*Vision Res.* Vol. 30, No. 9, pp. 1341–1362, 1990
Printed in Great Britain. 

0042-6989/90 $3.00 + 0.00


# PERCEPTION OF DIRECTIONAL SAMPLED MOTION IN RELATION TO DISPLACEMENT AND SPATIAL FREQUENCY: EVIDENCE FOR A UNITARY MOTION SYSTEM*

WALTER F. BISCHOF† and VINCENT DI LOLLO
University of Alberta, Edmonton, Alberta, Canada T6G 2E9

(*Received* 1 *August* 1989; *in revised form* 8 *January* 1990)

**Abstract**—Perception of directional motion was studied by displaying two images (F1 and F2) in rapid succession. The two images were identical except for a horizontal displacement of F2 with respect to F1. Observers reported the direction of horizontal motion over a wide range of displacements. The stimuli in Experiment 1 were one-dimensional gratings with spatial frequency between 0.125 and 6 c/deg. Motion was seen at all displacements to almost 0.5 cycles (counterphase) and remained invariant across spatial frequencies. In Experiment 2 the stimuli were band-pass filtered random-dot patterns. The bandwidth of the filters was 1 octave, and centre frequencies ranged from 0.75 to 12 c/deg. In every case, the response functions exhibited quasi-periodic oscillations related to structural properties of the images. One-dimensional analyses based on autocorrelation did not provide a satisfactory account of the data. By contrast, the data were fitted successfully by a two-dimensional analysis that integrated the responses of neighbouring motion detectors so as to yield a smooth motion flow field from which left–right directional motion could be derived. Practically and conceptually, the outcome supports a unitary motion system as distinct from separate systems subserving short-range and long-range motion.

Motion perception  Spatial filtering  Sinusoidal gratings  Random-dot kinematograms

Perception of motion in human vision is said to be mediated by two separate systems: the short-range and the long-range systems. According to this classification—first proposed by Braddick (1974)—the short-range system operates across shorter distances and briefer temporal intervals than the long-range system. Further, short-range processes are said to occur at an earlier stage of visual processing and to be based on simpler neurophysiological mechanisms than long-range processes. The characteristics of the two systems have been discussed in detail by Anstis (1980, 1986) and by Braddick (1980). The present work is concerned exclusively with effects usually associated with short-range processes.

Much of the psychophysical evidence concerning short-range motion has been gathered with random-dot stimuli displayed in two sequential frames separated by an inter-stimulus interval (ISI). The leading frame (F1) usually consists of dots distributed randomly within the viewing area. The corresponding trailing frame (F2) is the same as F1 except that each dot is displaced by a fixed extent in a uniform direction. Any dot in F2 that is displaced our of the viewing area is "wrapped around" so that it reappears at the opposite end. In this fashion, all apparent motion is confined to the individual

*Following the submission of this article, Dr O. Braddick kindly sent us a pre-publication copy of a paper (Cleary & Braddick, 1990a) that reported a virtual duplication of the research described in the present Experiment 2. The similarity of outcomes of the two independent investigations attests to the stability of the effects. Both revealed quasi-periodic oscillations of the response functions, as well as inverse scaling of $d_{max}$ with spatial frequency. Also, both papers offered similar accounts of Chang and Julesz' (1985) failure to find inverse scaling with $d_{max}$ at frequencies higher than about 4 c/deg. Finally, both investigations found the slope of the inverse-scaling function to be slightly shallower than might be expected on the basis of perfect scaling (Cleary & Braddick's Fig. 8, our Fig. 8).

The value of the present work, however, goes beyond mere replication. In Experiment 1 we estimated $d_{max}$ with one-dimensional gratings. This permitted useful comparisons with results obtained with two-dimensional band-pass filtered images in Experiment 2. Perhaps the most important difference between the two papers, however, is our suggested alternative model of motion perception based on integrative interactions among individual motion detectors. This model was added in a revision of the paper. Its formulation was prompted and facilitated by the theoretical discussion in Cleary and Braddick's paper.

†To whom reprint requests should be addressed.

dots, while the boundaries of the viewing area remain stationary.

With this type of display, it is possible to estimate the maximum displacement at which directional motion can be reliably perceived. In a seminal paper, Braddick (1974) estimated the extent of such maximum displacement (later named $d_{max}$, at approx. 15 min arc. Subsequent investigations examined variations in $d_{max}$ as a function of such variables as retinal eccentricity (Baker & Braddick, 1983; Bischof & Groner, 1985), state of retinal adaptation (Morgan & Ward, 1980), number of elements (Baker & Braddick, 1982), total viewing area (Baker & Braddick, 1982; Nakayama & Silverman, 1984), and temporal factors such as exposure duration and duration of ISI (Baker & Braddick, 1985a).

Among these and similar investigations, several have examined changes in $d_{max}$ in relation to the spatial-frequency contents of the stimuli. In most cases, however, the issue was approached indirectly. For example, it has been shown that increasing the angular size of the stimulus elements produces corresponding increments in $d_{max}$ (e.g. Lappin & Bell, 1976; Cavanagh, Boeglin & Favreau, 1985; Nakayama & Silverman, 1984). On the reasonable assumptions that perception of larger elements is mediated by larger receptive fields, and that larger receptive fields mediate perception of lower spatial frequencies, this result suggests that the magnitude of $d_{max}$ is greater at lower spatial frequencies. However, while increments in element size undoubtedly increased the relative power of the lower spatial frequencies in the displays, there was also a large spectral overlap among the stimuli. In terms of the relation between $d_{max}$ and spatial frequency, the major limitation of these studies lie in the fact that the spectral composition of the stimuli was not adequately controlled.

Limitations of this kind did not affect a pair of studies by Chang and Julesz (1983, 1985). In these studies, the stimuli were random-dot patterns band-pass filtered at different centre-frequencies. The spectral composition of these stimuli was unambiguous in that it was defined completely by the relevant filter. However, this work contained other ambiguities that, as seen below, do not permit firm conclusions to be drawn.

Spectral composition of the stimuli was clearly defined in two recent studies that used one-dimensional gratings to study the effects of spatial frequency on $d_{max}$ (Turano & Pantle, 1985) or on $d_{opt}$ (the optimal F1–F2 displacement for perceiving motion; Baker, Baydala & Zeitouni, 1989). An indirect approach was taken in both studies. That is, magnitude of $d_{max}$ (or $d_{opt}$) was inferred from the duration of the motion aftereffect generated by prolonged exposure to discontinuously moving stimuli. It was found that duration of the motion aftereffect and, by inference, magnitudes of $d_{max}$ and $d_{opt}$ expressed in fractions of cycles, were not affected by the spatial frequency of the stimuli. Of necessity, these estimates depend on a chain of inferences linking $d_{max}$ (or $d_{opt}$) to the duration of the motion aftereffect. The validity of at least some of these inferences (e.g. the assumption of communality of mechanisms) is questioned by some of the results reported below. A simpler approach, and one that avoids these assumptions, is to estimate $d_{max}$ directly from the motion seen in pairs of sequential gratings.

Just such a direct approach was taken by Nakayama and Silverman (1985) who employed pairs of sequential one-dimensional sinusoidal gratings with the trailing grating phase-shifted in respect to the first, so as to produce the appearance of horizontal motion. The research was aimed at the relation between contrast sensitivity and perception of motion. The contrast sensitivity functions (CSF) obtained for frequencies of 2, 4 and 8 c/deg were all identical upon normalization, showing maximum sensitivity at a phase shift of 90 deg and declining sensitivity as the extremes of the possible range (0 and 180 deg) were approached (Nakayama & Silverman, 1985, Fig. 5). From this it can be inferred that the magnitude of $d_{max}$, expressed in cycles, does not vary across spatial frequencies. This result is broadly in line with the findings of Baker et al. (1989) and of Turano and Pantle (1985); it is also in line with other findings (Baker & Braddick, 1982; Chang & Julesz, 1983, 1985; Lappin & Bell, 1976; Cavanagh et al., 1985; Nakayama & Silverman, 1984), suggesting that the magnitude of $d_{max}$, expressed in units of visual angle, decreases as the spatial frequency of the stimulus is increased.

As a statement of the relation between $d_{max}$ and spatial frequency, the inference based on Nakayama and Silverman's (1985) experiment is clearly the more appropriate in that the measurement was direct, and much finer control over spatial frequency was achieved with the gratings they employed than with the random-dot stimuli—whether filtered or

unfiltered—employed in other investigations. However, since Nakayama and Silverman's (1985) study had not been explicitly designed to investigate the relation between $d_{max}$ and spatial frequency, only three spatial frequencies were used, and the level of contrast of the stimuli was, perforce, at threshold. As a general rule, stimulus contrast is probably not a major determinant of motion perception (e.g. Barlow & Hill, 1963; Campbell & Maffei, 1981; Nakayama & Silverman, 1985). Nevertheless, if the relation between $d_{max}$ and spatial frequency is to be understood, it is first necessary to know whether Nakayama and Silverman's (1985) results can be obtained with stimuli that are clearly above threshold, and thus more readily comparable with the stimuli used in the studies that found a negative relation between $d_{max}$ and spatial frequency. Also, in a study explicitly designed for this purpose, $d_{max}$ should be examined over a greater range of spatial frequencies.

In the present work, we studied how perception of directional motion—and, specifically, $d_{max}$—are affected by the spatial frequency contents of suprathreshold stimuli. We controlled the spectral composition of the stimuli by employing sinusoidal gratings—in the manner of Nakayama and Silverman (1985)—and also by employing band-pass filtered random-dot patterns, in the manner of Chang and Julesz (1983, 1985). We found that identification of directional motion remains essentially invariant across spatial frequencies, thus confirming Nakayama and Silverman's (1985) results with suprathreshold stimuli. The results reported by Baker et al. (1989) and by Turano and Pantle (1985) were generally confirmed; but there were also significant differences between our results and theirs, attributable to the different ways of estimating $d_{max}$ (i.e. directly, or indirectly through movement aftereffects).

## EXPERIMENT 1

To investigate perception of motion in relation to the spatial-frequency contents of the stimuli, the present study employed the three-step paradigm (F1, ISI, F2) described above. In the display sequence, F1 was a vertical grating of given frequency and phase, and F2 was a grating of the same frequency, phase-shifted by varying amounts to produce the appearance of horizontal motion.

### *Method*

#### *Observers*

The two authors and one female student, naive as to the purpose of the study, served as observers. All had normal or corrected-to-normal vision.

#### *Visual displays*

All stimuli were displayed on a Hewlett-Packard 1333A oscilloscope equipped with P15 phosphor. Unless otherwise specified, the viewing distance was 57 cm, and the square images subtended an angle of 4 deg. The display screen was front-illuminated by a pair of shielded 15 cm Sylvania F4T5CW fluorescent tubes which produced an average screen luminance of 10 cd/m$^2$, as measured by a Spectra Spotmeter.

The stimuli were vertical sinusoidal gratings with space-average luminance of 38 cd/m$^2$ (including screen illumination), and Michelson contrast of 0.47. The spatial frequency of the gratings was either 0.125, 0.25, 0.5, 1, 2, 4 or 6 c/deg. The X, Y, and Z (intensity) coordinates of the images were held in the memory of a fast plotting buffer capable of transferring them to the oscilloscope at the rate of one dot per microsecond (Finley, 1985). Since there were 16,384 dots in an image, and each image was plotted four times in succession (to improve brightness and contrast), the total exposure duration of one image was just under 66 msec.

On any given trial, the sequence of events was as follows: the observer sat in a dimly-illuminated chamber and fixated on a cross shown in the centre of the screen. Upon a button-press by the observer, the fixation cross disappeared, and the first grating (F1) was displayed for 66 msec. The starting phase angle of F1 was randomized on each trial within the range of one cycle of the grating. The ISI between F1 and F2 was set to zero. The stimulus sequence ended with a display of the F2 grating for 66 msec. The gratings in F1 and F2 were identical except that F2 was displaced within the range $\pm 0.5$ cycle from F1 so as to produce the appearance of motion to the left or to the right, randomly. The amount of displacement on any trial ranged between zero and 0.5 cycle in steps of 0.05 cycles. It must be noted that no perception of directional motion was possible at either extreme of the displacement range. One extreme was the obvious case of zero displacement. At the other extreme, a displacement of 0.5 cycle produced an F2 grating which was in counterphase relation to

F1, a relation that yielded a totally ambiguous directional signal in that F1 could have moved in either direction to arrive at F2. The observer's task was to press one of two buttons to indicate whether the grating had moved to the left or to the right.

On trial runs, it was observed that, regardless of spatial frequency, perception of directional motion was impaired if a grating contained less than about three cycles per image. (See Hoekstra, van der Goot, van den Brink & Bilsen, 1974, for a similar effect on contrast sensitivity.) In order to investigate low spatial frequencies while maintaining a sufficient number of cycles in the images, we reduced the viewing distance and increased the horizontal dimension of the displays on the screen. Specifically, the dimensions of the images with the two lowest spatial frequencies (0.125 and 0.25 c/deg) were 16 deg horizontally and 4 deg vertically, at a viewing distance of 28.5 cm, and the dimensions of the grating with spatial frequency of 0.5 c/deg were 8 deg horizontally and 2 deg vertically at a viewing distance of 57 cm. To check whether these variations in display conditions had any affect on level of performance, we ran trials in which the spatial frequency of the displays remained constant but angular size and viewing distance were varied. The results showed no noticeable effects of angular size or viewing distance on performance.

There were 11 displacements (0–0.5 cycles, in steps of 0.05 cycles) in each of the seven conditions defined by the spatial frequencies of the gratings. Each observer participated in a total of 14 experimental sessions, two for each condition. One experimental session, which lasted approx. 15 min, comprised 25 randomly-ordered trials at each of the 11 displacements for a given grating. In total, each observer made 50 observations at every displacement-grating combination.

### *Results and Discussion*

Percentages of correct identifications of the direction of motion at each of the 77 combinations of displacement and spatial frequency are shown in Fig. 1 for Observer VDL. The

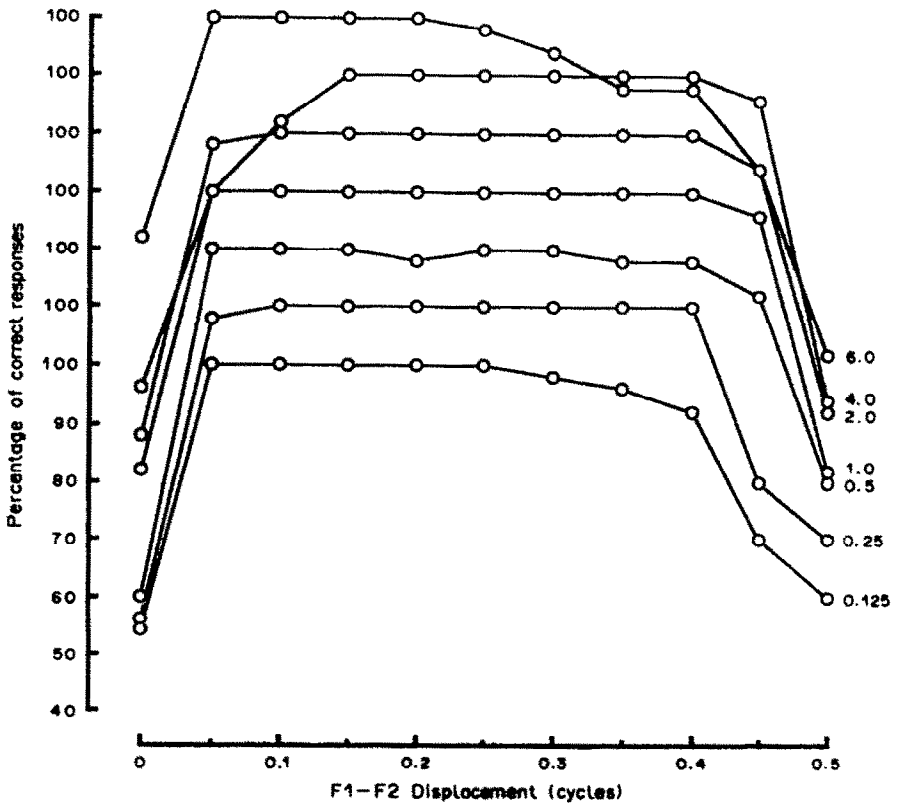


Fig. 1. Percentage of correct responses as a function of the horizontal displacement between two gratings presented sequentially. The two gratings were identical except for a phase-shift of magnitude indicated on the abscissa. Each curve represents results obtained with gratings of different spatial frequency, as indicated in the legend. The curves have been spaced vertically at intervals of 10 percentage points to avoid overlap.

other observers showed virtually identical patterns.

With some minor deviations, discussed below, the performance curves in Fig. 1 are essentially the same across spatial frequencies. From this it may be inferred that perception of direction of motion is invariant across spatial frequencies for all displacements between 0.05 and 0.45 cycles. That is, at all spatial frequencies, motion could be seen for displacements up to the Nyquist sampling limit, corresponding to half the period of the grating. The Nyquist limit represents the notional maximum value of $d_{max}$ in that displacements of 0.5 cycle yield directionally ambiguous signals, and displacements between 0.5 and 1 cycles yield reverse-motion signals.

Defined as the maximum displacement beyond which performance drops below 80% correct responses, $d_{max}$ shows little variation across spatial frequencies in Fig. 1. However, there are two slight deviations from the predominant pattern that require comment, even though neither had a significant effect on the value of $d_{max}$. The first concerns the curve for the highest spatial frequency (6 c/deg) which shows a gradual decrement at the larger displacements. This was true only for observer VDL and is attributable to a slight impairment in acuity. The second concerns the curves for spatial frequencies of 0.125 and 0.25 c/deg; the former begins to drop at a slightly smaller displacement than the latter, implying a lower value of $d_{max}$ at the lower frequency (0.43 vs 0.47 cycles). However, this is almost certainly an artefact of the total number of cycles in the image. With our equipment, a spatial frequency of 0.125 c/deg could be achieved only with gratings not exceeding two cycles per image. In such images, the largest displacements produced on overwhelming impression of centripetal or centrifugal horizontal motion which interfered with perception of uniform directional motion and produced a decrement in response accuracy.

To illustrate this point, we ran a subsidiary condition. From a distance of 57 cm, observers viewed a grating whose spatial frequency was fixed at 0.25 c/deg. The image was 8 deg wide and 2 deg high, and it contained two cycles of the grating. Figure 2 contains the results of the subsidiary condition as well as the curves for spatial frequencies of 0.125 and 0.25 c/deg from Fig. 1 for purposes of comparison. It is clear from Fig. 2 that the major determinant of the early decline in accuracy of

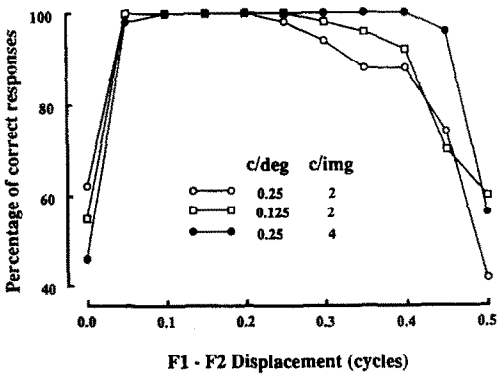


Fig. 2. Percentage of correct responses as a function of the horizontal displacement between two gratings presented sequentially. The pair of curves indicated by circles (both open and solid) were obtained with gratings of the same spatial frequency but unequal numbers of cycles per image. The pair of curves indicated by filled dots and by open squares were obtained with gratings of different spatial frequencies but equal number of cycles per image. The lower performance levels at the larger displacements can be attributed largely to the effect of number of cycles per image.

performance must be identified not with spatial frequency but with the total number of cycles per image.

Greater sensitivity for gratings containing more than two cycles per image has been taken as evidence that the corresponding receptive fields are multiple-oscillatory and narrowly tuned to the optimal frequency (De Valois, Thorell & Albrecht, 1985; Hoekstra, van der Goot, van den Brink & Bilsen, 1974). It must be stressed that the data in Fig. 2 cannot be interpreted in this fashion because contrast levels were far above threshold throughout. Rather, the effect appears to be more simply attributable to the spatiotemporal appearance of the two-cycle patterns themselves.

Irrespective of such minor variations (the difference in $d_{max}$ was only 0.04 cycles), the uniformity of the curves in Fig. 1 clearly indicates that the range of motion sensitivity was essentially the same at all spatial frequencies sampled. This pattern of results can be readily explained either in terms of *correlational* models of motion perception (e.g. Reichardt, 1961; van Santen & Sperling, 1985) or in terms of *gradient* models (e.g. Marr & Ullman, 1981). In either case, the input is provided by receptive fields linked to directionally selective units at subsequent stages. The present data befit the spatial frequency selectivity of the receptive fields postulated commonly by the two classes of models. The theoretical implications of these findings are considered in greater detail below.

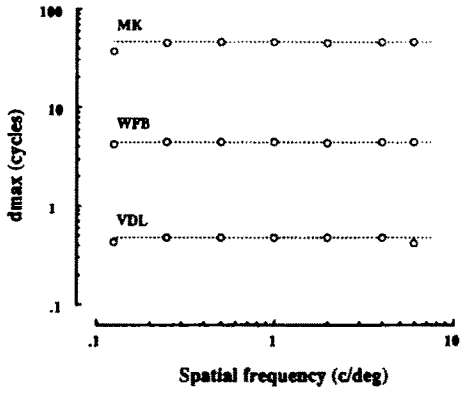


Fig. 3. Estimates of $d_{max}$, expressed in cycles, as a function of spatial frequency, separately for each observer. The lines have been spaced vertically at intervals of 1 log unit to avoid overlap.

*Estimates of $d_{max}$*

For each curve in Fig. 1 it is possible to determine a value of $d_{max}$, namely the maximum displacement beyond which performance falls below the 80% level. The estimated values are shown in Fig. 3, separately for each observer. It is clear from Fig. 3 that, expressed in cycles, the values of $d_{max}$ are invariant across spatial frequencies. This invariance agrees with the findings of Turano and Pantle (1985) and with a similar invariance that can be inferred from results reported by Nakayama and Silverman (1985, Fig. 5) and by Baker et al. (1985).

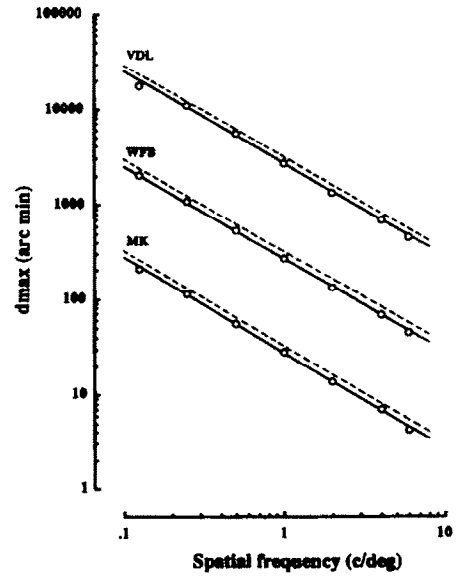


Fig. 4. Estimates of $d_{max}$, expressed in minutes of visual angle, as a function of spatial frequency, separately for each observer. In each individual plot, the segmented diagonal line indicates the displacement at which the grating is at counterphase, and no further increments in $d_{max}$ are possible. The individual plots have been spaced vertically at intervals of 1 log unit to avoid overlap.

Conventionally, $d_{max}$ is expressed in units of visual angle rather than in cycles as in Fig. 3. Conversion of the estimates of $d_{max}$ to units of visual angle yielded the data presented in Fig. 4. Also presented in Fig. 4 (as segmented lines with slope of $-1$) is the Nyquist limit representing the notional maximum value of $d_{max}$ at each spatial frequency. The lines fitted through the empirical points in Fig. 4 have the same slope as—and are tightly bound by—the Nyquist limit. That is, at all spatial frequencies, motion is seen up to a limiting value of $d_{max}$ that depends not on the ability of the visual system to see motion but on the maximum spatial displacement permitted by the structure of the stimuli.

These findings are in good agreement with those of Burr, Ross and Morrone (1986) who estimated the maximum displacement at which the perception of discontinuous motion induced by sinusoidal gratings displaced discontinuously (as in the present work) remained indistinguishable from smooth motion. In suit with the present outcome, the displacement values found by Burr et al. were bound by—but never reached—the notional limits corresponding to displacements of 0.5 cycle. It may be noted that the displacement values obtained by Burr et al. were consistently lower than the present estimates of $d_{max}$. This is so because Burr et al. estimated not absolute values of $d_{max}$—as in the present work—but relative displacement values that distinguished sampled from smooth motion.

*Direct and indirect estimates*

Estimates of $d_{max}$ obtained in the present work were based directly on the observer's performance at detecting directional motion. Indirect methods of estimating $d_{max}$ (Turano & Pantle, 1985) or $d_{opt}$ (Baker et al., 1989) have yielded results that are consonant with the present results in some—but not in other—respects. In broad agreement with the present findings, Baker et al. found that $d_{opt}$, measured in terms of duration of the movement aftereffect, and expressed in fractions of a cycle, remains invariant across spatial frequencies. However, Baker et al. failed to find a motion aftereffect (and hence a basis for estimating $d_{opt}$) at spatial frequencies higher than about 1.2 c/deg. Clearly, this limitation applies only to movement aftereffects, not to $d_{opt}$, because motion in one-dimensional gratings can certainly be seen, and $d_{max}$ (and, presumably, $d_{opt}$) can be estimated,

at spatial frequencies far exceeding 1.2 c/deg (Figs 3 and 4).

Similarly, Turano and Pantle (1985) found that, estimated in terms of duration of movement aftereffect, $d_{max}$ is invariant across spatial frequencies, and its magnitude is just shy of the Nyquist limit, just as was found in the present study. However, they also found that, with square-wave gratings, $d_{max}$ was only about 0.125 cycle because, at larger displacements, no motion aftereffect was obtained. To check on this result with a direct method, we re-ran Experiment 1 using square-wave instead of sinusoidal gratings with results identical to those in Figs 3 and 4. Again, the limitation found with square-wave gratings applies only to movement aftereffects, not to $d_{max}$.

Indirect estimates of $d_{max}$ in terms of duration of motion aftereffects depend critically on the assumption that $d_{max}$ is based on precisely the same mechanisms as the motion aftereffect. The different outcomes obtained with direct and indirect methods strongly suggest that $d_{max}$ and motion aftereffects are based on mechanisms that are, at least to some extent, different. Although more work on this issue is needed, the present evidence justifies the provisional conclusion that estimates based on motion aftereffects cannot be regarded as valid indices of $d_{max}$. The same argument applies to $d_{opt}$.

*Short-range and long-range motion*

It is notable that the values in Figs 3 and 4 are not bound by an absolute displacement limit, but extend well below and well above the range of 15–30 min arc that has often been regarded as encompassing the magnitude of $d_{max}$. Nor is there any evidence of a discontinuity in the range of values, suggestive of a transition from the "short-range" to the "long-range" motion systems. A similar observation has been made by Burr et al. (1986) who noted the lack of a dichotomy within a wide range of displacements—from a few minutes to several degrees—at which sampled motion remained indistinguishable from smooth motion. From the data in Figs 3 and 4, it can be safely inferred that the maximum displacement to which the directional motion units are sensitive is no less than just under 0.5 cycle of the lowest-frequency grating, namely, almost 4 deg of visual angle.

Taken collectively, the implications of the present and related work lead to a questioning of $d_{max}$ and of the range-based classification scheme as useful constructs for describing and explaining the perception of motion. In essence, motion can be seen over a broad range of spatial frequencies within limits that appear to be set exclusively by the structural composition of the stimuli. This pattern of events is consistent with a multi-channel conception of the visual system (e.g. Graham, 1980; Wilson & Bergen, 1979) that postulates several classes of directionally selective motion units, each tuned to a restricted range of spatial frequencies. The present data suggest that all units are equally capable of detecting motion within the range of frequencies to which they are tuned, without restrictions imposed by an absolute displacement limit ($d_{max}$) or by the confines of a short-range motion system.

## EXPERIMENT 2

Clear evidence was obtained in Experiment 1 that, expressed in cycles, $d_{max}$ is invariant across spatial frequencies. In the present experiment we set out to show that estimates of $d_{max}$ obtained with two-dimensional stimuli are subject to the same spatial constraints as the estimates obtained with one-dimensional stimuli.

Our approach can best be illustrated with reference to the work of Chang and Julesz (1983, 1985), in which the relation between $d_{max}$ and the spatial frequency contents of the stimuli was stated explicitly. As noted above, the stimuli employed by Chang and Julesz were random-dot patterns filtered through ideal band-pass filters centered at different spatial frequencies. The spectral composition of these stimuli was unambiguous in that it was defined completely by the appropriate filter. The studies showed that, down to a limit of about 4 c/deg, $d_{max}$ decreased steadily as the centre-frequency of the stimulus was increased. At all higher frequencies, $d_{max}$ remained constant at about 15 min arc.

On the face of it, the latter outcome may be taken as evidence that $d_{max}$ reflects a substantive property of the visual system. However, on the strength of the work reported thus far, the option must be entertained that the findings of Chang and Julesz might be indicative not of properties of the visual system but of limitations imposed by the structural configuration of the stimuli and by the psychophysical methodology.

Consider the appearance of the band-passed stimuli. A two-dimensional image filtered at a low centre frequency takes on the appearance of a set of large light blobs interspersed with

similarly large darker blobs [see Fig. 6(b)]. A spatial shift of a few min arc will be insufficient to place a light blob on top of an adjacent dark blob, just as a similarly small shift does not produce counterphase in a low-frequency grating. On the other hand, the same shift will suffice in the case of a grating of higher spatial frequency or in a two-dimensional image filtered at a higher centre-frequency. In general, the periodicity of a two-dimensional filtered image can be expected to function as a limiting factor in respect to $d_{max}$ in somewhat the same way as did the spatial frequency of the gratings in Experiment 1.

As would be true with gratings, if the spatial displacement exceeds the half-period of a two-dimensional image, aliasing will occur and reverse motion should be seen. If the displacement exceeds one period of the image, motion should again be seen in the correct direction but its quality should be poorer because of chance "mismatches" between corresponding blobs in two-dimensional filtered images originating from a field of random dots.

Fluctuations such as these are indeed obtained, but were not reported by Chang and Julesz (1983, 1985) because the psychophysical method they employed (a multiple-choice staircase procedure) yielded only a single value per image rather than a set of values showing variations in performance throughout the domain. As will be seen, this information is important for understanding the relation between $d_{max}$ and spatial frequency.

### *Method*

One of the authors and two female students, one of whom had served in preliminary trials with one-dimensional sinusoidal gratings, acted as observers. The equipment was the same as in Experiment 1. The process for constructing filtered images began with the computation of a 128 × 128 square matrix of dots in which each dot could be either black or white with a probability of 0.5. This image was filtered with one of eight ideal band-pass filters one octave wide. Expressed in c/deg, the widths of the filters were: 0.5–1, 1–2, 2–4, 3–6, 4–8, 4.75–9.5, 6–12 and 8–16. The corresponding centre frequencies, in c/deg, were: 0.75, 1.5, 3, 4.5, 6, 7.13, 9 and 12. In addition, the original 128 × 128 random-dot image was also used as a stimulus. The ideal filters with centre frequencies of 1.5 and 6 c/deg are illustrated in Fig. 5. The original 128 × 128 random-dot image and the filtered images with centre frequencies of 0.75, 4.5 and 9 c/deg are illustrated in Fig. 6 (a, b, c and d).

Each image contained up to 256 grey levels, and was normalized to a mean luminance of 38 cd/m$^2$ (including screen illumination of

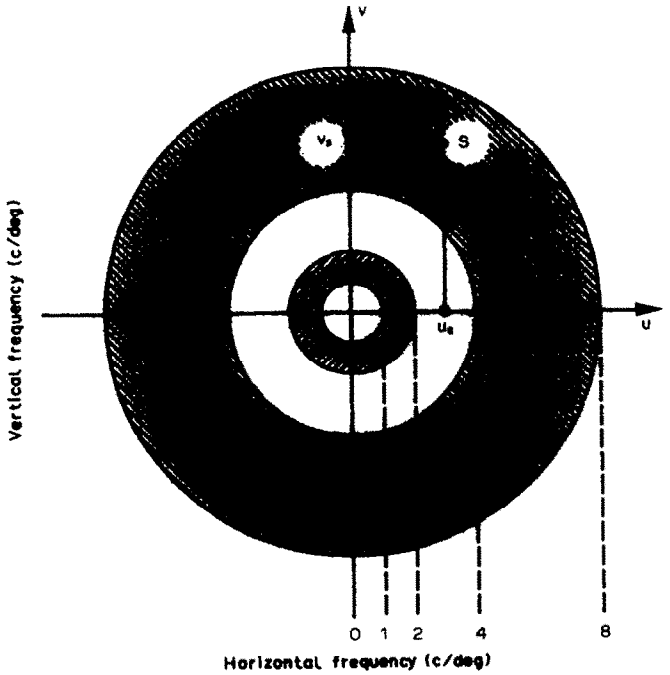


Fig. 5. Two-dimensional Fourier domain of an image, showing frequency bands of two ideal band-pass filters with bandwidth 1 octave and centre frequencies 1.5 and 6 c/deg. Also shown is a diagonal frequency component S with horizontal frequency $u_s$ and vertical frequency $v_s$.

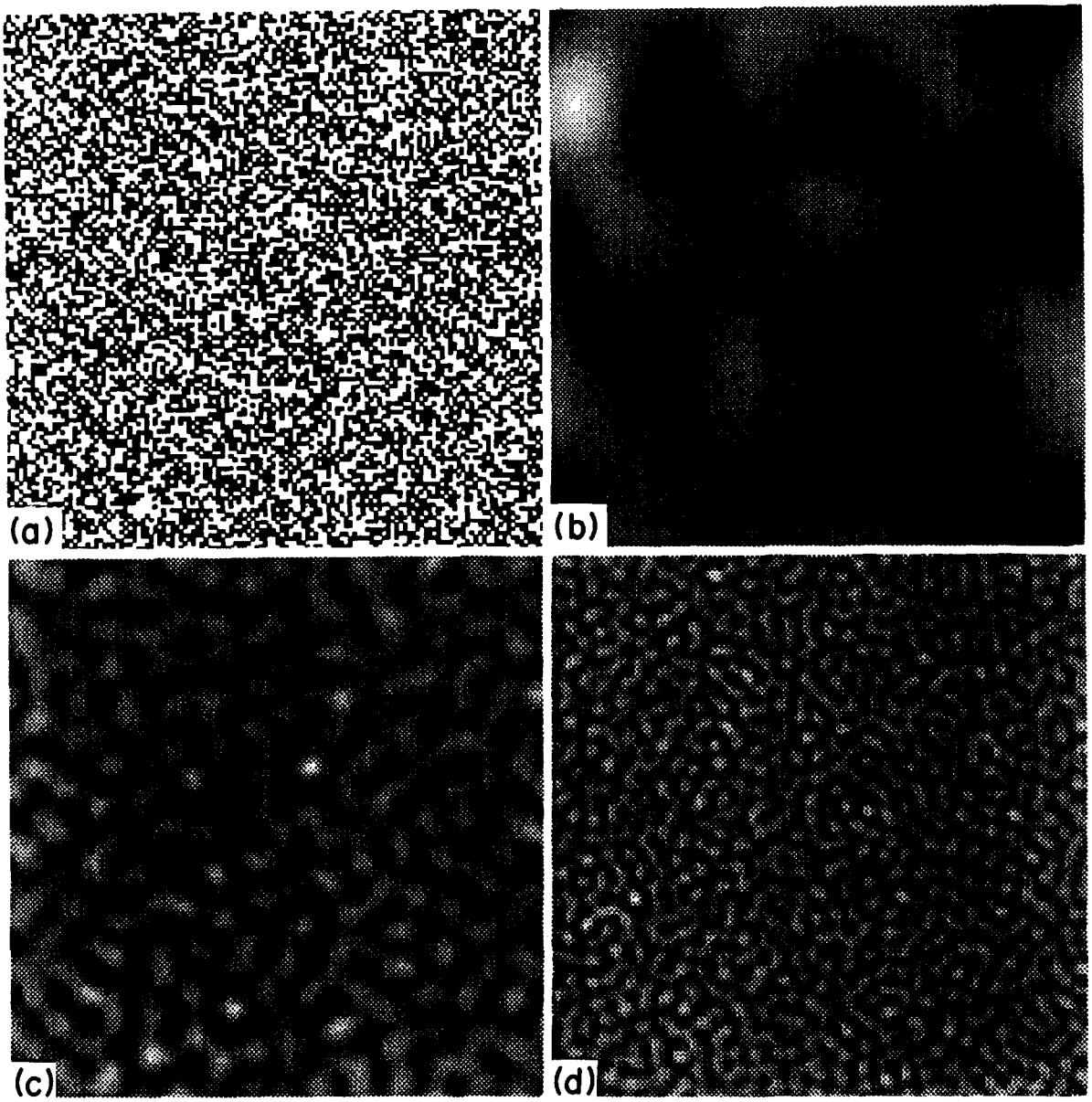


Fig. 6. Panel (a) shows an unfiltered 128 × 128 random-dot image. The remaining panels show band pass–pass filtered images as follows: (b) bandwidth: 0.5–1 c/deg, centre frequency: 0.75 c/deg. (c) bandwidth: 2–4 c/deg, centre frequency: 3 c/deg. (d) bandwidth: 4–8 c/deg, centre frequency: 6 c/deg.

10 cd/m$^2$) and peak Michelson contrast of 0.47, yielding a range of approx. 20–56 cd/m$^2$.

To avoid effects peculiar to the random structure of any one image, 20 different images were constructed for each centre frequency. On any one trial, the image to be displayed was chosen randomly from the pool of 20 images. As in Experiment 1, pairs of images were displayed sequentially, with the trailing image (F2) identical to the leading image (F1) except for a horizontal shift to the left or to the right, at ramdom. The parts of F2 that were displaced out of the viewing area by the horizontal shift were "wrapped around" to reappear at the opposite side of the image. The images were shifted by varying amounts, depending on centre frequency, as shown in the Results section. Exposure duration was 66 msec for each image, and the ISI was zero. Any one experimental session contained images of only one centre frequency, with F2 being shifted over the entire range of displacements, in random sequence. In total, each observer made 100 observations at each combination of displacement and centre frequency.

At the end of the experiment, the one-dimensional gratings from Experiment 1 were used to assess the CSF of each observer. The stimuli were pairs of one-dimensional gratings displayed sequentially for 66 msec each at an ISI of zero. The trailing grating was shifted to the left or to the right by 0.25 cycles, and the observer identified the direction of motion. The contrast of the gratings was adjusted by a dynamic tracking procedure (PEST, Taylor & Creelman, 1967) to a level that yielded approx. 80% correct responses, which was taken as the threshold for that spatial frequency. Threshold estimates were taken at spatial frequencies of 0.5, 1, 2, 3, 4, 5, 6 and 8 c/deg.

### *Results*

Percentages of correct identifications of the direction of motion for each centre frequency and for the unfiltered random-dot images are shown in Fig. 7 (a–h) for Observer TH, whose results are representative of the other observers. The functions are clearly oscillatory throughout the domain, as might be expected on the basis of the periodicity of the images discussed earlier. It should be noted that a dip in the response function below the level of 50% indicates perception of motion in the reverse direction. Values of $d_{max}$ (80% correct responses) were calculated for each observer at each centre frequency. Since the functions are clearly non-monotonic (Fig. 7), $d_{max}$ was defined as the smallest displacement at which the response function dipped below 80%. That is, the estimate of $d_{max}$ was determined by the first major dip in the response function and ignored the remaining oscillations. This definition of $d_{max}$ was adopted and justified by Cleary and Braddick (1990a). Figure 8 shows individual values of $d_{max}$ separately for each image. The significance of the lines in the individual plots is explained below.

The phenomenological appearance of the displays is worth noting. Small displacements always yielded uniform motion that was smooth, fluid and indistinguishable from real motion. At larger displacements, coherent one-directional motion was not seen in the whole image. Instead, some parts of the image were seen to move horizontally, while other parts appeared to move in directions other than horizontal. On occasions, particularly at the higher centre frequencies, some parts of the image appeared to move diagonally, while others appeared to rotate or to move in other diagonal directions. In these cases, observers attempted to detect the direction of the prevailing horizontal component of motion, and responded accordingly. Local, Brownian motion was also seen, particularly in the case of the unfiltered images.

Normalized CSFs are shown in Fig. 9, separately for each observer.

## DISCUSSION

### *Is there a floor level of "$d_{max}$"?*

Chang and Julesz (1985) studied variations in $d_{max}$ as a function of spatial frequency with stimuli similar to ours. Their results and ours are very similar, but only up to a centre frequency of about 4 c/deg, that is, in both experiments, $d_{max}$ declined as frequency was increased. However, beyond a frequency of 4 c/deg, Chang and Julesz's estimates reached a floor at a $d_{max}$ value of about 15 min arc and remained at that level for all higher frequencies. The two parts of the $d_{max}$ function (sloping up to 4 c/deg, and flat at higher frequencies) were regarded as arising from two different visual processes known as "noncooperative" and "cooperative", respectively (cf. Marr & Poggio, 1976).

No evidence of a floor level was obtained in our experiment, either at 15 min arc or at

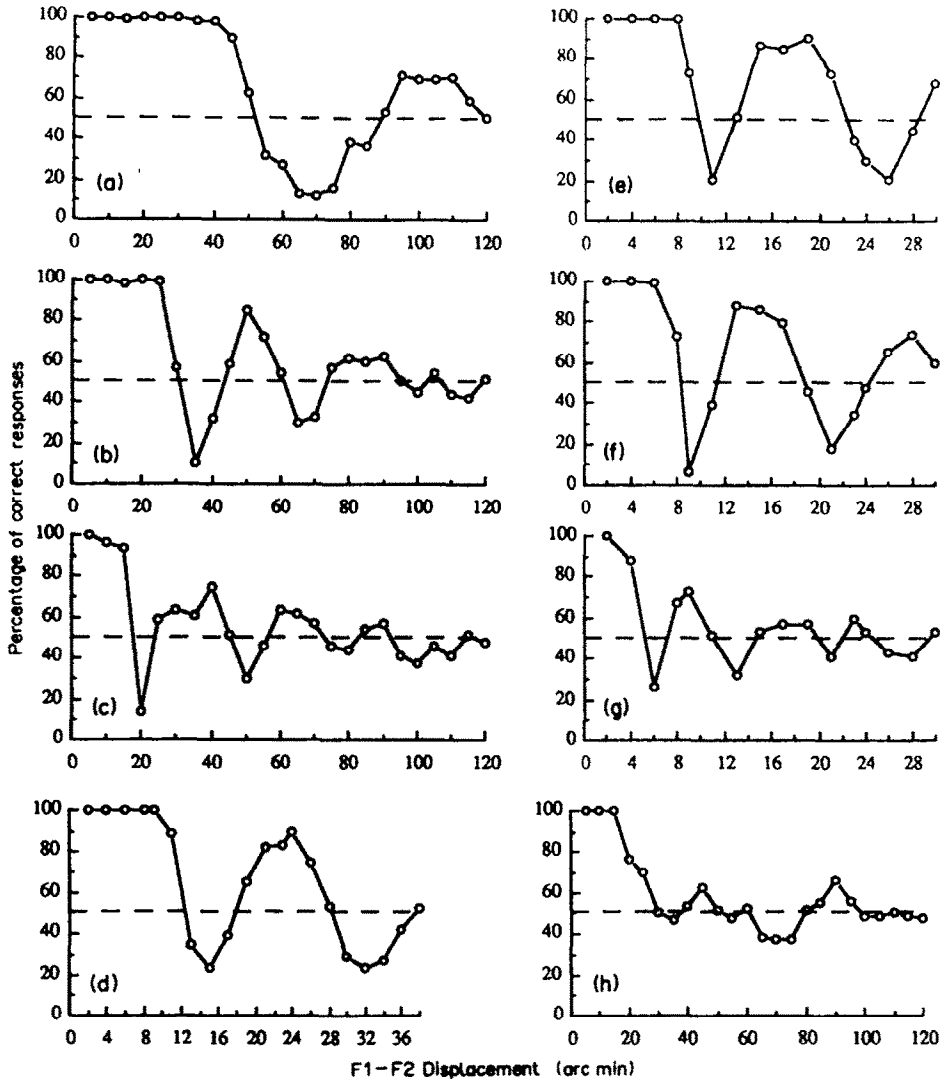


Fig. 7. Percentage of correct responses as a function of the horizontal displacement between two band-pass filtered images in Experiment 2. The data are for Observer TH. Scores >50% indicate motion in the correct direction; scores <50% indicate motion in the reverse direction. Each panel shows the results obtained with a different band-pass filter with bandwidth 1 octave and centre frequency as follows (in c/deg): (a) 0.75, (b) 1.5, (c) 3, (d) 4.5, (e) 6, (f) 7.13, (g) 12. Panel (h) shows the results for the unfiltered images. Note different scales on the abscissae.

any other value of $d_{max}$. As can be seen in Fig. 8, $d_{max}$ decreased continuously to values much lower than 15 min arc as spatial frequency was increased over a range comparable to that of Chang and Julesz (1985).

What is the source of the discrepancy between the two studies? There are several procedural and substantive differences that may underlie the discrepant outcomes. For example, control over the intensity of the stimuli was achieved differently in the two experiments, and there were 256 grey levels in our study against eight levels in the study of Chang and Julesz. But it is difficult to see how factors such as these might bring about a floor level in one study and not in the other. Rather, we believe that the discrepancy is related to the psychophysical methods employed in the two studies.

Estimates of $d_{max}$ were obtained by Chang and Julesz with a staircase procedure. Although efficient, staircase procedures yield a single threshold value instead of a function, as in the present study. Moreover, to be valid, staircase procedures depend critically on the assumption that the dependent variable (percentage of correct responses, in this case) varies monotonically throughout the domain. As can be seen in Fig. 7, percentage of correct responses does not vary monotonically with magnitude of displacement of band-pass filtered stimuli. In view of the

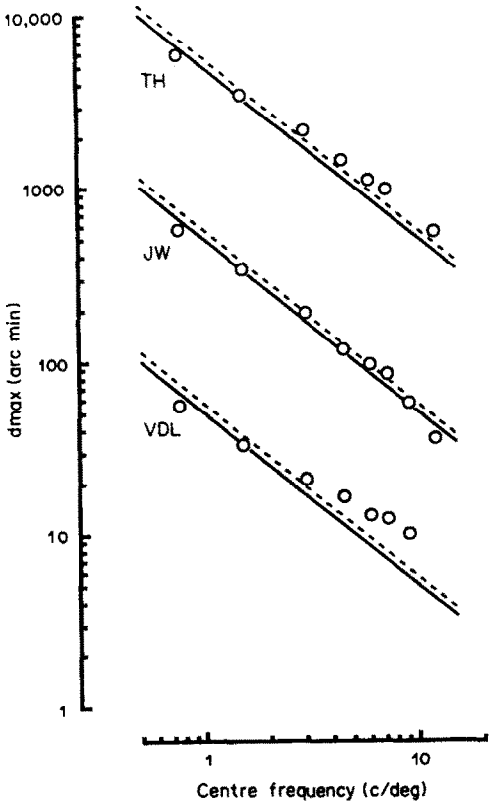


Fig. 8. Values of $d_{max}$ obtained in Experiment 2 at each centre frequency of the band-pass filtered images, separately for each observer. In each individual plot, the continuous lines represent the average values of $d_{max}$ that had been obtained in Experiment 1. The segmented lines indicate the limiting values of $d_{max}$ as defined by the displacement that yields counterphase in a grating with spatial frequency corresponding to the lowest frequency of the filter indicated on the abscissa. The individual plots have been spaced vertically insteps of 1 log unit to avoid overlap.

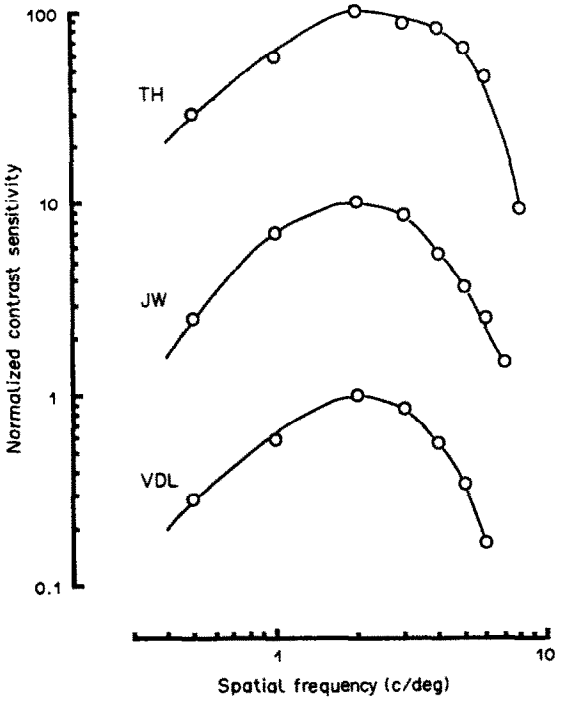


Fig. 9. Normalized contrast sensitivity functions for each observer. The curves have been spaced vertically in steps of 1 log unit to avoid overlap.

pronounced oscillations of the functions in Fig. 7, staircase procedures would clearly be in violation of the fundamental assumption of monotonicity; hence any estimate of $d_{max}$ obtained by such procedures must be regarded as invalid. As already noted, Cleary and Braddick (1990a) have offered a similar explanation for the nonlinear results of Chang and Julesz (1985).

We hasten to add that, at the time of publication of Chang and Julesz's research, there was no pertinent evidence suggestive of nonmonotonicity. Indeed, the assumption of monotonicity was entirely plausible, and the use of staircase procedures quite justifiable. Nevertheless, in light of the data in Fig. 7, it is not difficult to see how a staircase procedure might yield the pattern of results reported by Chang and Julesz (1985). At the lower frequencies (e.g. Fig. 7a–c), the oscillations are relatively slow, hence there is a sufficient range of displacements—both preceding and following the first major dip of the response function—to enable the staircase procedure to converge on a value within the ambit of the first dip. At the higher frequencies, however, the oscillations are faster, and convergence would be forced towards later peaks of the response function, beyond which the oscillations are dampened below the threshold criterion level. Of course, the precise outcome would depend not only on the oscillatory shape of the response function but also on such parameters of the staircase procedure as the size of the step, the change criterion, and the stopping criterion. But these issues are obviously moot: given the pronounced nonmonotonicity of the functions in Fig. 7, valid outcomes cannot be expected from a staircase procedure with these stimuli.

### *Upper limits of $d_{max}$*

All things being equal, it should be possible to account for variations in $d_{max}$ obtained with two-dimensional filtered stimuli in Experiment 2 in terms of the one-dimensional stimuli employed in Experiment 1.

Comparisons between Experiments 1 and 2 can be drawn by considering the range of spatial frequencies of the stimuli in Experiment 2 in

relation to the results obtained with gratings of the corresponding frequencies in Experiment 1. As a first approximation, it can be supposed that the value of $d_{max}$ obtained for any given filtered image in Experiment 2 could not exceed the $d_{max}$ obtained in Experiment 1 for a grating whose spatial frequency matched the lowest frequency contained in the filtered image. The reasoning is as follows: suppose that we employed a compound grating made up of three superimposed gratings whose frequencies were 1, 2 and 3 c/deg, respectively. As was shown in Experiment 1, motion can be seen for displacements up to 0.5 cycle regardless of spatial frequency. In the compound image, the limit would correspond to 30, 15 and 10 min arc for gratings of 1, 2 and 3 c/deg, respectively. On this basis, the largest possible value of $d_{max}$ for the compound image would be just under 30 min arc, namely, just under 0.5 cycle of the lowest frequency contained in the image. Estimates based on the two higher frequencies should yield smaller angular values of $d_{max}$. This is not to say that the addition of higher frequencies has no effect on $d_{max}$. Indeed, it has been shown by Cleary and Braddick (1990b) that the addition of higher frequencies can reduce (but never increase) the value of $d_{max}$. On these premises, it seems reasonable to uphold the working hypothesis that the value of $d_{max}$ for any filtered image in Experiment 2 should not exceed the value of $d_{max}$ obtained in Experiment 1 for the grating corresponding to the lowest spatial frequency component of the filtered image.

A direct comparison between the two experiments is made in Fig. 8. In addition to the data obtained in Experiment 2, each individual record in Fig. 8 contains two lines. The solid lines represent the values of $d_{max}$ obtained in Experiment 1 for gratings with frequency equal to the lowest frequency component of the corresponding images in Experiment 2. The segmented lines indicate the displacement corresponding to 0.5 cycle of the lowest spatial frequency in the image. For example, for an image with centre frequency of 9 c/deg, $d_{max}$ for observer VDL was 8 min arc. The lowest frequency component of that image was 6 c/deg, with a corresponding 0.5-cycle displacement limit of 5 min arc. This is shown by the segmented line. Similarly, the average value of $d_{max}$ obtained with a grating of 6 c/deg in Experiment 1 was about 4.2 min arc. This is indicated by the solid line.

Contrary to expectations based on a one-dimensional analysis, most values of $d_{max}$ in Fig. 8 are above the solid lines and thus exceed $d_{max}$ for the lowest frequency component as estimated in Experiment 1. Moreover, many of the $d_{max}$ values go beyond the 0.5-cycle displacement limit for the lowest spatial frequency defined by the filter. The seemingly paradoxical implication of these findings is that the observers responded on the basis of spatial frequencies that were lower than the lower bound of the relevant filter. Below, we show that the band-pass filtered images did contain such lower frequencies.

In the foregoing, perception of horizontal motion of two-dimensional filtered images was analyzed exclusively in terms of horizontal spatial-frequency components. For example, given a filter-band between 4 and 8 c/deg, the frequency components included in the analysis were only those that satisfied the conditions

$$4 \leqslant u \leqslant 8, \quad v = 0$$

where $u$ and $v$ are the horizontal and vertical frequencies, as shown in Fig. 5. Totally ignored in this analysis were all spatial frequency components with nonzero vertical frequency. However, the filters used to construct the images for Experiment 2 passed frequency components at all orientations within the relevant frequency band. As illustrated in Fig. 5, this included the frequency component S with a horizontal frequency $u_s$ much lower than 4 c/deg, the nominal lower bound of the filter. The results illustrated in Fig. 8 (where many values of $d_{max}$ are seen to exceed the notional maximum) suggest that the observers responded to the presence of such diagonal components with horizontal frequency lower than the notional minimum.

Although plausible, these notions are only qualitative. Whether the role of diagonal components is either necessary or sufficient in explaining the excessive values of $d_{max}$ cannot be decided without a quantitative model. One such model is described below.

*Models of motion perception: one-dimensional analysis*

It has been noted in Experiment 1 that, in sinusoidal gratings, motion is seen in the correct direction for displacements of up to 0.5 cycle, and in the reverse direction for displacements between 0.5 and 1 cycle. A similar—though not identical—rule should be expected to hold for band-pass filtered images. This follows from

the periodicities contained in band-pass filtered images, as was noted in the introduction of Experiment 2.

Whether correct motion, reverse motion, or no motion is seen in a sequential display will depend on the similarity of the two images across the displacement. Similarity can be defined as the cross-correlation between the images. In the case of a single translating image, this is equivalent to the autocorrelation which can be defined as

$$R_{II}(x, y) = \sum_{x'} \sum_{y'} I(x', y') I(x + x', y + y') \quad (1)$$

where $I$ is the image. In principle, image autocorrelation could be used as an explanatory basis for the results illustrated in Fig. 7. However, despite its intuitive appeal, this approach is not successful when quantitative predictions are attempted, as is shown below.

Autocorrelation functions (weighted in the manner described below) were fitted to the results of Experiment 2. As shown in Fig. 10, the functions fitted the data remarkably well throughout the range of displacements. It should be noted that the results with unfiltered images [Fig. 7(h)] did not exhibit the cyclic oscillations obtained with bandpass filtered images. In all probability, this is related to the flat power spectrum on the unfiltered images.

Two crucial points must be noted in respect to the fitted functions. First, the energies for all spatial frequencies within the range of the relevant filter was weighted according to the CSF of the observer (see Fig. 9). More specifically, instead of using the autocorrelation of the image as in equation (1), we computed the autocorrelation of the CSF-weighted images as follows:

$$R_{SS}(x, y) = \sum_{x'} \sum_{y'} S(x', y') S(x + x', y + y') \quad (2)$$

where

$$S(x, y) = F^{-1}[\tilde{I}(u, v) CSF(\sqrt{u^2 + v^2})]$$

$$\tilde{I}(u, v) = F[I(x, y)]$$

and where $F[\cdot]$ and $F^{-1}[\cdot]$ denote the Fourier and inverse Fourier transforms, respectively.

This procedure equalized the effective energy at all frequencies and produced a better fit to the data. Second, each autocorrelation function was shifted horizontally to obtain the best visual fit

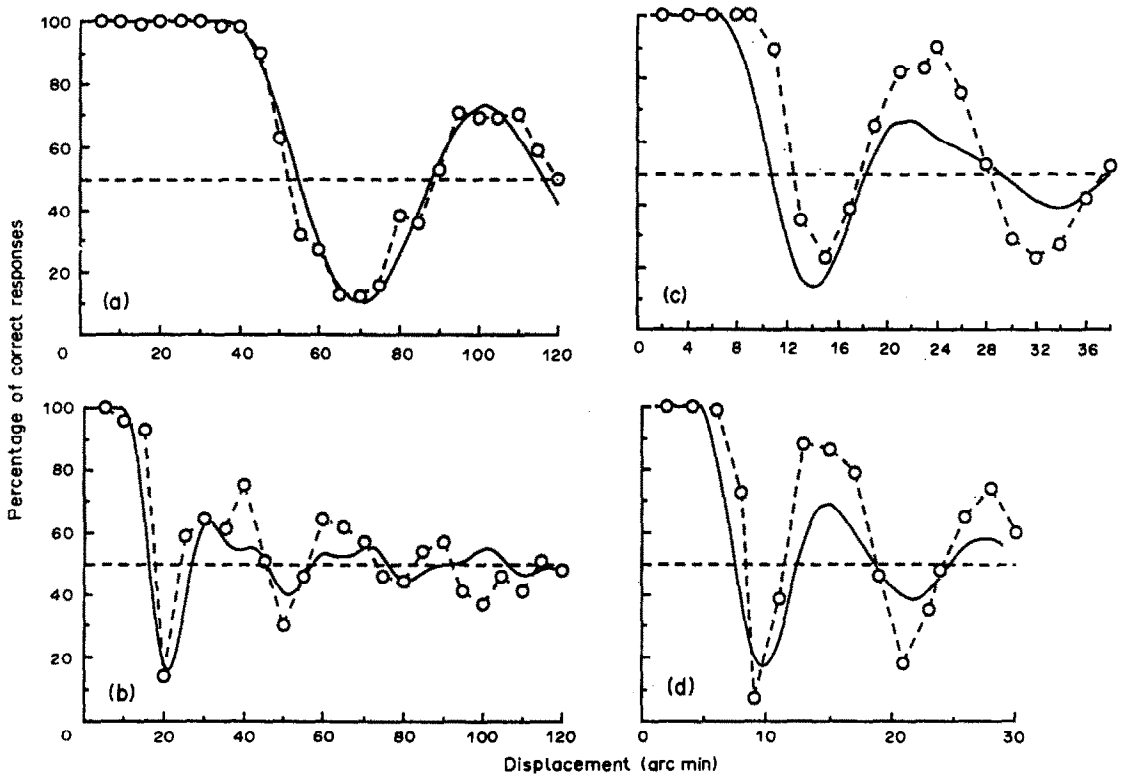


Fig. 10. The empirical data in panels a, b, c, and d correspond to the data in panels a, c, d, and f in Fig. 7, respectively. Within each panel, the continuous curve is the 1-dimensional autocorrelation function for the band-pass filter that had been used to produce the images that yielded the data in that panel. In the autocorrelation functions, positive values correspond to motion in the correct direction, and negative values to motion in the reverse direction. Note different scales on the abscissae.

to the data. There was no simple rule that appeared to govern the amount of shift required. Unlike the weighting of the frequency components in terms of the CSF, the horizontal shift of the autocorrelation function is totally arbitrary, and requires justification. The shift and its implications for models of motion perception are discussed below.

A rationale for using autocorrelation for fitting the data could be sought within the frameworks of the two major classes of models of motion perception—correlation and gradient models—in that both classes are closely tied to the autocorrelation of the images. Correlational models (e.g. Reichardt, 1961) or energy models (e.g. Watson & Ahumada, 1985; Adelson & Bergen, 1985) postulate directionally selective motion units that compute a cross-correlation—or an equivalent index (cf. van Santen & Sperling, 1985)—over some spatio-temporal separation. In the case of translation, such as employed in Experiment 2, the response of these units is therefore determined by the autocorrelation of the image. On the other hand, in gradient models (e.g. Marr & Ullman, 1981), the combined output of all motion units depends on the distribution of zero-crossings in the $\nabla^2$G-filtered image. In turn, the distribution of zero-crossings is related to the autocorrelation of the image (e.g. Leadbetter, 1969).

We first consider an explanation in terms of correlational (or energy) models. In these models, the basic motion detector consists of two spatially separate receptor units converging on a correlational unit that combines the two signals by multiplying them together. The output from one of the receptor units is delayed so as to enable detection of directional motion through concurrent activation of the correlational unit.

An explanation based on autocorrelation must account for two things: the horizontal shift and the quasi-periodic oscillations of the response functions. We begin by examining how a single motion detector with a given separation between sensors ($\Delta$) might respond to different displacements of the image. Performance of this unit is perfect for image displacements equivalent to $\Delta$, and varies according to the autocorrelation of the image for displacements larger or smaller than $\Delta$. This is illustrated in Fig. 11(a) and 11(b) for filters of bandwidth equal to 1 octave and 0.5 octave, respectively. With reference to empirical results, performance of this unit can account for the observer's responses to image displacements larger than $\Delta$, as shown by the fitted data in Fig. 10. However, image displacements smaller than $\Delta$ create a problem. Instead of following an autocorrelation function as in Fig. 11(a), (b), the observers' performance at image displacements smaller than $\Delta$ is errorless. Clearly, the output from a single motion detector provides an adequate account of the quasi-periodic performance at image displacements greater than $\Delta$, but not of the errorless performance at displacements smaller than $\Delta$.

To account for errorless performance at small displacements, one could consider a population of motion units with input sensors covering a broad range of separation. Cleary and Braddick (1990a) have proposed a population of such units with input separations covering the range from $d_{min}$ (i.e. the minimum visible image displacement) to $d_{max}$. Cleary and Braddick also assumed that perception of directional motion is determined by the unit with the strongest positive peak in the range $d_{min}$ to $d_{max}$ (for further details, see Cleary & Braddick, 1990a, particularly Fig. 6).

We implemented this model with the parameters suggested by Cleary and Braddick (1990a). Namely, we examined the combined performance of a population of units whose input separations varied continuously in the range from 0.2 cycle ($d_{min}$) to 1 cycle ($d_{max}$) of the centre frequency of the filter. Figure 11(c) and (d) illustrate the model's performance for bandpass filtered kinematograms with bandwidths of 1 octave (as used in the present work) and 0.5 octave (as used by Cleary and Braddick). In Fig 11(c) and (d), the probability of motion detection is assumed to be proportional to the strength of the corresponding autocorrelation.

To evaluate the model's performance, Fig. 11(c) should be compared with any of the response functions in Fig. 7. Upon comparison, two things are immediately obvious: first, the pronounced oscillations of the curves in Fig. 7 are virtually absent in Fig. 11(c). Second, a sizable reverse-motion effect is predicted for image displacements slightly smaller than $d_{min}$. An intuitive understanding of this effect may be gained by reference to Cleary and Braddick's (1990a) Fig. 6. An image displacement of 0.1 cycles to the right does not produce a motion signal to the right because it leaves the main peak of the autocorrelation function below $d_{min}$, but it is an effective reverse-motion signal because it brings the first major side-peak of the function within the sensitivity range that signals

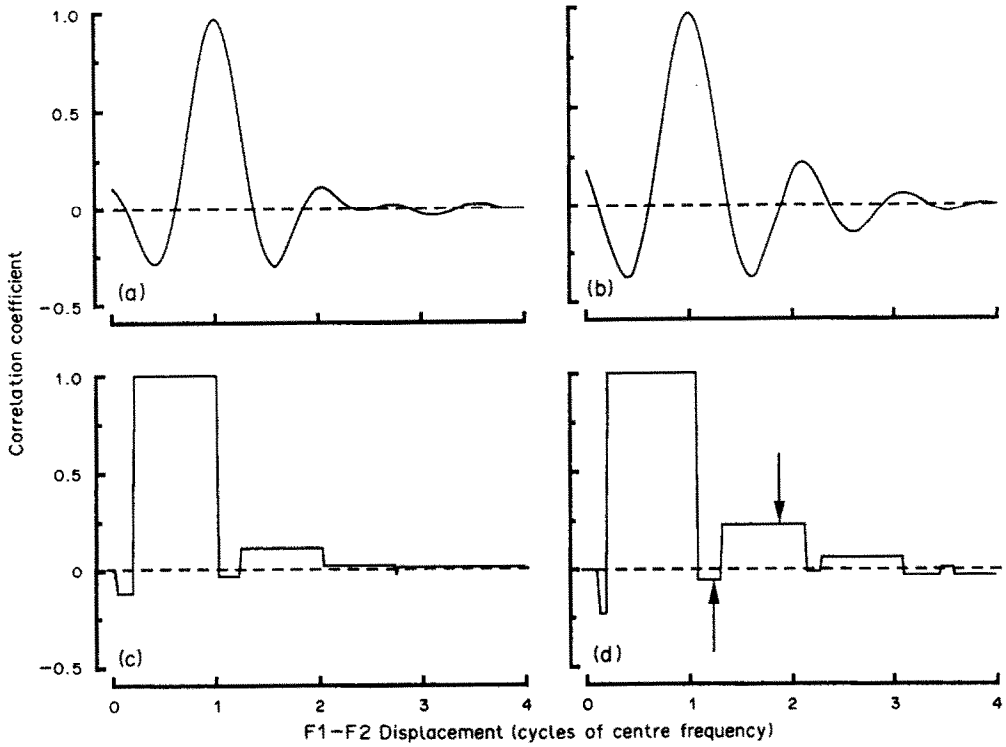


Fig. 11. Predicted percentage of correct responses as a function of horizontal displacement of bandpass filtered kinematograms. In panels a and c the filter's bandwidth was 1 octave; in panels b and d it was 0.5 octave. Panels a and b represent predictions based on a single Reichardt-type detector with input separation equal to $\Delta = 1$ cycle of centre frequency. Panels c and d represent predictions based on the qualitative model suggested by Cleary and Braddick (1990a); that is, predictions are based on a population of Reichardt-type detectors with input separations ranging from 0.2 to 1 cycle of centre frequency. In panel d, the two arrows indicate the displacement illustrated in Cleary and Braddick's (1990a) Fig. 6(a), (b).

motion to the left. Needless to say, such reverse-motion effect has not been observed in empirical investigations. In sum, this model cannot account either for the errorless performance over displacements smaller than $d_{max}$ or for the oscillatory performance at displacements greater than $d_{max}$.

Performance of the model, which is less than satisfactory with bandpass filtered stimuli, is even more wanting with sinusoidal gratings. It is assumed in the model that the separation between the sensors of a motion unit is such that correct motion signals can be produced by image displacements up to about 1 cycle of centre frequency. This is clearly violated by sinusoidal grating stimuli which, by their structure, must produce reverse-motion signals for displacements above 0.5 cycle. Although this limitation has been noted by Cleary and Braddick (1990a), it clearly requires theoretical justification if the model is to perform consistently across the stimulus domain.

A note of caution is required. Failure of the present implementation of the model cannot be taken as indicating complete failure of the qualitative suggestions made by Cleary and Braddick (1990a). It may well be that some variant of the present implementation may produce improved predictions. Indeed, some improvement was obtained in a simulation that replaced the position of the peaks of the autocorrelation function with the positions of just the highest levels of autocorrelation as the criterion for directional motion detection. Some improvement was also obtained by implementing a population of elaborated Reichardt detectors of the form proposed by van Santen and Sperling (1985). But the problem of predicting direction of motion with gratings remained in every case. We believe that at least some of the difficulties arise from the fact that this type of analysis is limited to just one dimension. Notably ignored are the contributions to motion perception made by the non-horizontal spatial frequency components of the two-dimensional filtered images. We now proceed to describe a simulation—briefly mentioned in the foregoing—designed to overcome this limitation.

*Models of motion perception: two-dimensional analysis*

The major objective of the present simulation was to predict horizontal directional motion by taking into account signals indicating motion not only horizontally but in any other direction. For this purpose, we implemented a general model capable of detecting motion in arbitrary directions. The model produces an estimated flow field representing all motion signals produced by a two-frame kinematogram as used in the present experiments. To predict horizontal motion, we estimated the probability of seeing left–right motion from all motion signals contained in the flow field.

Specifically, we adapted the model described by Horn and Schunk (1981) to analyze directional motion of the stimuli in Experiment 2. Horn and Schunk's is a "gradient" model closely related to that of Marr and Ullman (1981). In terms of this model, directional motion signals are generated at every point of the image, in the direction of the local brightness gradient. The motion signals are then integrated into a smooth velocity field using an iterative technique. A more comprehensive description of the model and its implementation is provided elsewhere (the present Appendix; Horn, 1986; Horn & Schunck, 1981).

Predicted and obtained outcomes are compared in Fig. 12 for filtered images with centre frequency of 1.5 c/deg. The empirical data are those of observers TH and VDL in Experiment 2. Except for a small underestimation, discussed below, the model provides an excellent fit to the quasi-periodic response function. More important, the theoretical fit of the errorless performance at small image displacements is not achieved by an arbitrary horizontal shift of the fitted function (cf. Fig. 10); rather, it is a predicted outcome of the model. In addition, the model performs equally well with bandpass filtered images as with sinusoidal gratings, which were problematic for the account suggested by Cleary and Braddick (1990a).

As illustrated in Fig. 12, the fitted function underestimates $d_{max}$ by about 5 min arc. It must be noted that the fit was achieved with the simplest possible implementation of the model. All spatial frequency components in the image were given equal weight, although differential weighting would be required with unfiltered images. By the omission of such weightings, the model's predictions were determined entirely by the structural properties of the images, as distinct from being determined by characteristics of the visual system. The model's fit could be improved by taking into account properties of the visual system known to affect motion perception. For example, as presently implemented, the model assumes a receptor surface of uniform sensitivity to motion. However, the images covered non-homogeneous retinal areas extending well into the parafovea where $d_{max}$ is known to increase (e.g. Baker & Braddick, 1985b). The

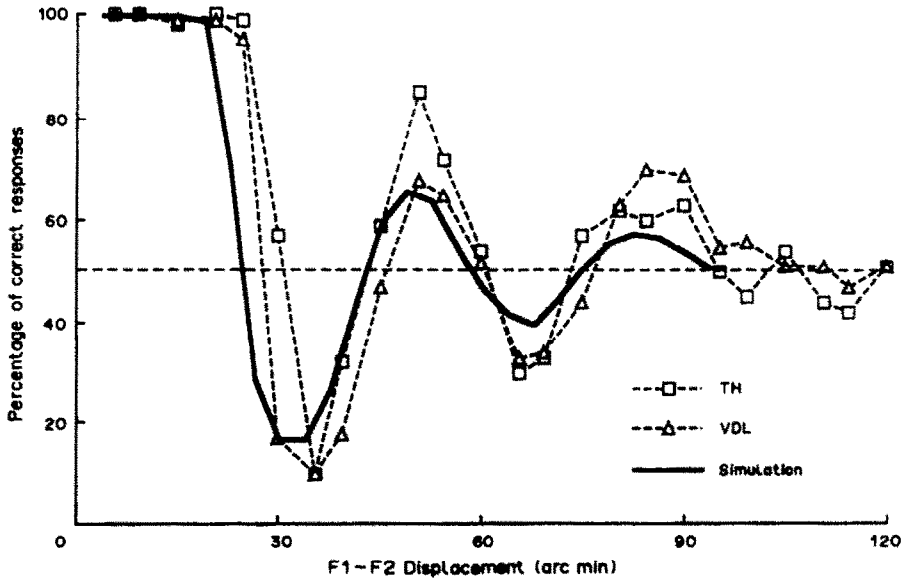


Fig. 12. The empirical data correspond to the data in Fig. 7(b) for observers TH and VDL. The continuous curve represents the prediction based on the modified Horn and Schunck (1981) model of motion perception.

model's performance could be easily improved by introducing appropriate non-uniformities to the receptor surface and suitable CSF weightings to the frequency contents of the images. In the present simulation, we have deliberately avoided such special procedures in order to show the basic strength of the model in its simplest form.

It should be emphasized that the predictions illustrated in Fig. 12 do not hinge on specific properties of Horn and Schunck's model. We have obtained virtually identical results with a very different model (Uras, Girosi, Verri & Torre, 1988; Reichardt & Schlögl, 1988). The crucial requirement is that the responses of neighbouring motion detectors be suitably combined into an estimated flow field from which perception of directional motion may be derived. By contrast with the present approach, the one-dimensional analyses described earlier did not include any form of motion integration. It is to such lack of integration that we ascribe the relatively poor performance of these models.

Despite appearances, the predicted curve in Fig. 12 is not an autocorrelation function. The curve simply represents percentage of correct responses, and requires no justification for its use, as was the case with the autocorrelation function in the one-dimensional approach.

## CONCLUDING COMMENTS

Do random-dot images, one-dimensional sinusoidal gratings, and two-dimensional filtered images address common mechanisms of motion perception within the visual system? This question has been asked—and answered affirmatively—by others (e.g. Nakayama & Silverman, 1985). The hypothesis of communality of mechanisms is supported, *inter alia*, by the results of dichoptic motion studies in which the two stimuli in a motion pair are presented separately, one to each eye. Under conditions usually regarded as yielding "long-range" motion, motion is seen with dichoptic stimulation. By contrast, neither sinusoidal gratings nor random-dot images produce the appearance of motion when displayed dichoptically (Braddick, 1974; Green & Blake, 1981). This has been taken as evidence consistent with the notion that perception of motion in sinusoidal gratings and in random-dot images is mediated by common mechanisms. The weight of the evidence (e.g. Baker et al., 1989; Chang & Julesz, 1983, 1985; Green & Blake, 1981; Nakayama & Silverman, 1985; Turano & Pantle, 1985) favours the conclusion that the same visual processes are indeed being studied with these stimuli.

What conclusions can be drawn regarding $d_{max}$ and the distinction between short- and long-range motion systems? An answer can best be given in an historical context. Although the term $d_{max}$ was not explicitly used by Braddick in his original work, we use it here to denote the absolute displacement limit of motion perception, as first proposed by Braddick (1974). Initially, $d_{max}$ was assigned a value of approx. 15 min arc which defined the absolute spatial limit of a system involved in the perception of directional motion. The system was characterized by the epithet "short-range" to distinguish it from the "long-range" system which was said to respond to displacements greater than $d_{max}$.

Before long, it was realized that $d_{max}$ was not an absolute but a relative limit. Its magnitude was found to change as a function of several variables, notably retinal eccentricity (e.g. Baker & Braddick, 1985b; Bischof & Groner, 1985). Acknowledging these findings, and wishing to retain $d_{max}$ as a basis for defining the short-range system, Baker and Braddick (1985b) redefined it as a valid measure that remains invariant at any given eccentricity. In this usage, $d_{max}$ ceased to be an invariant property of a unitary short-range motion system. Instead, it assumed a range of values, each of which was an invariant property of a separate short-range system (or sub-system) defined in terms of retinal eccentricity. The weight of the evidence now indicates that $d_{max}$ does not remain invariant even within a given eccentricity. Rather, expressed in units of visual angle, it varies continuously as a function of the spatial frequency of the stimuli.

Should $d_{max}$ be further redefined to cater to this source of variability? We think not. Systematic variations in $d_{max}$ can be regarded simply as representing corresponding differences in the sizes of receptive fields that provide the inputs to directionally-selective motion units at all retinal eccentricities (cf. Marr & Ullman, 1981; van Santen & Sperling, 1985; Wilson & Bergen, 1979). The results of the present and of related experiments (e.g. Baker et al., 1989; Nakayama & Silverman, 1985; Turano & Pantle, 1985) are explained naturally in terms of frequency-tuned mechanisms of the type described above, capable of mediating perception of motion over broad ranges of displacements and spatial frequencies. No construct akin to $d_{max}$ is required to explain the experimental outcomes. Indeed,

such a construct may well have misleading implications because of superfluous meaning accumulated in past usage. At best, $d_{max}$ may be used as a purely descriptive short-hand for denoting variations in displacement limit as a function of variables such as eccentricity (e.g. Baker & Braddick, 1985b) or level of dark adaptation (e.g. Dawson & Di Lollo, 1989).

Perhaps the most important function of $d_{max}$ in earlier theorizing was as a basis for the definition and classification of short-range and long-range motion systems. It is now clear that $d_{max}$ can no longer be regarded as a sound base for such bipartite classification. In fact, such classification must be questioned both on empirical and on conceptual grounds. Empirically, no discontinuity suggestive of a separation between the two systems has been found in the present work or in related experiments (e.g. Baker et al., 1989; Burr et al., 1986). Conceptually, a short-range motion system necessitates the postulation of spatially-limited mechanisms such as proposed by Cleary and Braddick (1990a). The problems encountered by this and similar approaches have been discussed above. Additional problems have been discussed by Cavanagh and Mather (1989) who favour a unitary motion system that receives input from different types of receptors. In the same vein, a unitary motion system is compellingly suggested by the present experiments and simulations. The present analyses also show that the empirical data can be successfully predicted without recourse to spatially-limited mechanisms.

*Acknowledgements*—This work was supported by Operating Grants from the Natural Sciences and Engineering Research Council of Canada No. OGP38521 to WFB and No. A6592 to VDL.

## REFERENCES

Adelson, E. H. & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America, 2A*, 284–299.

Anstis, S. M. (1980). The perception of apparent movement. *Philosophical Transactions of the Royal Society of London, 290B*, 153–168.

Anstis, S. M. (1986). Motion perception in the frontal parallel plane: Sensory aspects. In Boff, K. R., Kaufman, L. & Thomas, J. P. (Eds.) *Handbook of perception and human performance: Sensory processes and perception* (Vol. 1, Chap. 16). New York: Wiley.

Baker, C. L. Jr & Braddick, O. (1982). The basis of area and dot number effects in random dot motion perception. *Vision Research, 22*, 1253–1260.

Baker, C. L. Jr & Braddick, O. (1983). The effect of eccentricity on random dot motion perception. *Investigative Ophthalmology and Visual Science, Supplement, 24*, 277.

Baker, C. L. Jr & Braddick, O. (1985a). Temporal properties of the short-range process in apparent motion. *Perception, 14*, 181–192.

Baker, C. L. Jr & Braddick, O. (1985b). Eccentricity-dependent scaling of the limits for short-range apparent motion perception. *Vision Research, 25*, 803–812.

Baker, C. L. Jr, Baydala, A. & Zeitouni, N. (1989). Optimal displacement in apparent motion. *Vision Research, 29*, 849–859.

Barlow, H. B. & Hill, R. M. (1963). Selective sensitivity to direction of motion in ganglion cells in the rabbit's retina. *Science, 139*, 412–414.

Bischof, W. F. & Groner, M. (1985). Beyond the displacement limit: An analysis of short-range processes in apparent motion. *Vision Research, 25*, 839–847.

Braddick, O. (1974). A short-range process in apparent motion. *Vision Research, 14*, 519–527.

Braddick, O. (1980). Low-level and high-level processes in apparent motion. *Philosophical Transactions of the Royal Society of London, 290B*, 137–151.

Burr, D. C., Ross, J. & Morrone, C. (1986). Smooth and sampled motion. *Vision Research, 26*, 643–652.

Campbell, F. W. & Maffei, L. (1981). The influence of spatial frequency and contrast on the perception of moving patterns. *Vision Research, 21*, 713–721.

Cavanagh, P. & Mather, G. (1989). Motion: The long and the short of it. *Spatial Vision, 4*, 103–129.

Cavanagh, P., Boeglin, J. & Favreau, O. E. (1985). Perception of motion in equiluminous kinematograms. *Perception, 14*, 151–162.

Chang, J. J. & Julesz, B. (1983). Displacement limits for spatial frequency filtered random-dot cinematograms in apparent motion. *Vision Research, 23*, 1379–2385.

Chang, J. J. & Julesz, B. (1985). Cooperative and non-cooperative processes of apparent movement of random-dot cinematograms. *Spatial Vision, 1*, 39–45.

Cleary, R. & Braddick, O. J. (1990a). Direction discrimination for band-pass filtered random dot kinematograms. *Vision Research, 30*, 303–316.

Cleary, R. & Braddick, O. J. (1990b). Masking of low frequency information in short-range apparent motion. *Vision Research, 30*, 317–327.

Dawson, M. & Di Lollo, V. (1990). Effects of adapting luminance and stimulus contrast on the temporal and spatial limits of short-range motion. *Vision Research, 30*, 415–429.

De Valois, R. L., Thorell, L. G. & Albrecht, D. G. (1985). Periodicity of striate-cortex-cell receptive fields. *Journal of the Optical Society of America, 2A*, 1115–1123.

Finley, G. (1985). A high-speed point plotter for vision research. *Vision Research, 25*, 1993–1997.

Graham, N. (1980). Spatial frequency channels in human vision: Detecting edges without edge detectors. In Harris, C. S. (Ed.), *Visual coding and adaptability* (pp. 215–262). Hillsdale, New Jersey: Erlbaum.

Green, H. & Blake, R (1981). Phase effects in monoptic and dichoptic temporal integration: Flicker and motion detection. *Vision Research, 21*, 365–372.

Hoekstra, J., van der Goot, D. P. A., van den Brink, G. & Bilsen, F. A. (1974). The influence of the number of cycles upon the visual contrast threshold for spatial sine wave patterns. *Vision Research, 14*, 365–368.

Horn, B. K. P. (1986). *Robot vision*. Cambridge, Mass.: MIT Press.

Horn, B. K. P. & Schunck, B. G. (1981). Determining optical flow. *Artificial Intelligence, 17*, 185–203.

Lappin, J. S. & Bell, H. H. (1(976). The detection of coherence in moving random-dot patterns. *Vision Research, 16,* 161–168.

Leadbetter, M. R. (1969). On the distributions of times between events in a stationary stream of events. *Journal of the Royal Statistical Society, B, 31,* 295–302.

Marr, D. & Poggio, T. (1976). Cooperative computation of stereo disparity. *Science, 194,* 283–287.

Marr, D. & Ullman, S. (1981) Directional selectivity and its use in early visual processing. *Proceedings of the Royal Society of London, 211B,* 151–180.

Morgan, M. J. & Ward, R. (1980). Conditions for motion flow in dynamic visual noise. *Vision Research, 20,* 431–435.

Nakayama, K. & Silverman, G. H. (1984). Temporal and spatial characteristics of the upper displacement limit for motion in random dots. *Vision Research, 24,* 293–299.

Nakayama, K. & Silverman, G. H. (1985). Detection and discrimination of sinusoidal grating displacements. *Journal of the Optical Society of America, 2A,* 267–274.

Reichardt, W. (1961). Autocorrelation, a principle for the evaluation of sensory information. In Rosenblith, W. A. (Ed.), *Sensory communication.* Cambridge, Mass.: MIT Press.

Reichardt, W. & Schlögl, R. W. (1988). A two dimensional field theory for motion computation. *Biological Cybernetics, 60,* 23–35.

van Santen, J. P. H. & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America, 2A,* 300–320.

Taylor, M. M. & Creelman, C. D. (1967). PEST: Efficiency estimates on probability functions. *Journal of the Acoustic. Society of America, 41,* 782–787.

Turano, K. & Pantle, A. (1985). Discontinuity limits for the generation of visual motion aftereffects with sine- and square-wave gratings. *Journal of the Optical Society of America, 2A,* 260–266.

Uras, S., Girosi, F., Verri, A. & Torre, V. (1988). A computational approach to motion perception. *Biological Cybernetics, 60,* 79–87.

Watson, A. B. & Ahumada, A. J. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America, 2A,* 322–341.

Wilson, H. R. & Bergen, J. R. (1979). A four mechanism model for threshold spatial vision. *Vision Research, 19,* 19–32.

## APPENDIX

Outlined below are the procedural details of the simulation reported in Fig. 12. This is not intended as a complete description of Horn and Schunck's (1981) model. Rather, we present a broad outline of the model, with full details of those parts where our implementation differs from the original.

### *Motion Detection*

Let $E(x, y, t)$ denote the brightness at the image point $(x, y)$ at time $t$ and let $\mathbf{v}(x, y)$ denote the flow field. The brightness of image points is assumed to remain constant over time, that is

$$\frac{dE(x, y, t)}{dt} = 0. \tag{A1}$$

By expanding equation (A1) we can derive the following constraint on the flow field $\mathbf{v}(x, y)$:

$$\nabla E(x, y, t) \cdot \mathbf{v}(x, y, t) + E_t(x, y, t) = 0 \tag{A2}$$

where $\nabla E(x, y, t) = [E_x(x, y, t), E_y(x, y, t)]$ is the brightness gradient and $E_x$, $E_y$ and $E_t$ denote partial derivatives with respect to $x$, $y$, and $t$. The component of the flow field in the direction of the brightness gradient is given by

$$-E_t(x, y)[E_x^2(x, y, t) + E_y^2(x, y, t)]^{-1/2}. \tag{A3}$$

The spatial derivatives $E_x$ and $E_y$ were computed exactly to avoid effects of approximation errors. That is, given the Fourier transform $\tilde{E}(u,v)$ of $E(x, y)$, $E_x$ was computed directly using the derivative theorem of the Fourier transform

$$E_x(x, y) = F^{-1}[2\pi i u \tilde{E}(u, v)] \tag{A4}$$

where $F^{-1}[\cdot]$ denotes the inverse Fourier transform. The temporal derivative $E_t$ was approximated by

$$E_t(x, y) \approx E(x, y, t_2) - E(x, y, t_1). \tag{A5}$$

For the two-frame displays with a horizontal displacement $d$, $E(x, y, t_2)$ is given by $E(x, y, t_2) = E(x - d, y, t_1)$.

### *Motion Integration*

Equation (A3) provides only one component of the flow field, namely the component in the direction of the brightness gradient (this is known as the aperture problem). To recover the full flow field, Horn and Schunck introduce the additional constraint that the resulting flow field $\mathbf{v}(x, y)$ should be smooth almost everywhere. Recovering the flow field can then be posed as a problem of minimizing the function

$$\iint [\nabla E(x, y) \cdot \mathbf{v}(x, y) + E_t(x, y)]^2 + \lambda [\nabla \mathbf{v}(x, y)\, \nabla \mathbf{v}(x, y)]^2 \, dx\, dy \tag{A6}$$

where the first term is from equation (A2), the second term measures the smoothness of the flow field and $\lambda$ gives the relative weight of the two terms. Equation (A6) can be approximated in discrete form and solved using an iterative scheme with the following updating formulas [the terms $(x, y)$ have been left out to simplify the formulas].

$$v_1^{n+1} = \bar{v}_1^n - \frac{E_x \bar{v}_1^n + E_y \bar{v}_2^n + E_t}{1 + \lambda(E_x^2 + E_y^2)} E_x$$

$$v_2^{n+1} = \bar{v}_2^n - \frac{E_x \bar{v}_1^n + E_y \bar{v}_2^n + E_t}{1 + \lambda(E_x^2 + E_y^2)} E_y \tag{A7}$$

where $n$ is the iteration number and $\mathbf{v} = (v_1, v_2)$. $\bar{v}_1$ and $\bar{v}_2$ are local averages of $v_1$ and $v_2$, respectively. More precisely,

$$\bar{v}_1^n(x, y) = \frac{\sum_{(x', y') \in N} \omega(x', y') v_1^n(x', y')}{\sum_{(x', y') \in N} \omega(x', y')} \tag{A8}$$

where $N$ is some small neighbourhood of $(x, y)$ and

$$\omega(x', y') = |\nabla E(x', y')|^k \tag{A9}$$

with $k = 2$. The weighting factor $\omega$ ensures that motion estimates at image positions with a small brightness gradient, where the estimate in equation (A3) is unreliable, contribute only little to the local flow field estimate. Note also that for $k \to \infty$ only motion estimates at zero-crossings are used (Marr & Ullman, 1979).

In our simulation the flow field was estimated using 30 iterations of the updating scheme (A7).

### *Prediction of Directional Motion*

For large image displacements, the flow field estimate $\mathbf{v}(x, y)$ obtained for a given kinematogram is by no means uniform. This agrees with the appearance of patchy directional motion seen in Experiment 2. To derive an estimate of the probability of "seeing" motion to the right for a given flow field estimate, the following formula is used:

$$p_{right} = \frac{1}{N^2} \sum_{x=1}^{N} \sum_{y=1}^{N} r(x, y) \tag{A10}$$

where $N^2$ is the number of pixels in the image, and $v_1(x, y)$ is the horizontal component of the flow vector, as before. $r(x, y)$ is defined as

$$r(x, y) = \begin{cases} 1 & v_1(x, y) > 0 \quad \text{and} \quad |\mathbf{v}(x, y)| > T \\ 0 & \text{otherwise} \end{cases} \tag{A11}$$

that is, $r(x, y) = 1$ if the horizontal component of $\mathbf{v}$ is pointing to the right and the magnitude of the flow vector exceed the threshold $T$, which is defined as 1/2 pixel in our implementation.

Finally, the predictions shown in Fig. 12 were obtained by applying the procedure presented in this Appendix to 50 different kinematograms at all displacements and averaging the estimates shown in equation (A10).