



Reversed stereo depth and motion direction with anti-correlated stimuli

Jenny C.A. Read ^{a,*}, Richard A. Eagle ^b

^a *University Laboratory of Physiology, University of Oxford, Parks Road, Oxford OX1 3PT, UK*

^b *Department of Psychology, University of Oxford, Parks Road, Oxford OX1 3PT, UK*

Received 8 August 1999; received in revised form 6 April 2000

Abstract

We used anti-correlated stimuli to compare the correspondence problem in stereo and motion. Subjects performed a two-interval forced-choice disparity/motion direction discrimination task for different displacements. For anti-correlated 1d band-pass noise, we found weak reversed depth and motion. With 2d anti-correlated stimuli, stereo performance was impaired, but the perception of reversed motion was *enhanced*. We can explain the main features of our data in terms of channels tuned to different spatial frequencies and orientation. We suggest that a key difference between the solution of the correspondence problem by the motion and stereo systems concerns the integration of information at different orientations. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Stereo motion; Anti-correlated stimuli; Orientations

1. Introduction

The reconstruction of depth from retinal disparity, and the perception of movement in two-frame stimuli, present the brain with analogous problems. In each case, the visual system needs to solve a correspondence problem: which region in the left retina (first frame) corresponds to a region in the right retina (second frame). This problem is particularly stark in the case of random-dot patterns. These present a multitude of false matches, since any two black dots could potentially be partners. Many models have been developed to solve this problem (e.g. Marr & Poggio 1976; Poggio & Poggio, 1984; Blake & Wilson, 1991).

When stereopsis and motion have been studied together, the results have generally been similar. Properties such as contrast sensitivity, D_{\max} and D_{\min} , and dependence on dot size and density in random dot patterns are all similar between the two systems (Glennerster, 1998). In both systems, there is physiological and psychophysical evidence for spatial frequency and orientation selective channels (de Valois, Albrecht, &

Thorell, 1982; DeAngelis, Ohzawa, & Freeman, 1991; Mansfield & Parker, 1993; Eagle, 1997; Freeman & Ohzawa, 1990; Prince, Eagle, & Rogers, 1998; though see also Yang & Blake, 1991). However, there is one result in the literature which suggests there may be an important difference between the two. This concerns the response to anti-correlated stimuli, in which the contrast in one image is reversed. Anti-correlated kinematograms cause a clear perception of motion in the reversed direction (Anstis, 1970; Sato, 1989). In contrast, no depth reversal has been reported with anti-correlated stereograms in which the left eye's image is simply an inverted, disparate version of the right eye's. When sparse anti-correlated stimuli are presented, depth is perceived in the correct direction (von Helmholtz, 1909; Cogan, Kontsevich, Lomakin, Halpern, & Blake, 1995). For dense anti-correlated stimuli, no clear depth is perceived: the stimuli appear strange and 'lustrous', and the subjects perform at chance (Julesz, 1971; Cogan, Lomakin, & Rossi, 1993; Cumming, Shapiro, & Parker, 1998). In contrast, the disparity-tuning curve of cells in primary visual cortex is inverted, rather than abolished, when the stimulus is anti-correlated (Cumming & Parker, 1997). With this observation, the absence of reversed depth in anti-cor-

* Corresponding author. Tel.: +44-1865-272530.

E-mail address: jenny.read@physiol.ox.ac.uk (J.C.A. Read).

related stereograms takes on a new significance. Understanding why observers do not perceive reversed depth may throw new light on whether and how these neuronal responses are related to perception.

The phenomenon of reversed perception can be qualitatively understood if we assume that the brain analyses an image within channels tuned to different spatial frequencies and orientations (Campbell & Robson, 1968; Blakemore & Campbell, 1969). Consider a single vertical Fourier component: that is, a sine-wave grating of spatial period λ . Evidently, displacing this pattern a distance d to the right is indistinguishable from displacing it a distance $(d + m\lambda)$, where m is any integer. In cases such as this, where the periodicity of the stimulus means that many different displacements are equally possible, the visual system probably prefers the match with the smallest displacement; McKee and Mitchison (1988) present evidence that this is the case for stereopsis. For a single Fourier component, this smallest displacement is $\min_m(d + m\lambda)$. For $d < \lambda/2$, therefore, we obtain the correct displacement d . For anti-correlated stimuli, we notice that moving a sine-wave grating a distance d to the right is formally equivalent to displacing it a distance $(m\lambda + \lambda/2 - d)$ to the left and reversing its contrast. The preferred match is therefore $\min_m(m\lambda + \lambda/2 - d)$. For relatively small displacements ($d < \lambda/2$), this corresponds to movement of $\lambda/2 - d$ in the *opposite* direction. This is an intuitive explanation of how reversed motion can occur.

In reality, channels within the visual system see more than just a single Fourier component. The bandwidth of spatial frequency channels is of order 1–2 octaves (de Valois & de Valois, 1988). The cross-correlation function (CCF) provides a helpful way to generalise these arguments to finite-bandwidth channels (Cleary & Braddick, 1990a,b). Fig. 1 shows the CCF for correlated and anti-correlated one-dimensional (1d) noise stimuli, filtered by 1 octave channels centred on three different spatial frequencies. For correlated stimuli, the CCF has a peak at the correct displacement d , irrespective of the preferred spatial period and orientation of the channel. For small displacements, the largest peak is the closest to the origin, so simply picking the largest peak gives the correct displacement d . As d increases, subsidiary peaks — which may be on the wrong side of the origin — become closer to the origin than the largest peak. So a preference for peaks closer to the origin could explain the perceptual reversal seen with narrow-band stimuli (Cleary & Braddick, 1990a; Prince & Eagle, 2000b). As d is increased further, the subsidiary peaks may become small enough to be lost in noise — hence the existence of an upper limit D_{\max} on direction discrimination. This qualitative model predicts that D_{\max} for each channel is proportional to the preferred period λ of that channel; this is consistent with experimental evidence across a range of spatial frequencies (Chang & Julesz, 1985; de Bruyn & Orban, 1989; Cleary & Braddick, 1990a,b; Smallman &

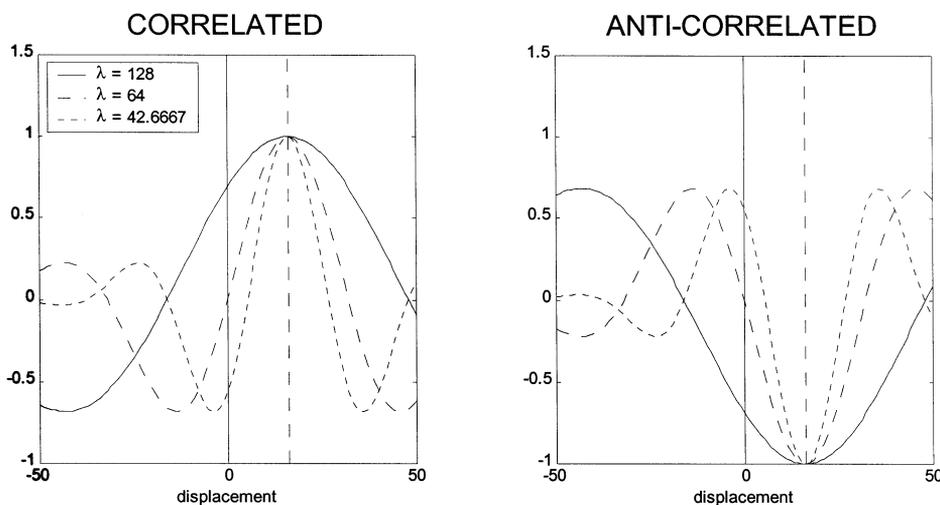


Fig. 1. Cross-correlation functions (CCFs) for a filtered disparate image pair, either correlated (left-hand plot) or anti-correlated (right-hand plot). The basic image is a set of 128 numbers uniformly distributed between ± 0.5 . The second image is the original image displaced 16 units to the right with wrap-around. In the right-hand plot, the second image is also anti-correlated. The images are then filtered by a Gabor function with bandwidth 1 octave and different spatial periods. Each filter represents a model spatial frequency channel. The graphs show the CCFs of the filtered images (normalised to have the same peak amplitude). In each case, the CCF of the anti-correlated image pair is the inverse of that for the correlated image-pair. Note that when the images are correlated, the CCF has its central peak at the true displacement for all filters (16 units, indicated by a vertical line). When the images are anti-correlated, the CCF has its central trough at 16 units, but its side-peaks occur at different positions ($d \pm \lambda/2$) for the different filters. The filters illustrated all have $\lambda > 2d$, so the side-peak closest to the origin is on the left of the origin. Thus, if these channels reported the position of the peak closest to the origin, they would agree on the sign of the displacement (which would be the reverse of the true value), but would differ on its value.

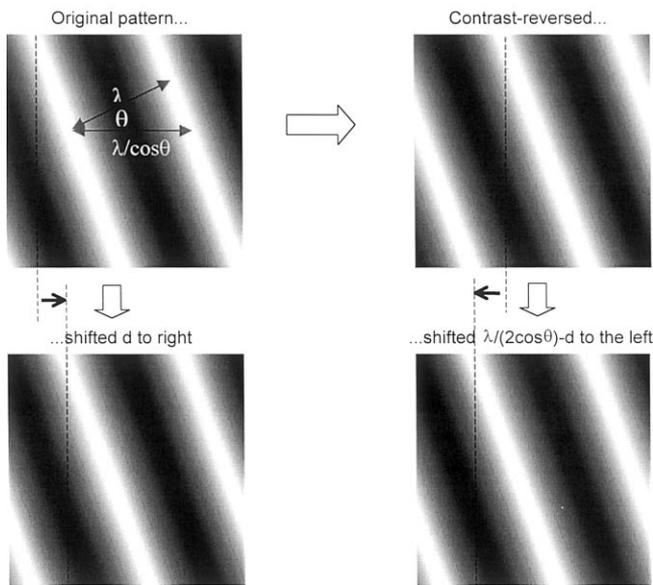


Fig. 2. A sinusoidal grating, of spatial period λ , oriented at θ to the vertical. Inverting the contrast of this pattern and shifting it $\lambda/(2 \cos \theta) - d$ to the left gives the same result as shifting the original pattern a distance d to the right.

MacLeod, 1994). For anti-correlated stimuli, the CCF is inverted. It has two peaks corresponding to the troughs on either side of the central maximum for the correlated stimulus, at $d \pm \lambda/2$. If the peak nearest the origin is the most likely to be chosen, and $d < \lambda/2$, this is in the wrong direction. Quantitative models based on these ideas can explain a wide range of data (Prince & Eagle, 2000a).

Thus we can understand why anti-correlated narrow-band stimuli could cause perceptual reversal in both motion and stereopsis. However, the literature concerning random-dot patterns suggests that anti-correlated broad-band stimuli cause perceptual reversal only for motion. These different responses provide a clue to the different mechanisms underlying stereopsis and motion discrimination. One possibility is that the two systems have different ways of handling conflicts between channels. Consider the qualitative model just outlined. For correlated stimuli, all channels have a peak at the correct displacement d . For channels with $\lambda/2 > d$, this correct match is the closest peak to the origin. Higher frequency channels have secondary peaks closer to the origin (cf. Fig. 1), but coarse-to-fine algorithms (e.g. Marr & Poggio, 1976) may be used to select the correct peak which agrees with other channels. For anti-correlated stimuli, the situation is very different. Now the position of the side-peak, $(\lambda/2 - d)$, is different for each spatial frequency and orientation channel. Thus, for anti-correlated stimuli, each channel will report a different amount of movement (although all those for which $\lambda/2 > d$ will agree on its direction). There is no displacement where all the channels agree on a peak. Thus one

explanation of why one could see reversed motion with anti-correlated kinematograms, but not reversed depth with anti-correlated stereograms, is that the motion system tolerates disagreement in reported magnitude between different channels, whereas stereoscopic depth perception requires all channels to agree. One reason why this might occur is that, for natural images, the motion system has access to information at a variety of temporal scales, and so perhaps can afford to be more tolerant of disagreement at any one temporal scale.

An important difference between stereopsis and motion concerns the dimensionality of the problem. Stereoscopic depth from horizontal disparity is essentially a one-dimensional problem, involving displacements only along the horizontal axis, whereas motion detection is inherently two-dimensional, involving displacements in all directions. The motion system may therefore be sensitive to components whose orientations are not orthogonal to the direction of displacement. These contribute additional horizontal scales. Consider a single Fourier component oriented at an arbitrary angle θ to the vertical. Moving such a component horizontally d to the right is formally equivalent to displacing it a distance $(m\lambda/\cos \theta + \lambda/(2 \cos \theta) - d)$ to the left and reversing its contrast (Fig. 2). The preferred match is therefore

$$\min_m \left(\frac{m\lambda}{\cos \theta} + \frac{\lambda}{2 \cos \theta} - d \right).$$

For relatively small displacements ($d < \lambda/(2 \cos \theta)$), this corresponds to movement of $\lambda/(2 \cos \theta) - d$ in the *opposite* direction. Thus, for anti-correlated two-dimensional stimuli, there is conflict not only between spatial frequency channels, but also between different orientation channels, as illustrated in Fig. 3. It would not be surprising if the stereo and motion systems had different ways of handling conflict between different spatial frequency and/or orientation channels. We therefore investigated the response to correlated and anti-correlated stereograms and kinematograms with a variety of spatial frequency and orientation bandwidths.

2. Methods

2.1. Generation of the stimuli

Our first set of experiments concerned one-dimensional (1d) stimuli, which contain only vertical orientations. We generated the images in Fourier space. Since only vertical orientations are present, the Fourier transform of each image is a function of the horizontal spatial frequency ω only. The phase of the Fourier transform was chosen at random (but constrained to be an odd function of frequency, so that the image is real). Its amplitude was defined to be inversely proportional

to the square root of the frequency. This ensures that the image contains *equal power in equal octaves*: a property approximately satisfied by some natural images (Field, 1987). Since both psychophysical and physiological evidence suggests that spatial-frequency channels have similar octave bandwidth across a range of frequencies (de Valois & de Valois, 1988), the 5-octave images present each channel with roughly equal spectral power. We hoped that this would maximise the potential for cross-channel conflict. However, this choice of frequency spectrum is not critical to our results. We also used random-dot patterns, whose frequency spectrum is statistically flat.

A rectangular spatial frequency filter was then applied to these images. In each case, the central fre-

quency of the filter was 5.6 cycles per image, corresponding to 3.2 cyc/deg. The bandwidth of the filter was either 1, 2, 3 or 5 octaves (Fig. 4). When displaying images of different bandwidths, we set the contrast so as to preserve the property of equal power in equal octaves across the whole stimulus set. By Parseval's theorem, the rms (global) contrast is equal to the square root of the total power summed over all frequencies. This means that our 5 octave stimuli must have $\sqrt{5}$ of the global contrast of the 1 octave stimuli.

In our second set of experiments, we used two-dimensional (2d) stimuli, containing the full range of orientations (Fig. 4). Once again, we chose the Fourier spectrum of our images so as to obtain equal power in equal octaves. However, the Fourier transform of a 2d

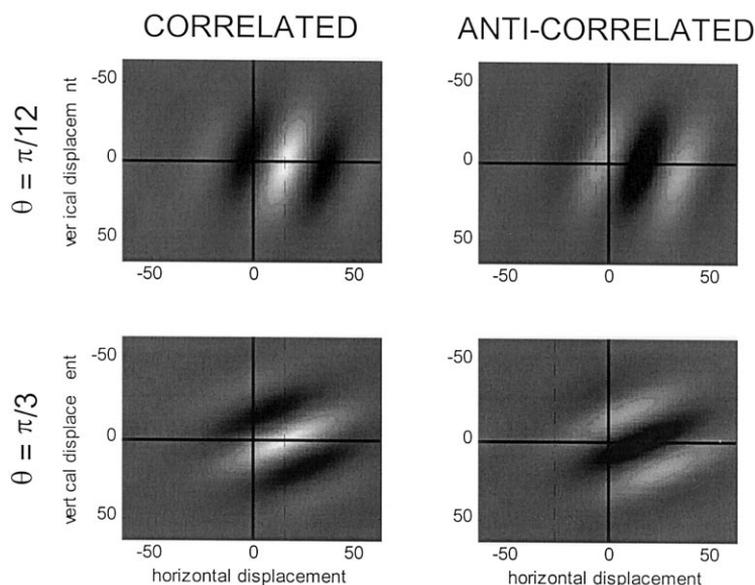


Fig. 3. Two-dimensional cross-correlation functions (CCFs) for a filtered disparate image pair, either correlated (left-hand plots) or anti-correlated (right-hand plots). (The function plotted is actually a 2d Gabor function, representing the mean CCF obtained with random noise images.) In each case, the filters have the same period (43 units), but different orientations: $\theta = \pi/12$ in the upper two plots, and $\theta = \pi/3$ in the lower. In each case, the CCF of the anti-correlated image pair is the inverse of that for the correlated image-pair. Note that when the images are correlated, the CCF has its central peak at the true displacement (16 units in the horizontal direction, indicated by a vertical line). When the images are anti-correlated, the CCF has its central trough at 16 units, but its side-peaks occur at different positions ($d \pm \lambda/(2 \cos \theta)$) for the different orientations. The position of the side-peak closest to the origin is marked with a vertical line.

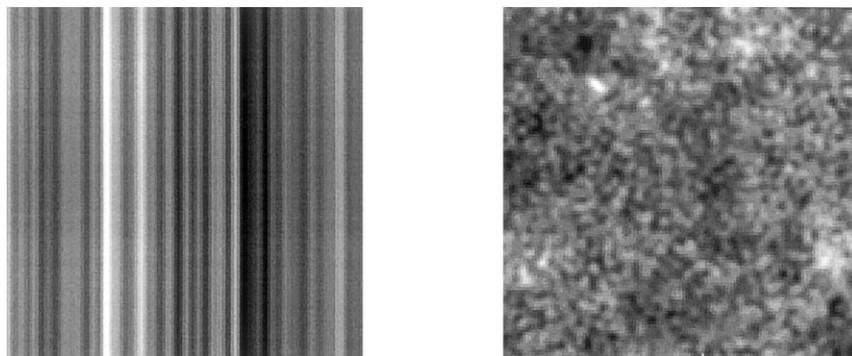


Fig. 4. Example one- and two-dimensional images. In each case, image size is 128×128 pixels, bandwidth 5 octaves centred on 5.6 cycles per image. On the screen, this corresponds to $1.7 \times 1.7^\circ$, central frequency 3.2 cyc/deg.

image is a function of two variables: the magnitude of the spatial frequency ω and its orientation θ . Thus to obtain equal power in equal octaves, the Fourier amplitude must depend on the inverse of the frequency. We wished a single orientation channel to be stimulated roughly equally by our 1d and by our 2d images. Psychophysical masking experiments (Campbell & Kulikowski, 1966; Mansfield & Parker, 1993) suggest that the orientation bandwidth of stereo and motion channels is approximately 30° . We therefore set the contrast of the 2d images so that the global power in a 30° orientation band was equal to the power in the 1d image with the same spatial frequency bandwidth.

2.2. Presentation of the stimuli

An Apple Macintosh drove two monitors in a modified Wheatstone stereoscope configuration. The monitors had a resolution of 85 dots per in. and a refresh rate of 70 Hz. Video attenuators (Pelli & Zhang, 1991) were used to combine the 8 bit signals from the three colour channels into a single 12-bit signal, allowing fine luminance control. The background luminance of the screen was 27 cd/m^2 . In all cases, the mean luminance of our stimulus was equal to this background luminance. It is therefore convenient to define the intensity of points in our images relative to the background. In these terms, anti-correlation corresponds to changing the sign of the intensity. The monitors were gamma-corrected using a Minolta LS-110 photometer. The viewing distance was fixed at 127 cm. At this distance, each pixel on the monitor subtended 0.8×0.8 minarc. Our images were 128 pixels square, i.e. $1.7 \times 1.7^\circ$. A small fixation cross (9×9 pixels, $7' \times 7'$) was permanently visible at the centre of each monitor. In the stereo experiment, nonius lines were also displayed in the intervals between stimulus presentations as a guide to correct vergence.

2.3. Experimental procedure

In both experiments, we used a two-interval forced-choice protocol. The stimulus was visible for a total of 130 ms in each interval (two 65 ms frames in the motion experiments), with a gap of 160 ms between intervals. This presentation is too brief for vergence movements. Observers pressed a keypad button to indicate which interval contained rightward motion/crossed disparity. No trial-by-trial feedback was given. Subjects performed runs of 140 trials, containing 20 trials at each of seven displacements, in pseudo-random order. Our psychophysical functions show the results of four runs, i.e. 80 trials at each displacement. Five subjects (four male, one female) participated in the experiments. All were experienced psychophysical observers with normal or corrected-to-normal vision. JR and RAE

were authors; the others were unaware of the aims of the experiment.

2.4. Image displacement

For each experiment, we generated five sets of random images (128×128 pixels). With each stimulus presentation, one of the five images was chosen at random. Then, a random number of columns of pixels were removed from the right edge of the stored image and added back onto its left edge; similarly, a random number of rows were removed from the top and added onto the bottom. This manipulation ensured that particular visual features did not appear in the same place time after time, so there was no monocular clue to the disparity. This modified image formed the stimulus to the left eye/first frame. This was then displaced by wrapping columns of pixels around to form the stimulus for the right eye/second frame. Therefore, the displaced target was presented directly on the gray homogeneous field. This differs from the traditional procedure with random-dot stereo- and kinematograms, in which a central target region is displaced over a zero-disparity random-dot background. It preserves the Fourier spectrum of the band-pass stimuli without introducing luminance discontinuities.

3. Results

3.1. Experiment 1: one-dimensional band-pass stimuli

Fig. 5 shows data for the stereopsis and motion experiments, using correlated and anti-correlated 1d stimuli. For these two subjects, experiments were conducted at four different spatial frequency bandwidths: 1, 2, 3 and 5 octaves. The results are similar between stereopsis and motion, and for both subjects. For correlated stimuli, subjects score close to 100% for small displacements, falling to chance as the displacement approaches half a cycle of the lowest frequency component. For small displacements of narrow-band anti-correlated stimuli, the subjects score close to zero, indicating that they are perceiving reversed motion and depth, although not as clearly as veridical depth with the correlated stimuli. At larger displacements, performance fell towards chance. The displacement at which this occurred increased with the bandwidth of the stimulus. It was therefore convenient to express the displacement as a fraction of the longest spatial period present in the stimulus.

If performance were limited solely by the longest period present in the stimulus, all four curves in Fig. 5 would coincide. In fact, the 5-octave curve, and to a lesser extent the 3-octave, depart from the narrow-band pattern. For correlated stimuli, the peak performance,

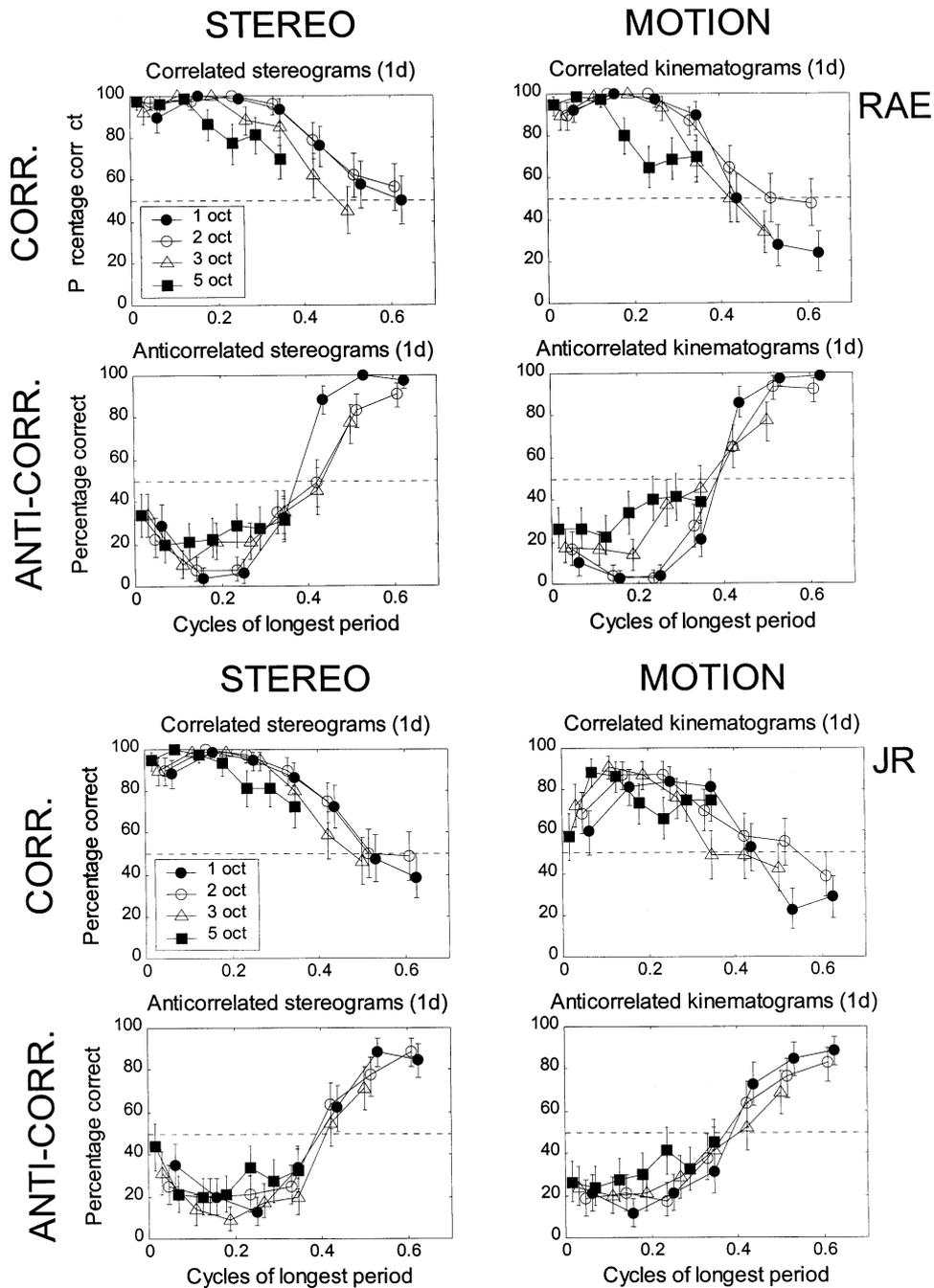


Fig. 5. Psychophysical functions for one-dimensional band-pass noise with four different spatial frequency bandwidths, for subjects RAE (upper set of four plots) and JR (lower set of plots). Percentage correct is plotted against displacement expressed as a fraction of the longest wavelength present in the stimulus (this wavelength, in arc-minutes, is given in the key.) In this and subsequent data figures, each point represents data from 80 trials. Assuming simple binomial statistics, the standard deviation of a data-point is $\sqrt{p(1-p)/n}$, where p is the probability of a correct response, and $n = 80$ is the number of repetitions. The horizontal lines represent chance (50% correct) and the positions of the 2.5 and 97.5% quantiles (38% and 62%, for simple binomial statistics with 80 repetitions). That is, if subjects were responding purely at random, each data-point would have only a 5% chance of lying outside the dashed lines.

attained for displacements around 0.05 cycles, is unimpaired. As the displacement is increased, however, the performance becomes noticeably poorer for larger bandwidth. For anti-correlated stimuli, performance for larger bandwidth stimuli is closer to chance throughout. However, the reversed depth/motion is never abolished:

even for the 5 octave stimuli, performance remains at 20–40% across the range of displacements studied. This departure from chance is highly significant. Assuming simple binomial statistics with 80 trials, the two-tailed 5% significance level is reached at 38% correct, and a score of 25% has significance $P < 0.001$. Thus almost all

the data-points in the 5-octave psychometric functions individually enable us to reject the null hypothesis that performance was at chance at that displacement. (The assumption of simple binomial statistics arguably underestimates the true variance, so the quoted significance levels may be excessively low. However, we can still unequivocally reject the null hypothesis: without

making any assumptions about the underlying stochastic process, the probability that all seven data-points would fall below chance, as observed, is 1 in 2^7 , less than 1%.)

For the broad-band stimuli, and for the narrow-band stereo stimuli, we repeated the experiments with three further subjects. Figs. 6 and 7 show these 1- and

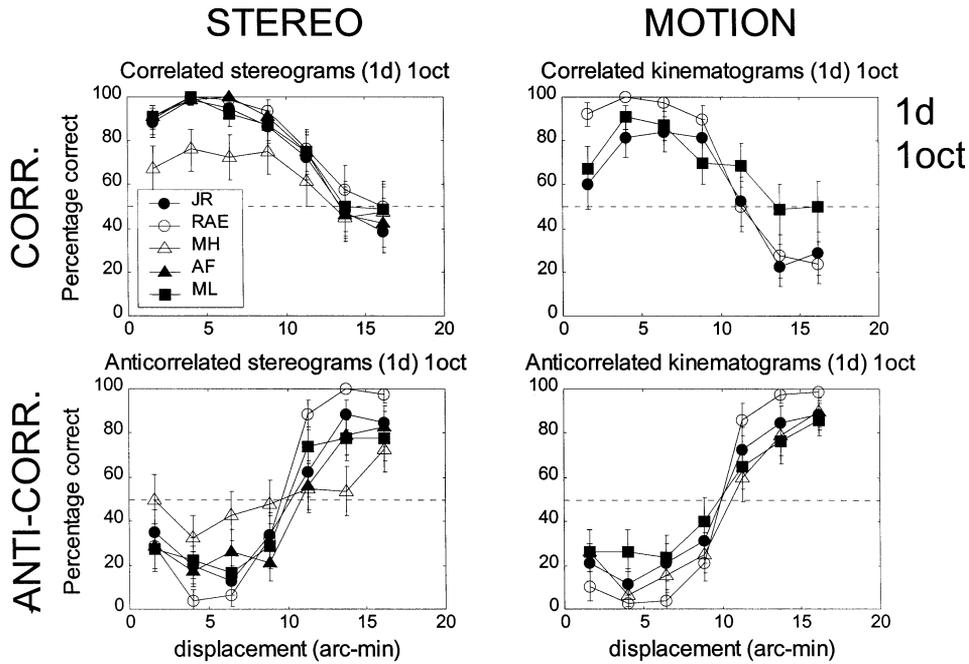


Fig. 6. Psychophysical functions for one-dimensional band-pass noise with spatial frequency bandwidth of 1 octave. In the stereograms, data for four subjects is superimposed; in the kinematograms, not all subjects performed this task.

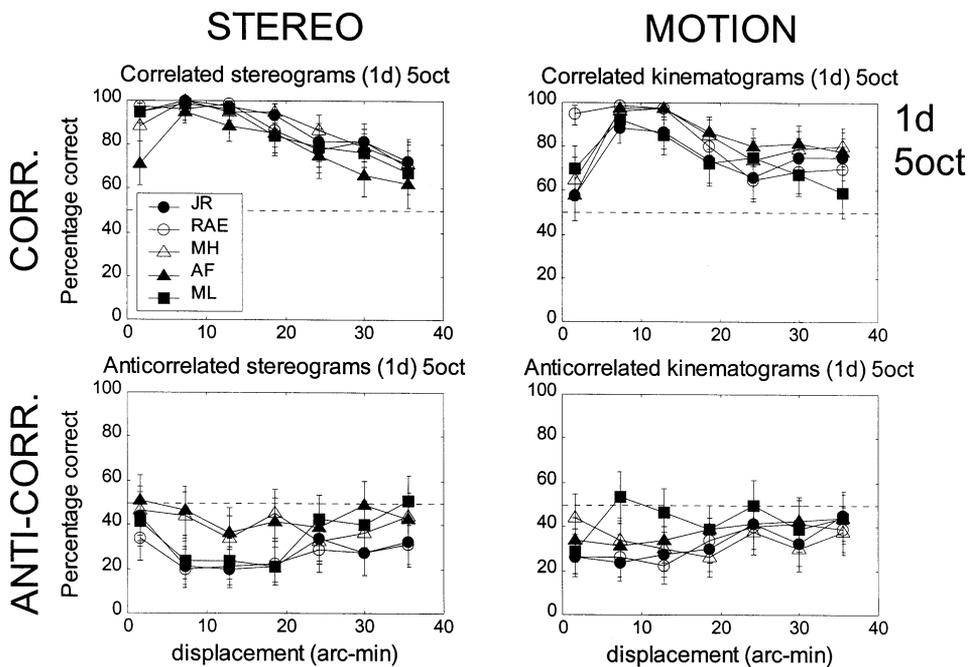


Fig. 7. Psychophysical functions for one-dimensional band-pass noise with spatial frequency bandwidth of 5 octaves. Data for four subjects is superimposed in each plot, as indicated in the key.

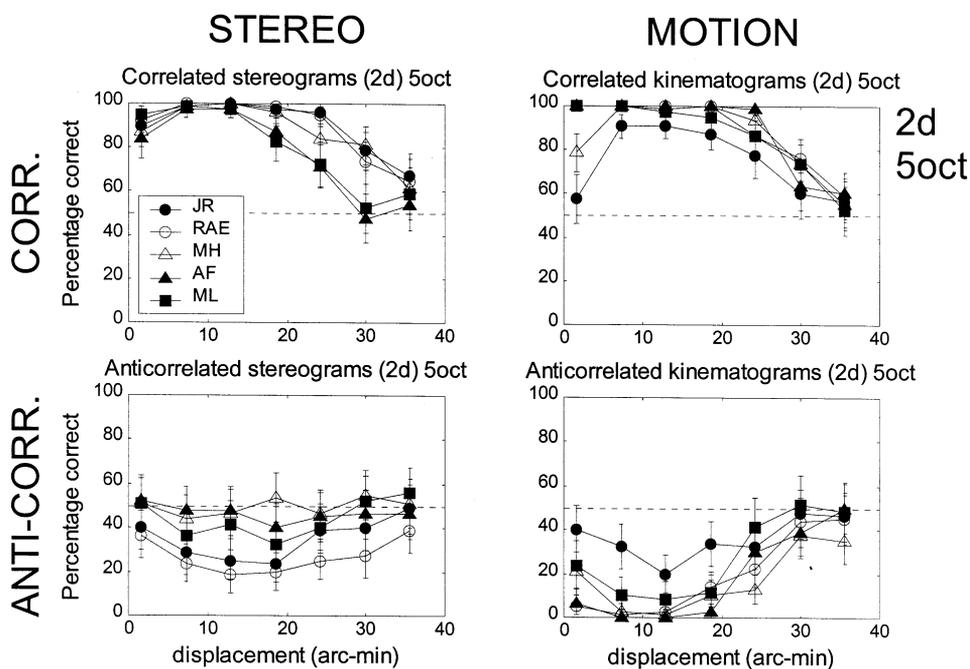


Fig. 8. Psychophysical functions for 2d band-pass noise with spatial frequency bandwidth of 5 octaves. Data for four subjects is superimposed in each plot, as indicated in the key.

5-octave results for all five subjects including the two shown earlier (JR and RAE). There are some inter-subject differences, which are not obviously related to prior experience or training. For instance, even after extensive practice subjects AF and MH could not demonstrate very strong reversed depth with the 5 octave stimuli, whereas ML was able to score 20% the first time he performed the task, with very little prior training. AF and MH thus remain slightly closer to chance for anti-correlated stereograms than for kinematograms, and subject ML vice versa. The two subjects who performed best throughout on our motion experiments (cf. Figs. 7, 9 and 10) were RAE and AF. RAE had extensive experience of motion experiments, but most of AF's prior experience had been with stereo. Thus the differences do not appear attributable to training differences. Subject MH performed poorly for the 1 octave stereograms, although he could do the motion task, and the later stereo experiments (e.g. Fig. 9). This may be due to the low contrast of the 1 octave images.

Despite the differences between subjects, within each subject the pattern is clear. Overall the broadband stereo and motion experiments give similar results, hovering just below chance for both anti-correlated stereograms and kinematograms. Our motion results in particular are strongly at odds with the existing literature, which reports performance close to 0% for anti-correlated kinematograms (Sato, 1989).

Our stimuli are different in several respects from those employed by the previous studies. Most obvi-

ously, they contain only vertical orientations, whereas other workers have used 'two-dimensional' stimuli. Hence we have avoided conflict between different orientation channels. To see if this is responsible for the difference between our results and those of previous studies, we must increase the orientation bandwidth.

3.2. Experiment 2: two-dimensional bandpass stimuli

We repeated our experiment using 5-octave stimuli which contained the full range of orientations. The results are shown in Fig. 8 and can be compared with the one-dimensional results in Fig. 7.

For correlated stimuli, the additional orientations make little difference. For anti-correlated stereograms, once again, subjects JR, RAE and ML are below chance, indicating a weak but consistent perception of reversed depth. However, it is clear that the activation of additional orientation channels has impaired performance. This impairment is enough to destroy the already weak reversed depth in subjects AF and MH; their performance is no longer significantly different from chance.

The surprise comes with anti-correlated kinematograms. Although we have argued that there is increased conflict between orientation channels, most subjects show a dramatic *improvement* in performance (in the sense that their results are further from chance) for 2d stimuli. They now score close to 0%, indicating a strong and reliable reversed motion perception, in ac-

cordance with Sato (1989) (The exception is subject JR, who is noticeably worse at motion than stereopsis throughout and who demonstrates no improvement here).

One could argue that our stereo experiments are not directly comparable to the motion, because in the stereo case the two images were presented dichoptically for 130 ms, whereas in the motion case the two images were presented in succession for 65 ms each. This might be important in view of the results of Pope et al. (1999), who suggested that only the transient stereo system could perceive depth in low-density anti-correlated stereograms. Two subjects therefore repeated the 2d stereo experiments, this time with the stimulus displayed for 65 ms in each interval. However, this did not appear to affect the results. In almost every case, the data-point with the 65 ms stimuli fell within the 95% confidence interval of the corresponding 130 ms data-point (data not shown).

We conclude, then, that the difference which previous workers had reported between the responses to anti-correlated stereo- and kinematograms occurs only for stimuli containing a range of orientations. The 1d anti-correlated stimuli, even of broad bandwidth, cause similarly weak perceptions both of reversed motion and of reversed depth. The 2d anti-correlated stimuli cause a strong perception of reversed motion, and little or no perception of reversed depth.

3.3. Experiment 3: random-dot stimuli

Our results are now much closer to those of previous workers who used 2d stimuli. Some subjects still perform significantly below chance for anti-correlated stereograms, which has not previously been reported. However, previous studies (e.g. Cumming et al., 1998) looked for a threshold, and would probably not have picked up such a weak perceptual reversal. Our 2d stimuli still differ in two respects from those used in other studies. Previous studies have generally used random-dot stimuli, as pioneered by Julesz (1960), and disparity has usually been introduced by displacing a target region across a background. For a more direct comparison, we explored the effect of these.

First, we repeated our experiments using random-dot patterns, in which each of the 128×128 pixels was independently made either black or white with equal probability. These were displaced using wrap-around within the target region, as before. The results (Fig. 9) were similar to those obtained with 2d band-pass noise (Fig. 8; note the considerably greater range of displacements there). For correlated stimuli, all subjects performed close to 100%. This acts as a check on the subjects' ability to discriminate stereo depth. In particular, it confirms that the poor performance of subject MH on an earlier experiment (Fig. 6) was due to a specific difficulty with this stimulus, rather than poor stereo ability in general.

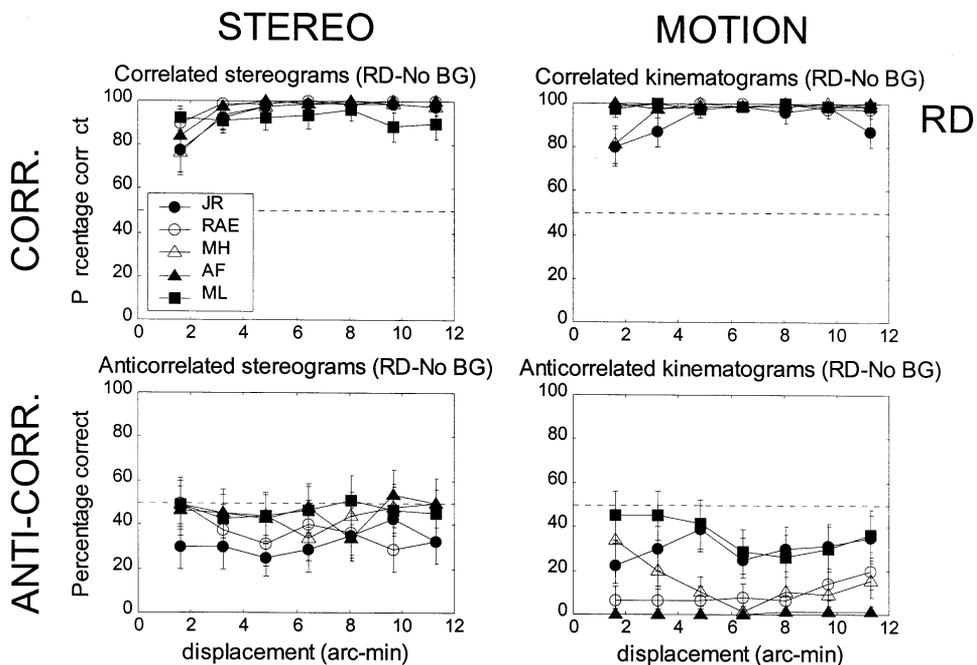


Fig. 9. Psychophysical functions for random-dot patterns (without a zero-disparity background). Data for four subjects is superimposed in each plot, as indicated in the key.

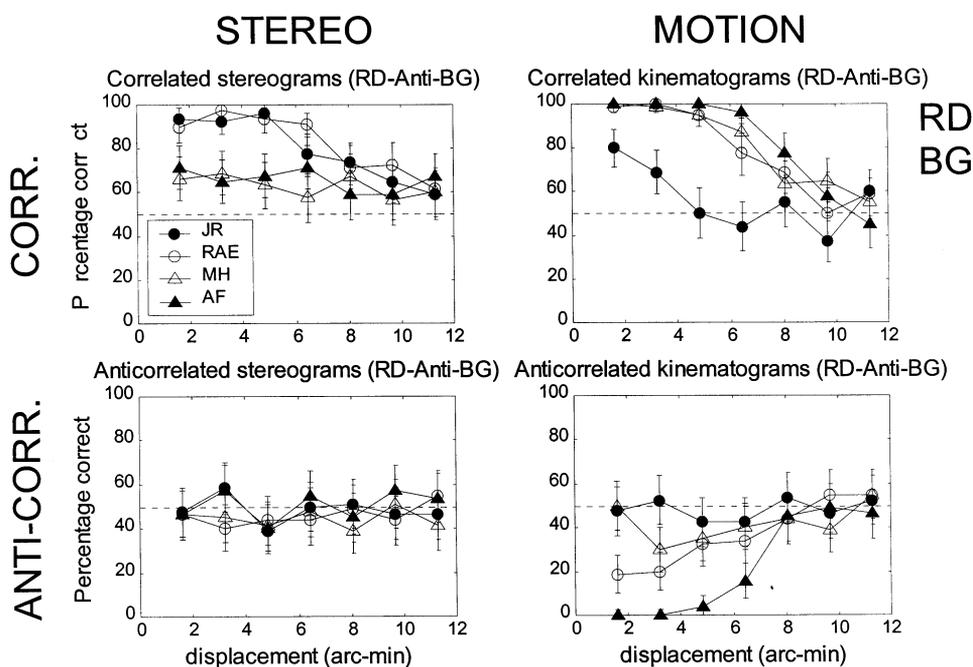


Fig. 10. Psychophysical functions for random-dot patterns on an anti-correlated zero-disparity background. Data for four subjects is superimposed in each plot, as indicated in the key.

Subjects JR and RAE continue to score significantly below chance with anti-correlated stereograms. With the other subjects, this effect is much weaker. As before, the discrimination of reversed motion with anti-correlated kinematograms is much more reliable than reversed depth. The exception is again subject JR, who scores around 30% for both anti-correlated stereograms and anti-correlated kinematograms. Subject ML also demonstrates only weak reversed motion. Despite these inter-subject differences, it is true for every subject that the perception of reversed motion is stronger than reversed depth. Random-dot patterns give results which are qualitatively similar to those obtained with 2d band-pass noise. Thus the inverse-frequency power spectrum of our band-pass images is not a critical factor in our results.

Previous workers (Cumming et al., 1998; Cogan et al., 1993; Julesz, 1971), in demonstrating that anti-correlated stereograms produced no sensation of depth, used an anti-correlated central region displaced over an anti-correlated background. We therefore ran a second set of random-dot experiments in which disparity was introduced by moving a target region across a zero-disparity background of $3.4 \times 3.4^\circ$ (as before, the target region was $1.7 \times 1.7^\circ$). We used an anti-correlated background in all experiments, even when the central, displaced region was correlated.

The results are plotted in Fig. 10. The addition of the anti-correlated background produces a striking deterioration in performance: all the psychophysical functions are considerably nearer chance. This deterioration is

particularly marked in the performance of subjects AF and MH with stereo. Even when the target region is correlated, they can barely perform the stereo task. On the motion task, the effect is less dramatic, but still present. Subject AF is the least affected, but even he is at chance for displacements larger than 10 arcmin, whereas in a previous experiment his perception of reversed motion persisted up to 30 or so arcmin (Fig. 8). With this stimulus, none of our subjects demonstrates any perception of reversed depth for anti-correlated stereograms. This agrees with previous studies (Julesz, 1971; Cogan et al., 1993; Cumming et al., 1998), in which disparity was introduced by shifting one region of the image across an anti-correlated zero-disparity background.

3.4. The wrap-around

We employed a slightly unorthodox means of introducing disparity. In the absence of a background, disparity was introduced by shifting the stimulus within a fixed region on the screen, using wrap-around as described in Section 2. This was done to avoid luminance discontinuities while preserving the Fourier spectrum of the band-pass stimuli. With a random-dot stimulus on a random-dot background, this artifice is no longer necessary. Disparity was introduced by simply shifting the stimulus across the screen. It is important to consider the possibility that this wrap-around of the stimulus is responsible for the reversed perception. When wrap-around is applied, the stimulus itself actually con-

tains reversed motion, since where we claim to have moved the image d pixels to the right, part of the image has jumped $(128-d)$ pixels to the left. However, we believe the wrap-around cannot explain the pattern of results seen here. In the random-dot experiments, we considered only small displacements. For instance, Fig. 9 shows that reversed motion was perceived at a displacement of 5 arcmin, or six pixels. So we have, for example, 122 pixels moving 5 min to the right, while six pixels jump 99 min to the left. It is hard to see how this could result in a perception of motion to the left — especially given the existing evidence indicating that, in the event of a conflict, perception of small disparities is more likely (McKee & Mitchison, 1988; Mallot & Bideau, 1990). Moreover, our experiments with correlated stimuli demonstrate that the perception in such a case is unambiguously of motion to the right (e.g. Fig. 9). Thus our correlated experiments act as a control on the anti-correlated stimuli, demonstrating that it is the anti-correlation which causes the perception of reversed motion, rather than the peculiarities of our wrap-around.

4. General discussion

4.1. Experiment 1: one-dimensional band-pass stimuli

For narrow-band stimuli, the results can be understood based only on the lowest frequency present. For small displacements, performance rises rapidly to nearly 100% for correlated stimuli, while with anti-correlated

stimuli, perception is reversed, as expected from considering a single sine-wave component. As the displacement approaches half a cycle, performance falls to chance, and subsequently starts to reverse. We can explain most features of the data with reference to the cross-correlation function (CCF) for this single frequency. Fig. 11 shows model CCFs for correlated (left-hand plots) and anti-correlated (right-hand) stimuli, in the case where the stimulus displacement is small compared to the channel period (upper plots), and where the stimulus displacement is just over half a period (lower plots). As before, we assume that the channel reports a displacement corresponding to one of the peaks in the CCF. Which peak is chosen represents a compromise between the largest peak and the peak closest to zero displacement (Prince & Eagle, 2000a).

First, we consider small displacements of correlated stimuli (upper left plot). Here, the largest peak is the closest to the origin. For very tiny displacements, we presume that noise makes it difficult to distinguish which side of the origin the peak is, explaining the existence of a D_{\min} , a lower bound on the displacements which can be reliably discriminated. For displacements slightly larger than D_{\min} , there is no ambiguity, resulting in a strong veridical perception.

For small displacements of anti-correlated stimuli, however, (upper right plot of Fig. 11), the side-peak closest to the origin is on the opposite side of the true displacement, resulting in reversed perception. Our data suggest that the reversed perception obtained with anti-correlated stimuli is not quite as strong as veridical perception with correlated stimuli. With 1 octave anti-

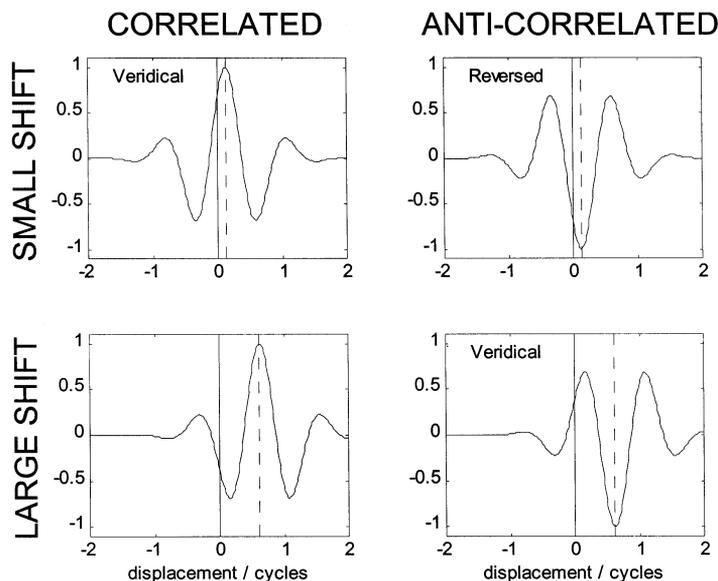


Fig. 11. A qualitative model for the response of a single channel. The figures represent the cross-correlation function of a pair of filtered images. In the left-hand plots, the images are correlated; on the right, anti-correlated. In the upper two plots, the stimulus displacement is 1/8th of the period of the channel; in the lower two plots, it is 5/8th of the period. Vertical lines mark on zero displacement and the correct stimulus displacement. The CCF plotted is a Gabor function with period 43 units, bandwidth 1 octave.

correlated stimuli, performance falls to only 5–10%, whereas with correlated stimuli it rises to 100% (Figs. 5 and 6). The weaker reversed depth with stereograms is especially noticeable for the very smallest displacement. Correlated performance then drops to around 90%, but the anti-correlated performance is much closer to chance, around 30%. Considering the CCF for small-disparity anti-correlated stimuli (Fig. 11, upper right plot), we can easily understand why this should be so. The channel is faced with two equally-large side-peaks, both far from the origin. Presumably, noise makes it difficult for the system to distinguish which side-peak is in fact marginally closer to the origin, pushing performance closer towards chance for these very small displacements. As displacement increases, performance moves towards 0%, perhaps because now one side-peak is significantly closer to the origin than the other. An additional factor may be that the side-peaks are smaller than the main peak for correlated stimuli, and thus may have a lower signal-to-noise ratio.

As the displacement is increased to half a cycle, the performance with correlated stimuli declines towards chance. With correlated kinematograms, there is some evidence of a perceptual reversal at displacements greater than half a cycle. However, with anti-correlated stimuli, this ‘reversal’ (back to veridical depth) occurs earlier. Performance is at chance at around 0.4 cycles, and at 0.5/0.6 cycles, we observe performance close to 100%, suggesting a veridical perception stronger than the reversed perception obtained with smaller displacements. Again, this can be simply understood in terms of the CCF. The lower two plots of Fig. 11 show CCFs for displacements just greater than half a cycle. For anti-correlated stimuli (lower right plot), it is easy to see why we obtain strong veridical perception at displacements just exceeding half a cycle. There are two equally-strong side-peaks, both on the veridical side of the origin. Equally, for correlated stimuli, it is easy to see why perception might be more ambiguous. Here, the two peaks closest to the origin are on opposite sides. The veridical peak at the correct disparity is larger, but it is also further from the origin. Presumably, in the stereo system, these conflicting cues are more or less equally opposed, since performance is at chance for displacements just over half a cycle (Figs. 5 and 6). In the motion system, performance here is slightly below chance for most subjects. This could be explained by a slightly stronger preference for smaller matches in motion than in stereo, resulting in a higher tendency to pick the side-peak on the wrong side of the origin.

At larger bandwidths, the curves in Fig. 5 depart from the narrow-band results. This immediately suggests that the brain is not performing the task merely on the basis of a single channel, such as that activated by the Fourier component with the longest period.

Rather, there are interactions between the different spatial frequency channels (Eagle, 1997). First we consider the results with correlated broad-band stimuli. At small displacements, d is less than half a period for most channels, so they unambiguously signal the correct displacement d (since the peak at d is both the largest and the closest to the origin, cf. Fig. 11 (upper left plot)). Thus performance is close to 100% at displacements which are small relative to the longest period. At the very smallest displacements, performance is slightly better for the broad-band stimuli than for narrow-band, suggesting that the activation of several channels, all in agreement, results in an improved signal-to-noise ratio and a lower D_{\min} . As the displacement increases beyond around 0.1 cycles, the performance declines relative to that for narrower-band stimuli at the same number of cycles of the longest period (Fig. 5, although note that the performance at a given displacement is better with broader-band stimuli, cf. Figs. 6 and 7). We can explain this decline in performance in terms of cross-channel conflict. As d increases, it exceeds half a cycle for more channels. For instance, in a 5 octave stimulus, when d is a quarter of the longest period present, d exceeds 0.5 cycles for 4 of the 5 octaves present. In this case, the largest peak in the CCF is no longer the closest to the origin (Fig. 11, left lower plot), and may be on the opposite side of the origin. Perhaps this introduces ambiguity in the channel’s response, which is the reason for the decline in performance for 5-octave stimuli once the displacement exceeds around 0.1 cycles of the longest period. The veridical signal from the lower frequency channels is partially masked by incoherent responses from channels tuned to higher frequencies (de Bruyn & Orban, 1989; Cleary & Braddick, 1990b).

For anti-correlated stimuli, even at small displacements we expect conflict between channels as to the absolute value of the displacement, though not its direction (Fig. 1). This discrepancy in absolute value may explain why the perception of depth/motion is especially weaker with anti-correlated stimuli than with correlated for broader-band stimuli.

4.2. Experiment 2: two-dimensional bandpass stimuli

For correlated stimuli, the additional orientations make little difference, apart from a slight increase in D_{\max} for kinematograms. Bischof and Di Lollo (1991) have argued that this effect occurs because oblique orientations contribute longer horizontal periods, enabling larger horizontal displacements to be distinguished. In terms of the 2d CCF (Fig. 3), this is because the distance along the horizontal axis between successive peaks of the CCF is $\lambda/(2 \cos \theta)$. That is, the effective horizontal period is larger for components oriented further from the vertical.

For anti-correlated stereograms, the move to 2d stimuli again makes little difference, apart from a slight decline in performance, readily explicable in terms of increased cross-channel conflict (cf. Fig. 2). However, for anti-correlated kinematograms, we find a startling *improvement* in the perception of reversed motion, indicated by performance much closer to 0%. This is precisely the opposite of what a simple argument based on cross-channel conflict would predict.

Our results may indicate a difference in the way the stereo and motion channels combine information from different orientation channels. Recall that for small displacements of anti-correlated stimuli, each Fourier component appears to move through a different amount ($\lambda/(2 \cos \theta) - d$), although the *sign* of the displacement agrees for all components with $d < \lambda/(2 \cos \theta)$. Thus if the motion system simply compares the *direction* reported by different orientation channels, it could end up with a higher signal-to-noise ratio for 2d stimuli (since more channels are active), explaining the clearer perception there. In contrast, if the stereo system relies predominantly on the results from channels tuned to vertical orientations, this would explain why performance is barely changed from 1d to 2d stimuli. The similar results for 1d stereo- and kinematograms suggest that the two systems have a similar method for combining information across different spatial scales, although they seem to differ in how they combine information at different orientations.

4.3. Experiment 3: random-dot stimuli

The results with random-dot stimuli agree with those with 2d band-pass noise. One striking feature is that the addition of an anti-correlated background severely impairs perception even when the target region is correlated, although it has been amply documented that subjects perform excellently for correlated targets on a correlated background (Julesz, 1971; Sato, 1989). This may be because each Fourier component of a zero-displacement anti-correlated background can be perceived equally as moving in either direction. Thus rather than perceiving a target moving across a stationary background, the brain is presented with a target moving across a background which gives an impression of incoherent motion with no clear direction. It is understandable that, under these circumstances, it becomes harder to judge the absolute direction of motion. In fact, subjects found it difficult even to distinguish the central region. Subjects reported that, most of the time, the whole $3.4 \times 3.4^\circ$ image seemed to move incoherently. Even when the central square was correlated, it was only occasionally visible — presumably for the narrow range of displacements in which performance was not at chance. This agrees with Sato's (1989) experiments with anti-correlated kinematograms. He found that subjects were much

poorer at judging the shape of a target region than its direction.

4.4. Alternative explanations in terms of feature-matching

A qualitative model based on Fourier channels with spatial-frequency and orientation tuning has been able to explain most features of our data. Purely feature-matching mechanisms cannot easily explain reversed perception with broad-band stimuli. Mechanisms sensitive to polarity are unlikely to work with anti-correlated images, whereas if they are insensitive to polarity, they produce veridical depth signals (Sato, 1998). Although non-Fourier mechanisms cannot explain the basic features of our results, they may contribute to some aspects of our data. For instance, a non-Fourier mechanism reporting veridical displacement (such as the amplitude-modulated (AM) component; Kovacs & Feher, 1997; Hess & Wilcox, 1994) may contribute to the weakness of reversed depth with anti-correlated stereograms and reversed motion with 1d kinematograms. If so, we must conclude that 2d kinematograms do not activate the non-Fourier sensors. It is unclear why this should be so.

5. Conclusions

We have compared motion and stereoscopic perception of anti-correlated stimuli using the same images and the same subjects. We find that the difference between stereopsis and motion is by no means so clear-cut as suggested by the existing literature. Rather, much of this difference was due to the stimuli used by previous investigators. With 1d vertical patterns, stereo and motion appear similar. Weak reversed motion and reversed depth is apparent in both cases. The main features of these data can be explained assuming the image is analysed by channels tuned to different spatial frequencies. Major differences between stereo and motion emerge when all orientations are present. The perception of reversed depth is then impaired, whereas the reversed motion in anti-correlated kinematograms is greatly enhanced. We suggest that a key difference between stereo and motion may be how information from different orientation channels is combined in solving the correspondence problem.

Acknowledgements

JCAR is supported by a Wellcome Training Fellowship in Mathematical Biology, RAE by a Royal Society Fellowship and the Wellcome Trust. We thank our subjects Maarten Hogervorst, Alexander Foulkes and Matthew Lipson for their patience. JCAR is profoundly grateful to Bruce Cumming and Simon Prince for their generous help and advice after the death of RAE.

References

- Anstis, S. M. (1970). Phi movement as a subtraction process. *Vision Research*, *10*, 1411–1429.
- Bischof, W. F., & Di Lollo, V. (1991). On the half-cycle displacement limit of sampled directional motion. *Vision Research*, *31*, 649–660.
- Blake, R., & Wilson, H. (1991). Neural models of stereoscopic vision. *Trends in Neuroscience*, *14*, 445–451.
- Blakemore, C., & Campbell, F. W. (1969). On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology*, *203*, 237–260.
- Campbell, F. W., & Kulikowski, J. J. (1966). Orientation selectivity of the human visual system. *Journal of Physiology*, *187*, 437–445.
- Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology*, *197*, 551–566.
- Chang, J. J., & Julesz, B. (1985). Cooperative and non-cooperative processes of apparent movement of random-dot cinematograms. *Spatial Vision*, *1*(1), 39–45.
- Cleary, R., & Braddick, O. (1990a). Direction discrimination for band-pass filtered random dot kinematograms. *Vision Research*, *30*, 303–316.
- Cleary, R., & Braddick, O. (1990b). Masking of low frequency information in short-range apparent motion. *Vision Research*, *30*, 317–327.
- Cogan, A. I., Lomakin, A. J., & Rossi, A. F. (1993). Depth in anti-correlated stereograms: effects of spatial density and interocular delay. *Vision Research*, *33*, 1959–1975.
- Cogan, A. I., Kontsevich, L. L., Lomakin, A. J., Halpern, D. L., & Blake, R. (1995). Binocular disparity processing with opposite-contrast stimuli. *Perception*, *24*, 33–47.
- Cumming, B. G., & Parker, A. J. (1997). Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature*, *389*, 280–283.
- Cumming, B. G., Shapiro, S. E., & Parker, A. J. (1998). Disparity detection in anti-correlated stereograms. *Perception*, *27*, 1367–1377.
- DeAngelis, G. C., Ohzawa, I., & Freeman, R. D. (1991). Depth is encoded in the visual cortex by a specialized receptive field structure. *Nature*, *352*, 156–159.
- de Bruyn, B., & Orban, G. A. (1989). Discrimination of opposite directions measured with stroboscopically illuminated random-dot patterns. *Journal of the Optical Society of America*, *A*, *6*, 323–328.
- de Valois, R. L., Albrecht, & Thorell (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, *545*–559.
- de Valois, R. L., & de Valois, K. K. (1988). *Spatial vision*. Oxford, UK: Oxford University Press.
- Eagle, R. (1997). Independent processing across spatial frequency in moving broadband patterns. *Perception*, *26*, 961–976.
- Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America*, *A*, *4*, 2379–2393.
- Freeman, R. D., & Ohzawa, I. (1990). On the neurophysiological organization of binocular vision. *Vision Research*, *30*, 1661–1676.
- Glennister, A. (1998). d_{max} for stereopsis and motion in random dot displays. *Vision Research*, *38*, 925–935.
- von Helmholtz, H. (1909). *Handbuch der physiologischen optik*. Hamburg, Germany: Voss.
- Hess, R. F., & Wilcox, L. M. (1994). Linear and non-linear filtering in stereopsis. *Vision Research*, *34*, 2431–2438.
- Julesz, B. (1960). Binocular depth perception of computer generated patterns. *Bell Systems Technical Journal*, *39*, 1125–1162.
- Julesz, B. (1971). *Foundations of cyclopean perception*. Chicago: University of Chicago Press.
- Kovacs, I., & Feher, A. (1997). Non-Fourier information in bandpass noise patterns. *Vision Research*, *37*, 1167–1175.
- Mallot, H. A., & Bideau, H. (1990). Binocular convergence influences the assignment of stereo correspondences. *Vision Research*, *30*, 1521–1523.
- Mansfield, J. S., & Parker, A. J. (1993). An orientation-tuned component in the contrast masking of stereopsis. *Vision Research*, *33*, 1535–1544.
- Marr, D., & Poggio, T. (1976). Cooperative computation of stereo disparity. *Science*, *194*, 283–287.
- McKee, S. P., & Mitchison, G. J. (1988). The role of retinal correspondence in stereoscopic matching. *Vision Research*, *28*, 1001–1012.
- Pelli, D. G., & Zhang, L. (1991). Accurate control of contrast on microcomputer displays. *Vision Research*, *31*, 1337–1350.
- Poggio, G. F., & Poggio, T. (1984). The analysis of stereopsis. *Annual Review of Neuroscience*, *7*, 379–412.
- Pope, D. R., Edwards, M., Schor, C. S. (1999). Extraction of depth from opposite-contrast stimuli: transient system can, sustained system can't. *Vision Research*, *39*, 4010–4017.
- Prince, S. J. D., Eagle, R. A., & Rogers, B. J. (1998). Contrast masking reveals spatial-frequency channels in stereopsis. *Perception*, *27*, 1345–1355.
- Prince, S. J. D., & Eagle, R. A. (2000) Weighted directional energy model of human stereo correspondence. *Vision Research* (in press).
- Prince, S. J. D., & Eagle, R. A. (2000b). Stereo correspondence in 1d Gabor stimuli. *Vision Research*, *40*, 913–924.
- Sato, T. (1998). D_{max} : relations to low and high-level motion processes. In: High-level motion processing, ed. T. Watanabe, Boston, MIT press, pp. 115–151.
- Sato, T. (1989). Reversed apparent motion with random dot patterns. *Vision Research*, *29*, 1749–1758.
- Smallman, H. S., & MacLeod, D. I. A. (1994). Size-disparity correlation in stereopsis at contrast threshold. *Journal of the Optical Society of America* (A), *11*, 2169–2183.
- Yang, Y., & Blake, R. (1991). Spatial frequency tuning of human stereopsis. *Vision Research*, *31*, 1177–1189.