affect perceived slant judgements for stimuli that subtend less than 10 deg of visual angle, or for stimuli contained within a visible zero disparity surround. The stimuli used in this study subtended less than 7 deg. Moreover, the surrounding monitor was dimly visible. We thus concluded that cyclovergence was unlikely under these viewing conditions.

Test stimuli were single sinusoidal gratings or plaids. For the horizontal shear and rotation transformations, subjects were asked to classify surfaces as either a "ground plane", with the bottom appearing closer in depth, or a "sky plane" (top closer). For the horizontal compression transformation, subjects classified each stimulus as a "left wall" (left closer) or a "right wall" (right closer). The magnitude of slant or inclination was varied between trials using an adaptive probit algorithm (APE; Watt & Andrews, 1981). Discrimination thresholds were defined as the standard deviation of the psychometric function, which was estimated by APE. Each psychometric function was based on 64 individual trials. The data points, and standard error bars shown were obtained from three separate estimates for each condition and for each subject.

EXPERIMENT 1

The purpose of the first experiment was to determine how contrast and spatial frequency influence inclination thresholds for both plaid patterns and sinusoidal gratings. We conducted these experiments to consider the influence of these factors in subsequent experiments. We measured inclination thresholds for vertical sinusoidal gratings and for plaids. The orientations of the plaid components were fixed at ±25 deg relative to horizontal during all sessions.

For manipulations of contrast, the spatial frequencies of the plaid components and vertical sinusoidal gratings were fixed at 3.2 cycles/deg. For manipulations of spatial frequency, both the sinusoidal grating and the plaid beat's spatial frequency ranged between 0.5 and 2.0 cycles/deg.

Results

Figure 4 shows thresholds for both sinusoidal gratings and plaids as a function of contrast. The data are plotted on log–log axes. One can see from the data that the slopes for each subject are similar across conditions. Linear regression fits to the data across subjects yielded a mean slope of −0.47 for gratings, and −0.78 for plaids. These results suggest that both grating and plaid inclination thresholds show an inverse power relationship with contrast.

Figure 5 shows inclination thresholds as a function of spatial frequency for both sinusoidal gratings and for plaids. Results for the plaids are plotted against the spatial frequency of the plaid beats. While some

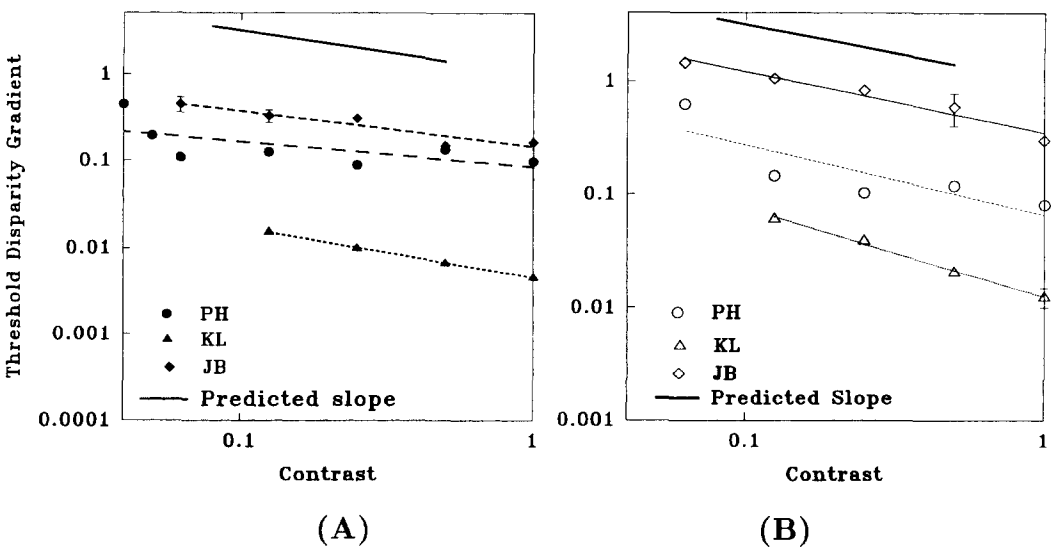

FIGURE 4. Inclination thresholds plotted as a function of contrast. (A) Gratings. (B) Plaids. The thick lines represent the slope predicted by the expected square-root relationship. Error bars (shown here and elsewhere) represent 1 standard error of the mean.

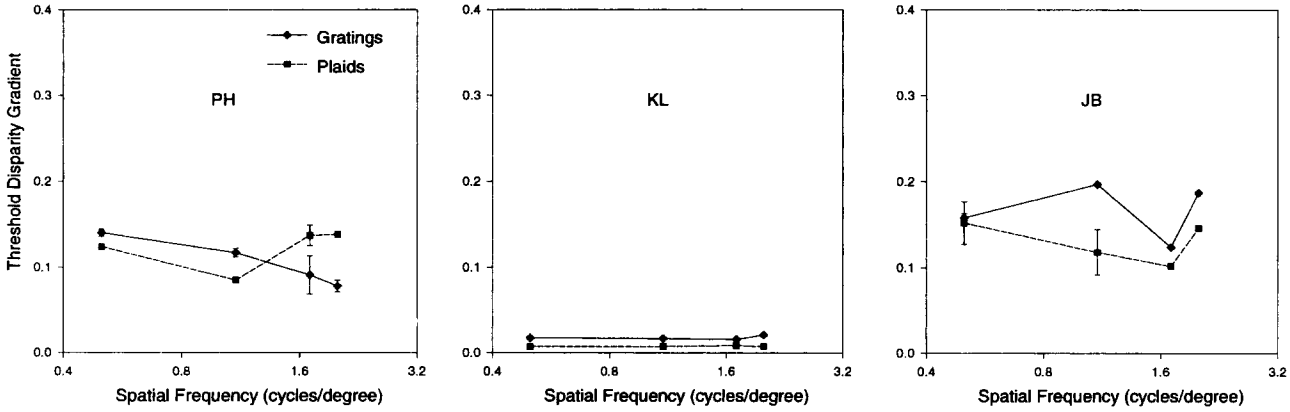

FIGURE 5. Inclination thresholds (posed as disparity gradients) are plotted as a function of sinusoidal grating and plaid beat spatial frequency. Plaid thresholds were found on the whole to be lower than grating thresholds. To convert the vertical axis from disparity gradients to angles of slant and inclination, one should refer to Mitchison and McKee (1990).

variability in the results both between and within subjects can be seen, we observed no consistent trends across either subjects or stimuli. For the conditions used here, plaid thresholds were generally lower than grating thresholds.

Discussion

The solid lines depicted in Fig. 4 were drawn with a slope of −0.5, and reference a relationship by which inclination thresholds vary inversely with the square-root of contrast. The mean slope taken across the three subjects was found to be −0.47, which is close to −0.5. Similar results have been reported in experiments that considered contrast thresholds for positional disparity cues only (Schor & Wood, 1983; Legge & Gu, 1989; Kontsevich & Tyler, 1994). Plaid thresholds, however, showed a greater dependence on contrast. The estimated slope for plaids taken across subjects' means was determined to be −0.78. The increased magnitude of the slope may be explained by the perceived plaid contrast. Plaids are generally perceived at a lower contrast than their Michelson contrast (Georgeson & Shackleton, 1994). To test for this, we recalibrated thresholds for plaids, based on the data of Georgeson and Shackleton. The slope of the regression fit after this manipulation was found to be −0.64 and closer to the predicted slope of −0.5. This suggests that plaid inclination thresholds were influenced by perceived contrast.

For both grating and plaid stimuli, no consistent effects of spatial frequency were found across the three subjects. Although Schor and Wood (1983) and Kontsevich and Tyler (1994) both reported that disparity discriminability was linearly related to spatial frequency below 2.5 cycles/deg, Hess and Wilcox (1994) showed that this relationship only holds for broadband stimuli. The stimuli used here were narrowband. Therefore, our data appear consistent to those studies that have investigated positional disparity cues only.

EXPERIMENT 2

In the second experiment we measured slant and inclination thresholds for sinusoidal gratings, while manipulating the absolute orientation of the gratings. Three binocular transformations were used: horizontal shear, horizontal compression and rotation. Details of these transformations are given in the Appendix. The rotation condition was used to assess the role of binocular orientation differences in the absence of binocular spatial-frequency differences. For this condition subjects were asked to respond in a manner consonant to an inclined surface.

Absolute grating orientations relative to horizontal ranged between 10 and 90 deg. For each transformation and condition, three sessions were run at separate times. The grating contrast was fixed at 99.8%. The grating's spatial frequency (prior to binocular transformation) was also fixed at 3.2 cycles/deg. Three models were fit to the data. The derivations of these models are given in the Appendix. The models assumed that thresholds could be predicted from positional disparity cues only, or a combination of orientation and one of either positional disparity cues or frequency difference cues.

We should mention here an unavoidable difficulty with the stimuli used. From A(1) (see Appendix) one should note that an affine model of slant has up to six unknown parameters. Hence, image stimuli composed solely from sinusoidal gratings are degenerate under such a transformation. Hence, surface tilt was ambiguous. We found, however, that subjects were able to classify the image stimuli according to the required slant and inclination transformations. We therefore reasoned that the degeneracy in these stimuli could be ignored because manipulations for inclined and slanted surfaces were run as separate sessions.

Results

Results are shown in Fig. 6 for the three binocular transformations. Data are presented as disparity gradient thresholds plotted against absolute grating orientation.

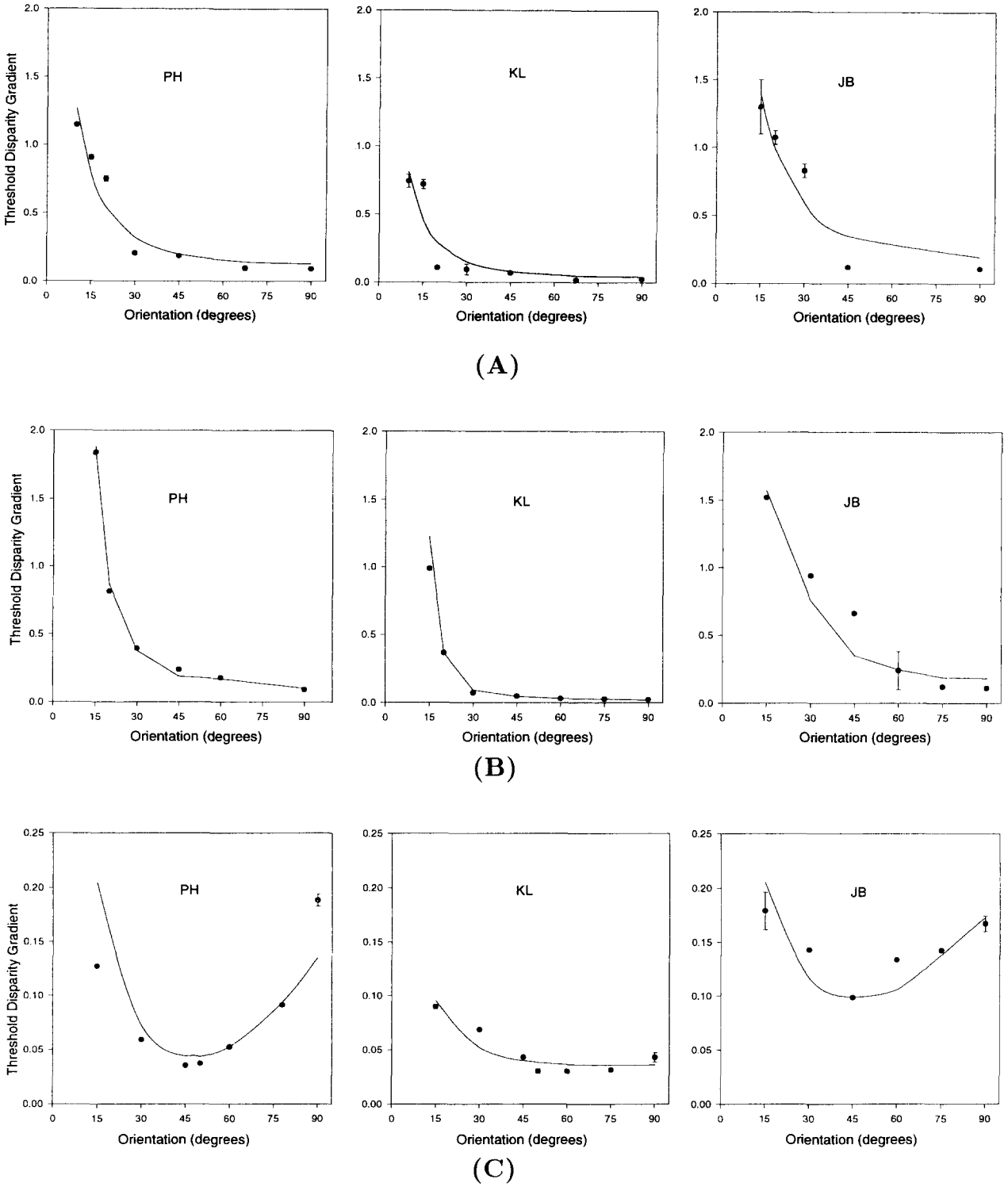

FIGURE 6. Thresholds for grating stimuli plotted as a function of orientation. (A) Thresholds for horizontal shear. (B) Thresholds for rotation. (C) Thresholds for horizontal compression. The solid curves represent the thresholds predicted from a model that assumes that slant and inclination are mediated via a combination of orientation and spatial-frequency differences, but with different sensitivities (see Fig. 7). In (C) a different vertical scale has been used to highlight the shape of the curve fits and the data.

We found that it was not possible to measure thresholds for grating orientations below 10 deg for subjects PH and KL, and below 15 deg for subject JB. For grating orientations less than these angles, subjects were unable to detect slant or inclination. For all three transformations, thresholds showed a marked dependence upon absolute grating orientation. Superimposed on the data are the thresholds that are predicted from the model with the highest R^2 value: the model that combined orientation and spatial-frequency disparities. From Fig. 7, one can

	Position	Orientation/Frequency	Orientation/Position
PH	$R^2 = 0.782$	$R^2 = 0.943$	$R^2 = 0.940$
KL	$R^2 = 0.676$	$R^2 = 0.873$	$R^2 = 0.873$
JB	$R^2 = 0.707$	$R^2 = 0.951$	$R^2 = 0.949$

FIGURE 7. R^2 values for the fitted models. Combining orientation with phase or spatial-frequency differences gave a better fit than the model that predicted thresholds from phase differences only. The model combining orientation and spatial-frequency disparity gave a slightly better fit overall.

see that this model gave slightly higher R^2 values than the model that combined orientation and position disparities. R^2 values for these two models were higher than the model which considered positional disparities only. Hence, the R^2 values from the fitted models indicate that binocular orientation differences are an important cue required to account for the variation in slant and inclination thresholds for the grating stimuli.

Discussion

From Fig. 6, one can see that subjects were unable to discriminate slant and inclination for grating stimuli whose absolute orientation was less than 10 deg. For the inclination condition, each subject attempted the task at one lower grating orientation: 5 deg (KL, PB) and 10 deg (JB). To these orientations we extrapolated the fitted curves and measured each subject's disparity gradient. We found, for these orientations that disparity gradients were 1.1 (KL), 1.8 (PH) and 2.1 (JB). Burt and Julesz (1980) proposed a value of 1 as an upper limit on discernible disparity gradients. Hence, slant and inclination thresholds would not be expected for orientations less than approx. 10 deg.

Our major result arises from the estimated sensitivities to each difference cue. For the model including binocular orientation differences (BOD) and binocular spatial-frequency differences (BSFD), sensitivity to BOD was found to be 1.75 times greater than sensitivity to BSFD for subject JB, 1.72 for subject KL, and 1.73 for subject PH. Subjects were more sensitive to BOD than BSFD. These sensitivities are surprisingly similar given the differences in thresholds across subjects and transformations.

For the horizontal shear and rotation transformations, each subject's data showed a similar trend. Thresholds were small for vertical gratings, and increased as the absolute orientation of the gratings approached horizontal. These threshold curves appear a close approximation to the reciprocal of the curves that show the magnitude of orientation disparities as a function of orientation, as depicted in Fig. 1(A). For gratings approaching horizontal in orientation, thresholds for the rotation condition were higher than those for the horizontal shear condition. This result provides some confirmation that BSFD play a role in the detection of surface inclination.

For the horizontal compression condition, thresholds for subjects PH and JB were again reciprocal curves to those shown in Fig. 1(A). For subject KL, the reciprocal relationship was less pronounced: thresholds for vertical gratings at 90 deg for both slanted and inclined surfaces were similar. Over a wide range of subjects, Rogers (1997, pers. comm.) found that 10% of the subjects studied appear isotropic with respect to slant and inclination thresholds. However, it is interesting to note that the relative sensitivity to BOD and BSFD for subject KL was similar to those measured for the other two subjects.

For each subject, thresholds for horizontal compression were lower than those for the horizontal shear condition for grating orientations less than 45 deg (Fig. 6). This is a reversal of the anisotropy reported by Rogers and Graham (1982). This reversal may be predicted again by assuming that BOD determine slant and inclination thresholds. This can be seen from Fig. 1(A), where for inclination the magnitude of BOD exceeds that of slant for orientations less than 45 deg. Here, it is interesting to note that in these conditions subject KL also produced lower slant than inclination thresholds. It may be the case that for vertically oriented stimuli, subject KL was able to exploit positional disparities. This might also explain the somewhat poorer R^2 values to the curve fits found for this subject and the isotropic thresholds measured for vertically oriented gratings. However, the reversal to the anisotropy found here suggests that a critical factor that allows one to measure the relative sensitivity between BOD and BSFD is to consider binocular image stimuli composed of spatial-frequency components that are close to horizontal in absolute orientation. This prediction is further borne out in the next experiment.

EXPERIMENT 3

The third experiment measured slant and inclination thresholds for plaid stimuli and was intended to test whether plaid thresholds could be predicted from sensitivity to disparities in their component gratings. The spatial orientation of plaid components was manipulated in such a way that the plaid beats were always vertical (see Figs 2 and 3). We reasoned that if thresholds for plaid stimuli depend on disparities in their individual sinusoidal components, then we should be able to predict plaid thresholds from the component grating thresholds (see also Welch, 1989).

The range of orientations studied was limited to

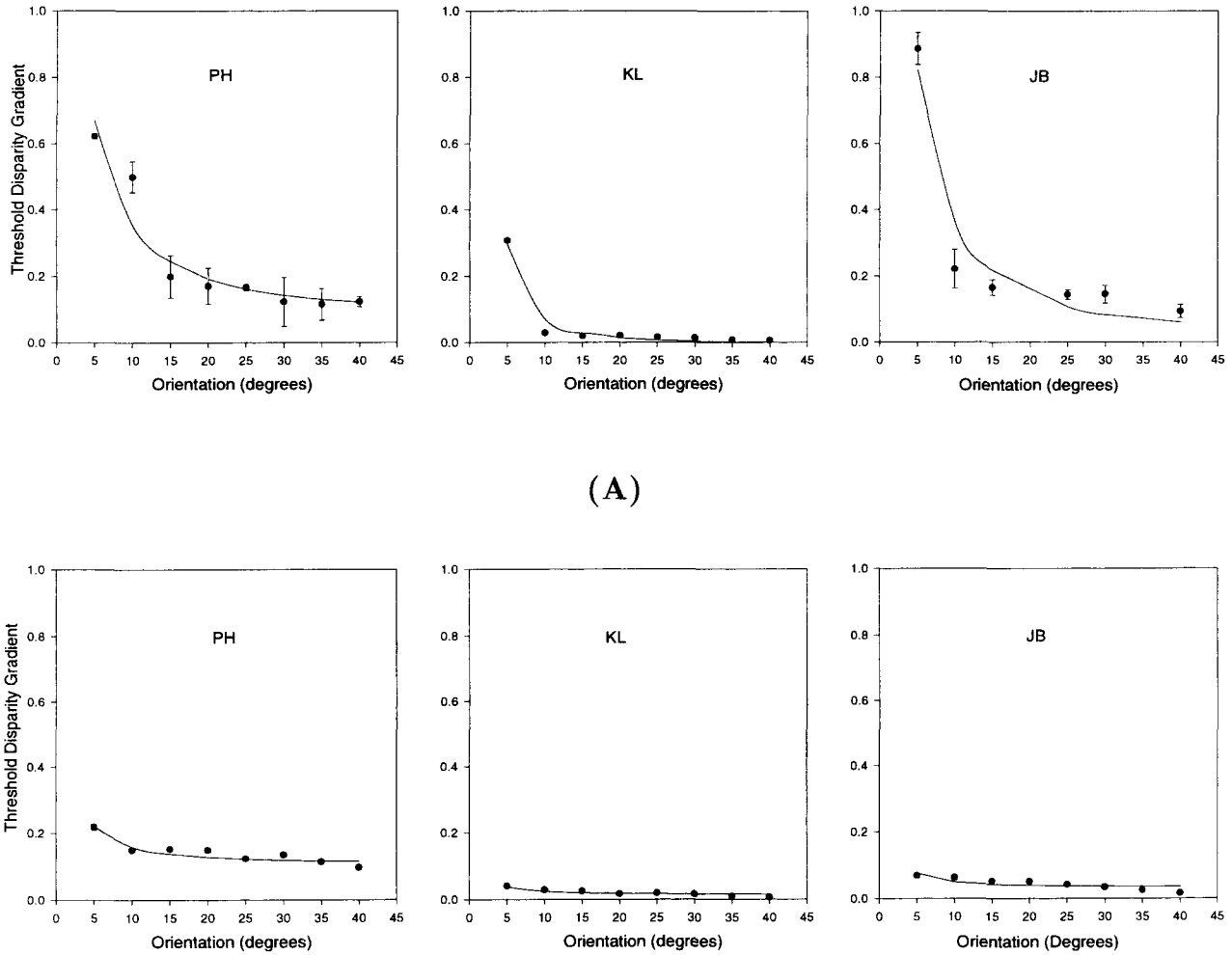

FIGURE 8. Thresholds for plaid stimuli plotted as a function of component orientation differences. (A) Thresholds for horizontal shear. (B) Thresholds for horizontal compression. The solid curves represent predictions obtained from the model that combined orientation and spatial-frequency differences.

between ±5 and ±40 deg. Reducing component angular differences also reduces the spatial frequency of the contrast beat. This places a limit on the number of visible cycles of the beats for a given stimulus window. We ensured that at least two full cycles of the contrast beat were always visible. Plaids were presented to subjects using the same protocol used in the previous experiments, for both horizontal shear and horizontal compression transformations.

If thresholds for plaids depend on disparity cues present in their individual components, we should be able to model this using the results taken from Experiment 2. To test this, curves were fitted using the sensitivities measured from the model based on orientation and spatial-frequency disparities found in Experiment 2. We did not expect plaid thresholds to have the same magnitude as grating thresholds: from Experiment 1, plaid thresholds were generally lower than grating thresholds. This may be because plaids are composed of two sinusoidal gratings, and therefore more information is available to the visual system. To take this into account in modeling the data, we fixed the relative sensitivities to orientation and spatial-frequency disparities for each subject according to those found in Experiment 2, while we allowed the magnitude of sensitivity to vary during the curve fitting procedure.

Results

Results for the three subjects are shown in Fig. 8. In each case, thresholds increased sharply as the orientation of components approached horizontal for both inclined and slanted surfaces as the orientation difference between plaid components reduced. Each fitted curve, shown by the solid lines in Fig. 8, can be seen to fit the data well. R^2 values for the curve fits were 0.95 (KL), 0.90 (PH) and 0.98 (JB). These R^2 values indicate that plaid thresholds

could be predicted from their component gratings. A final point to note is that thresholds for slanted surfaces were smaller than those for inclined surfaces. This was especially marked for plaids whose components had small differences in orientation and were therefore close to horizontal in absolute orientation.

Discussion

The curve fits suggest that plaid thresholds may be predicted from their component disparities. To confirm this result we also considered the additional factors of plaid contrast, and the plaid beats themselves. Decreasing the separation of the plaid components increases their perceived contrast (Georgeson & Shackleton, 1994). In Experiment 1, we found that plaid inclination thresholds decreased with increasing contrast. This trend would suggest that thresholds would decrease with decreasing component orientation differences. The opposite trend was found here. Hence, we have reasoned that the variation in thresholds found in our experiments did not arise from changes in perceived plaid contrast.

The plaid beats were always oriented vertically prior to the binocular transformation. From this, one might argue that a second, nonlinear binocular channel, perhaps tuned to the spatial frequencies of the contrast envelope itself, may have contributed to the perception of slant or inclination. Such ideas are central to some models of visual motion detection for plaid patterns (e.g. Wilson, Ferrera, & Yo, 1992), and have also been suggested in stereopsis (Sato & Nishida, 1994; Hess & Wilcox, 1994; Fleet & Langley, 1994). However, in the first experiment, there was no evidence to show that changes in the beat spatial frequency affected inclination thresholds, for the range of spatial frequencies used. Second, we note that both Sutter, Sperling and Chubb (1995) and Langley, Fleet and Hibbard (1996) report that sensitivity to plaid beats is itself spatial-frequency tuned with a peak sensitivity at about 1/8th–1/16th of the carrier spatial frequency. From this, one might predict that, if a second nonlinear channel were to influence thresholds, a local threshold minimum would have occurred for component orientation differences of 10 deg. Figure 8 shows no consistent evidence for such a local minimum.

The observation that slant thresholds are higher than inclination thresholds for small component orientation differences is consistent with the results of Experiment 2. This again shows a reversal of the anisotropy reported by Rogers and Graham (1983). The high R^2 values confirm that a model of detection thresholds that combines different sensitivities to BOD and BSFD provides a good account of the data. The results also demonstrate that, by using stimuli comprising gratings oriented close to horizontal, we were able to further test the notion that the anisotropy results from different sensitivities to BOD and BSFD in sinusoidal component gratings. The reversal to the anistropy shown in Fig. 8 further argues against the idea that nonlinear channels were solely responsible for the detection of binocular disparities in contrast beats.

GENERAL DISCUSSION

Slant and inclination thresholds for both grating and plaid stimuli can be accounted for by a model which makes use of interocular differences in both orientation and either frequency and/or position. Our data suggest that sensitivity to BOD is higher than sensitivity to BSFD. These data are consistent with those already reported by Rogers and Graham (1983) and Cagenello and Rogers (1993). However, our results advance these studies in two ways. First, we have suggested that a critical factor that allows one to assess the contribution of orientation disparities towards slant and inclination is to study either plaids or sinusoidal gratings with absolute orientations that approach horizontal. Second, we have tested whether the perception of slant and inclination might arise from either gradients of positional differences, or from the processing of higher order binocular differences, such as orientation and spatial frequency.

Our experiments cannot determine whether orientation or spatial-frequency disparities are represented explicitly by the visual system. This can be deduced from A(1)A(2)A(3)A(4)A(5) in the Appendix. Transformations between spatial gradients of disparity, and binocular differences of orientation and spatial frequency are linear and may be inverted. Therefore, orientation and spatial-frequency differences should be viewed as a convenient representation to explain our results. Hence, the underlying factor accounting for the differences in sensitivity remains to be discovered. We will, however, discuss here how these might occur, in the context of the model of stereopsis forwarded by Fleet, Wagner and Heeger (1996).

The model of Fleet *et al.* (1996) takes the established view that the responses of simple cells located in the primary visual cortex are summed linearly and combined by binocular complex cells (e.g. Heeger, 1992). This combination involves the rectification of simple cells of different phases and positions to yield a binocular energy response. Under this responses model, disparity is encoded by the peaks in the binocular energy response. The dominance* of the peaks in binocular energy predicts the model's ability to discriminate changes in disparity. This model is especially attractive to us as the underlying processes are energy-based, and a single mechanism suffices to represent both luminance and contrast changes. Hence, binocular transformations of both grating and plaid patterns would be processed by a common mechanism, consistent with the results of Experiment 3.

Fleet and colleagues' model encodes disparity using binocular energy neurons exhibiting both interocular position and phase shifts. This idea is again attractive because such an encoding scheme would be consistent with Morgan and Castet's (Morgan & Castet, 1995) finding that disparity thresholds for grating stimuli varying in orientation demonstrate a constant sensitivity

*Dominance here refers to the sharpness of the binocular energy point spread function about the peak.

to phase disparity. However, phase-based disparity encoding requires knowledge of the local phase-gradient, and is therefore subject to error (Fleet *et al.*, 1996). This error would predict disparity thresholds to be dependent on spatial frequency. Such a dependency was not found here, nor is it a general feature reported in other studies (Schor & Wood, 1983; Legge & Gu, 1989; Halpern & Blake, 1988; Kontsevich & Tyler, 1994; Hess & Wilcox, 1994). Hence, we reason that sensitivity to BOD and BSFD does not arise from a mechanism encoding phase disparities alone.

Through multiple stages of processing, it is not unreasonable to suppose that the visual system is subject to transmission and coding errors. Such errors could arise because neurons are limited in transmission bandwidth (Laughlin, 1989). Here, Fleet and Jepson (1993) show us that the variance of a linear filter's response in the presence of noise is proportional to frequency bandwidth. The orientation bandwidths of simple cells have been found to be smaller than their spatial-frequency bandwidths (e.g. Movshon, Thompson & Tolhurst, 1978). Hence, for the reasons mentioned the visual system would be expected to encode orientation more robustly than spatial frequency. This view is supported by the finding that, in spatial discrimination tasks, judgements of orientation differences are 2–3-times more accurate than those of spatial-frequency differences (e.g. Bowne, 1990).

However, if we were to include these ideas into the current model of Fleet *et al.* (1996), then we would predict that sensitivity to BSFD would be greater than that to BOD. This is because the model combines responses from cells with the same orientation and spatial-frequency tuning in left- and right eyes. Now introducing BOD and BSFD will both reduce the dominance of binocular energy peaks by reducing the correlation between interocular image signals. Here one would expect that the reduction in the sharpness of binocular energy peaks arising from BOD would be greater than that for equivalent BSFD, given that the encoding of orientation is more robust than spatial frequency. This prediction would be inconsistent with the anisotropy in slant and inclination thresholds.

Alternatively, let us consider a revision of the model of Fleet *et al.* (1996): a binocular energy model that combines both position and orientation shifts. This idea has a physiological basis. For example, Nelson *et al.* (1977) found binocular cells in the cat primary visual cortex whose responses appeared to depend on both BOD and binocular position differences. This model would be expected to yield a strong peak in binocular energy when the interocular summation of monocular mechanisms occurs for stimuli with the correct BOD and position difference. Now if we suppose that orientation is more reliably encoded than spatial frequency, then we would expect the dominance of the binocular energy response to be less affected by BOD than BSFD because of the interocular combination of binocular energy neurons at different orientations. This could then account for the anisotropy in slant and inclination thresholds.

To summarise, while we are unable to determine precisely why different sensitivities to BOD and BSFD occur, our data are, however, important to models of binocular stereopsis similar to the one proposed by Fleet *et al.* (1996) in three ways:

- Plaid patterns with the beats oriented vertically and the carrier oriented horizontally most stimulate binocular processes supported by linear filtering mechanisms tuned to horizontal orientations (Fleet & Langley, 1994). Our data are, therefore, consistent with the idea that the visual system pools information across spatial-frequency and orientation-tuned channels (Langley, Atherton, Wilson & Larcombe, 1991; Fleet *et al.*, 1996).
- No evidence was found for a separate binocular nonlinear channel tuned to the beat spatial frequency for coherent plaids.
- Our data are consistent with a binocular energy model that exploits both position and orientation shifts towards the detection of spatial changes in binocular differences in disparity.

Finally, our experiments are consistent with (and based upon) an analogous study in visual motion detection (Welch, 1989). Welch found that plaid motion discrimination thresholds could be predicted from the velocities of the plaid's sinusoidal components. Interestingly, in Welch's study the motion information was encoded in a direction that one might expect to be detected by the nonlinear channel of the model of Wilson *et al.* (1992). It would be possible to adapt a model of this class to account for binocular slant and inclination thresholds for plaid stimuli. However, both Welch's results and those of the current study suggest that for plaid stimuli in which carrier and beat frequencies together signal a single coherent structure, motion and stereo processing are based upon the component gratings. It is difficult to see how this dependency could be explained by a two-stage model that explicitly separates both plaid carrier and plaid beat information (e.g. Wilson *et al.*, 1992). This then implies that models of visual processing that rely upon the clean separation of visual information into separate linear and nonlinear channels must be revised to take into account circumstances where both contrast and luminance changes in coherent surfaces are processed by a common mechanism.

REFERENCES

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

Bowne, S. F. (1990). Contrast discrimination cannot explain spatial frequency, orientation or temporal frequency discrimination. *Vision Research, 303*, 449–461.

Burt, P. & Julesz, B. (1980). Modifications of the classical notion of Panum's fusional areas. *Perception, 9*, 671–682.

Cagenello, R. B. & Rogers, B. J. (1993). Anisotropies in the perception

of stereoscopic surfaces: the role of orientation disparity. *Vision Research, 33*, 2189–2201.
Fleet, D. J. & Jepson, A. D. (1993). Stability of phase information. *IEEE Transactions of PAMI, 15*, 1253–1268.
Fleet, D. J. & Langley, K. (1994). Non-Fourier channels in stereopsis and motion. *Perception, 23 (Suppl)*, 83.
Fleet, D. J., Wagner, H. & Heeger, D. J. (1996). Neural encoding of binocular disparity: energy models, position shifts and phase shifts. *Vision Research, 36*, 1839–1857.
Georgeson, M. A. & Shackleton, T. M. (1994). Perceived contrast of gratings and plaids: non-linear summation across oriented filters. *Vision Research, 34*, 1061–1075.
Gillam, B., Chambers, D. & Russo, T. (1988). Postfusional latency in stereoscopic slant perception and the primitives of stereopsis. *Journal of Experimental Psychology: Human Perception and Performance, 142*, 163–175.
Gillam, B., Flagg, T. & Finlay, D. (1984). Evidence for disparity change as the primary stimulus for stereoscopic processing. *Perception and Psychophysics, 36*, 559–564.
Halpern, D. L. & Blake, R. R. (1988). How contrast affects stereoacuity. *Perception, 17*, 483–495.
Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience, 9*, 427–443.
Hess, F. R. & Wilcox, L. M. (1994). Linear and non-linear filtering in stereopsis. *Vision Research, 3418*, 2431–2438.
Howard, I. P. & Kaneko, H. (1994). Relative shear disparities and the perception of surface inclination. *Vision Research, 34*, 2505–2517.
Koenderink, J. J. & Van Doorn, A. J. (1976). Geometry of binocular vision and stereopsis. *Biological Cybernetics, 21*, 29–35.
Kontsevich, L. L. & Tyler, C. W. (1994). Analysis of stereothresholds for stimuli below 2.4 c/deg. *Vision Research, 34*, 2317–2329.
Langley, K., Atherton, T. J., Wilson, R. G. & Larcombe, M. H. E. (1991). Vertical and horizontal disparities from phase. *Image and Vision Computing, 5*, 296–303.
Langley, K., Fleet, D. J. & Hibbard, P. B. (1996). Linear filtering precedes nonlinear processing in early vision. *Current Biology, 6*, 891–896.
Laughlin, S. B. (1989). Coding efficiency and design in visual processing. In D. G. Stavenga & R. C. Hardie (Eds), *Facets of vision* (pp. 213–234). Berlin: Springer.
Legge, G. E. & Gu, Y. (1989). Stereopsis and contrast. *Vision Research, 29*, 989–1004.
Mitchison, G. J. & McKee, S. P. (1990). Mechanisms underlying the anisotropy of stereoscopic tilt perception. *Vision Research, 30*, 1781–1791.
Morgan, M. J. & Castet, E. (1995). Stereoacuity for oblique gratings predicted by phase shifts, not disparities. *Investigative Ophthalmology and Visual Science, 36 Suppl*, S231.
Movshon, J. A., Thompson, I. D. & 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.
Nelson, J. I., Kato, H. & Bishop, P. O. (1977). Discrimination of orientation and position disparities in binocularly activated neurons in cat striate cortex. *Journal of Neurophysiology, 40*, 260–283.
Rogers, B. J. & Graham, M. (1982). Similarities between motion parallax and stereopsis in human depth perception. *Vision Research, 22*, 216–270.
Rogers, B. J. & Graham, M. (1983). Anisotropies in the perception of three-dimensional surfaces. *Science, 221*, 1409–1411.
Sato, T. & Nishida, S. (1994). Second order depth perception with texture-defined random check stereograms. *Investigative Ophthalmology and Visual Science, 34 Suppl*, 1438.
Schor, C. M. & Wood, I. (1983). Disparity range for local stereopsis as a function of luminance spatial frequency. *Vision Research, 23*, 1649–1654.
Sutter, A., Sperling, G. & Chubb, C. (1995). Measuring the spatial-frequency selectivity of 2nd-order mechanisms. *Vision Research, 35*, 915–924.
Treisman, M. & Watts, T. R. (1966). Relation between signal delectability theory and the traditional procedures for measuring thresholds: estimating d' from results given by the method of constant stimuli. *Psychological Bulletin, 66*, 438–454.
Watt, R. & Andrews, D. P. (1981). Adaptive probit estimation of psychometric functions. *Current Psychological Reviews, 1*, 205–214.
Welch, L. (1989). The perception of moving plaids reveals two motion-processing stages. *Nature, 337*, 734–736.
Wilson, H. R., Ferrera, V. P. & Yo, G. (1992). A psychophysically motivated model for two-dimensional motion perception. *Visual Neuroscience, 91*, 79–97.
Wonnacott, R. J. & Wonnacott, T. H. (1979). *Econometrics*. New York: Wiley.

Acknowledgements—P. B. Hibbard is grateful to the Department of Psychology, UCL, the UCL graduate school, and the Tregaskis Bequest for supporting the research presented here. We would also like to acknowledge helpful comments and suggestions of Prof. O. J. Braddick, Prof. M. Georgeson and Dr Mark Bradshaw, and the loan of computing facilities by Prof. P. Howell.

APPENDIX

In our experiments we used three binocular transformations: rotation, horizontal shear, and horizontal compression. Figure 9 shows each transformation as applied to a single sinusoidal grating, represented in the Fourier frequency domain. The introduction of binocular changes in both orientation and spatial frequency is clear from the relative position in the frequency domain of the left and right gratings. These transformations approximate the disparity map obtained from slanted and inclined surfaces and were introduced according to the affine model:

$$\begin{bmatrix} D_x(x,y) \\ D_y(x,y) \end{bmatrix} = \begin{bmatrix} D_x(0,0) \\ D_y(0,0) \end{bmatrix} + \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad \text{(A1)}$$

where $D_x(x,y)$, $D_y(x,y)$ refer to the binocular disparities in the vertical and horizontal directions, respectively. Disparity is assumed to vary with spatial image position such that $D_y(x,y) = D_x(0,0) = D_y(0,0) = c = d = 0$. These latter conditions imply that disparities are primarily horizontal. The constants a, b depict spatial gradients of disparity in the horizontal and vertical directions, respectively. Transforming A(1) into polar coordinates via the Jacobian, we obtain:

Horizontal shear:

$$\Delta\theta = \frac{1}{2}\tau(1 - \cos 2\theta)$$

$$\frac{\Delta f}{f} = \frac{1}{2}\tau \sin 2\theta \quad \text{(A2)}$$

Horizontal compression:

$$\Delta\theta = \frac{-1}{2}\tau \sin 2\theta$$

$$\frac{\Delta f}{f} = \frac{1}{2}\tau(1 + \cos 2\theta) \quad \text{(A3)}$$

Rotation:

$$\Delta\theta = \tau \quad \text{(A4)}$$

$$\frac{\Delta f}{f} = 0, \quad \text{(A5)}$$

where $\Delta\theta$ and Δf refer to the difference in orientation and spatial frequency between the binocular images. These transformations are depicted in Fig. 9. The parameter τ refers to the magnitude of disparity gradient, which is a constant and equal to either a, or b depending upon the transformation that we are considering. One should note that under this transformation, small increments of spatial frequency are orthogonal to small increments in orientation [see Fig. 9(C)]. Morgan and Castet (1995) found that disparity sensitivity for gratings can be

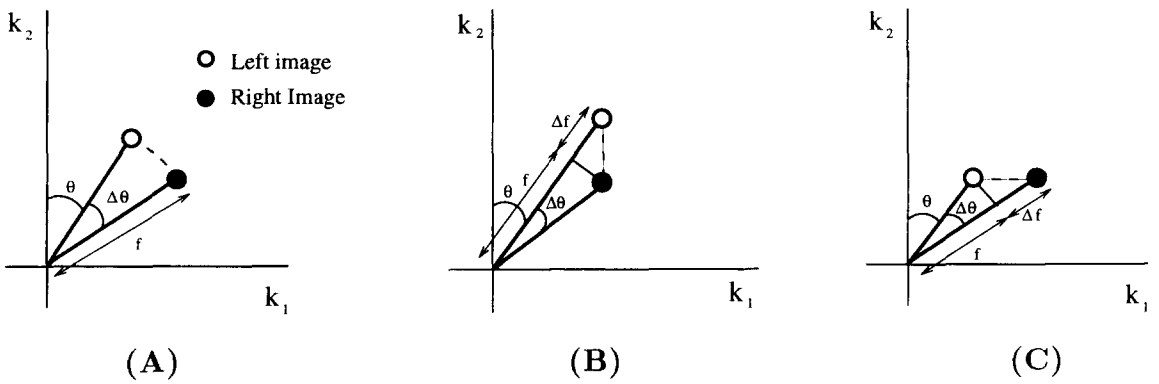

FIGURE 9. The three binocular transformations used. (A) Rotation. (B) Horizontal shear. (C) Horizontal compression. Each transformation is expressed in the Fourier frequency domain. Horizontal and vertical spatial frequencies are denoted by k_1, k_2, respectively. One can see from the figures that both horizontal shear and horizontal compression introduce orientation and spatial-frequency disparities. One should also note that a horizontal shear in the spatial domain is manifested as a vertical shear in the frequency domain.

predicted directly from binocular phase differences [one should note that phase-differences in two-dimensional (2-D) images suffer from the aperture problem and we have hence assumed them to be taken normal to a grating's spatial orientation (Langley *et al.*, 1991)]. The horizontal shear and expansion compressions defined above introduce gradients of interocular phase differences, ($\nabla[d,\phi]$), given by:
Horizontal shear:

$$\nabla[d_\phi] = \tau f \sin\theta \tag{A6}$$

Horizontal compression:

$$\nabla[d_\phi] = \tau f \sin\theta \tag{A7}$$

Predicting threshold disparity gradients from disparities

A(2)A(3)A(4)A(5)A(6)A(7) give the magnitude of orientation disparity, spatial-frequency disparity and phase disparity gradient for the three transformations used. These equations may be used to predict slant thresholds if we assume that the cues are independent, and that measurement errors are uncorrelated. We also assume that sensitivity to a linear combination of disparity cues is given by the Euclidean summation of the sensitivities to the individual cues. This leads to:

$$d'_{\text{comb}} = [d'_1 2 + d'_2 2 + \ldots + d'_n 2]^{\frac{1}{2}}. \tag{A8}$$

Following Treisman and Watts (1966), who advocate that discrimination thresholds are inversely proportional to sensitivity, we obtain:

$$\tau = [\tau_1^{-2} + \tau_2^{-2} + \ldots + \tau_n^{-2}]^{-\frac{1}{2}}, \tag{A9}$$

where we have used $\tau_i^{-2} \propto d_i'^2$. Using the relationships given by A(2)A(3)A(4)A(5)A(6)A(7), threshold models that combine spatial gradients of phase, or BSFD with BOD for sinusoidal gratings are determined to be:

$$\tau_p = \begin{cases} \left[\left(\frac{1-\cos 2\theta}{\sigma_\theta^2}\right)^2 + \left(\frac{\sin 2\theta}{\sigma_f^2}\right)^2\right]^{-\frac{1}{2}} & \text{for horizontal shear} \\ \left[\left(\frac{\sin 2\theta}{\sigma_\theta^2}\right)^2 + \left(\frac{1-\cos 2\theta}{\sigma_f^2}\right)^2\right]^{-\frac{1}{2}} & \text{for horizontal compression} \end{cases} \tag{A10}$$

$$\tau_{sf} = \begin{cases} \left[\left(\frac{1-\cos 2\theta}{\sigma_\theta^2}\right)^2 + \left(\frac{\sin 2\theta}{\sigma_{d_\phi}^2}\right)^2\right]^{-\frac{1}{2}} & \text{for horizontal shear} \\ \left[\left(\frac{\sin 2\theta}{\sigma_\theta^2}\right)^2 + \left(\frac{1-\cos 2\theta}{\sigma_{d_\phi}^2}\right)^2\right]^{-\frac{1}{2}} & \text{for horizontal compression} \end{cases} \tag{A11}$$

where τ_p and τ_{sf} refer to the predicted relationship between thresholds and orientation, based upon a combination of position and orientation, and spatial frequency and orientation disparity cues, respectively. The parameters σ_ϕ^2,$\sigma_{d_\phi}^2$ and σ_f^2 are estimated by the curve fits and yield the sensitivity to orientation, position and spatial-frequency differences, respectively. Threshold data for both plaids and gratings were fitted by these models using a Sigma-Plot nonlinear curve fitter. The method of dummy variables (Wonnacott & Wonnacott, 1979) was used to allow curves to be fit to the data for both horizontal compression and horizontal shear simultaneously. This then allowed us to estimate the relative sensitivity to each cue for each stimulus, independent of the binocular transformation used.