Vision Res. Vol. 27, No. 1, pp. 61–75, 1987
Printed in Great Britain. All rights reserved

0042-6989/87 $3.00 + 0.00

ERRORS IN DIRECTION-OF-MOTION DISCRIMINATION WITH COMPLEX STIMULI

ANDREW M. DERRINGTON* and G. BRUCE HENNING†
Department of Psychology, University of Durham, Durham DH1 3LE, England

(*Received* 31 *October* 1985; *in revised form* 29 *May* 1986)

Abstract—The direction of apparent motion in a complex pattern comprising a static 1-cycle/degree (c/deg) grating and a moving 3-c/deg grating changes with stimulus duration. At durations longer than about 150 msec, motion is seen almost veridically; the motion of the 3-c/deg grating, which is seen correctly, merely induces in the 1-c/deg grating a weak apparent motion in the opposite direction. At shorter durations, however, the only motion seen is in the opposite direction from that which, in fact, occurs. The reversed apparent motion is both compelling and consistent; it is reported both by naive and by experienced observers, and, although it only occurs for certain ranges of spatial frequency, contrast and duration, the ranges are substantial. The reversal appears to be almost independent of the temporal frequency and the spatial phase of the stimulus; it occurs both for discrete and for continuous motion. It seems likely that the apparent motion with short duration stimuli reveals properties of local visual movement detection previously unknown and difficult to account for within the framework of current models of motion perception.

INTRODUCTION

Visual detection of motion is thought to involve two different types of mechanism (Anstis, 1980; Braddick, 1980). The first, with which we are principally concerned, is thought to operate at a relatively low level in the visual system and to perform relatively local operations from which the direction of motion of small regions of the image is determined. The output of this mechanism may be used to identify features of the stimulus defined by their common motion (Braddick, 1974). The second type of mechanism operates at a higher level, and encodes the direction of motion of features of the image. Such features, often corresponding to visual objects, may themselves require substantial processing before they can be extracted. Examples of features thought to be dealt with exclusively by the high-level system are those defined by variations in texture (Ramachandran *et al.*, 1973), and local variations in contrast caused by beats between sinusoidal gratings of slightly different spatial frequency (Derrington and Badcock, 1985).

Recent attempts to model the low-level system (Van Santen and Sperling, 1984; Watson and Ahumada, 1985; Adelson and Bergen, 1985a) all use variants of the "sequence-detection" or correlational model introduced by Reichardt (1961). The basic principle, illustrated in Fig. 1, is that each motion detecting sub-unit receives inputs from two elements with spatial receptive fields that differ in the relation between their response and the spatial phase (or position) of the stimulus. By delaying the output of one element (or by introducing a temporal phase shift in the response) and then comparing the output of the two elements (most simply by multiplying them) differential responses to motion in different directions can be produced. Different models differ in the characteristics of the receptive field of the initial elements, in the delay mechanism, and in the procedure by which responses are combined, but their principle of operation is the same; indeed, Van Santen and Sperling (1985) have argued that relatively minor changes in the three models make them equivalent.

In this paper, under stimulus conditions thought to expose the operation of the low-level system, we show that movement of one component in a one-dimensional display containing two spatial components with different spatial frequencies is systematically mis-perceived. The pattern of errors we observe is not easily pre-

*Present address: Department of Physiological Sciences, The Medical School, Framlington Place, Newcastle upon Tyne NE2 4HH, England.
†Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, England. Present address: RCA/David Sarnoff Research Center, Princeton, NJ 08540, U.S.A.

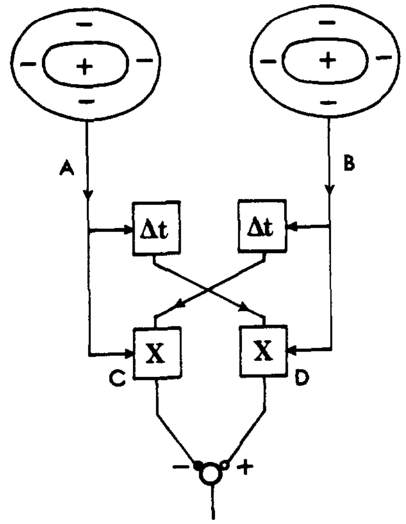

Fig. 1. Schematic diagram of a version of a Reichardt motion-detector. The input channels (A,B) have identical on-centre receptive fields in slightly different positions (in fact one would expect them to overlap somewhat). The multipliers for each sub-unit (C,D) multiply signals from the two elements after delaying one of them by Δt. Thus (D) gives a positive response for rightward motion and (C) gives a positive response for leftward motion. The output of the detector is the difference between the signals from (C) and (D), and is thus positive for rightward motion and negative for leftward motion.

dicted by any variant of Reichardt's motion detector that we have been able to find. We then consider a number of possible "design flaws" in a motion detector of Reichardt's type which might occur for our results and show that none of them produces the same pattern of errors that our observers show. Finally we consider an alternative to Reichardt-type detectors which exhibits some of the behaviour we find in our Observers.

METHODS

Stimuli

Stimuli were luminance patterns displayed by a computer (PDP 11/73), using four 12-bit digital-to-analogue-converters, two multipliers, an adder, and a logarithmic attenuator to produce voltage waveforms for display [using the method of Schade (1953)] on a television display with a linear z axis (Joyce Electronics, Cambridge, England). Each pattern was the sum of two sinusoidal gratings so that luminance as a function of position, $L(x)$, was described by the equation

$$L(x) = L_m\{1 + m_1 \sin[2\pi(f_1 x + z_1)] + m_2 \sin[2\pi(f_2 x + z_2)]\}, \quad (1)$$

where L_m, the mean luminance of the display, was 160 cd/m^2 and constant throughout. The spatial frequencies of the two components of the complex grating, f_1 and f_2, are expressed in c/deg of visual angle (c/deg). The Michelson contrasts of the two gratings, m_1 and m_2, were always Gaussian functions of time determined by the parameter σ and truncated by the 1-sec observation interval.

The spatial phases of the two components, $2\pi z_1$ and $2\pi z_2$, were expressed in radians, relative to the left-hand edge of the display. Motion was introduced into the display *either* by making z_1 or z_2 an appropriate linear function of time (incremented at the start of each display frame, i.e. 125 times/sec), *or* by changing z_1 or z_2 once between a pair of frames when the stimulus contrast was maximal, i.e. at the mid-point of the observation interval. We also measured Vernier acuity in a subsidiary experiment. Stimuli for measuring Vernier acuity were generated by displaying two patterns of the sort described by equation (1) in the top and bottom halves of the display. (This was accomplished by switching between the two waveforms using an analogue switch synchronised to the vertical displacement signal.) For Vernier acuity judgments, the patterns in the top and bottom halves of the display differed *only* in the value of z_1 or z_2 both of which remained constant throughout each observation interval. Plate 1 illustrates one such stimulus.

The entire visual display subtended 7.5 (horizontal) by 6.25 deg at the viewing distance of 2.4 m, and had a dark surround; we shall neglect spatial frequency components introduced by this truncation. High-frequency components introduced by the sampling of the waveform (68 samples/deg of visual angle) were attenuated by filters in the adder.

Observers

The phenomenon under study came to our notice when we were acting as our own Observers in an experiment attempting to study the spatial-frequency selectivity of the receptive fields of the putative elements in motion-detecting units. For this reason the Observers in our first experiment, which simply demonstrates the phenomenon, were naive colleagues and research students. They had little or no experience as Observers but readily saw motion and

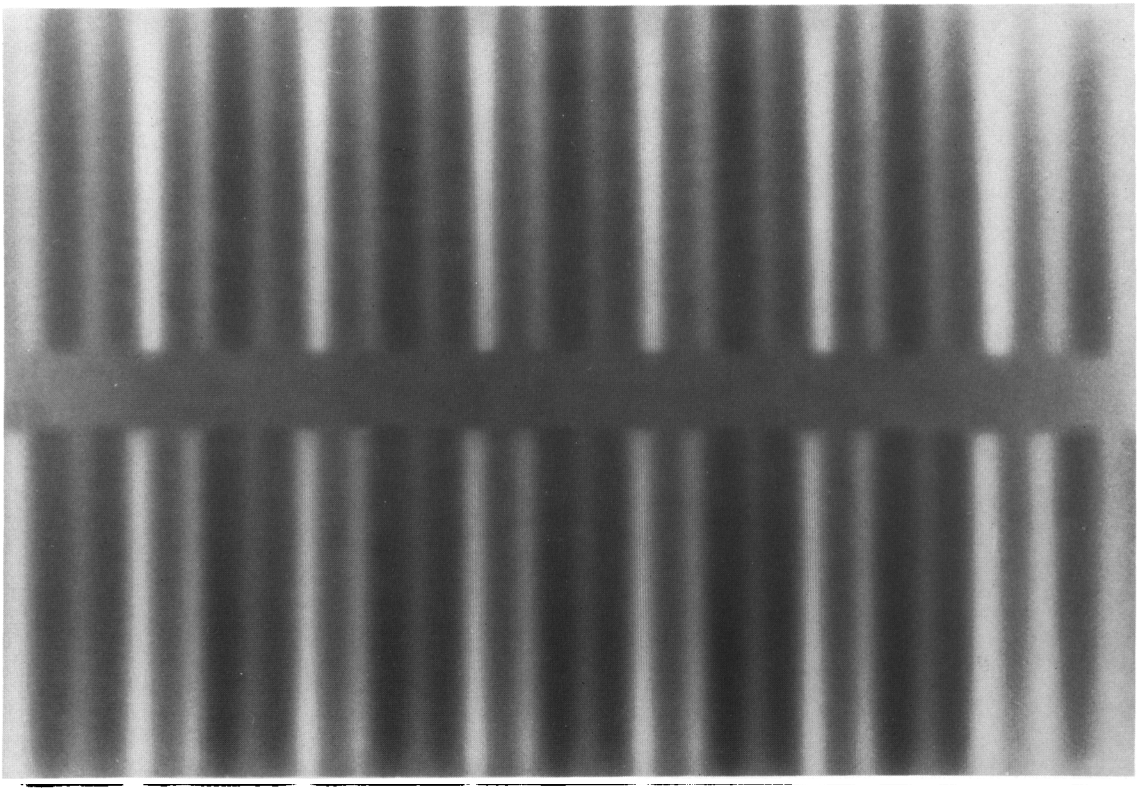

Plate 1. Example of a Vernier display in which each half field comprised two sinusoidal components having a frequency ratio of 1:3. In the upper half of the display the higher-frequency component has been shifted to the left by one quarter of its period; the location of the lower frequency component is the same in both the top and bottom halves of the stimulus.

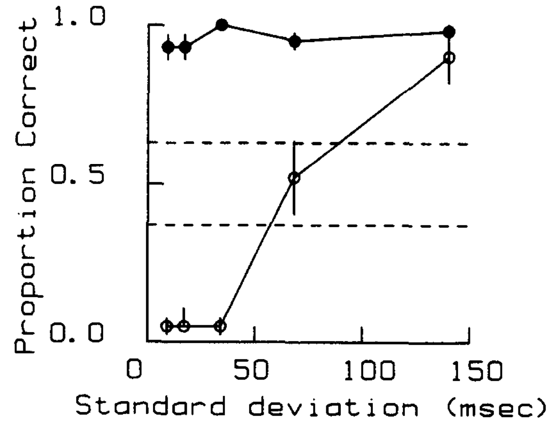

Fig. 2. Mean direction-discrimination performance of six naive observers plotted as a function of stimulus duration (expressed as the standard deviation of the Gaussian temporal envelope). The moving stimulus was a 3-c/deg sinusoidal grating, of contrast 0.12, moving at 12.2 Hz. It was either presented alone (solid symbols) or added to a static 1-c/deg grating of the same contrast (open symbols). The addition of a static pattern reverses the apparent direction of motion at short durations. Each point represents the mean number of correct responses based on 10 trials per observer; error bars show ±1 SEM. Dashed lines mark the limits, assuming binomial variability, of ±2 standard deviations from the mean performance expected by chance.

were able to press the keys to their own satisfaction after a few minutes instruction and practice. All wore appropriate correction. We served as Observers in the remaining experiments.

Experimental procedure

All experiments were self-paced, temporal, two-alternative forced-choice (2-AFC) tasks requiring discrimination of direction of motion (or Vernier offset). Each trial, initiated by the observer, contained two observation intervals, each one second in duration and marked by bursts of audible noise. During one of the intervals, chosen at random, the pattern under investigation was presented moving in one direction (or with one sign of Vernier offset). During the other observation interval, the same pattern was presented with the opposite direction of motion (or Vernier offset). The observers' task was to indicate, using a keyswitch, which observation interval had contained the leftward-moving pattern (or in which the upper pattern had the leftward Vernier offset). The observers were not informed whether they had judged correctly. The number of observations per point, which varied between 10 and 100, is indicated in the figure legends.

The pattern to be presented on each trial was chosen at random from a set of between 5 and 12 with the constraint that no pattern was presented for the *n*th time until the whole set had been presented $n - 1$ times. The spatial phase of each component at the start of each observation interval (or in the reference half of each Vernier display) was selected at random.

RESULTS AND DISCUSSION

This section is divided into two parts: first we describe the basic phenomenon and explore the effect of changing a number of stimulus parameters. The parameters that we manipulated were chosen with a view to evaluating explanations based on design flaws in a Reichardt motion detector. We then evaluated the performance of a model in which motion is detected by comparing the temporal and spatial partial differentials of the luminance pattern (Limb and Murphy, 1973; Fennema and Thompson, 1979). This type of model predicts many of our results.

(1) *The basic phenomenon*

Figure 2 shows the average proportion of correct responses of 6 naive Observers discriminating direction of motion plotted against the duration for which the stimulus was displayed. Duration is expressed as the standard deviation of the Gaussian temporal envelope containing the stimulus. Solid circles show performance when the moving component, a 3-c/deg grating of contrast 0.12 and moving past a fixed point at 12.2 Hz, was presented alone. In this case, changing the duration of the display has no effect on performance, which is almost errorless. The slight fluctuations in performance are attributable to the occasional errors made by naive observers unable to remember which key to press.

Open circles show the performance of the same observers with the same moving component (a 3-c/deg grating moving at 12.2 Hz) which, in this case, was added to a static sinusoid of the same contrast but lower spatial frequency (1 c/deg). With this stimulus, performance is strongly dependent on duration. Performance is almost errorless at the longest duration but deteriorates rapidly at shorter durations. The dashed horizontal lines indicate the two standard deviation confidence interval either side of the mean expected for chance performance on the assumption of binomial variability. Note that performance is well below this range at the three shortest durations. This means that the observers do not simply fail to perceive motion; rather they show a systematic error in perceived direction. They clearly see motion but motion in the opposite direction from that in which the 3-c/deg component actually moves.

One possible explanation of this result is that the addition of a static component to the moving grating introduces a new feature into the luminance profile, a new feature that moves in a direction opposite to that of the moving component. Such an explanation also requires that we be more sensitive to the motion of the hypothetical new feature than to the motion of 3-c/deg component at short—but not at long—exposure durations. Inspection of sample luminance profiles for two relative phases and two movement amplitudes (Fig. 3) lends no great support to this hypothesis: when the 1-c/deg component is present, there are no obvious features that shift to the right when the 3-c/deg component is shifted leftward. In particular the peaks, troughs, and edges of the profile all move

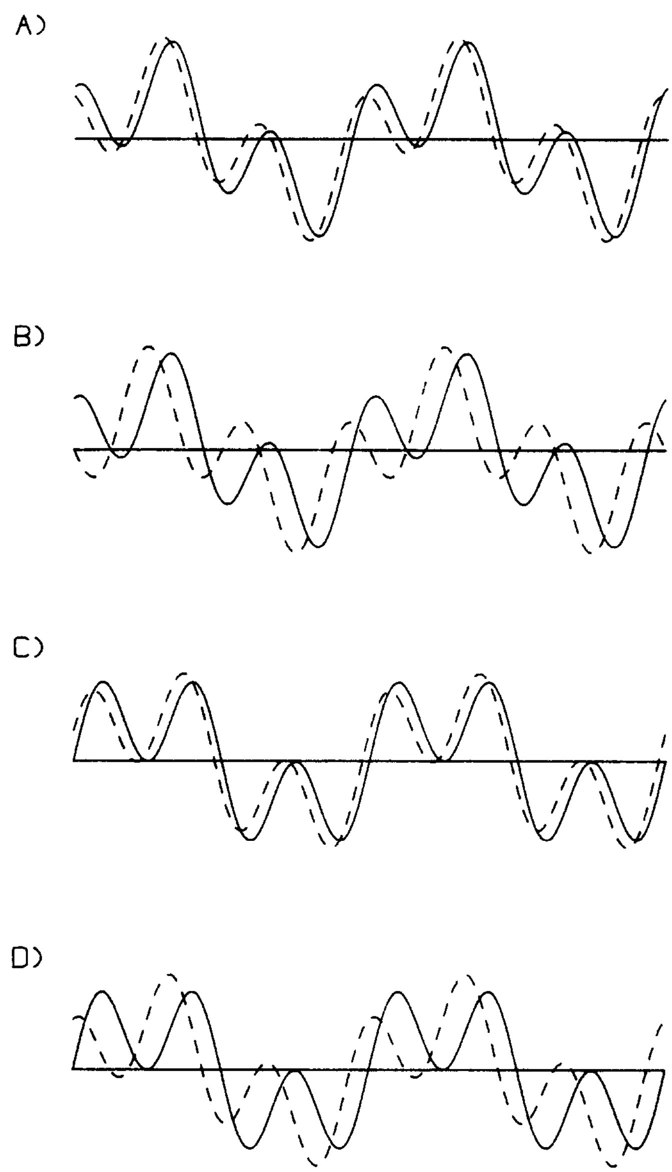

Fig. 3. Solid lines show luminance profiles of patterns composed by adding sinusoidal gratings of 1 and 3 c/deg and equal contrast. Each dashed line shows the profile produced by moving the 3-c/deg component to the left by 0.1 cycle (A,C) or 0.25 cycles (B,D). In (A) and (B) the 1-c/deg component is in sine phase and the 3-c/deg component (unshifted) is in cosine phase at the left-hand side. In (C) and (D) both components start in sine phase.

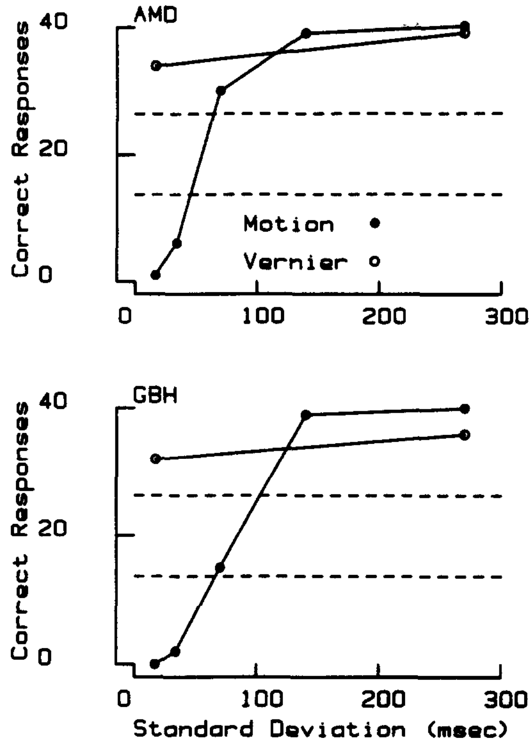

Fig. 4. Open circles show performance in discriminating the direction of Vernier offset between two patterns containing a vertically aligned 1-c/deg component and a 3-c/deg component by 0.25 cycles. Solid circles show direction discrimination performance for apparent motion induced in the same pattern by the same phase shift. At short exposure durations apparent direction of motion reverses but apparent offset does not change. Data points were based on 40 trials per point.

to the left. Other possibilities are that a nonlinearity early in the visual system introduces a distortion product which moves in the opposite direction from the 3-c/deg component or that asymmetry in the time-course of the responses to the two gratings introduces features moving in the opposite direction from the 3-c/deg component. The former hypothesis is improbable as the largest likely distortion product is at the difference frequency (2 c/deg in this case) and moves in the same direction as the moving component if (as in this case) the moving component has the higher spatial frequency (Badcock and Derrington, 1985). The latter is certainly plausible given the results of Watson and Nachmias (1977). However, since both hypotheses are based on the assumption that some feature in the luminance profile (or its internal representation) moves in the opposite direction from the 3-c/deg grating, both predict that the errors in discriminating the direction of motion should also occur in a task where observers discriminate the direction of a static displacement. Therefore the models can be tested by using the same patterns in direction-discrimination and in Vernier offset-discrimination tasks and comparing the performance obtained. Figure 4 allows such a comparison.

In Fig. 4, solid circles show performance in a direction discrimination task in which the 3-c/deg component was displaced instantaneously (by 0.083 deg; i.e. 0.25 of its spatial period) half-way through the observation interval. As with continuous motion when the static low-spatial-frequency component is present (Fig. 2), observers report movement in the wrong direction at short durations and in the correct direction at long durations. Note that the perception of motion still depends on exposure duration even though the actual motion is an "instantaneous" step in the temporal centre of the observation interval.

Although it is surprising in this sort of motion that the reversal of apparent direction shows a similar dependence on stimulus duration as does continuous motion, it is no surprise that the reversal occurs with brief exposures; with brief exposures, motion produced by a step cannot be distinguished from continuous motion, as one might expect from comparison of the temporal frequency spectra produced in the two cases (Watson and Ahumada, 1985).

The open circles in Fig. 4 show performance of the same Observers using the same patterns in a Vernier task. Here the two patterns, which if presented successively give an impression of motion, are presented simultaneously, one above the other, and the subject judges the direction of the offset. Although performance declines somewhat at the shorter exposure duration, performance is always substantially above chance: there is certainly no reversal.

The lack of a reversal in performance in the Vernier task shows that there are no systematic errors in perception of a *static* displacement at short exposure durations: the error lies in the perception of motion. Thus, unless we are prepared to assert that the receptive fields providing signals to the motion system are not used to signal static position, we must reject both the notion of incorrectly located distortion products or inappropriately located features as explanations of our results. Consequently we must look for an explanation in terms of motion-detectors, rather than position-detectors.

One simple explanation, which also takes account of the difference in behaviour at short and long exposure durations, uses the notion of two systems for signalling visual motion. Long exposures are known to favour the high-level system (Anstis, 1970) which correctly signals the motion in our complex stimulus. When the exposure is too brief for the high-level system to operate, we see the effect of the low-level system, which signals motion in the correct direction for a simple grating but in the wrong direction for our complex stimulus. All that remains is to explain why adding a static pattern throws the low-level system, but not the high-level system, into such complete confusion.

There are two ways in which a single Reichardt-type detector might signal the direction of motion of our stimuli incorrectly. The first possibility is that the non-linear operation of multiplication [which occurs in different ways in two of the three models (Adelson and Bergen, 1985a; Van Santen and Sperling, 1985)], will add components to the internal representation that are not in the stimulus itself. For appropriate combinations of relative phase and relative amplitude of the stimulus components, these distortion products will generate motion signals of the wrong sign. Under normal circumstances these components are removed by averaging in time (Van Santen and Sperling, 1985); however this strategy may be ineffective at short stimulus durations. Nevertheless this does not seem to be a plausible explanation of the reversed motion perception that we observe, as it predicts that the phenomenon should be sensitive to the relative phases of the two components. Although we have not explored the effects of varying phase systematically, the fact that we observe consistent performance when phase is varied randomly suggests that phase has little effect.

The linear motion sensor (Watson and Ahumada, 1985) will not generate such distortion products, as it has no multiplication stage. Nor will it produce systematic errors for any combination of a moving and a static sinusoid. This can be seen from Watson and Ahumada's representation of their sensor as a spatio-temporal filter: the spatio-temporal frequency representation of our stimuli is dominated by components representing motion in the correct direction, so filters responding to motion in that direction will dominate the response.

The second possible failure in a Reichardt detector is some form of aliasing caused by spatial or temporal frequencies outside the range for which the detector is designed. Two possible types of aliasing, spatial and temporal, are discussed in some detail by van Santen and Sperling (1984) who carefully include in their model spatial and temporal filters designed to prevent processing of components which might cause aliasing. The next two experiments examine the possibility that aliasing in a single Reichardt-type detector causes the errors we observe.

Spatial aliasing occurs when the distance between the two elements of the detector exceeds 0.5 cycles of the input spatial frequency (van Santen and Sperling, 1984). Detectors which rely on differences in the phase response of their input elements (Watson and Ahumada, 1985; Adelson and Bergen, 1985a) should not be vulnerable to spatial aliasing. In a detector like that illustrated in Fig. 1, spatial aliasing will occur if the mechanism attempts to transmit spatial frequencies so high that their spatial period is less than twice the separation between the two receptive fields. Aliasing need not always give rise to incorrect responses, but if the spatial frequency of the stimulus is such that the responses in the two elements of the detector are 180° out-of-phase, a response based on the aliased component will be incorrect. Thus, if spatial aliasing is causing our Observer's errors, a plot of performance as a function of the spatial frequency of the moving grating might be expected to oscillate between correct and incorrect with a period equal to the reciprocal of the distance between the elements of the detector.

Before we can test for spatial aliasing, we need a first guess at the separation between the elements of the detectors giving errors since this dictates the period of the oscillation we expect to find. The most likely value is 1 degree, since errors are induced by a 1-c/deg pattern which, apart from any direct effect, will cause a modest 1-c/deg modulation in the contrast of the 3-c/deg grating. In preliminary experiments we showed that it was unlikely that the effect was caused by elements closer than 1 degree apart by confirming that direction-discrimination errors occur when a moving grating of 2, 3, 4, 5 or 6 c/deg is added to a static grating of 1 c/deg. Figure 5 shows performance as a function of the frequency of the moving grating in the range between 3 and 4 c/deg when the temporal frequency was held constant at 12.2 Hz. Performance is consistently below chance, thus

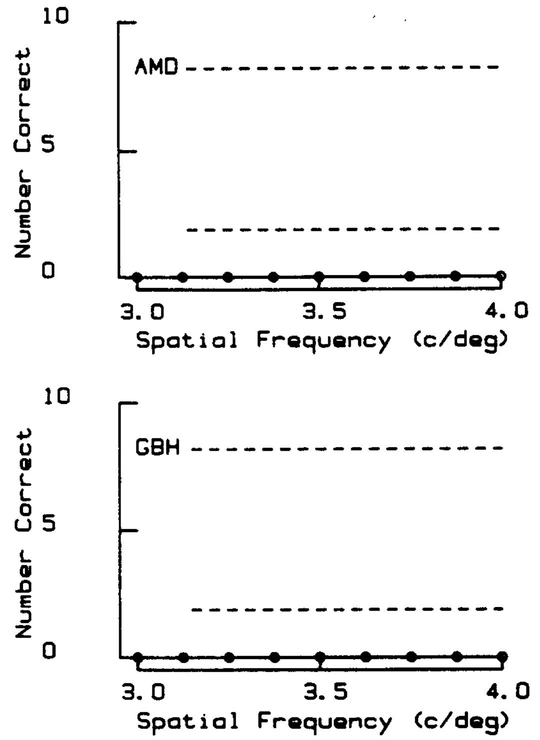

Fig. 5. Direction discrimination as a function of the spatial frequency of the moving grating added to a static 1-c/deg grating. Apparent direction is reversed at all spatial frequencies between 3 and 4 c/deg. The signal duration was 9 msec and each data point was based on 10 observations; other details as Fig. 2.

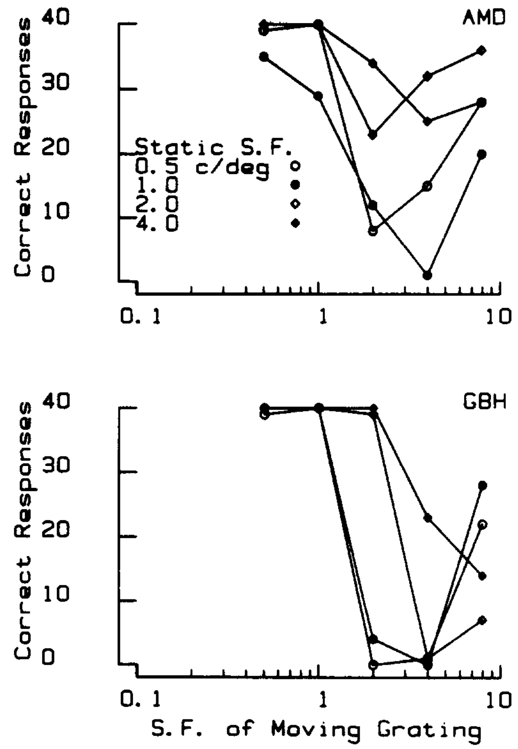

Fig. 6. Direction discrimination as a function of the spatial frequency of a moving grating added to static gratings of the following different spatial frequencies: open circles, 0.5 c/deg; solid circles, 1.0 c/deg; open diamonds, 2.0 c/deg; solid diamonds, 4.0 c/deg. The moving grating always had a temporal rate of 12.2 Hz. As the spatial frequency of the static grating increases, the range of moving gratings which give reversal is reduced. There were 40 observations per point; other details as Fig. 5.

excluding spatial aliasing by detectors with receptive fields more widely separated than 1 deg; spatial aliasing does not account for our results.

In spite of large individual differences, Fig. 6 shows that motion discrimination is not, of course, independent of the spatial frequencies of the moving and static gratings. This is shown in Fig. 6, where the number of correct responses is plotted as a function of the spatial frequency of the moving grating. The parameter is the spatial frequency of the static grating. The contrast of both gratings was 0.12. The region of reversal is indicated by performance below 20 correct responses and depends on the spatial frequencies of both the static and the moving grating; the higher the spatial frequency of the static grating, the smaller is the range of spatial frequency of the moving grating for which induction of the incorrect apparent motion occurs. With static gratings of sufficiently high spatial frequency, neither observer sees a reversal although performance near chance is produced. Further, reversed apparent motion does not occur when the spatial frequency of the moving grating is below that of the static one. This observation is consistent with those of Levi and Schor (1984) for induced motion.

Temporal aliasing will occur when the temporal frequency produced by the grating moving past each element is such that the delayed response lags the non-delayed response by more than 0.5 cycles (van Santen and Sperling, 1984). Although it is difficult to imagine a simple mechanism whereby the addition of a static pattern would alter the temporal properties of the motion-detector, we felt it worth examining performance as a function of temporal frequency. Figure 7 shows performance as a function of the temporal frequency of the moving grating for two observers and two display durations. Increasing temporal frequency causes performance to decline monotonically from a starting value not significantly different from chance; performance is consistently below chance for all temporal frequencies above 2 Hz.

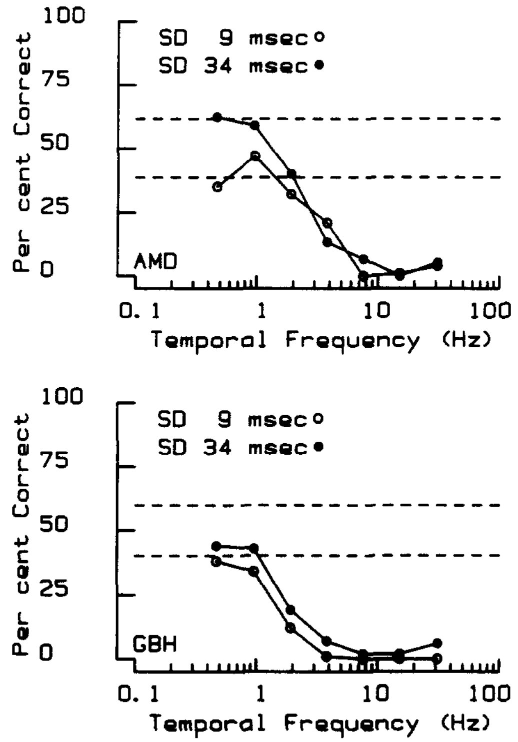

Fig. 7. Direction discrimination for a moving 3-c/deg grating added to a static 1-c/deg grating is plotted as a function of the temporal frequency of movement for two exposure durations. Data for G.B.H. were based on 100 observations per point; 75 observations per point for A.M.D. Other details are as in Fig. 2.

Although the temporal frequency axis is not sufficiently densely sampled to detect oscillations in performance that might be expected if temporal aliasing caused by a pure delay produced the failure in discrimination performance, the fact that the failure begins at such a low temporal frequency makes it seem unlikely that failure is produced by temporal aliasing. Rather, the results suggests that the only motion visible with these spatial patterns is motion in the wrong direction. Increasing the duration of the stimulus from 9 to 34 msec has no discernible effect on the pattern of performance.

It appears, then, that simple design flaws in Reichardt's motion detector are unlikely to account for our observers' behaviour. We consider the properties of an alternative model in the next section.

(2) *Motion detection by comparison of spatial and temporal gradients*

When an object defined by spatial variation in luminance moves, the direction and velocity of motion can usually be computed from the joint knowledge of the temporal change in luminance at each point and the spatial profile of the luminance. For a pattern varying in only one spatial dimension the velocity of its motion, V, at any location, x_0, is given directly by dividing the temporal luminance derivative at x_0 by the spatial derivative at the same point

$$V(x_0, t) = -\{[\partial f(x, t)/\partial t]/[\partial f(x, t)/\partial x]\}|_{x_0} \quad (2)$$

(Limb and Murphy, 1973). [The sign inversion reflects the fact that an increase in phase causes leftward (negative) motion.]

Limb and Murphy, and Fennema and Thompson (1979) have developed two-dimensional motion-detection systems based on this quantity. Marr and Ullman (1979) use the same principle, but impose physiological constraints. In this section we show that, if motion detection in man were based on Limb and Murphy's system, then errors like those we observe with our stimuli arise naturally.

The cross-sectional luminance profile of our stimulus is given as a function of space (x) and time (t) by

$$L(x, t) = A_1 \sin[2\pi(f_1 x + g_1 t)] + A_2 \sin[2\pi(f_2 x + g_2 t)] + L_m, \quad (3)$$

where g_1 and g_2 are the rates of change in location of the two components of our stimuli with time and f_1 and f_2 are their spatial frequencies; g_1 was always zero.

The temporal derivative, $\partial f(x, t)/\partial t$, is thus given by

$$\partial f(x, t)/\partial t = 2\pi A_2 g_2 \cos[2\pi(f_2 x_2 + g_2 t)] \quad (4)$$

and the spatial derivative is given by

$$\partial f(x, t)/\partial x = 2\pi A_1 f_1 \cos[2\pi(f_1 x + g_1 t)] + 2\pi A_2 f_2 \cos[2\pi(f_2 x + g_2 t)]. \quad (5)$$

In the case of the one-component display

$$A_1 = 0 \quad (6)$$

so that $\partial f/\partial x$ reduces to

$$\partial f/\partial x = 2\pi A_2 f_2 \cos[2\pi(f_2 x + g_2 t)] \quad (7)$$

and

$$V(x, t) = -g_2/f_2 \quad (8)$$

independent of location and time; that is, with only one component present all points of the display appear (correctly) to have the same velocity.

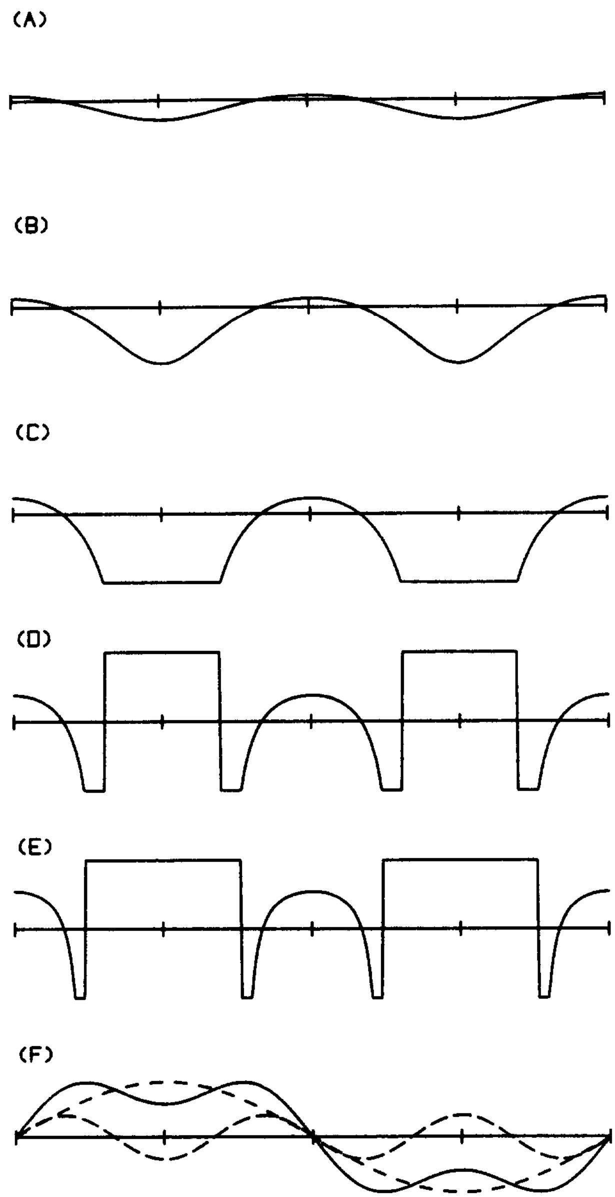

Fig. 8. (A–E) Apparent velocity, computed from the ratio of the temporal and spatial derivatives of the luminance profile [equation (9)], as a function of position within a pattern comprised of a moving 3-c/deg grating (of different contrast in each figure) added to a static 1-c/deg grating. Upward deflection indicates motion in the same direction as the moving 3-c/deg component, downward deflection indicates motion in the opposite direction. The velocity scale is linear up to the point of truncation. The contrast of the 3-c/deg component relative to that of the 1 c/deg component is: (A) 0.025, (B) 0.05, (C) 0.1, (D) 0.2, (E) 0.4. The solid line in (F) shows the luminance profile of the pattern whose velocity profile is shown in (E). The dashed lines show the two components separately.

In the case of the complex display

$$V(x,t) = -g_2 A_2 \cos[2\pi(f_2 x + g_2 t)]/ \{A_1 f_1 \cos[2\pi(f_1 x + g_1 t)] + A_2 f_2 \cos[2\pi(f_2 x + g_2 t)]\}. \quad (9)$$

This is a bit more complicated: velocity is no longer calculated to be uniform across the display, although our observers never report seeing other than uniform motion; further, the distribution of the estimates of the velocity depends on the ratio of A_1 and A_2. We do not know this ratio for the internal representation of the stimuli. Figure 8 shows vertically truncated plots of velocity as a function of position within one cycle of a pattern where f_1 is 1 c/deg, f_2 is 3 c/deg,

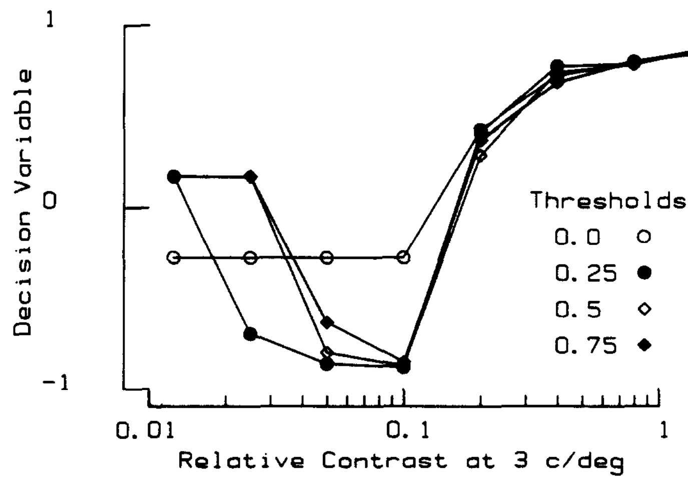

Fig. 9. Decision variable for estimating direction of motion (computed from equation 9 as described in the text) as a function of the relative contrast of the 3-c/deg component in a pattern containing a moving 3-c/deg component and a static 1-c/deg component.

and A_2 takes on values (relative to A_1) that range between 0.025 (plot A) and 0.4 (plot E). Positive values of the variable indicate motion in the same direction as the moving 3-c/deg component; negative values indicate motion in the opposite direction.

Where the spatial gradient of the pattern is large, motion is always in the direction of actual motion. Further, when A_2 is large, almost all the pattern has a positive velocity: most of it should appear to go the same way as a simple pattern. However, when A_2 is small, the reverse is true: most of the pattern has negative velocity and should appear to go in the wrong direction. Thus, Limb and Murphy's model predicts a disruption of motion perception similar to that we observe with short duration stimuli.

These reversals in direction reflect the fact that, in a pattern constructed by adding a moving and a static component, the ratio of spatial and temporal luminance derivatives no longer gives a valid estimate of velocity. This is not surprising if we remember that the spatial luminance derivative in such a pattern contains contributions from both the moving and the static components, whereas the temporal derivative is determined only by the moving component.

With the intention of making systematic comparisons among stimuli, and to illustrate how one might begin to develop a model of human motion perception based in Limb and Murphy's idea, we constructed a decision variable from $V(x, t)$ as follows: first we use equation (9) to evaluate velocity at 100 points, i.e. at intervals of 0.01 degree across one full cycle of the pattern. Then we count the number (F_c) of points where the velocity is both correct and above some threshold magnitude, and the number (F_i) where it is incorrect and greater than the threshold. The decision variable is given by

$$D = (F_c - F_i)/(F_c + F_i + F_k). \qquad (10)$$

F_k is simply a constant introduced to prevent the decision variable reaching large values when extremely small fractions of the display are moving. We expect that positive values of the decision variable would lead to correct discrimination, and negative values would lead to incorrect discrimination. Values close to zero would lead to performance close to chance.

Figure 9 shows how the decision variable changes as a function of the relative contrast of the moving 3-c/deg grating in a display which is the sum of a moving 3-c/deg grating and a static 1-c/deg grating of fixed contrast. The decision variable first falls with contrast, and then rapidly climbs as contrast increases further. Changing the threshold speed used to compute the decision variable affects the width of the negative region of the curve, but has little effect on the overall shape.

The open circles in Fig. 10 show the effect of changing the contrast of a moving 3-c/deg grating, presented together with a static 1-c/deg grating (contrast 0.12) on the performance of our observers. Performance falls with increasing contrast of the moving grating until it is about twice that of the static grating, after which it rises rapidly. The superficial agreement between

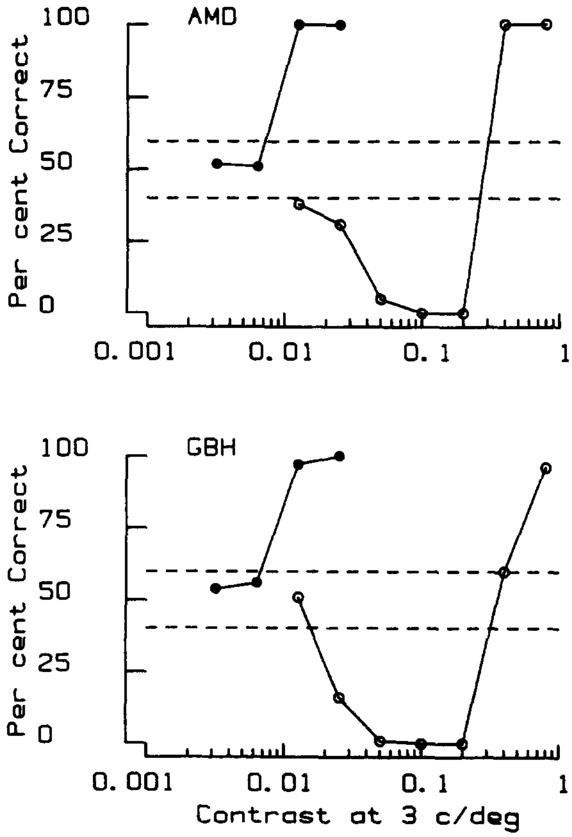

Fig. 10. Open circles show performance in discriminating direction of motion of a 3-c/deg component of varying contrast, added to a static 1-c/deg component of contrast 0.12. Performance is plotted against the contrast of the 3-c/deg component. Solid circles show direction discrimination as a function of contrast for the 3-c/deg grating presented alone. The stimulus duration was 9 msec and each data point was based on 100 observations. Other details as for Fig. 2.

the patterns of change in psychophysical performance and in the decision variable is encouraging. The difference of about an order of magnitude between the range of relative amplitudes where reversal occurs psychophysically, and the range where it occurs in the model is not a problem; it simply implies that the motion detector is preceded by a stage of spatial low-pass filtering.

Of course, the simple decision variable we use cannot be expected, as it stands, to account for all characteristics of motion discrimination. For example, performance with a 3-c/deg moving grating presented alone (solid circles in Fig. 10) varies with contrast and increases rapidly from chance levels as the contrast rises. The decision variable, however, is not affected by absolute contrast and so for a single grating its value depends only on the rate of displacement; the model takes no account of signal-to-noise ratio. Further, models sensitive to gradients are very sensitive to noise whenever the gradients become small (Adelson and Bergen, 1985b).

GENERAL DISCUSSION

Our main finding is that at short stimulus durations, the addition of a static 1-c/deg grating to a moving 3-c/deg grating makes a stimulus appear to move in the wrong direction. The reversal is both compelling and consistent; it is reported both by naive and by experienced observers and appears to be a fundamental characteristic of human motion sensitivity, which, as far as we can tell, is not consistent with most current models of motion detection. One possible route to a successful model would be to incorporate appropriate interactions between Reichardt-detectors sensitive to different bands of spatial frequency. On the other hand, there is a distinct possibility that a plausible and more sophisticated model based on comparison of spatial and temporal gradients might be more successful in predicting the behaviour we observe.

Spatial-frequency selective Reichardt detectors

It is commonly assumed that individual Reichardt detectors are relatively selective for spatial frequency and the usual crudely tuned spatial-frequency selectivity of the visual system has certainly been demonstrated in motion detection (Anderson and Burr, 1985). Selectivity results because of the spatial configuration of the receptive fields of the elements that drive motion detectors (van Santen and Sperling, 1984; Watson and Ahumada, 1985; Adelson and Bergen, 1985a). The two components of our stimulus would be detected by different sets of detectors, each of which should correctly signal the motion of the component to which it was most sensitive, thus giving rise to a conflict. If this conflict were resolved by taking the output of the detector tuned to high spatial frequencies as a reference signal, then the low-frequency grating might be treated as moving in the opposite direction.

Such a scheme would undoubtedly account for the results, but there are two obvious weaknesses. The first is that it is fundamentally implausible in three respects: there seems to be no advantage in making the system vulnerable to error in this way; there is no obvious reason for choosing high frequencies to form the frame of reference, there is no reason for motion discrimination to depend on duration. The second is that our observers see no ambiguity in the motion of brief stimuli, except at the limits of the spatial-frequency range in which reversal occurs, although they do observe it at long exposures, where the motion of the 3-c/deg grating (itself correctly perceived) induces a weak apparent motion of the 1-c/deg grating in the opposite direction.

An alternative form of interaction between spatial frequency selective motion detectors has been suggested to us (Adelson and Bergen, 1985b). They point out that when our stimuli are presented briefly, their temporal-frequency spectra will be smeared to such an extent that the static 1-c/deg grating may be expected to stimulate motion detectors tuned to both directions of motion. Antagonistic and asymmetric interactions between detectors tuned to *different* spatial frequencies but to the *same* direction of motion could obliterate the response of detectors tuned to 1-c/deg gratings moving in the same direction as the high spatial-frequency component, thus leaving us with an overall impression of motion of the low spatial-frequency component in the opposite direction. It will be interesting to see what additional predictions for motion perception such a model possesses.

Possible gradient-intensity transform models

The procedure we use to compute a decision variable for direction-discrimination takes no account of the physiological constraints imposed by known elements capable of approximating both temporal and spatial derivatives. Such constraints, however, are easily imposed. Thus, although we take encouragement from the fact that a model based on Limb and Murphy's or Fenema and Thompson's devices predicts a reversal of apparent motion with complex stimuli, it is clear that such a model would require extensive development in order to predict all our results. In addition to incorporating the spatial and temporal properties of the neural mechanisms likely to perform the relevant processing, such a model would have to take account of the effects of contrast, and of the variation in predicted velocity in different regions of the stimulus.

Finally, it is worth noting that a model which uses the principle of comparing temporal and spatial gradients, and which makes some concessions to physiological data (Marr and Ullman, 1979) does not appear to account for the reversal of apparent movement with complex stimuli. It fails because it only operates at zero-crossings in the (effectively) twice-differentiated and filtered luminance profile. In the simulations in Fig. 8, the direction of motion at the "zero-crossings" in the spatial luminance profile is always positive, indicating motion in the same direction as the moving component of the stimulus. Although we have not simulated the two spatial differentiations and filtering postulated by Marr and Ullman, they will have a negligible effect on the locations of the zero crossings. Thus we can conclude that Marr and Ullman's motion detector will not make errors with these stimuli.

CONCLUSIONS

The available models of visual direction selectivity based on Reichardt's notion fail to predict the systematic errors that human observers make with briefly exposed complex stimuli. We conclude that this is because they are inadequate as models of the "short-range" motion detector (Braddick, 1974).

Acknowledgements—This research was supported by MRC project grant 82/19394N and by the Addison Wheeler Fund. A.M.D. was an Addison Wheeler Fellow.

REFERENCES

Adelson E. H. and Bergen J. R. (1985a) Spatio-temporal energy models for the perception of motion. *J. opt. Soc. Am.* **A2**, 284–299.

Adelson E. H. and Bergen J. R. (1985b) Personal communication.

Anderson S. J. and Burr D. C. (1985) Spatial and temporal selectivity of the human motion detection system. *Vision Res.* **25**, 1147–1154.

Anstis S. M. (1970) Phi movement as a subtraction process. *Vision Res.* **10**, 1411–1430.

Anstis S. M. (1980) The perception of apparent movement. *Phil. Trans. R. Soc.* **B290**, 153–168.

Badcock D. R. and Derrington A. M. (1985) Detecting the displacement of periodic patterns. *Vision Res.* **25**, 1253–1258.

Braddick O. J. (1974) A short-range process in apparent motion. *Vision Res.* **14**, 519–529.

Braddick O. J. (1980) Low-level and high-level processes in apparent motion. *Phil. Trans. R. Soc.* **B290**, 137–151.

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

Fennema C. L. and Thompson W. B. (1979) Velocity determination in scenes containing several moving objects. *Comput. Graphics* **9,** 301–315.

Limb J. O. and Murphy J. A. (1973) Estimating the velocity of moving images in television signals. *Comput. Graphics* **4,** 311–327.

Levi D. M. and Schor C. M. (1984) Spatial and velocity tuning of processes underlying induced motion. *Vision Res.* **24,** 1189–1196.

Marr D. and Ullman S. (1979) Direction selectivity and its use in early visual processing. *Proc. R. Soc.* **B211,** 151–180.

Ramachandran V. S., Rao V. M. and Vidyasagar T. R. (1973) Apparent movement with subjective contours. *Vision Res.* **13,** 1399–1401.

Reichardt W. (1961) Autocorrelation, a principle for the evaluation of sensory information by the nervous system. In *Sensory Communication* (edited by W. A. Rosenblith). Wiley, New York.

van Santen J. P. H. and Sperling G. (1984) A temporal covariance model of motion perception. *J. opt. Soc. Am.* **A1,** 451–473.

van Santen J. P. H. and Sperling G. (1985) Elaborated Reichardt detectors. *J. opt. Soc. Am.* **A2,** 300–321.

Schade O. H. (1956) Optical and photoelectric analog of the eye. *J. opt. Soc. Am.* **46,** 721–739.

Watson A. B. and Nachmias J. (1977) Patterns of temporal interaction in the detection of gratings. *Vision Res.* **17,** 893–902.

Watson A. B. and Ahumada A. (1985) Model of human visual motion sensing. *J. opt. Soc. Am.* **A2,** 322–342.