THE ROYAL SOCIETY

Linear and nonlinear transparencies in binocular vision

Keith Langley[1], David J. Fleet[2] and Paul B. Hibbard[1]

[1]*Department of Psychology, University College London, Gower Street, London WC1E 6BT, UK*
[2]*Department of Computing and Information Science, Queen's University, Kingston, Canada K7L 3N6*

When the product of a vertical square-wave grating (contrast envelope) and a horizontal sinusoidal grating (carrier) are viewed binocularly with different disparity cues they can be perceived transparently at different depths. We found, however, that the transparency was asymmetric; it only occurred when the envelope was perceived to be the overlaying surface. When the same two signals were added, the percept of transparency was symmetrical; either signal could be seen in front of or behind the other at different depths. Differences between these multiplicative and additive signal combinations were examined in two experiments. In one, we measured disparity thresholds for transparency as a function of the spatial frequency of the envelope. In the other, we measured disparity discrimination thresholds. In both experiments the thresholds for the multiplicative condition, unlike the additive condition, showed distinct minima at low envelope frequencies. The different sensitivity curves found for multiplicative and additive signal combinations suggest that different processes mediated the disparity signal. The data are consistent with a two-channel model of binocular matching, with multiple depth cues represented at single retinal locations.

Keywords: envelope tuning; additive transparency; multiplicative transparency; depth asymmetry; binocular rivalry

1. INTRODUCTION

A central issue in stereopsis is the correspondence problem: how one determines the retinal locations in the left and right eyes that are projections of the same three-dimensional points. Conventional (first-order) models solve for correspondence by matching linearly filtered versions of the views of the left and right eyes. One might consider cross correlation or phase differences between small neighbourhoods of the filtered left and right signals to establish the binocular matches (e.g. Fleet *et al.* 1996). It is common for these techniques to use a uniqueness assumption (Marr 1982) to constrain the matching process. With binocular transparency, however, the correspondence problem becomes more difficult because of the need to recover several depth planes, and to deal with different types of transparency (Weinshall 1991; Kersten 1991).

Transparency occurs naturally in two basic forms (Kersten 1991). One is linear, where the retinal intensities are composed from the sum of two surface reflectance patterns. This may occur, for example, when looking at a river-bed through a reflection on the surface. The second form is nonlinear, occurring when one looks through a translucent surface. In this case, the retinal intensities stem from the product of the material transmittance function of the overlaying surface and the reflectance from the underlying surface. How the visual system interprets surface depth for both additive and multiplicative transparency remains an interesting issue because the two types of transparency are incompatible.

Kersten (1991) (see also Frisby & Mayhew 1978) reported that some multiplicative transparencies are asymmetrical, where one's percept depends on the left and right ordering of the binocular images. Kersten believed that binocular asymmetries reflect competition between contradicting monocular and binocular depth cues. This view originated from two constraints on one's percept of monocular transparency as proposed by Metelli (1974): (i) no matter how a multiplicative transparency is produced, the overlaying transparent surface must not change the values of the order of luminance reflected from the underlying surface; and (ii) when values of lightness are attenuated by a transparent surface, local differences in lightness seen through the transparent surface must be less than those seen without the transparent surface.

Beck (1984) reasoned that Metelli's constraints were incomplete. From the physics of multiplicative transparency, Beck showed that the surface transmittance and reflectance functions for both the overlaying and the underlying surfaces must be positive-valued. Studying Metelli's constraints with overlapping square patches, Beck *et al.* (1984) found that when the luminance values of overlapping squares satisfied Metelli's two constraints, subjects reported transparency; even when values of the luminance of one of the squares violated the assumption that its transmittance function was positive-valued. However, when values of the luminance violated Metelli's constraints, subjects did not report transparency. To explain their results, Beck *et al.* supposed that the visual system cannot access surface reflectance functions. Rather, it processes

perceived lightness, defined as a nonlinear function of surface reflectance.

There are two predictions from Metelli's constraints that could influence binocular depth perception. One is that images composed of a product of two positive-valued signals may be perceived symmetrically; either signal may be seen in front of, or behind, the other at different depths depending upon the disparity cue present in each signal. In a companion paper (Langley *et al.* 1998), we have confirmed these predictions by using similar stimuli to those manipulated here. The other prediction is that second-order binocular signals, like a positive-valued contrast envelope and a mean-zero carrier, will be perceived asymmetrically in depth. This is because Metelli's monocular constraints would be violated when the disparity cues imply that the carrier is the signal nearest in depth to the binocular observer, thus leading to conflicting binocular and monocular depth relations.

Metelli's constraints make no predictions concerning additive (first-order) transparencies. To help explain the first-order case, Weinshall (1991) showed that binocular cross correlation of the left and right images of additive transparent signals gave two peaks at the disparities of the individual signals. One could modify this cross-correlation model to account for multiplicative transparency by introducing an early retinal nonlinearity (e.g. Burton 1973) or a later nonlinearity in the cortex (e.g. Langley 1997). One nonlinearity that might be considered is logarithmic. It is well known that the logarithm of a product of two signals is equal to the sum of the logarithms of the individual signals. With second-order (non-Fourier) binocular signals, the nonlinearity will introduce power at the frequencies of the second-order envelope so that a single-channel model of transparency, such as the one proposed by Weinshall (1991), could then detect disparities for both first- and second-order signal combinations. Unfortunately, this model may encounter difficulties when presented with first-order transparencies, because the logarithm of the sum of two signals is no longer equal to a superposition of the two signals.

Instead of a single-channel model, many researchers favour a two-channel model. Two-channel models include a separate nonlinear (second-order) channel to detect the second-order disparity signal for binocular matching. Models such as this have been proposed to explain the perception of second-order motion stimuli (Chubb & Sperling 1988) and stereopsis (Wilcox & Hess 1996).

In analysing properties of second-order stimuli and models, Fleet & Langley (1994) noted that idealized second-order signals have a simple characterization in terms of oriented power concentrations in the Fourier domain. They showed that orientated power occurs with multiplicative signal combinations such as those caused by multiplicative transparency and occlusion boundaries. Fleet & Langley further showed that the bandpass-filtered image signal may be separated into (first-order) phase and (second-order) amplitude signals by a logarithmic transformation, which is consistent with the two-channel hypothesis. This approach is attractive because the binocular cross correlation of phase and amplitude may be used to detect transparencies for first- and second-order signal combinations. Hence, two-channel models may be motivated by a computational strategy that leads to the recovery of binocular depth for both first- and second-order signal combinations for the two types of transparency mentioned.

In this paper, we examine the one- and two-channel models, and one's perception of binocular transparency with additive and multiplicative signal combinations. It is shown that depth perception for contrast envelopes is asymmetrical, consistent with Metelli's constraints. By varying the frequency of the contrast envelope, we also show that minima in transparency thresholds for contrast envelopes occur for frequencies approximately 2.4 octaves below the carrier frequency. By comparison, no minima were found for first-order signals which were, broadly speaking, perceived symmetrically. The binocular asymmetry found for second-order signals suggests that different constraints are used in processing first- and second-order signal combinations, implicating a two-channel model.

2. METHODS

(a) *Apparatus and procedure*

Monocular images of the binocular stimuli were presented on the left and right sides of a Sony monitor with a refresh rate of 76 Hz and 256 grey scales. The luminance of the monitor was linearized by taking luminance measurements with a photometer, to which a logarithmic curve was fitted. This was then used to generate a linear look-up table. The residual error from the fitted curve at any one luminance was no more than 0.2% of the fitted curve at any one of the sampled points. The mean luminance of the monitor was 37.7 cd m^{-2}.

Experiments were done in a darkened room. The only visible illumination originated from the monitor. A modified Wheatstone stereoscope was used to view the binocular images on a single monitor. The distance from the screen to the stereoscope was 44 cm. The stereoscope was adjusted so that the stereo image pairs were correctly aligned using a parallel viewing geometry. The entire visual extent of each monocular image was 7.9°. Image pixels were square with a width of 2 min. A fixation spot was used as a reference point to help keep vergence fixed. Subjects were seated with their heads stabilized in a chin rest in front of the stereoscope. They responded by using a computer mouse in forced-choice discrimination tasks. Subjects were asked to respond as quickly as they could, but were not otherwise constrained by the viewing time.

(b) *Stimuli*

Both contrast-modulated sinusoidal gratings and contrast-modulated noise signals were used. Some examples are shown in figure 1. The raw envelope signal, $E(x, y)$, was an approximation to a vertically orientated, square-wave grating. It was formed by summing the first and third harmonics of the square-wave:

$$E(x,y) = \alpha[\sin(\omega x + \phi) + (1/3)\sin 3(\omega x + \phi)], \quad (1)$$

where ω was the fundamental frequency of the envelope, and α was chosen so that $E(x, y)$ ranged between 1 and -1. The phase of the envelope, ϕ, was randomized between trials.

Given the envelope $E(x, y)$, offset to make it strictly positive, along with carrier signals for the left and right eyes, $C_l(x, y)$ and $C_r(x, y)$, the contrast-modulated second-order stimuli were given by

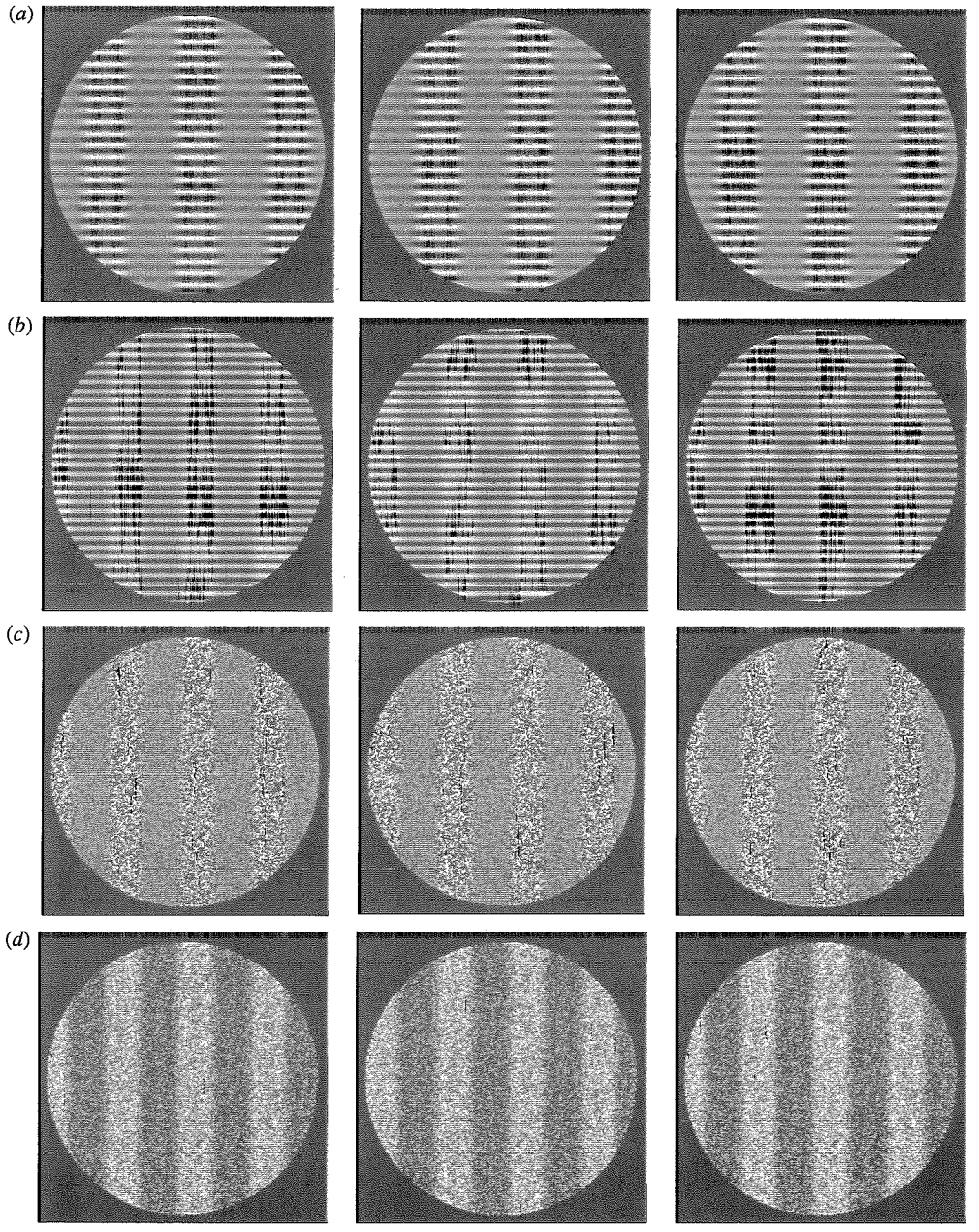

Figure 1. Examples of the binocular stimuli used. (*a*) A multiplicative (square-wave) contrast envelope multiplied by a sinusoidal grating. (*b*) The superposition (addition) of a square-wave and sinusoidal grating. (*c*) A square-wave contrast envelope multiplied by a mean-zero noise pattern. (*d*) The superposition of a square-wave grating and a noise pattern. For each image triplet, cross-eyed fusion of the left and centre images reveals a vertical structure perceived transparently in front of the plane of fixation. However, cross-eyed fusion of the centre and right images may yield either a coherent (*a*), transparent (*b*, *d*) or diplopic (*c*) square-wave grating. Note that nonlinearities probably caused by the photographic procedure have reduced the saliency of the diplopic envelope shown in (*c*).

$$M_l(x,y) = \mu[1 + (a/2)(1 + mE(x + d/2, y))C_l(x,y)]$$
$$M_r(x,y) = \mu[1 + (a/2)(1 + mE(x - d/2, y))C_r(x,y)], \quad (2)$$

where μ is the mean illumination, m is the depth of contrast modulation, a is the contrast of the carrier, and d is the positional disparity. In addition, the subscripts l and r are used to denote the left and right eyes. The stimuli were visible only within a circular window as shown in figure 1. In each condition, the signal $C(x, y)$ was either a horizontal or vertical sinusoidal grating, or a mean-zero random-dot noise pattern with zero disparity. The mean value of $C(x, y)$ was always zero. When $C(x, y)$ was a grating, its spatial frequency was 2.25 or 4.5 cycles per degree (cpd). When $C(x, y)$ was a noise pattern, each pixel was randomly assigned a value of ± 1. The depth of modulation m was 0.75, and the contrast a was 0.98.

Additive combinations of the same signals were also used, and are given by

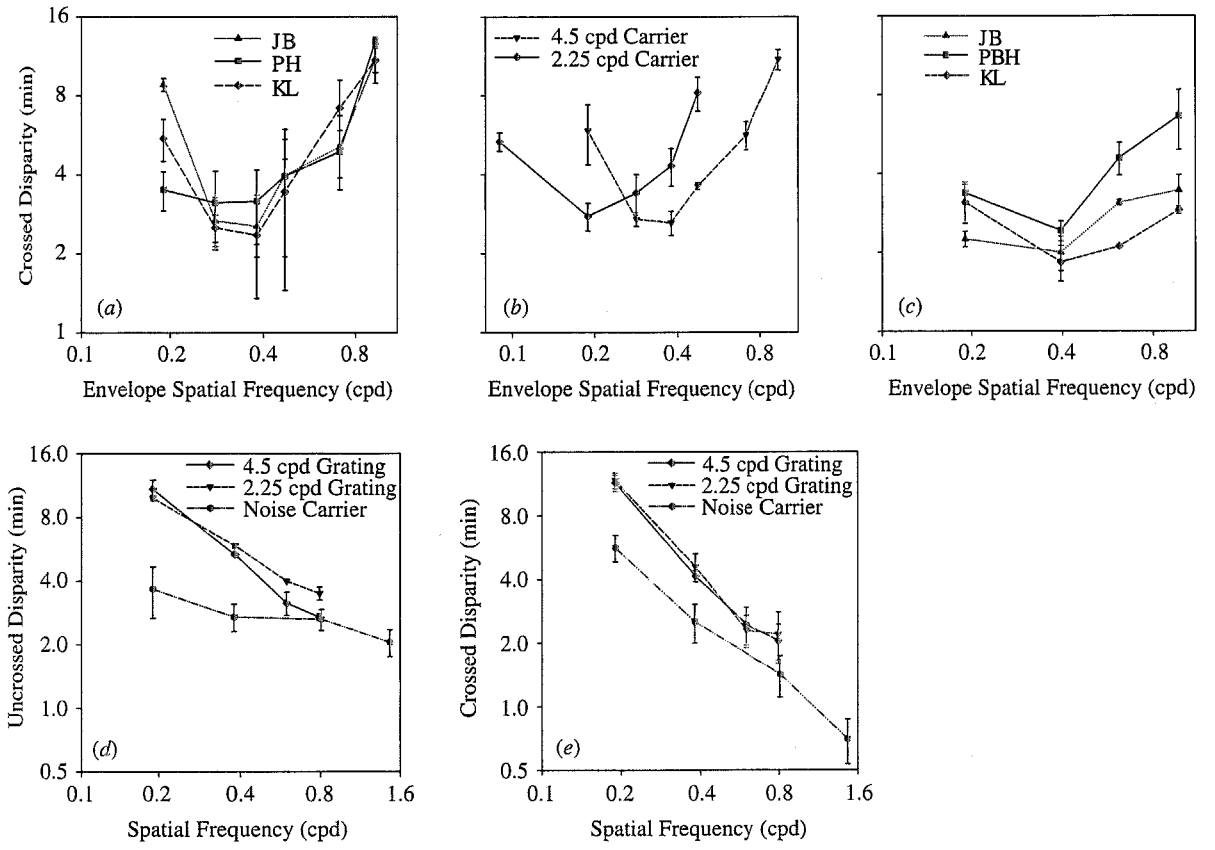

Figure 2. Transparent disparity thresholds are shown as a function of the fundamental frequency of $E(x,y)$. (*a*) A multiplicative condition is shown for three subjects. $C(x,y)$ was a 4.5-cpd grating. (*b*) Mean transparency thresholds for two different frequencies of $C(x,y)$ are shown for the multiplicative condition. (*c*) Individual subject's responses when $C(x,y)$ was noise. (*d*, *e*) Thresholds for additive combinations of $E(x,y)$ and $C(x,y)$, showing that transparency could be reported for both crossed and uncrossed disparities. In (*a*) and (*c*) the error bars represent standard errors for each subject across different sessions. In (*b*), (*d*) and (*e*) the error bars represent the standard error for each subject's thresholds across different sessions.

$$A_l(x,y) = \mu[1 + (a/2)C_l(x,y) + (b/2)E(x + d/2, y)]$$
$$A_r(x,y) = \mu[1 + (a/2)C_r(x,y) + (b/2)E(x - d/2, y)]. \quad (3)$$

For additive signal combination conditions, $a/2$ and $b/2$ were fixed at 0.45. The spatial frequencies of the additive gratings matched those used in the multiplicative conditions.

3. EXPERIMENT 1

The first experiment examined differences between the processing of first- and second-order transparencies, by quantifying the symmetry of the transparent percept. We measured the smallest disparity, d in equations (2) and (3), that was required by subjects to perceive transparency. Disparity was varied between trials using APE, an adaptive probit algorithm (Watt & Andrews 1981). After each session, a psychometric function was fit to the subject's responses by APE. The disparity at which subjects reported transparency on 50% of trials was deemed the disparity threshold for perceived transparency. Each session consisted of 64 trials, and was repeated three times.

The signal $C(x, y)$ was either a horizontal sinusoidal grating or a noise pattern. The other signal, $E(x, y)$, was always vertical (as in equation (1)). Subjects reported transparency when the two signals $E(x, y)$ and $C(x, y)$ were perceived to be at different depths. In the multiplicative condition, when $C(x, y)$ was a horizontal grating, we tested both crossed and uncrossed disparities. When it was noise, we used only crossed disparities. We did this because, when the disparity of the envelope $E(x, y)$ was uncrossed, the signal appeared diplopic and could not be fused. For the additive signal combinations, crossed and uncrossed conditions were run as separate sessions because there were separate disparity thresholds for each.

Finally, note that when $C(x, y)$ was a horizontal grating, the two monocular images in both additive and multiplicative conditions were shifted versions of one another. They were, therefore, consistent with a single coherent surface under binocular viewing. In these cases the percept was often bistable so that coherent and transparent percepts could be reported at different instances. When perception was bistable, subjects were instructed to report transparency.

(a) *Results and discussion*

We summarize the data from the three subjects by showing how transparent disparity thresholds (TDTs) varied as a function of the spatial frequency of $E(x, y)$. Figure 2*a* shows the TDTs for each subject in the multiplicative condition when $C(x, y)$ was a horizontal

sinusoidal grating. The data for each subject show similar behaviour, with the lowest threshold occurring near 0.4 cpd, approximately two octaves lower than the carrier frequency.

Figure 2*b* shows the results collapsed across the three subjects for two different carrier frequencies, one octave apart, at 2.25 and 4.5 cpd. The threshold functions in these two cases appear to be shifted versions of one another. In particular, for the lower frequency, the minimum TDT was one octave lower, at about 0.2 cpd instead of 0.4 cpd This suggests that the envelope TDTs depended upon the carrier frequency.

For the noise carrier in the multiplicative condition, shown in figure 2*c*, the TDTs also show a minimum at *ca.* 0.4 cpd, as in figure 2*a*. However, they do not vary as markedly as a function of envelope frequency. This may be due to the broadband nature of the carrier, in which case the thresholds measured here could reflect the combined effects of several spatial frequency tuned channels.

With second-order stimuli, transparency was only reported when the depth ordering of the envelope $E(x, y)$ appeared nearer to the observer than the carrier $C(x, y)$; i.e. when the envelope disparities were crossed. Although not reported here, when $C(x, y)$ was a vertical grating, the percept of transparency (versus coherence) was found to depend upon the relative phase between $C(x, y)$ and $E(x, y)$ for both additive and multiplicative conditions. Similar observations have been reported in the context of motion capture (e.g. Gurney & Wright 1996). Capture and bistable percepts are common occurrences with transparent stimuli. They probably reflect competition between different visual processes (models) when the input stimuli do not provide a sufficient number of constraints (von Grunau & Dube 1993; Langley 1997).

Finally, figure 2*d,e* shows the data from the additive condition collapsed across the three subjects for both crossed and uncrossed disparities. The curves show that for additive transparency the TDTs decreased as a function of the spatial frequency of $E(x, y)$. No minimum was evident over the range of frequencies tested. Figure 2*d,e* also shows a difference between crossed and uncrossed disparity cues when $C(x, y)$ was additive noise. The slope of transparency thresholds as a function of frequency was less steep for the uncrossed than for the crossed disparities.

The transparency asymmetry observed with contrast envelopes may be explained by Metelli's two constraints. The contrast envelope, $1+mE(x, y)$, is consistent with a transparent overlaying surface because it is a positive-valued function. The envelope does not change the sign of the underlying grating or noise carrier, and the intensity differences of the carrier at the location of the envelope troughs are smaller than those at the crests. Allowing the carrier to be the overlaying surface violates Metelli's second constraint, because a contrast-modulated carrier is not a positive-valued function.

These results favour a two-channel model over a single-channel model. A single-model of transparency, such as that proposed by Weinshall (1991), that incorporates a nonlinearity may detect two disparity signals, but it would not be able to determine the origin of each disparity signal (i.e. whether it is first- or second-order). Hence, this single-channel model would not be consistent with the asymmetry. A two-channel model could explain the asymmetry because each disparity signal would have been processed by a separate channel.

To explain the symmetrical transparency observed in the additive condition, note that additive transparencies define the boundary condition of Metelli's second inequality constraint (see Beck 1984). This is because the intensity differences taken in a vertical direction for both the dark and light regions of the vertical square-wave grating are equal. In addition, the additive transparencies may be decomposed into a sum of two positive-valued intensity functions. Therefore, there are no simple monocular cues that constrain either signal to be an overlaying or underlying surface, and so the disparity cues may be used to specify the depth ordering of the binocular signals.

Figure 2*d,e* shows another difference in the additive noise condition. Although subjects reported transparency for both crossed and uncrossed disparities, the slope of the curves were less steep in the case of the uncrossed disparities compared with the crossed. Noise patterns are, however, broadband stimuli. For these binocular signal pairs, there will be interference between the disparity cues present in $E(x, y)$ and $C(x, y)$, especially when their individual frequency spectra overlap and their binocular disparities are different. The distortion products introduced into the image signal by an additive combination of signals may be second-order. These distortion products may occur as a result of either early visual nonlinearities, or image transformations that follow bandpass filtering and binocular cross correlation. Therefore, the different slopes found for crossed and uncrossed disparities from the additive (broadband) signal combinations further implicate second-order processes in binocular depth asymmetries.

Finally, it is interesting that in the additive condition when $C(x, y)$ was random noise, subjects reported transparency when the frequency of $E(x, y)$ was greater than 0.8 cpd. In the other conditions, subjects were unable to report transparency reliably at such higher frequencies. Wilson *et al.* (1991) reported that two added signals must differ in scale by two or more octaves in order for binocular depth transparency to be perceived. Our data are consistent with these reports but also suggest that frequency constraints apply to both first- and second-order signals. The dependence of transparency upon spatial frequency is likely to reflect the properties of the mechanisms used for binocular matching by the visual system. This is because a dependency upon spatial frequency may not be predicted from Metelli's (1974) monocular constraints. A two-channel model of transparency may explain these trends. This model would posit that first-order transparencies are detected by bandpass processes tuned to different scales and/or orientation, whereas second-order transparent signals are detected by a spatial-frequency-selective second-order channel (e.g. Wilson & Kim 1994).

4. EXPERIMENT 2

Experiment 1 showed that the TDTs for multiplicative transparency were minimal when the frequency of $E(x, y)$ was relatively low, near 0.4 cpd. For the additive condition, however, no minimum was found over the

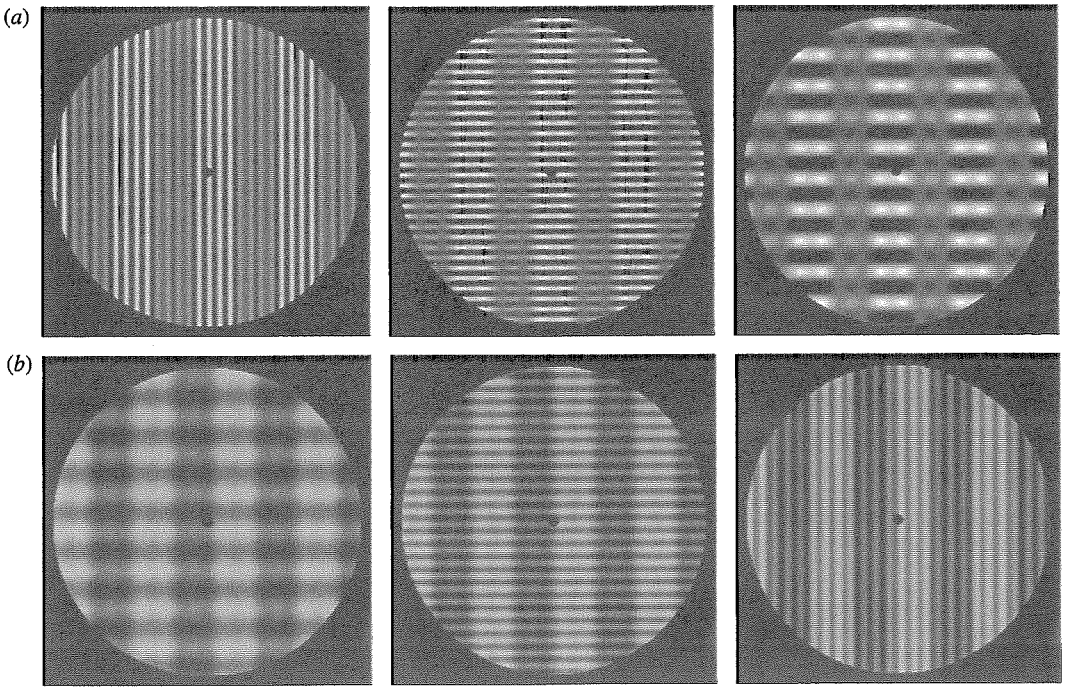

Figure 3. Example of the binocular stimuli used for experiment 2. (*a*) Multiplicative: crossed-eyed fusion yields a strong percept of depth when the centre and right images are fused. When the carrier gratings for the envelope are orientated orthogonally, the percept of depth is much weaker, but disparity discrimination thresholds may still be determined. For these stimuli, the envelope was not seen in depth behind its carrier grating. (*b*) Additive: the matching vertical gratings could be seen both in front of and behind the rivalling gratings. The perceived depth of the rivalling gratings may be seen to lie in the plane of the fixation spot (see also Wurger & Landy 1989).

same frequency range. One possibility is that the minimum found for the multiplicative condition is indicative of the peak sensitivity of a second-order mechanism (see Sutter *et al.* 1995; Langley *et al.* 1996).

This possibility was the focus of the second experiment. In this experiment we manipulated the disparity d of $E(x, y)$ in a disparity discrimination task. However, in this case the signals $C_l(x, y)$ and $C_r(x, y)$ were allowed to differ. We measured the percentage of trials on which subjects reported that the signal $E(x, y)$ was seen in front of a fixation spot. As in the previous experiment, d was varied using APE. Disparity discrimination thresholds (DDTs) were determined from the slope of the psychometric function (fitted by APE) at about the point where subjects reported 'in front of' on 50% of trials. Each session consisted of 64 trials and was repeated three times.

In multiplicative conditions, $C_l(x, y)$ and $C_r(x, y)$ differed in three ways. In a matched condition, the left and right gratings were both horizontal with the same spatial frequency. In a Vert–Hor condition, the two signals differed in orientation (one vertical and one horizontal), but had the same spatial frequency (cf. Liu *et al.* 1992). Finally, in an F–3F condition, the two signals were both horizontal, with frequencies of 1.5 and 4.5 cpd. For the additive case, only two conditions were reported, namely the matched and F–3F conditions, because the curve shapes found for the additive conditions were similar. The contrast and spatial frequency for one of $C_l(x, y)$ or $C_r(x, y)$ was always the same as that used in experiment 1. Figure 3 shows examples of the stimuli.

(a) *Results and discussion*

Results for this experiment are shown in figure 4. With multiplicative signal combinations, the DDTs as function of the frequency of $E(x, y)$ show a minimum near 0.4 cpd for both F–3F and Vert–Hor conditions. For the matched carrier condition, the DDTs flattened out at frequencies above 0.4 cpd. In the F–3F condition, subject DS was unable to discriminate disparities for low or high spatial frequencies of $E(x, y)$. At these frequencies this subject reported that he was only able to see the contrast envelope in front of the fixation spot and was, therefore, unable to perform the discrimination task.

When $E(x, y)$ and $C(x, y)$ were added together, figure 4*b* shows no minimum in DDTs. Rather, the DDTs decreased as a function of increasing frequency. This trend may be explained by a first-order model of stereopsis based upon phase differences (Fleet *et al.* 1996). To transform an interocular phase-difference into a binocular disparity, one must divide the phase difference by the horizontal frequency of the signal. This leads to a disparity error that is inversely proportional to the spatial frequency of the signal. The predictions from this model are shown in the left panel of figure 4*b*. The DDTs shown in figure 4*b* are broadly similar to those predicted by such a model (PBH's thresholds were consistently flatter across all conditions reported here.)

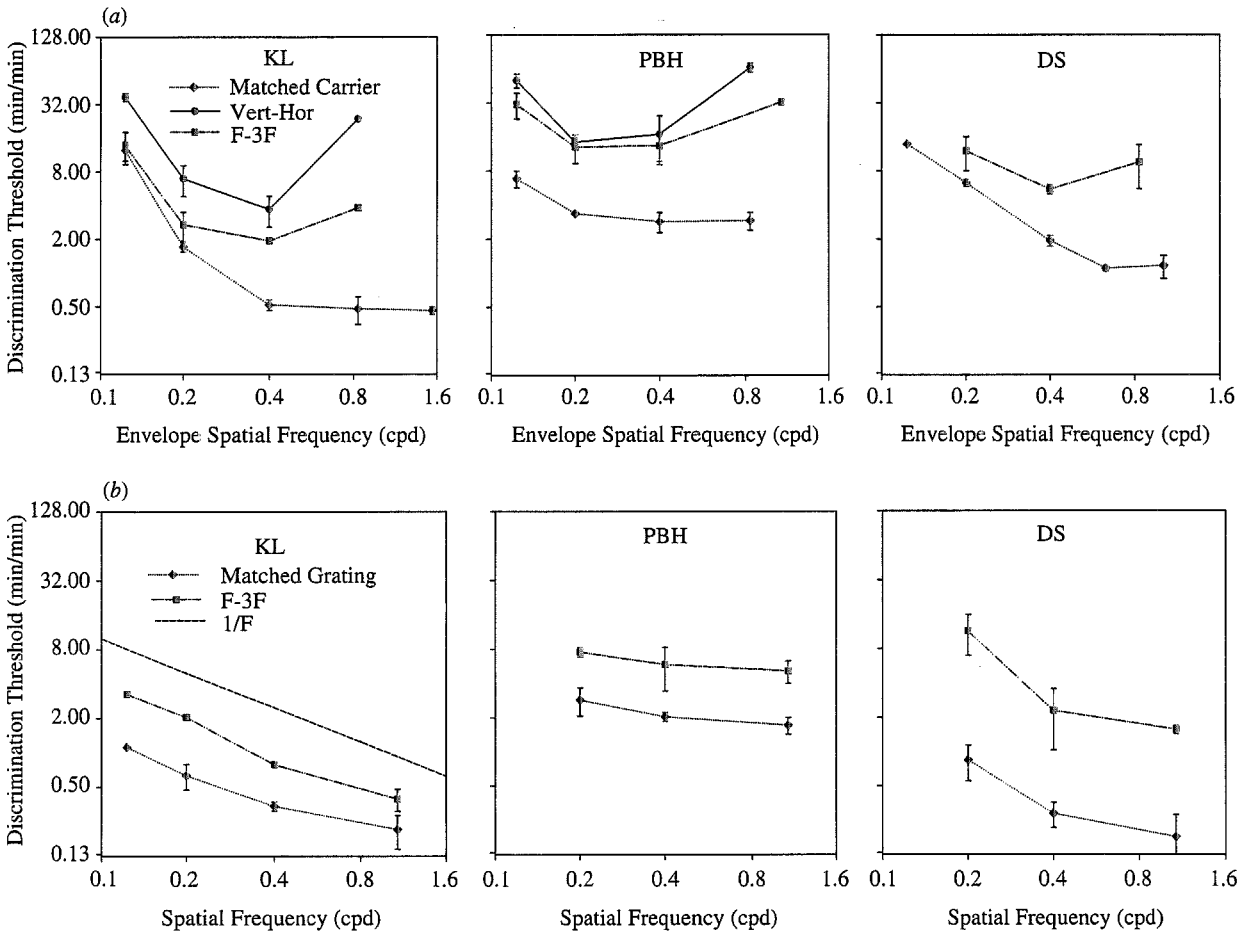

Figure 4. (*a*) DDTs for the multiplicative condition are plotted as a function of envelope spatial frequency on log–log axes. (*b*) Shows each subject's data for the additive condition. The dashed line labelled 1/F shown in the left panel of (*b*) is the relationship predicted if DDTs are proportional to wavelength in a horizontal direction. In these figures, the error bars represent the standard error of each subject's thresholds taken across different sessions.

Schor *et al.* (1984) have measured DDTs for first-order stimuli. They used difference-of-Gaussian (DoG) signals and found that the DDTs decreased as the peak frequency of the DoG functions increased (up to 2.5 cpd). Therefore, the decreasing DDTs found over this range of frequencies are consistent with those expected from a first-order mechanism, as explained above. (The flattening of DDTs above 2.5 cpd may reflect position-shifted disparity detectors (cf. Fleet *et al.* 1996).) These trends were found in figures 2*d*,*e* and 4*b* when $E(x, y)$ was an additive (first-order) signal. The similarity between the disparity thresholds for DoG stimuli, transparency thresholds (experiment 1) and DDTs (experiment 2) suggest that the binocular disparities for additive conditions were detected by first-order processes.

When $E(x, y)$ was multiplied (second-order) with $C(x, y)$, the thresholds for transparency and DDTs showed a different trend. Figures 2*a*–*c* and 4*a* show that a minimum occurred near 0.4 cpd. Such a minimum does not reflect the curve shapes that would be expected from a first-order model of binocular matching. The similarities between the curve shapes for the envelope signal $E(x, y)$ across the two experiments, and the differences between first- and second-order conditions suggest that there was a common, and probably second-order, process that mediated the disparity signal.

In both experiments 1 and 2, our results show that transparency thresholds for contrast envelopes first decreased, and then increased as a function of the frequency of $E(x, y)$. This trend was pervasive in our data, and suggests that the processing of disparity from contrast envelopes was frequency selective. Consistent with this, Langley *et al.* (1996) found that sensitivity for the detection of envelope spatial orientation was maximal at approximately one-tenth of the carrier spatial frequency. Sutter *et al.* (1995), using bandpass-filtered random-dot noise patterns, reported that sensitivity to envelope frequencies was maximum when the envelope was approximately one-eighth to one-sixteenth of the centre frequency of bandpass noise carriers. In comparing the results of Sutter *et al.* and of Langley *et al.*, note that narrowband filtering itself introduces contrast variations, which will probably affect the resultant contrast-envelope sensitivity curves. This may explain why the envelope sensitivity curves reported by Langley *et al.* were more tightly concentrated about their minimum than those reported by Sutter *et al.* A similar feature can be seen by comparing figure 2*b* with 2*c*. The curves for the perceived

transparency thresholds are more tightly tuned about their minimum for the grating condition than the noise condition. These data could be explained if the second-order channel pooled the responses from different spatial-frequency-tuned channels, because random-dot noise patterns are broadband patterns.

Finally, figure 4*a* shows that the DDTs for the (second-order) matched condition differ from the other multiplicative conditions. Here the DDTs decreased as a function of the spatial frequency of $E(x, y)$ and then flattened out at frequencies above 0.4 cpd. Morgan & Castet (1997) found that the DDTs for sinusoidal gratings as a function of orientation, were constant when measured by interocular phase differences. For the matched condition, this model would posit that DDTs vary as a function of envelope spatial frequency with a slope of 1/F, as in the additive conditions mentioned earlier. This is because an increase in the envelope spatial frequency will decrease the orientation of the envelope's sidebands or linear frequency components. This prediction is inconsistent with the trend reported for this matched condition. One can also note that the flattening of the DDTs occurred at around 0.4 cpd rather than at 2.5 cpd, as in first-order signals. This difference could reflect an average taken between first- and second-order processes. For higher spatial frequencies, the increasing DDTs from a second-order process may have been offset by decreasing first-order thresholds (cf. Lin & Wilson 1995).

5. GENERAL DISCUSSION

There is growing evidence for first- and second-order processing in binocular stereopsis. Hess & Wilcox (1994) showed that envelope disparities can influence stereo-acuity, especially when the envelope and the carrier have different disparities. They concluded that stereo-acuity depends on the envelope size when the stimulus bandwidth is smaller than 0.5 octaves. Sato & Nishida (1993) presented subjects with second-order random-dot stereograms, much like some of the stimuli used in studies of second-order motion. They found that the upper limits on disparity were lower with the second-order stimulus than with conventional random-dot stereograms. Liu *et al.* (1992) using similar stimuli to the ones manipulated here, reported that, when presented with binocular Gabor stimuli in which the left- and right-eye sinusoidal carriers were perpendicular to one another, subjects perceived depth correctly from the envelope while the carrier components were in binocular rivalry.

Although many of these data support the idea of a two-channel hypothesis, they do not entirely rule out the possibility that a single-channel model of transparency, which exploits a nonlinearity introduced into the binocular pathway, could explain binocular depth perception for second-order signals. A key feature of our results in favour of the two-channel model is the transparent asymmetry reported for second-order signals (see also Frisby & Mayhew 1978; Kersten 1991). It would be difficult to explain the depth asymmetry by a single-channel model of transparency because the origin of the two disparity signals would be undiscernible. On the other hand, a two-channel model may explain the depth asymmetry because second-order disparity signals would be represented separately. The frequency tuning curves reported here for contrast envelopes resemble those found in spatial vision tasks (Sutter *et al.* 1995; Langley *et al.* 1996). Again, it would be difficult to explain how different tuning curves between first- and second-order signals could arise from a single-channel model of transparency.

Our data also support the idea that Metelli's (1974) constraints on monocular transparency affect binocular depth perception (Kersten 1991). However, our data implicate one simple strategy by which Metelli's constraints may be introduced into visual processes, namely, a second-order channel. This is because a contrast envelope, as detected by a second-order channel, is a positive-valued signal (Fleet & Langley 1994). The two-channel model posits that first-order transparencies may be detected by multiple peaks in the interocular cross-correlation function of a first-order channel as in Weinshall (1991), whereas second-order signals may be detected by interocular cross correlations of the second-order channel. Hence, a two-channel model could reflect a combined strategy exploited by the visual system that leads to the detection of binocular disparities for both first- and second-order signal combinations. This view supports the notion that the visual system may represent multiple depth cues at common image locations and that the motivation for a two-channel model stems from the incompatible nature of additive and multiplicative signal combinations.

Portions of this research were presented at the European Conference on Visual Perception (ECVP) in 1994 and 1995. D.J.F. was funded in part by NSERC Canada, and by an Alfred P. Sloan Research Fellowship.

REFERENCES

Beck, J. 1984 Perception of transparency in man and machine. *Hum. Mach. Vision* **2**, 1–12.

Beck, J. Prazdny, K. & Ivry, R. 1984 The perception of transparency with achromatic colors. *Percept. Psychophys* **5**, 407–422.

Burton, G. J. 1973 Evidence for non-linear response process in the visual system from measurements on the thresholds of spatial beat frequencies. *Vision Res.* **13**, 1211–1255.

Chubb, C. & Sperling, G. 1988 Drift-balanced random-stimuli: a general basis for studying non-Fourier motion perception. *J. Opt. Soc. Am.* **5**, 1986–2007.

Fleet, D. J. & Langley, K. 1994 Computational analysis of non-Fourier motion. *Vision Res.* **34**, 3057–3079.

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

Frisby, J. P. & Mayhew, J. E. W. 1978 The relationship between apparent depth and disparity in rivalrous-texture stereograms. *Perception* **7**, 661–678.

Gurney, K. N. & Wright, M. J. 1996 A model for the spatial integration and differentiation of velocity signals. *Vision Res.* **36**, 2939–2955.

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

Kersten, D. 1991 Transparency and cooperative computation of scene attributes. In *Computational models of visual processing* (ed. M. Landy & J. A. Movshon), pp. 209–228. London: MIT Press.

Langley, K. 1997 Degenerate models of multiplicative and additive motion transparency. In *Proc. British Mach. Vision Conf. Colchester*, vol. 2, pp. 440–450.

Langley, K., Fleet, D. J. & Hibbard, P. B. 1996 Linear filtering precedes nonlinear processing in early vision. *Curr. Biol.* **6**, 891–896.

Langley, K. Fleet, D. J. & Hibbard, P. B. 1998 Stereopsis from contrast envelopes. *Vision Res.* (Submitted.)

Liu, L. Schor, C. W. & Ramachandran, V. S. 1992 Positional disparity is more efficient in encoding depth than phase disparity. *Invest. Ophthalmol. Vis. Sci.* **33**(Suppl.), 1373.

Lin, L. & Wilson, H. R. 1995 Stereoscopic integration of Fourier and non-Fourier patterns. *Invest. Ophthalmol. Vis. Sci.* **36**(Suppl.), 364.

Marr, D. 1982 *Vision*. New York: W. H. Freeman & Co.

Metelli, F. 1974 The perception of transparency. *Scient. Am.* **230**, 90–98.

Morgan, M. J. & Castet, E. 1997 The aperture problem in stereopsis. *Vision Res.* **37**, 2737–2744.

Sato, T. & Nishida, S. 1993 Second-order depth perception with texture-defined random-dot stereograms. *Invest. Ophthalmol. Vis. Sci.* **34**(Suppl.), 1438.

Schor, C. M., Wood, I. C. & Ogawa, J. 1984 Spatial tuning of static and dynamic local stereopsis. *Vision Res.* **24**, 573–578.

Sutter, A., Sperling, G. & Chubb, C. 1995 Measuring the spatial-frequency selectivity of 2nd-order mechanisms. *Vision Res.* **35**, 915–924.

von Grunau, M. & Dube. S. 1993 Ambiguous plaids: switching between coherence and transparency. *Spatial Vision* **7**, 199–211.

Watt, R. & Andrews, D. P. 1981 Adaptive probit estimation of psychometric functions. *Curr. Psychol. Rev.* **1**, 205–214.

Weinshall, D. 1991 Seeing 'ghost' planes in stereo vision. *Vision Res.* **31**,1731–1749.

Wilcox, L. M. & Hess, R. F. 1996 Is the site of nonlinear filtering in stereopsis before or after binocular combination. *Vision Res.* **36**, 391–399.

Wilson, H. R. & Kim, J. 1994 A model for motion coherence and transparency. *Vis. Neurosci.* **11**, 1205–1220.

Wilson, H. R., Blake, R. & Halpern, L. 1991 Coarse spatial scales constrain the range of binocular fusion on fine scales. *Vision Res.* **8**, 229–236.

Wurger, S. M. & Landy, M. S. 1989 Depth interpolation with sparse disparity cues. *Perception* **18**, 39–54.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.