Pergamon

0042-6989(95)00315-0

Vision Res., Vol. 36, No. 16, pp. 2489–2495, 1996

Printed in Great Britain
0042-6989/96 $15.00 + 0.00

How Similar must the Fourier Spectra of the Frames of a Random-dot Kinematogram be to Support Motion Perception?

TIM LEDGEWAY*†

Received 21 June 1995; in final form 30 October 1995

Direction-discrimination performance was measured for two-frame random-dot kinematograms in which one or both frames were spatial frequency filtered with a one octave band-pass filter and the centre frequency of this filter was varied in the range 0.75–9 c/deg independently for each frame. When both frames were filtered so that they contained common (overlapping) spatial frequencies direction discrimination was extremely good but it deteriorated rapidly as the degree of spectral overlap between the two frames decreased. These results are consistent with previous findings that suggest that the mechanisms that mediate the initial stages of motion detection are narrowly tuned for spatial frequency and cannot combine information conveyed at disparate frequencies in order to compute an unambiguous estimate of the direction of local motion. However, when only one of the frames was band-pass filtered and the other was unfiltered (broadband), the correct direction of stimulus motion could be discriminated reliably for a broad range of filter centre frequencies. Performance was best when the centre frequency of the filtered frame was at medium spatial frequencies and tended to deteriorate as the centre frequency approached either extreme of the spatial frequency range examined. This basic pattern of results may be attributed to the visual system's differential sensitivity to the Fourier components present in the unfiltered frame. Copyright © 1996 Elsevier Science Ltd.

Motion perception Random-dot kinematograms Band-pass filter

INTRODUCTION

Random-dot kinematograms (RDKs) have been used widely in vision research as a tool for studying the mechanisms involved in human motion perception. In such a stimulus, a dense array of random dots is displaced over both space and time and the ability to discriminate the correct direction of coherent dot motion is measured as a function of some stimulus parameter (e.g., spatial and temporal displacement, dot size, dot density, contrast, spatial frequency content, etc.) in order to elucidate the nature of the underlying processes mediating motion perception. As a consequence of their dense spatial structure it is presumed that RDKs isolate low-level motion mechanisms (e.g., Braddick, 1974, 1980) and minimise high-level motion-detecting strategies involving the tracking and matching of image features (e.g., Anstis, 1980; Cavanagh, 1991, 1992; Ullman, 1979) over space and time.

The majority of previous studies that have examined direction discrimination using RDKs have measured the limits of performance in order to make inferences concerning the spatial and temporal properties of motion-detecting mechanisms in human vision. In particular, the maximum spatial displacement ($D_{\max}$) over which the direction of coherent dot motion can be reliably discriminated (Braddick, 1974) has received a great deal of attention. Although many studies of $D_{\max}$ have been conducted, surprisingly little research has focused on the question of how similar the individual frames (images) of a RDK must be, in terms of their spatial frequency content, in order to support coherent motion detection. This issue is interesting because it may provide useful information concerning the underlying selectivity of the motion sensitive mechanisms which govern performance in two-dimensional (2D) stimuli. Several studies are relevant to this issue and these are discussed below.

Morgan & Mather (1994) measured $D_{\max}$ for two-frame RDKs in which each frame was independently low-pass spatial frequency filtered and the cut-off point of the filter applied to one of the frames was fixed while the cut-off for the other frame was varied. Although they reported that $D_{\max}$ decreased as the spatial frequency

---

*Department of Psychology, University of London, Royal Holloway College, Egham, Surrey, TW20 0EX, U.K.
Address all correspondence to the Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford, OX1 3UD, U.K.

content of one of the frames increased (i.e., as the filter cut-off point increased) and that motion perception collapsed entirely when the cut-off frequencies of the two frames differed by a factor of about 4, it is not known how similar the Fourier spectra of the two frames of a RDK must be in order to support motion perception. Recently Yang & Blake (1994) used a masking technique involving the spatial superposition of band-pass (0.4 octaves wide) filtered RDKs with similar band-pass filtered random noise masks to estimate the bandwidths of the mechanisms underlying the perception of coherent motion. The ratio of mask to signal contrasts was varied in order to measure thresholds for discriminating the coherent RDK plus mask stimulus from incoherent noise. The resulting masking functions were relatively broad with a full-width tuning of approximately 2.4 octaves at the 3 dB roll-off point. In contrast, several previous studies that have attempted to measure the spatial frequency bandwidths of motion-detecting mechanisms in the human visual system using one-dimensional (1D) masking paradigms (e.g., Anderson & Burr, 1985; Burr *et al.*, 1986) suggest that tuning is generally much narrower than the results of Yang and Blake would suggest. For example, Anderson & Burr (1985), using sinusoidal gratings, estimated the full-width, half-height bandwidth of motion-detectors to be approximately 1 octave for mechanisms sensitive to frequencies above 3 c/deg. Narrow tuning of motion sensors for spatial frequency is also consistent with the results of studies (e.g., Watson, 1986) that have measured the likelihood of perceiving apparent motion between adjacent, sinusoidal stimuli that differ in terms of their spatial frequencies. For example, Green (1986) found that observers failed to perceive motion between neighbouring stimuli when their spatial frequencies differed by 0.5 to 1 octaves. Furthermore, current computational models of local motion detection that assume the existence of narrowly tuned spatial frequency filters in the visual system (e.g., Adelson & Bergen, 1985; Van Santen & Sperling, 1985; Watson & Ahumada, 1985) can successfully accommodate a wide range of motion phenomena. Although filtered RDKs may not be particularly well suited for making estimates of the spatial frequency tuning of motion-detecting mechanisms because of their relatively broad Fourier spectra, for example when compared with sinusoidal gratings, the results obtained with gratings and RDKs are clearly discrepant. One possible explanation of the discrepancy between these results is that the broadly tuned filter implicated by the results of Yang & Blake (1994) does not reflect the initial stages of motion detection where individual local dot motions are extracted, but rather a subsequent processing stage involving the integration of local motion signals at different spatial scales (and hence at different spatial frequencies). There is some evidence in support of this suggestion. For example, Smith (1992) found that plaid patterns composed of two sinusoidal gratings could support the perception of coherent (2D) pattern motion even when the spatial frequencies of the component gratings differed by 3–4 octaves (especially at low drift speeds and high contrasts). These results are consistent with the view that the visual system pools information about motion across spatial frequencies in order to compute the overall direction of image motion.

The aim of the present experiment was to clarify this issue by measuring direction-discrimination performance for two-frame RDKs in which each frame was independently band-pass filtered (or unfiltered) and the centre frequencies of the filters were systematically varied. If the direction of motion can be reliably perceived even when the two-frames of a RDK contain widely disparate spatial frequencies, this would imply that the mechanisms that initially detect the local motions within the stimulus are broadly tuned for spatial frequency. If, however, unambiguous motion is only perceived when the two frames of the motion sequence contain overlapping spatial frequencies, this would suggest that any broad tuning of motion phenomena reflects integration of motion signals derived from different spatial frequencies.

METHOD

Observers

Two observers participated in the experiment and both had normal or corrected to normal acuity. Observer T.L. was the author and observer T.F. was a paid volunteer who was unaware of the purpose of the experiment.

Apparatus and stimuli

All motion stimuli were composed of 8-bit images generated by an Apple Macintosh LC475 computer and were precomputed and stored on disk. The stimuli were displayed on an Apple monochrome monitor with a refresh rate of 66.7 Hz. The monitor was carefully gamma-corrected using a look-up-table so that luminance was a linear function of the digital representation of the images. The mean luminance of the display was approximately 48 cd/m^2 and was viewed binocularly at a distance of 1.36 m. All stimuli were presented within a 2D Gaussian contrast envelope (standard deviation 0.33 deg truncated at ± 1 deg) at the centre of the display, in order to avoid artefacts due to abrupt luminance edges.

Three types of RDK stimulus were employed and these are described in detail below:

1. *Unfiltered RDKs*. These were composed of two frames, each containing circular black dots of diameter 7.5 arcmin superimposed on a white background. The dot density was 75 dots/deg^2 and the Michelson contrast of the dots, prior to spatial windowing, was 95% (measured using a spot photometer). The dots in the second frame were all displaced horizontally (either leftward or rightward) relative to the dots in the first frame by a constant step size of 1.875 arcmin. The duration of each frame was 90 msec and there was no inter-stimulus-interval (ISI) between the presentation of the two frames. These particular spatial and temporal parameters were chosen on the basis of

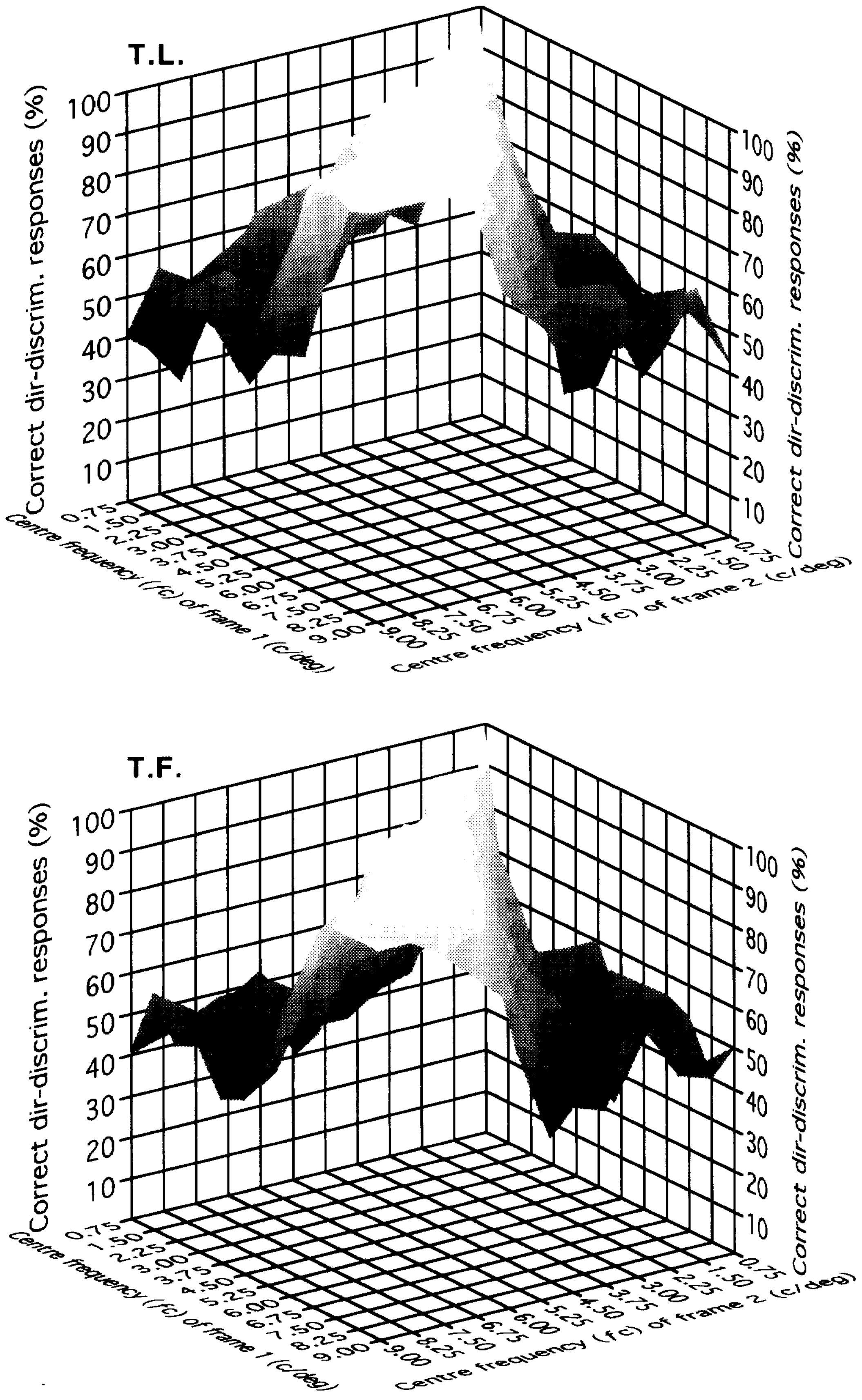

FIGURE 1. Results for two observers depicted as three-dimensional surfaces showing the percentage of correct direction-discrimination responses as a function of both the centre spatial frequency (f_c) of the individual frames in each band-pass filtered RDK and the order in which the two frames were presented. The brightness of the shading of each surface represents the percentage of correct direction-discrimination responses obtained for each condition, such that white indicates 100% correct performance and progressively darker shades indicate progressively lower levels of performance. The bandwidth of the isotropic filter applied to each frame was 1 octave and the spatial displacement between frames was 1.875 arcmin. The mean standard errors for observers T.L. and T.F. were 7.05% (range 0–18.1%) and 7.26% (range 0–16.23%), respectively.

pilot studies which revealed that under these conditions observers could reliably identify the correct direction of displacement for RDKs in which both frames contained the same range of spatial frequencies (i.e., D_{max} was never exceeded —see Results and discussion).

2. *Filtered RDKs*. These were identical to the unfiltered RDKs described in (1) above, with the exception that each frame was independently band-pass filtered using conventional Fourier techniques (e.g., Chang & Julesz, 1983, 1985; Cleary & Braddick, 1990a; Yang & Blake, 1991, 1994). Each frame was band-pass filtered with an isotropic filter with a bandwidth of 1 octave. The centre frequency (f_c), defined as half the sum of the highest and lowest frequencies passed by the filter, was systematically varied for each frame of the RDK. Twelve centre frequencies (f_c) were employed: 0.75, 1.5, 2.25, 3.0, 3.75, 4.5, 5.25, 6.0, 6.75, 7.5, 8.25 and 9.0 c/deg and all possible combinations (i.e., of f_c and presentation order) of the filtered pairs of frames were used to construct a total of 144 RDKs of this type. The step size of 1.875 arcmin between frames [see (1) above] ensured that whenever the two frames of any filtered RDK contained common spatial frequencies these were always displaced by less than 0.5 cycles of their spatial periods in order to prevent any possible confounding effects of aliasing. In line with previous studies that have examined motion perception for filtered RDKs (e.g., Cleary & Braddick, 1990a,b; Morgan, 1992; Morgan & Mather, 1994) the filtered frames were scaled in the present experiment to have the same peak-to-trough contrast as the unfiltered images in order to cover the whole available range (0–255). Pilot studies revealed that this manipulation was not crucial since it made little difference to the results whether the filtered stimuli were re-normalised in contrast or not.
3. *Mixed RDKs*. These were composed of one frame that was unfiltered [see (1) above] and one that was band-pass filtered [see (2) above]. By varying the order of presentation of the frames and the value of f_c for the filtered frame a total of 24 RDKs of this type were constructed.

For all three types of motion stimulus the order in which the two frames of any RDK were presented was independent of the direction of motion, which could be either leftward or rightward with equal probability. Thus, for all RDKs direction judgements could not be based on the temporal order of the individual frames.

PROCEDURE

Each trial consisted of the presentation a single RDK (composed of two frames) which was then followed by the presentation of a homogeneous blank field (luminance 48 cd/m^2) and a tone to indicate to the observer that a response was required. The observer's task was to indicate, using one of two response buttons, the perceived direction of motion of the RDK (either leftward or rightward). There was then a 3 sec delay before the next trial commenced, and so on. Within each run of trials the observer was presented with a total of 1690 RDK stimuli (10 unfiltered RDKs, 10 × 144 filtered RDKs and 10 × 24 mixed RDKs) and the order in which any RDK was presented within any one run of trials was randomised. Each observer completed a total of four runs.

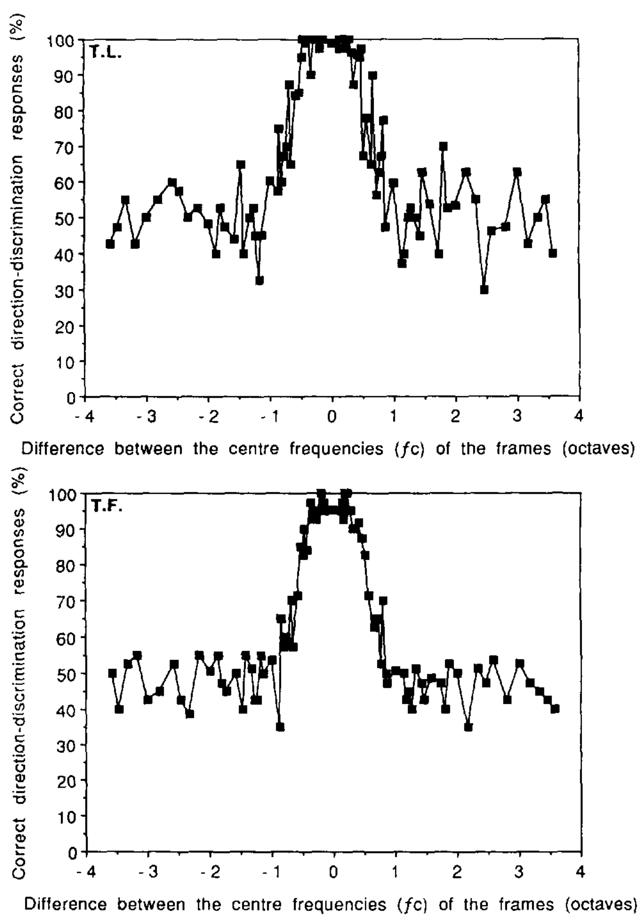

FIGURE 2. Data from Fig. 1 replotted as the percentage of correct direction-discrimination responses as a function of the difference (in octaves) between the centre spatial frequencies (f_c) of the two band-pass filtered frames in the filtered RDKs.

RESULTS AND DISCUSSION

The resulting data were pooled across the two directions of motion and the percentage of correct direction-discrimination responses was calculated separately for each type of RDK and for each observer.

1. *Unfiltered RDKs*. When both of the frames in the RDKs were unfiltered (spatially broadband) direction-discrimination performance for the two observers was perfect. These results demonstrate clearly that the spatial displacements of the individual dots between the two frames of the RDKs never exceeded D_{max} for these stimuli.
2. *Filtered RDKs*. In Fig. 1 the results for each observer are plotted as a three-dimensional (3D) surface showing the percentage of correct direction-

discrimination responses as a function of both the centre spatial frequency (f_c) of the individual frames in each filtered RDK and the order in which the two frames were presented. The surfaces are very similar for the two observers and it is evident that each is basically composed of an elevated central plateau that rapidly descends on both sides to form flanking regions that are essentially flat. The characteristic shape of these surfaces indicates that when the two frames of a RDK contained similar spatial frequencies (the elongated plateau region of each surface) observers were readily able to discriminate the correct direction of stimulus motion. However, direction-discrimination performance fell rapidly to chance levels for both observers when the individual filtered frames of the RDKs contained progressively disparate spatial frequencies (the flat regions surrounding the central plateau). This basic pattern of results was found across the entire spatial frequency range examined. It is apparent that the order in which the two frames of each RDK was presented was unimportant, as evidenced by the symmetry of each surface about a vertical plane that bisects the intersections of the abscissae. It is clear, therefore, that correct direction-discrimination judgements were only possible when the two frames of each RDK contained common spatial frequencies (i.e., had similar Fourier spectra). This is readily apparent from Fig. 2 which shows the percentage of correct direction-discrimination responses plotted as a function of the difference (in octaves) between the centre spatial frequencies (f_c) of the two frames in the filtered RDKs. When the centre spatial frequencies (f_c) of the individual frames differed by more than about ±1 octave (the bandwidth of the band-pass filters applied to the individual frames) the perceived direction of motion became ambiguous and performance for the two observers fell to chance levels (50% correct). Thus, the results clearly support the view that the mechanisms that mediate the initial stages of motion detection in RDKs cannot usefully combine information conveyed at widely different spatial scales in order to yield unambiguous estimates of the direction of image motion. That is, it appears that at least for the purposes of motion extraction they only integrate (combine) signals over a limited range of spatial frequencies. This is not meant to imply that broadly tuned mechanisms may not exist in the visual pathways mediating motion perception, but rather that such mechanisms must operate either before motion is extracted (c.f. Morgan, 1992; Morgan & Fahle, 1992; Morgan & Mather, 1994) at each of several spatial scales or at later stages that presumably integrate motion signals derived from different spatial frequencies (Smith, 1992; Yang & Blake, 1994).

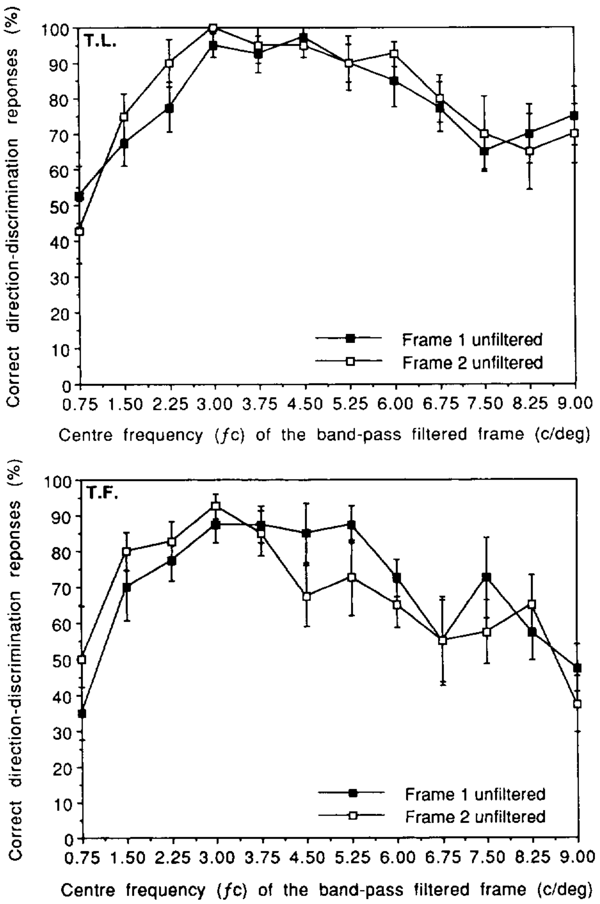

FIGURE 3. Direction-discrimination performance for two observers for the mixed RDK motion sequences in which one frame of each pair was unfiltered (spatially broadband) and the other band-pass filtered with a 1 octave isotropic filter. The centre spatial frequency (f_c) of the filtered frame is shown on the abscissa. The filled squares depict the results obtained when the unfiltered frame in each RDK was presented before the filtered frame and the open squares represent performance when order of presentation of the pairs of frames was reversed.

3. *Mixed RDKs*. Figure 3 shows the results for two observers when each RDK contained one frame that was unfiltered (spatially broadband) and one frame that was band-pass filtered. The percentage of correct direction-discrimination responses are plotted as function of the centre spatial frequency (f_c) of the filtered frame. The results obtained when the unfiltered frame was presented before the band-pass filtered frame are plotted separately, in Fig. 3, from those obtained with the opposite presentation order. For both conditions the two observers could discriminate readily the correct direction of motion for the majority of RDKs employed, although performance is noticeably worse for RDKs in which the centre frequency (f_c) of the filtered frame approached either extreme of the spatial frequency range examined (especially the low frequencies where performance is close to chance for both observers). From Fig. 3 it is also apparent from the similar shapes of the two functions that direction-discrimination performance was little affected by the order in which the two frames of the mixed

> RDKs were presented. Morgan & Mather (1994) have also recently shown that for RDKs composed of one unfiltered (spatially broadband) frame and one low-pass filtered frame the temporal order in which the two frames are presented is unimportant for most combinations of frames. However, when the cut-off point of the low-pass filter was very low (i.e., less than about 2 c/deg) direction-discrimination performance (as measured by D_{max}) was sometimes dependent on the order in which the two frames were presented.

Since the observers could readily report the correct direction of motion for RDKs in which both frames were band-pass filtered and contained common spatial frequencies restricted to either extreme of the range examined (see Fig. 1) the deterioration in performance found when only one of the frames was filtered at these frequencies needs to be explained. A plausible explanation for this pattern of results is that they reflect the differential sensitivity of the visual system to motion conveyed at different spatial frequencies. It is apparent that the shapes of the functions shown in Fig. 3 are remarkably similar to the shape of the spatial frequency contrast sensitivity function (CSF) in that performance is best at mid-range spatial frequencies (2–6 c/deg) and falls off rapidly for frequencies beyond this range. Fourier analyses of the unfiltered frames of the mixed RDKs revealed that although their spectra were essentially flat (approximately equal power at all spatial frequencies), the amplitude of any particular spatial frequency component was small (i.e., approximately 0.5–1%), especially when compared to its corresponding component in the band-pass filtered frames. Since the sensitivity of the visual system to the low and high spatial frequencies will be much less than that to frequencies close to the peak of the CSF, many of the low and high spatial frequency components present in the unfiltered frame are likely to be at or below their respective detection thresholds. As a result performance would be expected to decline when these frames are paired with filtered frames containing frequencies confined to these regions of the spectrum. In support of this suggestion Morgan & Mather (1994) found that motion discrimination became impossible for RDKs in which one frame was unfiltered and the other low-pass filtered by a Gaussian blurring function when the half-amplitude cut-off point of the filter fell below about 1 to 2.0 c/deg, a finding that could also be attributed to a lack of sensitivity to the low spatial frequency components present in the unfiltered frame.

An alternative explanation is that the results shown in Fig. 3 reflect the operation of a broadly tuned mechanism in the visual system that operates after motion has been initially detected at each of several different spatial scales. The function of such a mechanism could be to integrate motion signals conveyed at different spatial frequencies in order to compute an overall estimate of the direction of image motion. Within such a scheme direction discrimination will be impaired for mixed RDKs that have coherent motion confined to low or high spatial frequencies because of the visual system's relative insensitivity to these motion signals. In support of this suggestion Yang & Blake (1994) found that not only did their masking functions all peak at approximately the same value of 4 c/deg, but also that sensitivity to coherent motion signalled at different spatial frequencies (measured by varying the ratio of signal to noise dots in a filtered motion sequence) was broadly tuned and greatest at 4 c/deg, implicating the operation of a broadly tuned mechanism for detecting global motion. Although the present results using mixed RDKs cannot resolve this issue, they clearly demonstrate that studies employing broadband motion stimuli need to take account of the visual system's differential sensitivity to the spatial frequencies present in such stimuli if meaningful inferences are to be made concerning the mechanism(s) that determine direction-discrimination performance.

SUMMARY

The results of the present study clearly show that (i) direction-discrimination performance for two-frame RDKs is only possible when the two frames contain common (overlapping) bands of spatial frequencies; and (ii) that the visual system's differential sensitivity to information conveyed at different spatial frequencies may constrain direction judgements when broadband stimuli are employed in psychophysical experiments. The results suggest that, at least at the initial stages of motion detection, the processes responsible for encoding direction information do not integrate information across widely different spatial frequencies in order to compute motion. In this respect the present findings are in agreement with current models of motion detection (e.g., Adelson & Bergen, 1985) as well as the results of several previous studies (e.g., Anderson & Burr, 1985; Burr *et al.*, 1986; Cleary & Braddick, 1990a,b) in that they support the idea that motion-detection is likely to be carried out within restricted bands of spatial frequency. However, the results do not rule out the existence of broadly tuned mechanisms in the visual pathways mediating motion perception but rather imply that such mechanisms may operate prior to (Morgan, 1992; Morgan & Fahle, 1992; Morgan & Mather, 1994) and/or after motion computations that utilise information within narrow ranges of spatial frequency.

REFERENCES

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

Anderson, S. J. & Burr, D. C. (1985). Spatial and temporal selectivity of the human motion detection system. *Vision Research, 13*, 1147–1154.

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

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

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

Burr, D. C., Ross, J. & Morrone, M. C. (1986). Seeing objects in motion. *Proceedings of the Royal Society of London, Series B, 227*, 249–265.

Cavanagh, P. (1991). Short-range vs long-range motion: not a valid distinction. *Spatial Vision, 5*, 303–309.

Cavanagh, P. (1992). Attention-based motion perception. *Science, 257*, 1563–1565.

Chang, J. J. & Julesz, B. (1983). Displacement limits for spatial frequency filtered random-dot kinematograms in apparent motion. *Vision Research, 23*, 1379–1385.

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.

Green, M. (1986). What determines correspondence strength in apparent motion? Vision Research, 26, 599–607.

Morgan, M. J. (1992). Spatial filtering precedes motion detection. *Nature, 365*, 344–346.

Morgan, M. J. & Fahle, M. (1992). Effects of pattern element density upon displacement limits for motion detection in random binary luminance patterns. *Proceedings of the Royal Society of London, Series B, 248*, 189–198.

Morgan, M. J. & Mather, G. (1994). Motion discrimination in two-frame sequences with differing spatial frequency content. *Vision Research, 34*, 197–208.

Smith, A. T. (1992). Coherence of plaids comprising components of disparate spatial frequencies. *Vision Research, 32*, 393–397.

Ullman, S. (1979). *The interpretation of visual motion*, Cambridge, MA: MIT Press.

Van Santen, J. P. & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A, 2*, 322–341.

Watson, A. B. (1986). Apparent motion occurs only between similar spatial frequencies. *Vision Research, 26*, 1727–1730.

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

Yang, Y. & Blake, R. (1991). Spatial frequency tuning of human stereopsis. *Vision Research, 31*, 1177–1189.

Yang, Y. & Blake, R. (1994). Broad tuning for spatial frequency of neural mechanisms underlying visual perception of coherent motion. *Nature, 371*, 793–796.

Acknowledgements—I thank Andy Smith for his support and insightful discussions during the writing of this paper. The work was in part supported by a Science and Engineering Research Council grant awarded to Dr Smith.