Mechanisms for dynamic stereomotion respond selectively to horizontal velocity components.

MICHAEL J. MORGAN[1] AND CHRISTOPHER W. TYLER[2]

[1]Department of Visual Science, Institute of Ophthalmology, Bath Street, London EC1V 9EL, U.K.
[2]Smith-Kettlewell Eye Research Institute, 2232 Webster St, San Francisco, California 94115, U.S.A.

SUMMARY

When dynamic visual noise such as the 'snow' on a detuned television receiver is inspected with a delay between the signals in the two eyes, the noise appears to rotate in depth around a vertical axis. We propose that this dynamic noise stereophenomenon arises because channels tuned jointly to depth and motion respond selectively to horizontal motion components in the noise. We used spatially filtered stimuli to reduce the strength of vertically oriented Fourier components in the noise, and found that this reduced the strength of the stereophenomenon. Reducing the strength of horizontally oriented components had a lesser effect. Our evidence is consistent with the conclusion that stereo-motion is predominantly detected by oriented receptive fields with an aspect ratio similar to those of cortical simple cells.

1. INTRODUCTION

In the 'Pulfrich effect' a swinging pendulum appears to move in an elliptical path in depth when viewed with a sunglass lens over one eye. The explanation (see figure 1) is thought to be that the reduction in illumination causes a delay in the response of the covered eye and thus introduces a velocity-dependent stereoscopic disparity between the covered and uncovered eye (Pulfrich 1922; Morgan & Thompson 1975). However, this 'delay line' explanation (originally proposed by Fertsch, an optical engineer at Zeiss Jena) does not obviously apply to an effect seen with dynamic visual noise such as the 'snowstorm' on a detuned television. If dynamic noise is inspected binocularly with a neutral density filter over one eye (Tyler 1974) or with a pure interocular delay (Ross 1974), the noise segregates into a variety of depth planes moving in opposite horizontal directions in front of and behind the fixation point (see figure 1).

Why is depth seen, and why should the motion be horizontal? Although the effect is in the same direction as the Pulfrich phenomenon, the problem with the delay-line explanation is that an interocular time delay in viewing an uncorrelated noise stream should result in binocular rivalry, not in stereopsis. As there is no correlation between the noise at time t and time $t-i$ a delayed image of noise will by definition have no correlation with an undelayed image.

Here we suggest a resolution of the paradox. Random spatio-temporal noise contains spatio-temporal energy in all directions: it is the equivalent in the motion domain of a random spatial texture, containing all orientations and spatial frequencies. We propose that, because the eyes are horizontally separated, it is the horizontal disparity between receptive fields in the two eyes that is the controlling variable for the perception of depth, regardless of the receptive field orientation. Receptive fields jointly tuned to motion and horizontal disparity are stimulated by the corresponding noise energy in the stimulus in the same way as they are by a horizontally moving target at their preferred orientation, and thus produce a Pulfrich effect. The depth impression for the motion components will fall with their angular deviation from the vertical according to the postulate that only the horizontal disparity contributes to depth perception. In particular, horizontally oriented components should give no depth impression at all.

The question we address is the nature of the neural substrate for stereomotion. Both physiological (Hubel & Wiesel 1959) and psychophysical (Nakayama et al. 1985) data suggest that a high proportion of primate receptive fields are direction-selective and are elongated in the direction orthogonal to their preferred direction of motion. As the perceived direction of motion in the DVN stereophenomenon is in the horizontal direction, such results suggest that the mediating receptive fields should be elongated vertically. If such receptive fields are of the simple, linear type, their optimal stimuli should be similarly elongated. The strength of the perceived stereomotion would be expected to increase with stimulus length up to the length of one receptive field and then show no further improvement.

We therefore devised stimuli that would maximize the selectivity for length in this neural substrate. The stimuli are elongated strips of alternating light and dark patches on a grey background, resembling tigertails (see figure 2).

When the patches are small, a linear elongated receptive field would sum the contrast energy to zero along their length, receiving no net stimulation. Increasing the elongation of the patches within each strip would progressively increase the stimulation for elongated receptive fields without changing the overall
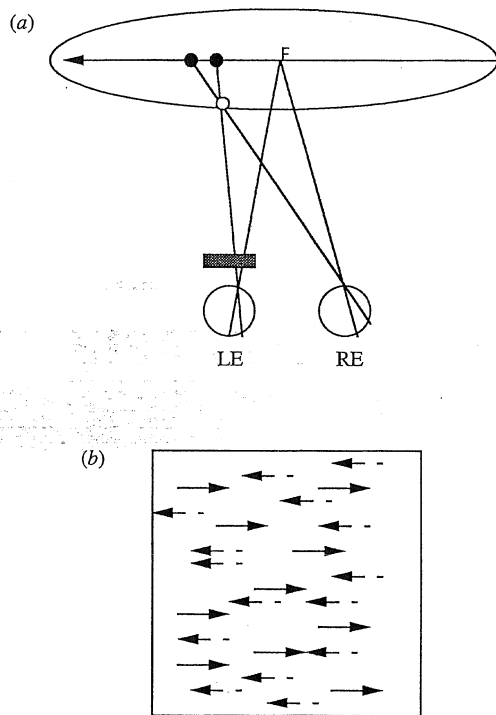
Proc. R. Soc. Lond. B (1995) 262, 371–376
Printed in Great Britain

371

(a)

LE   RE

(b)

Figure 1. Pulfrich's pendulum effect is seen when the observer views a horizontally oscillating target with a neutral density filter over one eye. *(a)* The schematic plan view from above the observer illustrates Fertsch's 'delay line' explanation. The filtered eye signals the instantaneous position of the moving target with a time delay, because of the inverse relation between luminance and the latency of the visual response, and thus signals the target position at an earlier point on the motion trajectory. This gives rise to an instantaneous disparity between the two eyes' views, which causes the target to be mislocated in depth. *(b)* Appearance of dynamic noise with LE filtered. A similar effect is seen when the observer views dynamic visual noise such as the snowstorm on a detuned television receiver. A filter over one eye causes the random motion of noise to be replaced by horizontal streaming, with opposite directions of streaming in front of and behind the fixation point (x), depending on which eye is filtered.

⟶ denotes seen behind; ◀- - denotes seen in front.

contrast energy. Such stimuli could therefore be used as a probe for the operation of linear, elongated receptive fields in mediating a perceptual task such as the stereomotion task of our study.

On the other hand, stereomotion could be mediated by local circular receptive fields, which are present in large numbers in monkey cortex (Baizer *et al.* 1977). Such isotropic receptive fields typically have poor motion selectivity, but are characterized as being uniformly binocular. They could, therefore, provide a substrate for disparity selectivity, from which the motion information was subsequently extracted from the spatiotemporal disparity signals. According to this hypothesis, the orientation distribution of the energy should have little impact on the strength of the stereomotion percept. In fact the tigertails with isotropic patches might produce the strongest stimu-

lation because they are flanked on either side by patches of opposite sign, which would provide preferential stimulation of an inhibitory surround if one were present.

We therefore used such tigertail stimuli to investigate whether the spatiotemporal filter properties for lateral stereomotion are isotropic in spatial selectivity or whether they show some preference for the orientation of the elements in the stimulus.

2. METHODS

(a) Apparatus

The stimuli were generated on the 14″ colour display of a Macintosh II computer, using an 8-bit scale of grey levels. Observations were carried out in a darkened room, with the only illumination supplied by the screen, which had a constant mean luminance of 50 cd m^{-2}.

(b) Subjects

The subjects were M.J.M. and C.W.T, both experienced psychophysical observers. The basic finding that there was no perceived depth effect with vertically balanced stimuli was confirmed by informal tests with several naive observers.

(c) Stimuli

To filter the dynamic noise we presented a temporal sequence of frames each of which was spatially filtered by composing it of a set of elongated patches. Each patch consisted of a 4.0 cycles per degree of horizontally oriented sinewave grating (Michelson contrast 0.14) multiplied by a vertically oriented elliptical Gaussian envelope with a width of approximately 0.33 deg (± 2 s.d) and a height:width aspect ratio of 16:1 (see figure 2). The grey background had the same mean luminance (50 cd m^{-2}) as the grating. The stimuli were composed of closely spaced dots, which were not resolved at the viewing distance used in the experiment. Patches could overlap spatially, as shown in figure 2. The whole image in every frame was therefore band-limited in frequency and orientation in exactly the same way as its component patches. Note that stimulus differed from the normal 'Gabor patch' used in psychophysical and physiological experiments by having its grating component oriented at 90° to the envelope. The resulting pattern of stripes resembles a 'tigertail' and we refer to it as such for simplicity. In the control stimuli all the black or all the white half-cycles in each blob were set to the background luminance, so that the patches were now no longer luminance-balanced in the vertical direction

Each frame consisted of a different set of randomly positioned vertically oriented tigertails. The average luminance along a single oriented tigertail is the same as the background, and it is therefore invisible to a linear filter oriented in the same direction as the patch, provided that at least one cycle of the grating falls within the filter's receptive field. Thus a stimulus consisting of vertically oriented tigertails should be invisible to vertically oriented linear filters when the sinewave component is of a sufficiently high frequency, but still would stimulate circularly symmetric filters tuned to a half cycle of the grating. The patch would, however, be visible to a variety of second-order filters, for example those collecting input only from black bars of the grating or only from white (half-wave rectifiers) or those adding black and white bars after squaring (Chubb & Sperling 1989; Cavanagh & Mather 1990). Any loss in
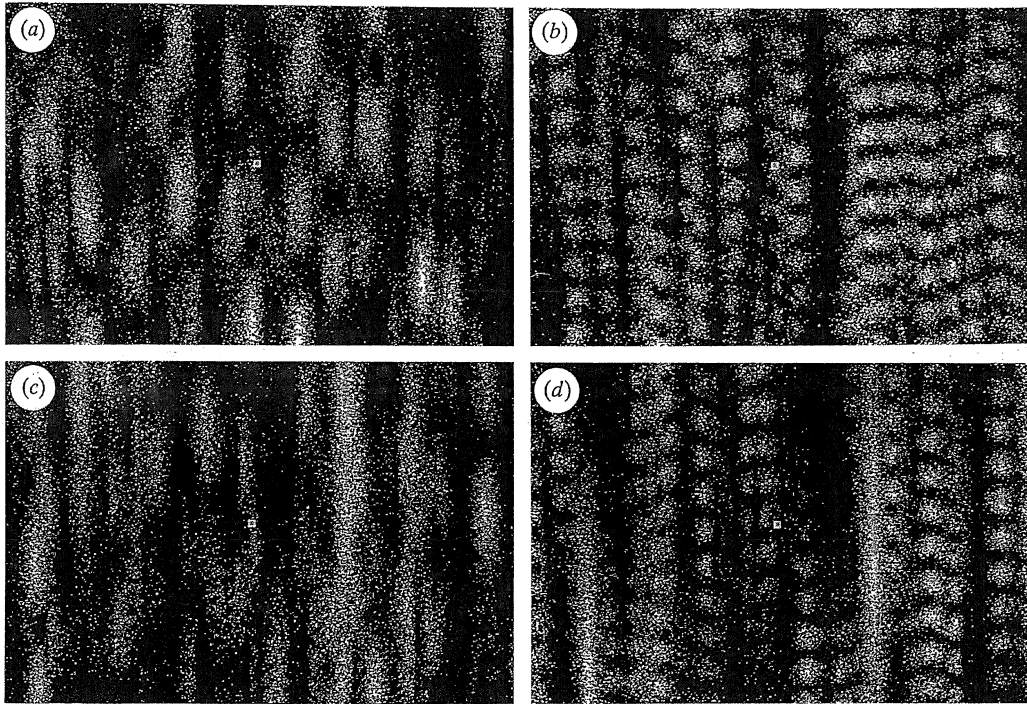
Figure 2. The figure illustrates single frames of the stimuli used in the experiment to determine the role of vertically oriented Fourier components in the dynamic noise stereophenomenon. In the examples shown in (*a*) and (*b*) each frame was composed of a set of randomly scattered tigertail patches, each patch consisting of a sinewave grating (Michelson contrast 0.14) multiplied by an elliptical (16:1) Gaussian envelope. The grey background had the same mean luminance (50 cd m^{-2}) as the grating. The stimuli were composed of closely spaced dots, which were not resolved at the viewing distance used in the experiment. Patches could overlap spatially, accounting for the variability seen in the examples. In (*a*) the grating had a low frequency relative to the patch height (2 cycles per envelope); in (*b*) the frequency is 8 cycles per envelope. We found that high frequency gratings could not support stereomotion, but if exactly the same patches were oriented horizontally, or if the average luminance of the patches differed from the gray background the stereomotion effect returned. (*c*) and (*d*) illustrate control stimuli, in which all the black or all the white half-cycles in each blob were set to the background luminance. The patches are now no longer luminance-balanced in the vertical direction and the stereophenomenon returns.

stereomotion sensitivity as the modulating spatial frequency is increased would imply that neither circularly symmetric nor second-order filters are involved in the effect.

(d) Psychophysics

The strength of the stereo-depth effect with a 1.0 log unit neutral density filter was measured by determining the duration of the stimulus required to see clear clockwise or counterclockwise motion (the integration time: see figure 3).

The threshold duration was determined by a staircase procedure in which the duration was decreased every time the subject pressed the button indicating that they had seen the rotation in depth, and increased if the 'no motion in depth' button was pressed.

3. RESULTS

The results for vertically oriented tigertails showed that the strength of the effect fell off rapidly as the vertical frequency of the modulating grating was increased, although there was little effect of grating frequency on the unbalanced, control stimulus (see figure 3*a*,*b*). At the highest spatial frequency (16 cycles per envelope), the stereomotion effect could not be

seen at any duration, although at 8 cycles per envelope it was only barely visible at the cutoff duration of 2 s. One observer (M.J.M.) carried out a second experiment (see figure 3*c*) which showed that doubling the viewing distance made little difference to the frequency limit for the effect: therefore, it is the frequency of the stimulus in relation to its width that is important, not the frequency *per se*.

The results imply that the neural mechanisms for stereomotion must be anisotropic, because isotropic receptive fields would be equally stimulated at all spatial frequencies. To estimate the aspect ratio of the filter responsible for the decrease in stereomotion with grating frequency, we calculated the ratio between the balanced and control integration times and fitted the data to linear equations in log–log coordinates (see figure 3*d*). The fitted lines were used to derive the sinewave frequency at which performance had fallen to 50 % of its maximum value. Values of approximately 2.0 cycles were found for this corner frequency, corresponding to an aspect ratio of about 4.0 (i.e. 16/4) for each half cycle of the modulation, a figure not dissimilar from the aspect ratio of simple cells in primary visual cortex.
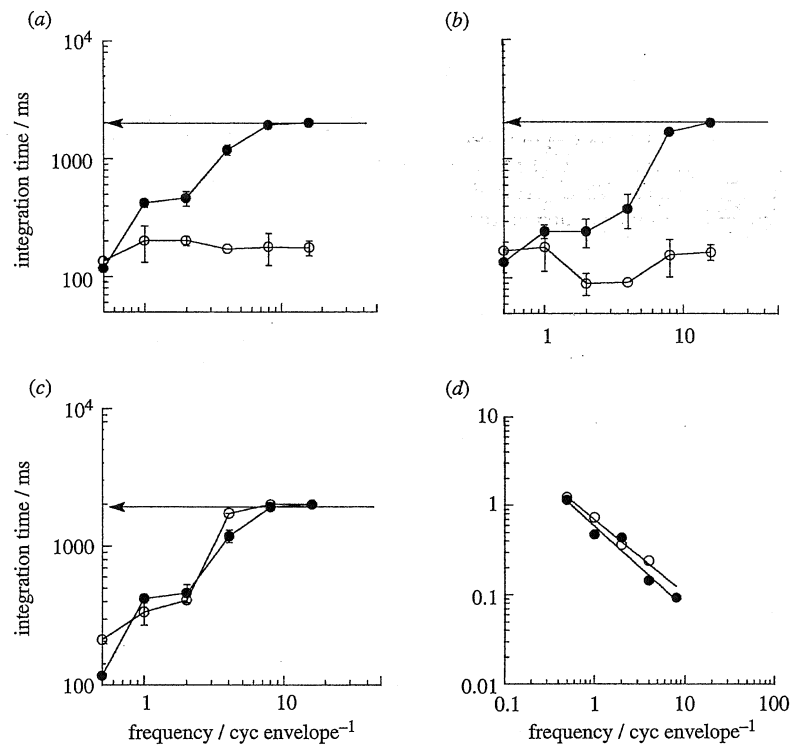
Figure 3. (a) and (b) show the results of an experiment using the stimuli described in the legend to figure 2. The time required to see the stereophenomenon from the time the stimulus appeared (integration time: vertical axis) was determined over a series of trials in which the observer pressed one button to reduce the exposure if stereomotion was seen, and another button to increase the exposure if stereomotion was not seen. When the observer was satisfied that the duration was just at the boundary for seeing stereomotion, a third button was pressed, and the duration recorded. If the effect was not seen with a duration of 2 s, a notional integration time of 2 s was recorded. Four independent observations were made, two with the left eye filtered and two with the right eye; the data presented are the means and standard deviations of these four readings. The data show that the integration time for the effect increases with the grating spatial frequency in the balanced (see figure 2 a,b filled circles) but not in the control (see figure 2 c,d open circles) condition. (c) Doubling the viewing distance makes little difference to the effects of frequency. Filled circles denote 1 × dist, open circles denote 2 × dist. (d) The ratio balanced/control integration time was calculated and the ratios fit by the equation $\log(y) = a$. Log $(x) + c$. The equation was used to calculate the frequency at which integration time was doubled, and the aspect ratio of one half-cycle of the grating at this 50 % performance point was thus estimated at about 4:1, similar to the aspect ratio of cortical simple cells. Filled circles denote M.J.M. ration, open circles denote C.W.T. ratio.

The crucial variable in weakening stereomotion appears to be the luminance balance of the patches in the vertical direction. To verify this we altered the balance by increasing the luminance of the white bands while keeping the black constant. The result was an increase in the strength of stereomotion with the degree of imbalance (see figure 4a). Further, we predicted that horizontally oriented tigertails would support a larger effect than vertically oriented stimuli since horizontal patches are not balanced in the vertical direction; and this was found to be so (see figure 4b). The last finding also shows that reduction in the motion effect in tigertails cannot be due completely to a decrease in their detectability. Indeed, it cannot be due to any loss in their detectability because the maximal spatial frequency was 4.0 cycles per degree, close to the peak of the contrast sensitivity function for human vision.

4. DISCUSSION

Our findings support the view that the dynamic noise stereophenomenon depends upon horizontal motion signals embedded in the noise (Tyler 1974; Morgan & Ward 1980). Detection of these components could be carried out by linear, oriented mechanisms tuned to vertically oriented Fourier components. These mechanisms are additionally tuned to specific disparities between the two eyes. We could detect no second-order input into stereomotion, even though horizontal movement was clearly visible in the vertically balanced, high-frequency stimuli, and informal observation suggested that such stimuli could support horizontal smooth pursuit, just like broad-band dynamic noise (Ward & Morgan 1978).

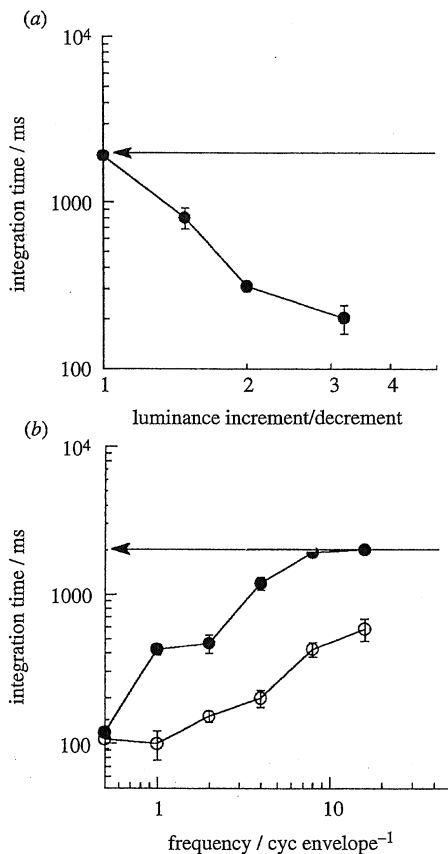Neurones tuned to both direction of movement and disparity have been described in the monkey prestriate

Figure 4. (a) M.J.M. 16:1, shows the effects of changing the luminance balance of the gratings illustrated in figure 2 (a) (b). With a balanced grating (luminance increment above background = luminance decrement below background) no stereomotion was seen at the longest (2 s) duration. As the size of the increment was increased, the effect increased, as shown by the reduced stimulus duration required to see it. (b) M.J.M. 16:1 replots the data for the balanced condition in figure 3, along with the results for exactly the same stimuli, horizontally rather than vertically oriented. The stereomotion effect is seen at much shorter durations with the horizontally oriented stimuli, which are not balanced in the vertical direction, and which therefore contain vertically oriented Fourier components.

area MT (Maunsell & Van Essen 1983 a,b). These units respond better to wide-field moving random dot patterns than to single bars, but many show orientational selectivity to single bars usually, but not always, orthogonal to their preferred motion direction (Maunsell & Van Essen 1983 a,b). More recently, tuning to motion and disparity has been found in area MST (Roy et al. 1992), a leading candidate for the cortical area involved in the computation of self movement. Most of these disparity selective units were 'near or far' selective relative to the fixation point, rather than to disparities near to the fixation plane; of particular interest is the fact that 40 % of the units changed their direction preference when the stimulus

changed from near to far or *vice versa*. It is not clear whether they were orientationally specific. Our results suggest that the mechanisms involved in the joint computation of movement and disparity are strongly orientationally tuned, and that our balanced dynamic stimuli may be useful in probing the properties of units in areas MT and MST. The function of mechanisms jointly tuned to disparity and motion direction may be to provide the basis for computing the motion parallax field resulting from observer self-movement.

Our findings do not necessarily imply an absence of second-order motion to stereopsis. The dynamic noise stereophenomenon is not particularly strong even with unfiltered noise. Naive observers sometimes take several minutes to see it, even though once they do see it they find it compelling. The presence of motion energy at all orientations and velocities may mask the small subset of signals that can support horizontal motion in depth. In these circumstances, any second order signal may be too weak to support stereomotion at all.

In unmasked conditions there is evidence that the second-order signal may be adequate for stereopsis. Halpern (1991) reported stereopsis with contours defined by motion alone i.e. with no coherent matching of individual dots between frames. There remains the possibility that bias in the random disparity matches at the motion contour could be sufficient to explain Halpern's result without second-order mechanisms. Wright & Gurney (1992) reported that stereoscopic plaids composed of two sinusoidal components supported disparate motion when their components were near to vertical, provided that the plaid features moved horizontally. However, luminance nonlinearities could have generated vertically oriented sets of features similar to the Gabor patches of this study. Our claim that stereomotion is mediated by vertically oriented receptive fields may need to be amplified to include the involvement of second-order mechanisms, under conditions that are favourable to the operation of such mechanisms.

If the substrate for horizontal stereomotion were solely vertically oriented receptive fields with horizontal direction selectivity, we should expect their sensitivity to fall off according to the orientation bandwidth of their receptive fields as the stimuli are rotated towards horizontal. For sinusoidal grating stimuli, Morgan & Castet (1995) have reported a stereomotion percept even for gratings oriented at 45° from vertical. Because the orientation bandwidth for typical monkey receptive fields is such that they have a negligible response at 45° from their optimum orientation, the implication is that oblique receptive fields can also support stereomotion. These data therefore imply that, although the receptive fields underlying stereomotion are elongated and linear, they may have a range of orientations, as postulated in our original hypothesis. What is specifically excluded by these results is the role of isotropic receptive fields, which would be equally stimulated in all the conditions tested in our experiments (or more strongly stimulated at high spatial frequencies if there is surround inhibition).

REFERENCES

Baizer, J. S., Robinson, D. I. & Dow, B. M. 1977 Visual responses of area 18 in awake behaving monkey. *J. Neurophysiol.* **40**, 1024–1037.

Cavanagh, P. & Mather, G. 1990 Motion: The long and short of it. *Spatial Vis.* **4**, 103–129.

Chubb, C. & Sperling, G. 1988 Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception. *J. opt. Soc. Am.* **5**, 1986–2006.

Chubb, C. & Sperling, G. 1989 Two motion perception mechanisms revealed through distance-driven reversal of apparent motion. *Proc. natn Acad. Sci. U.S.A.* **86**, 2985–2989.

Halpern, D. L. 1991 Stereopsis from motion-defined contours. *Vis. Res.* **31**, 1611–1617.

Hubel, D. H. & Wiesel, T. N. 1959 Receptive fields of single neurons in the cat's striate cortex. *J. Physiol.* **148**, 574–591.

Maunsell, J. H. R. & Van Essen, D. C. 1983a Functional properties of neurons in the middle temporal visual area of the macaque monkey II. Binocular interactions and sensitivity to binocular disparity. *J. Neurophysiol.* **49**, 1148–1167.

Maunsell, J. H. R. & Van Essen, D. C. 1983b Functional properties of neurons in the middle temporal visual area of the macaque monkey. I. selectivity for stimulus direction, speed, orientation. *J. Neurophysiol.* **49**, 1127–1147.

Morgan, M. J. & Thompson, P. 1975 Apparent movement and the Pulfrich effect. *Perception* **4**, 3–18.

Morgan, M. J. & Ward, R. 1980 Interocular delay produces depth in subjectively moving noise patterns. *Q. Jl exp. Psychol.* **32**, 387–395.

Morgan, M. J. & Castet, E. 1995 Stereoacuity for oblique gratings predicted by phase shifts, not disparities. *Invest. Ophthal. vis. Sci.* **36**, 231. (Suppl.)

Nakayama, K., Silverman, G. H., Macleod, D. I. A. & Mulligan, J. 1985 Sensitivity to shearing and compressive motion in random dots. *Perception* **14**, 225–238.

Pulfrich, C. 1922 Die Stereoscopie im Dienste der isochromen und heterochromen Photometrie. *Naturwissenschaften* **10**, 553–564.

Ross, J. 1974 Stereopsis by binocular delay. *Nature, Lond.* **248**, 363–364.

Roy, J. -P., Komatsu, H. & Wurtz, R. H. 1992 Disparity sensitivity of neurones in monkey extrastriate area MST. *J. Neurosci.* **12**, 2478–2492.

Tyler, C. W. 1974 Stereopsis in dynamic visual noise. *Nature, Lond.* **250**, 363–364.

Ward, R. & Morgan, M. J. 1978 Perceptual effect of pursuit eye movements in the absence of a target. *Nature, Lond.* **274**, 158–159.

Wright, M. J. & Gurney, K. N. 1992 Dependence of stereomotion on the orientation of spatial frequency components. *Ophthal. Physiol. Opt.* **12**, 264–268.