

Motion Discrimination in Two-Frame Sequences with Differing Spatial Frequency Content

M. J. MORGAN,* G. MATHER†

Received 20 October 1992; in revised form 13 May 1993

We measured the upper threshold for directional motion discrimination (D_{\max}) in two-frame random binary luminance patterns (random dot kinematograms) in which either one or both frames was spatially low-pass filtered by convolution with a Gaussian filter. When both frames were low-pass filtered, D_{\max} increased as a function of the standard deviation of the Gaussian blurring function, in agreement with previous findings. However, when only one of the two frames was blurred, D_{\max} showed little change with blurring space constants below about 20 min arc, and at larger space constants motion discrimination became impossible. We take this as evidence against the proposal that D_{\max} is preferentially determined by motion signals from high spatial frequencies; and as evidence for the alternative that D_{\max} depends upon the mean spatial interval between features in the pattern after a single stage of spatial frequency pre-filtering. The breakdown in motion discrimination for space constants above about 20 min arc can be predicted from the computed effects of blurring upon the correlation between features (zero-bounded regions) in the broad-band and spatially filtered patterns. At values of blur where motion discrimination began to collapse there was a temporal order asymmetry: discrimination was easier when the low-pass pattern preceded the broadband pattern than when the broadband pattern appeared first. We propose that the temporally sustained high spatial frequency signal in the broadband pattern is delayed relative to the more transient low frequency signal; or alternatively, that the inhibitory surround of the spatial prefilter is switched in after a delay relative to the excitatory centre. The processing-delay interpretation was tested and confirmed in a second experiment by manipulating the frame duration.

Motion Spatial frequency M cells Zero-crossings Sustained and transient channels.

INTRODUCTION

Low-level motion discrimination mechanisms in human vision have been extensively investigated psychophysically using random binary luminance patterns (RBLPs; see Fig. 1 for examples). RBLPs are presented as a sequence of frames in which the pattern is moved through a constant displacement in the same direction between each frame. The observer's task is to report the direction of displacement. It is usually assumed that the use of random patterns containing many elements makes it impossible for the observer to track individual elements across frames by a shift in attention; thus RBLPs specifically isolate the low-level or "short range" motion discrimination process (Braddick, 1974; Anstis, 1970). Low-level motion discrimination is now generally held to depend upon a spatio-temporal filtering process (Morgan, 1980a, b; Ross & Burr, 1983; Burr, Ross & Morrone, 1986; Watson & Ahumada, 1985; van Santen

& Sperling, 1985; Adelson & Bergen, 1985; for an account of two-frame apparent motion see Watson, 1990). The main evidence for the "short-range" process is that motion direction discrimination in RBLPs has an upper limit for the size of the motion displacement (D_{\max}), beyond which coherent motion is replaced by incoherent motion. It was originally proposed that D_{\max} reflected a limited spatial range for low-level motion detector units (Braddick, 1974) but there is evidence that this is not a complete explanation. The evidence against a fixed size of motion detector is that low-pass filtering of RBLPs increases D_{\max} in inverse proportion to the centre frequency of the filter (Chang & Julesz, 1983; Cleary & Braddick, 1990; Bischof & Di Lollo, 1990). There are at least two interpretations of the effects of low-pass filtering upon D_{\max} . Cleary and Braddick (1990) suggested that there are multiple motion detector channels tuned to spatial frequency (see also Fahle & Poggio, 1981; Anderson & Burr, 1985; van den Berg & van de Grind, 1991; van de Grind, van Doorn & Koenderink, 1983), and that low-frequency tuned mechanisms support larger values of D_{\max} than high frequency tuned mechanisms. To account for the relatively small values

*To whom all correspondence should be addressed at: Department of Visual Science, Institute of Ophthalmology, Bath Street, London EC1V 9EL, England.

†University of Sussex, Falmer, Brighton BN1 9RH, England.

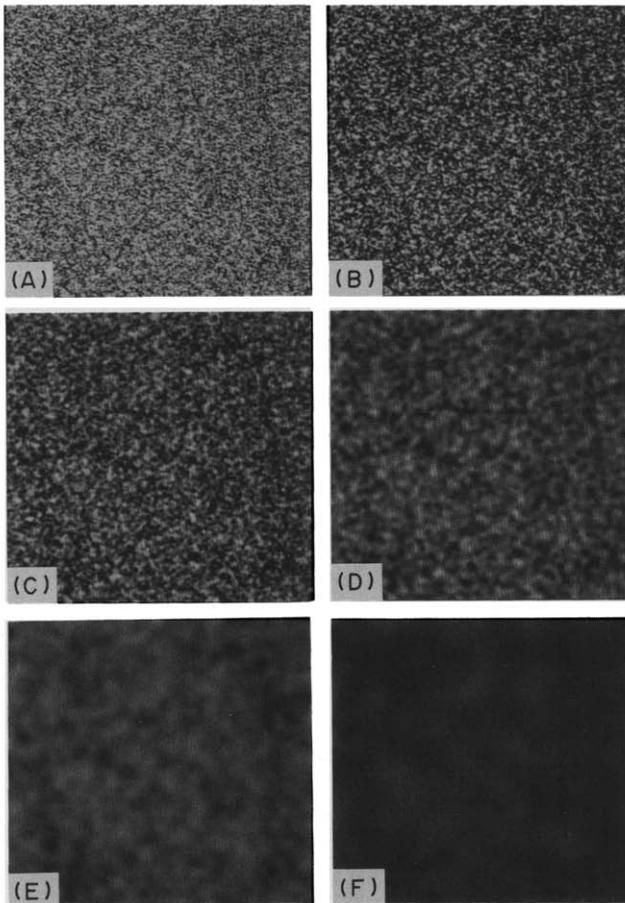


FIGURE 1. Example of the stimuli used in the experiments. (A) A broadband random binary luminance pattern with an element size of 4.5 min arc. Black and white pixels have equal probability. To investigate thresholds for motion discrimination the pattern was flashed twice in rapid succession, with a spatial displacement of the pattern within the stationary frame. The task was to detect the direction of displacement (upwards or downwards). Thresholds for motion discrimination were also investigated by pairing the broadband stimulus (A) with low-pass filtered versions of the same stimulus. (B, C, D, E, F) Stimuli are Gaussian blurred with space constants of 2.25, 4.5, 9, 18 and 36 min arc respectively.

of D_{\max} in broad-band patterns, which of course contain low spatial frequencies, one can propose either that high frequency channels inhibit low (Cleary & Braddick, 1990) or that they mask the coherent responses at low spatial frequencies with incoherent signals.

An alternative interpretation of the effects of low-pass filtering is that it has the purely physical effect of increasing the mean spacing between features in the pattern (Bischof & Di Lollo, 1990; Morgan, 1992). This would be expected to increase D_{\max} if the latter is determined by the "wagon wheel" aliasing that results when a pattern is displaced through more than one half of the mean spacing between its elements (Lappin & Bell, 1976; Morgan, 1992). In Morgan's (1992) model, D_{\max} was predicted from the distance between zero-crossings in the output of a single spatial filter in the early visual pathway to RBLPs: we shall refer to this for simplicity as the "single-channel prefilter" theory. The spatial filter implicated by the data could be well modelled by a difference-of-Gaussian filter with a space constant in the

range 8–16 min arc, and it was argued that this relatively coarse filter represented neurons of the magnocellular pathway (M cells) in the central 5 deg of the visual field. Zero-crossings were chosen as features because of their computational convenience without any special implication that they are the primitives actually involved in edge location. In fact, Watt and Morgan (1985) have argued against zero-crossings as spatial primitives, and in favour of zero-bounded regions (ZBRs) in the filtered image, and in the present paper we shall use ZBRs to model the edge-matching process.

In this paper we describe an experimental test to decide between these two possible interpretations of the effects of low-pass filtering upon D_{\max} . Consider the predictions of the two theories in the case where not both but only one of the frames in a motion sequence is low-pass filtered. High frequencies are thus present in only one of the frames. The inhibition theory states that in normal broad-band patterns, D_{\max} is limited by the properties of high-frequency filters. However, if the high frequencies are present in only one frame there is no motion signal to be detected at high frequencies. The high frequency channels may well be activated by the transients between the two frames, but there cannot be any overall directional signal. Therefore, motion discrimination should collapse entirely, and D_{\max} values should be zero.

A plausible modification of the inhibition theory might make a different prediction. Suppose that the inhibition is not an automatic result of activity in high frequency channels, but depends upon a coherent signal in those channels. In other words, a coherent but not an incoherent signal at high frequencies inhibits motion discrimination at low frequencies. This version of the inhibition theory predicts that removing the high frequencies from one frame only will release the low frequency channels from inhibition, so that D_{\max} values will be the same as those found when both frames are low-pass filtered.

What would be difficult to explain by the inhibition theory is a pattern of results in which increasing the blur in one of the frames has little effect upon D_{\max} , until a point is reached where motion discrimination abruptly collapses. This is the actual result that we shall report.

According to the signal-channel prefilter theory, however, the presence of extra elements in the broad-band frame will prevent coherent motion discrimination with displacements greater than the mean interval between elements, and values of D_{\max} with such patterns should thus be similar to those found when both patterns are broadband. The broadband pattern will be blurred by *intrinsic* filtering before motion discrimination, which according to the data of Morgan (1992) can be modelled by a Laplacian-of-a-Gaussian (LoG) filter with a space constant of 8–16 min arc. The low-pass pattern will be subjected to additional *extrinsic* blur depending on the space constant of the blurring function. It will be shown below that features (ZBRs) in the low-pass pattern have a significant spatial correlation with features in the broadband pattern only as long as the blurring space

constant does not exceed about twice the value of the internal filter. Below this limit, when the magnitude of motion displacement is greater than half the mean interval between features in the broadband pattern, correspondence will be lost. This is exactly the same limit as applies to unfiltered sequences and thus values of D_{\max} should not change.

MATERIALS AND METHODS

Apparatus and stimuli

The apparatus was identical to that described by Morgan (1992) and Morgan and Fahle (1992). A Mantron 85 Hz monochrome monitor was controlled by an 8-bit graphics processor (Ikon Pixel Engine). In Expt 1 each of the two frames lasted for six video frames, equivalent to 72 msec, with no interval between frames; in Expt 2 the frame duration was varied. A linear look-up table was used to correct the gamma of the display. The stimuli were generated and filtered on a SUN SPARC/IPC workstation using HIPS image processing software. Examples of the patterns are illustrated in Fig. 1. The broadband stimulus consisted of a random binary luminance pattern with element size 4.5 min arc, contrast 100% and mean luminance 30 cd/m². Between trials the screen was filled with a mean luminance field of 30 cd/m². Viewed from 2.28 m the stimulus frame subtended 5 × 5 deg of visual angle. The low-pass stimuli were obtained by convolving the broadband pattern with a Gaussian filter. In the main experiment the luminance range of the filtered patterns was not re-normalized following convolution, with the result that their peak-to-trough amplitude declined in proportion to the space constant of the filter. In a subsidiary investigation, the patterns were re-normalized to the unfiltered contrast to cover the whole available range (0–255).

Each frame consisted of 45 × 45 elements drawn from a larger (256 × 256) matrix, with the origin of the smaller display randomized over trials. Each trial was preceded by a mean luminance screen which remained on until the observer initiated the trial by a button press.

Procedure

D_{\max} values were obtained by the method of constant stimuli. For each condition, eight magnitudes of motion displacement were randomly interleaved until there had been 10 presentations of each. On each trial the direction of displacement (up or down) was randomly determined and the observer's response classified as correct or incorrect. The resulting psychometric function (pmf) was used to find the 80% correct point by linear interpolation, and this was defined as D_{\max} . At least two independent determinations were made of D_{\max} in each condition and the pmfs pooled to find a value of D_{\max} based upon 8 × 20 trials.

Subjects

The observers were the two authors (MM and GM) and one other (JH), who had served in a previous

investigation of D_{\max} (Morgan, 1992) but who was not informed of the purpose of the investigation.

EXPERIMENT 1

The space constant of the Gaussian blurring filter was varied between 0.5 and 8.0 pixels, equivalent to a range 2.25–36.0 min arc. The frame duration, or stimulus onset asynchrony (SOA) was 72 msec. In the low–low condition both patterns in the two-frame sequence were blurred. In the broad–low condition the first frame was unfiltered and the second was blurred, and *vice versa* in the low–broad condition.

We also investigated the case where one of the frames (the first) was low-pass filtered (space constant 36 min arc) and the filter applied to the second frame was systematically varied in the range 4.5–36 min arc.

Results and discussion

Experimental. The results are shown in Fig. 2. When both the first and second frames were low-pass filtered (low–low condition), D_{\max} values increased with the space constant of the Gaussian blurring function. In agreement with Cleary and Braddick (1990) small amounts of blur ($\sigma < 1$ min arc) had little effect, presumably because they were smaller than the intrinsic blur. Thereafter, the function relating D_{\max} to blur was accelerating, with some evidence that it was approaching an asymptote where D_{\max} was proportional to blur. It appeared to make little difference whether the stimuli were re-normalized in contrast or not.

However, when only one of the frames was filtered (low–broad or broad–low), the result was different. D_{\max} showed little change with the filter space constant until the latter reached 10–20 min arc. At larger values of blur there was some evidence for an increase in D_{\max} , especially in observed MM, although values were still much smaller than those obtained when both frames were low-pass filtered. Finally, when the blur exceeded a critical value, about 10.5 min arc for GM, 18 min arc for MM, and 30 min arc for JH, motion discrimination collapsed entirely. The rightmost point on the broad–low and the low–broad curves was the last one at which D_{\max} could be determined. The order in which the two frames were presented was important: when the first frame was the low-pass pattern, motion discrimination could be obtained at larger values of blur than in the case where the broadband pattern was first.

Modelling. One-dimensional modelling was carried out on random binary patterns presented in a 256 element array with an element size of one. The justification for using one-dimensional rather than two-dimensional modelling is discussed by Morgan (1992), in which it is reported that elongating the elements in the direction orthogonal to the motion displacement has little effect upon D_{\max} values. One-dimensional modelling reveals the essential statistical features of the filtering process. Two-dimensional modelling might change some of the conclusions quantitatively but not qualitatively.

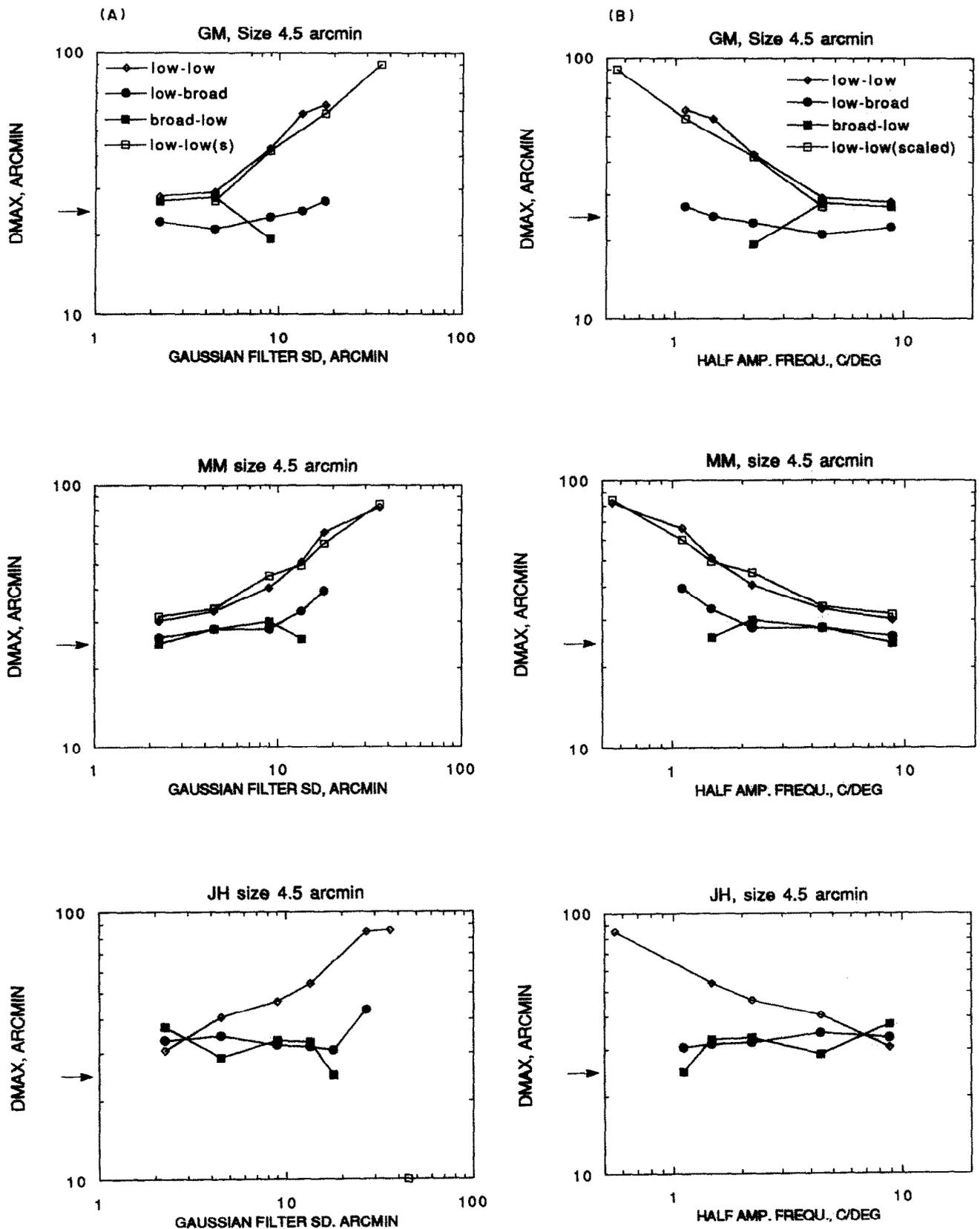


FIGURE 2. Results of Expt 1 in which the upper threshold for motion discrimination (D_{max}) was determined with low-pass-low-pass stimulus pairs and with mixed broadband-low-pass sequences. (A) The data plotted against the standard deviation of the Gaussian blurring function applied to the low-pass filtered member of the pair. (B) The same data plotted against the half-amplitude cut-off point of the Gaussian filter. Note that D_{max} rises with the blurring space constant in the low-low sequences, but not in the mixed sequences. Where data points are not shown in the low-broad and broad-low conditions, motion discrimination was impossible. Data for the low-low sequences was obtained both with stimuli that were not scaled in contrast following the convolution (low-low) and with stimuli that were scaled to have the same contrast as the broadband patterns (low-low scaled). The arrows on the figure show the level of performance attained with conventional broadband-broadband sequences. For further explanation see the text.

In previous versions of the model (Morgan, 1992; Morgan & Fahle, 1992) features were identified with zero-crossings in the band-limited image. In this paper we use the centroids of zero-bounded regions following Gaussian low-pass filtering, and removal of d.c., for the following reasons. First, we do not wish it to be thought that the prefilter model specifically favours a zero-crossing model of edge detection (Marr & Hildreth, 1980). Watt and Morgan (1985) have argued that zero-crossing position is too noise sensitive to account for the accuracy of human edge location, and that the centroids of zero-bounded regions are statistically more reliable. We initially attempted to model our data by locating and comparing the positions of ZBR centroids following filtering with a LoG filter, but encountered a difficulty which applies equally to zero-crossings following LoG filtering. The stimulus patterns used in the experiments were random, so it follows that the autocorrelation of the filtered image is given by the autocorrelation of the impulse response of the filter. Since the LoG filter is biphasic, it follows that short-range positive correlations are followed by longer-range negative correlations, predicting that beyond the D_{\max} limit, reverse-phi motion will be seen. We have looked carefully for evidence of this reversal in psychometric functions, without finding any. From the point of view of virtually any correlation model this could be considered as quite powerful evidence that the spatial filters involved in motion detection are low-pass rather than band-pass, or at least approximately so. The use of a low-pass filter for what is essentially a transient stimulus agrees with previous investigations (Robson, 1966; Wilson & Bergen, 1979) and can also be related to Watt's suggestion that the relative weighting given to low spatial frequencies declines rapidly following stimulus onset Watt (1987).

We therefore used a Gaussian prefilter with a standard deviation of 6.75 min arc, which we found the best replacement of the filter used by Morgan (1992) in predicting the knee-point in the functions relating D_{\max} to element size. In order to split the output into alternating ZBRs of positive and negative polarity, we removed the d.c. component, which produced some "ringing" and thus a small amount of predicted reverse-phi but not enough to be psychophysically detectable according to our model (see below).

The conjecture lying behind the model is that D_{\max} is determined by the displacement at which ZBRs in the two patterns no longer match correctly to signal the direction of motion. If only one of the patterns is low-pass filtered, the model shows that the spacing of ZBRs in the broadband pattern continues to have a dominating influence on D_{\max} . To investigate the degree of spatial correlation between the broadband and blurred patterns following a relative motion shift, we determined the mean distance between closest ZBRs of the same sign in the first and second patterns. We first verified that if the two patterns were identical except for a relative motion displacement of one element width, the mean separation between closest same-signed edges was also equal to one element width. In other words, the

statistic of mean edge distance between the patterns accurately encodes the actual motion displacement. We then investigated the effects of unequal spatial frequency filtering upon the accuracy of encoding a one-element motion displacement. Both patterns were blurred by a Gaussian filter with a space constant of 6.75 min arc. In addition, the low-pass filtered pattern of the pair was subjected to extra Gaussian blur and given a motion shift relative to the first pattern. The effect of motion shifts in the range 1–8 element widths was investigated.

Every ZBR in the low-pass pattern was matched to the nearest neighbour of the same sign in the broadband pattern, and the direction of the displacement between their centroids noted. If the displacement was in the

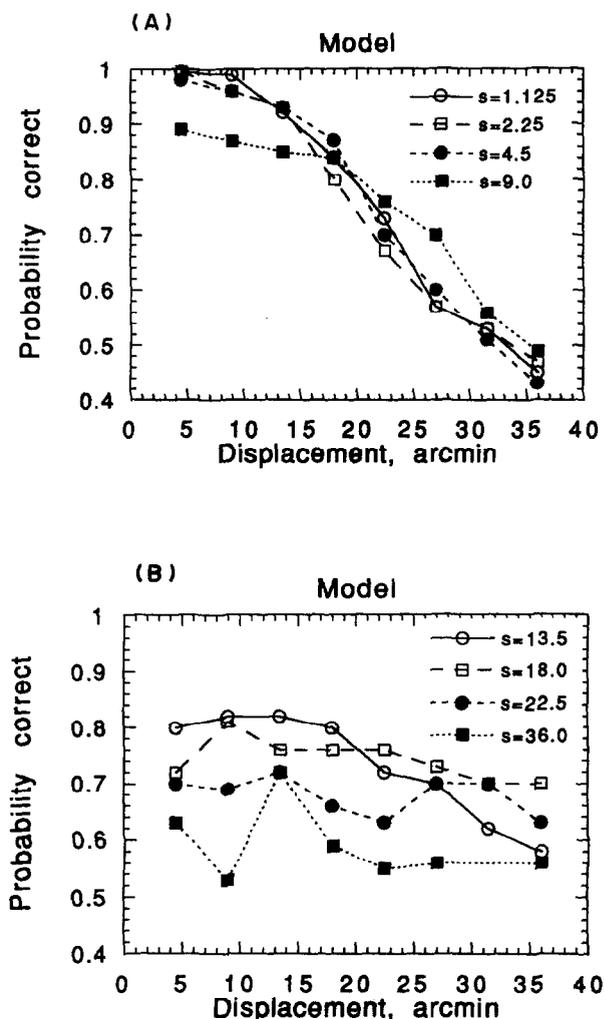


FIGURE 3. Results of the ZBR motion-matching model described in the text. One stimulus frame was broadband; the other was subjected to low-pass filtering with a Gaussian filter having the standard deviation shown as the curve parameter. In the model, both frames were subjected to intrinsic Gaussian filtering with a space constant of 6.25 min arc; and the resulting convolution profile was split into positive and negative zero-bounded regions (ZBRs) following removal of d.c. Following a spatial translation of one of the patterns (horizontal axis) each ZBR in the extrinsically-blurred pattern was then matched to the nearest ZBR of the same sign in the broadband pattern. The match was counted as "correct" if it was in the direction of the actual motion shift; otherwise as "incorrect". The probability correct (vertical axis) was calculated over all ZBRs and over 20 simulated repetitions. (A) The results with filters in the range 1.125–9.0 min arc; (B) the range 13.5–36 min arc.

actual direction of pattern displacement, a “+” score was given; if in the opposite direction, a “-” score. The probability of a + displacement was calculated over all ZBRs and an overall percent correct computed. It is this final statistic that we use in comparing the model and the data.

Psychometric functions computed from the model are shown in Fig. 3. Figure 3(A) shows the results with the four smallest added filters, and Fig. 3(B) shows the four largest. The two sets are separated for clarity. The first point to emerge is that the model accounts for the D_{\max} phenomenon: the probability of a correct match falls off as the displacement increases, eventually falling to a chance level. The D_{\max} values for observers in the experiment correspond to a point that gives a probability of about 0.7 correct identification according to the model. Any relationship between this and the 80%

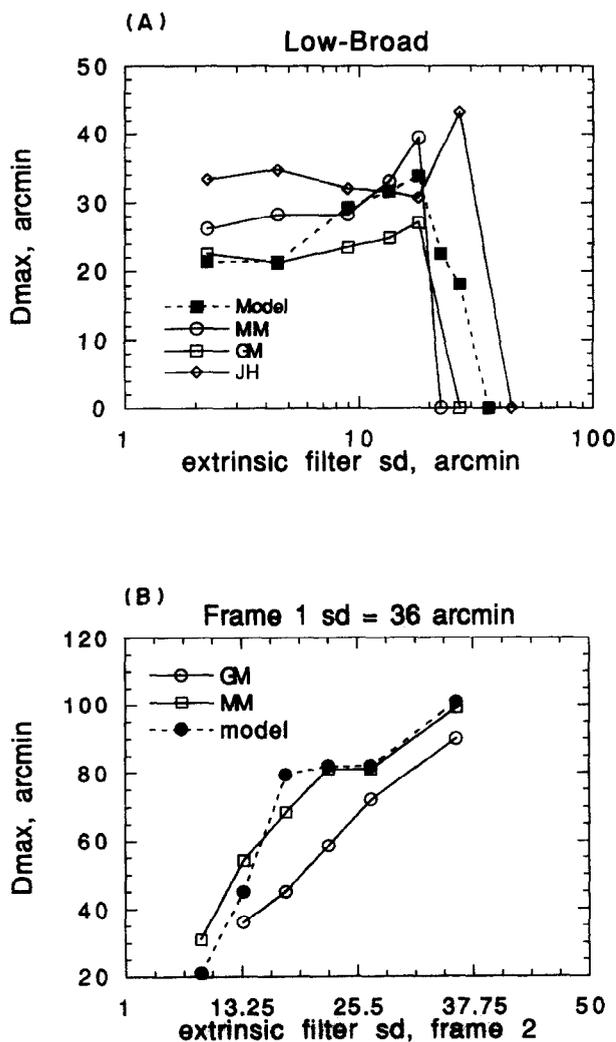


FIGURE 4. (A) Calculated “ D_{\max} ” values derived from the computed psychometric functions in Fig. 3. D_{\max} is defined as the displacement beyond which the probability correct falls below 0.7. The D_{\max} value predicted by the model is compared to the actual results for three observers in the case where the low-pass pattern preceded the broadband (cf. Fig. 2). (B) Results of another experiment in which the first frame was always low-pass filtered ($s = 36$ min arc) and the second frame was filtered with the space constant shown on the horizontal axis. The predicted D_{\max} values from the model (0.7 correct point) are compared to the data from two observers.

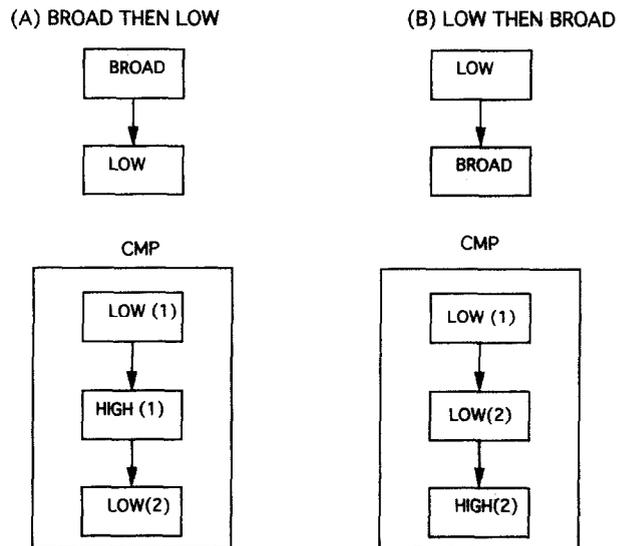


FIGURE 5. A model of the temporal order asymmetry described in the text. The boxes at the top represent the temporal order of the broadband stimulus and the low-pass filtered version of the same stimulus. The boxes at the bottom represent the hypothetical sequence of arrival of high and low frequencies from the first (1) and second (2) frames at the central motion processor (CMP).

criterion for D_{\max} in the experiments is, of course, entirely coincidental, since the probability of the observer's being correct cannot be directly predicted from the probability of a single ZBR moving in the correct direction.

The second point to emerge from Fig. 3 is that the model psychometric function show little change with an added filter to one frame with a space constant in the range 1.125–9.0 min arc. This is in agreement with the psychophysical data showing no change in D_{\max} in the same range. However, from about 9.0 min arc onwards the functions change their shape, becoming flatter. Initially this results in an improvement in detection at intermediate displacement levels, but with the largest filter values (22.5–36) performance does not rise above 0.7.

To compare the model with the observer's D_{\max} value, we used the psychometric functions in Fig. 3 to calculate the largest displacement at which performance met the criterion of $P = 0.7$ or greater, and plotted these displacements in Fig. 4 along with the observers' data for the low-broad condition, which they most resemble. The two sets of data have qualitative similarities. Performance is flat across an initial range of filter values, then improves slightly and then collapses. Figure 4(B) also shows the experimental results and results of the model when the first frame was low-pass (36 min arc) and the space constant of the filter in the second frame was systematically varied as shown by the values on the horizontal axis (4.5–36 min arc). The model successfully explains the gradual decline in D_{\max} as the blur difference between the two frames increases, and the complete failure of performance below values of 4.5–9.0 for the second frame.

The model as it stands is unable to explain the difference in the experimental results depending upon the

order of the frames. This must presumably depend upon a physiological rather than a statistical mechanism. It is curious that in the relatively easy case (low-to-broad) the low-pass pattern tended to be invisible, whereas in the relatively hard case (broad-low) it was clearly seen as persisting after the broadband stimulus.

To explain the temporal order asymmetry we propose that the high spatial frequencies from the broadband frame arrive at the central site of motion processing with a delay relative to the low spatial frequencies. The response to low-spatial frequencies is characterized by a rapid, transient impulse response (Tolhurst, 1975, human reaction time data). The consequences of a relative delay in processing between high and low frequencies in the present experiment is illustrated in Fig. 5. When the first frame is the broadband pattern, the low spatial frequencies from the frame arrive at the central motion processor (CMP) first, followed after a short delay by the high frequencies. We propose that the high and low frequencies are integrated at this stage and that the resulting pattern of ZBRs is the same as that of the broadband pattern. The second frame then arrives,

and since it is low-pass it will fail to correspond to the zero-crossing pattern of the first frame as soon as the blur exceeds a critical value, as we have shown in the model (Fig. 3). In other words the sequence of events at the CMP is: low(1) low(1) + high(1) low(2) + high(1) where low(1) represents the low frequency content for the first frame and low(2) the low frequency content of the second frame.

However, when the first frame is the low-pass pattern, the sequence of events at the CMP is: low(1) low(2) low(2) + high(2). There is thus an opportunity for the low frequencies in the first frame to be matched with the low frequencies in the second pattern without interference from the high frequencies, and this will allow motion discrimination to take place.

This account of the temporal order effect relies upon the interval between frame 1 and frame 2 being larger than the delay between high and low frequencies. Thus, if the inter-frame interval were shortened, it might be possible to reduce the asymmetry and for motion discrimination to become possible again in the broadband, low-pass sequence. This is because the sequence of events

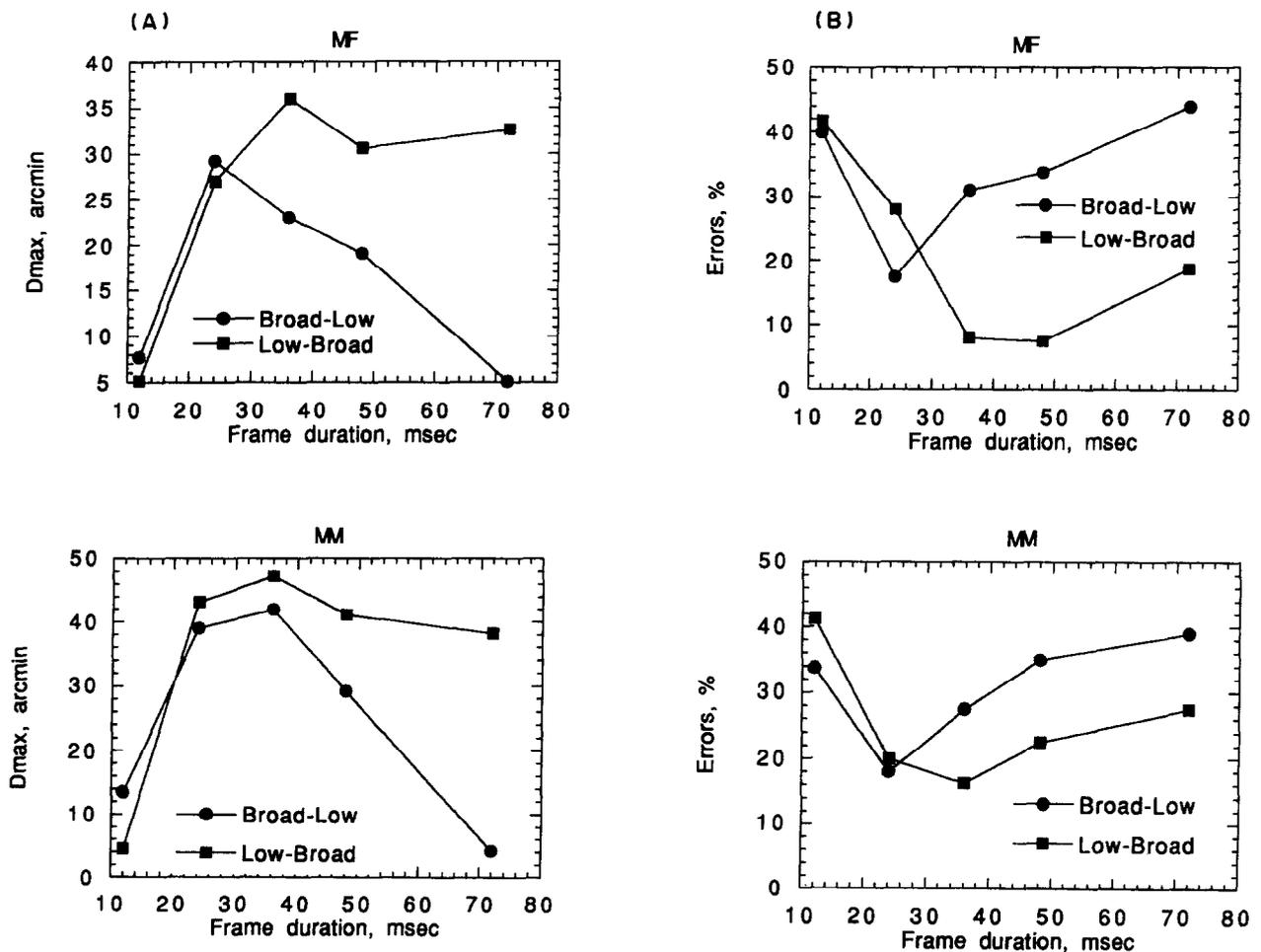


FIGURE 6. The figure shows the results of Expt 2, in which the frame duration was varied in two-frame motion sequences one frame of which was a broadband RBLP and the other of which was the same pattern subjected to low-pass filtering. The space constant of the Gaussian blurring filter was 18 min arc for observer MM and 13.5 min arc for observer MF. (A) D_{max} (vertical axis) varied as a function of frame duration (horizontal axis) and the order of presentation of the two frames (see the key against the symbols). (B) The error rate (vertical axis), calculated as the percentage of errors made over all displacements. Note that at a frame duration of 20 msec performance is independent of temporal order, while at longer durations performance is superior in the low-broad order than in the broad-low order.

at the CMP with a sufficiently short inter-frame interval would become: low(1), low(2), low(2) + high(1). This prediction was tested in Expt 2.

EXPERIMENT 2

As in Expt 1 one of the frames was an unfiltered 50% RBLP, and the other was a low-pass filtered version of the same pattern. Different values of the space constant for the filter were chosen for two observers (MM, MF) depending on the point where the asymmetry between discrimination and non-discrimination occurred using the frame duration of the first experiment (72 msec).

Materials and methods

Apparatus and stimuli. The apparatus and stimuli were identical to those in the first experiment. The space constant of the Gaussian blurring filter was 18 min arc for observer MM and 13.5 min arc for observer MF. The frame duration was varied between 12 and 72 msec.

Subjects. The observers were one of the authors (MM) and another psychophysically experienced observer (MF) who had served in a previous investigation of D_{max} with RBLPs (Morgan & Fahle, 1992).

Results and discussion

The results shown in Fig. 6 confirm the prediction that motion discrimination in the broad-low order would become easier when the frame duration was reduced. Indeed, at a frame duration of 24 msec there was no difference between the two temporal orders. It follows from the model that we have proposed that the delay between high and low spatial frequencies arriving at the central motion processor is >24 msec, and <72 msec.

One interpretation of the delayed processing of high spatial frequencies is that it depends upon a signal from the parvocellular pathway. There is evidence that the very early stages of motion discrimination depend upon an input from neurons in the magnocellular pathway, which because of their relatively large receptive field are insensitive to the high spatial frequencies in a broadband RBLP. This was the conclusion reached by Morgan (1992) on the basis of the lack of an effect of element size upon D_{max} in the range 1–10 min arc; it is also compatible with the finding of Cleary and Braddick that moderate low-pass filtering has no effect on D_{max} (cf. our Expt 1).

Other evidence, however, suggests that the parvo system can also have an input to motion discrimination. The main evidence for this is that chromatic information can influence motion discrimination. For example, Morgan and Cleary (1992) found that D_{max} was reduced if red-green elements in a random binary colour/luminance pattern changed colours between frames. More recently, Morgan and Ingle (1994) have shown that chromatic and achromatic information can either facilitate or hinder one another in motion discrimination. These findings strongly imply that chromatic information can have an input to low-level motion discrimination. Since neurons of the magno-pathway appear to be essentially unselective for colour, the

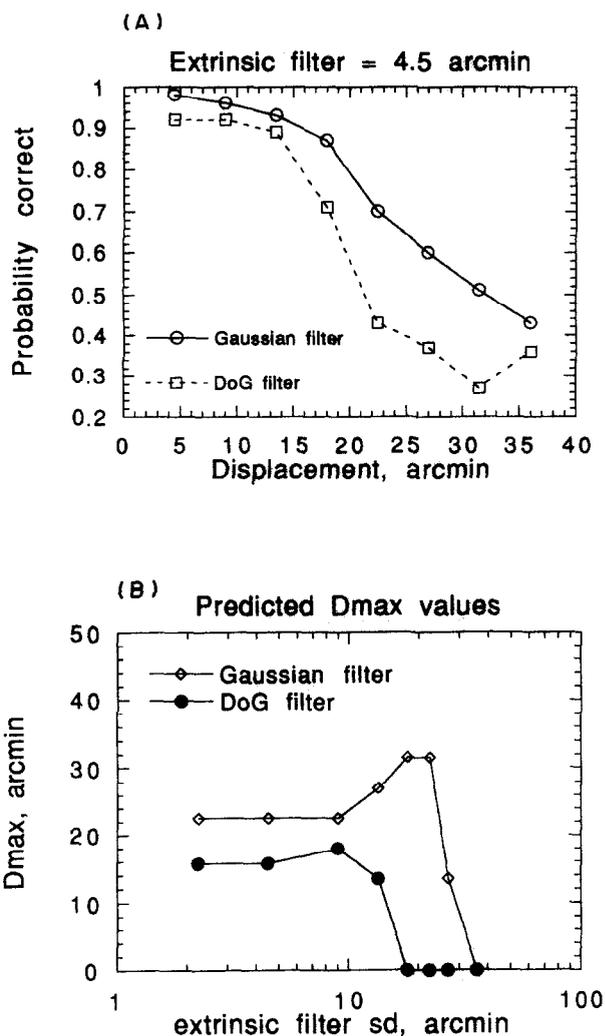


FIGURE 7. The figure compares the effects of Gaussian and Difference-of-Gaussian (DoG) filtering upon motion detection by the ZBR model described in Fig. 3. (A) Examples of computed psychometric functions in the case where the extrinsic filter (applied to the first frame) has the space constant 4.5 min arc. (B) Predicted D_{max} values derived from the 0.7 point in the computed psychometric function, for a range of values for extrinsic blur.

conclusion follows that the parvo-system has an input to low level motion discrimination (see also Merigan, 1991).

Our findings can now be explained if it is supposed that the parvo-signal arrives at the central motion processor after a delay with respect to the magno-signal. This could be either because the parvo-signal has a more sluggish, sustained temporal impulse response or because it arrives at the central motion processor by a less direct route than the magno-signal. We argue that when the delayed P signal arrives it is automatically integrated with the M signal, and that the latter is not independently accessible. This follows from the results of the present experiment, and from the effects of colour in reducing D_{max} (Morgan & Cleary, 1992).

In summary, we are suggesting: (1) that the higher spatial frequencies from our broadband RBLPs are carried by the parvo-subsystem, and the low by the magno-subsystem; (2) the parvo-signal arrives at the

central motion processor with a delay relative to the magno-signal; (3) the magno- and parvo-signals are integrated prior to motion discrimination, which explains why the values of D_{max} are smaller when only one frame is low-pass filtered than when both are; and finally, (4) the fact that the high frequency signal is delayed explains the temporal order asymmetry between the broad-low and the low-broad temporal orders.

We have considered an alternative but related explanation of the temporal asymmetry, based upon the idea that the surround inhibition of the spatial prefilter increases over time. The initial response of a filter with delayed inhibition would be low-pass, followed by band-pass. The different behaviour of the ZBR motion detection model of a low-pass filter (Gaussian) *vs* a band-pass filter obtained by adding an inhibitory surround (Difference-of-Gaussian filter, DoG) are shown in Fig. 7. Errors rise more rapidly with displacement in the case of the DoG filter, and over a range of extrinsic filter sizes, D_{max} values are smaller with the DoG filter. The larger D_{max} for the low-pass filter is predictable from the autocorrelation function, as pointed out earlier. Since the autocorrelation of a filtered random sequence is given by the autocorrelation of the impulse response of the filter, the autocorrelation of the Gaussian and DoG filtered noise sequences will be Gaussian and DoG respectively. If, as in the case analysed in Fig. 7 the Gaussian and DoG filters have the same excitatory component, it follows that the range of spatial shifts over which the patterns are correlated is higher in the Gaussian (low-pass) case.

The difference in correlation structure between Gaussian and DoG filtered random noise does not immedi-

ately offer an explanation of the temporal order effect. However, if the initial response of the intrinsic prefilter is low-pass, and then after a brief delay becomes band-pass, we can assume that the motion response triggered by the second frame will depend upon a comparison of the low-pass response to the second frame and the band-pass response to the first frame. When the first pattern has been externally blurred, then the comparison will be easier because it will involve two low-frequency patterns. By this argument, if the first pattern is low-pass filtered and the second is broadband, motion detection should be relatively easier. This is not the case if the broad band pattern comes first, thus the temporal asymmetry. This is an informal verbal argument, but we have modelled the temporal order effect by applying a Gaussian filter to one of the frames, and a DoG to the other, with results shown in Figs 8 and 9. The first of these figures (Fig. 8) shows computed psychometric functions according to whether the DoG filter is applied to the broadband pattern (solid symbols) or to the low-pass filtered pattern (open symbols). Note that at all values of extrinsic blur, but especially at intermediate values, less errors are made when the DoG filter is applied to the externally low-pass filtered pattern. These psychometric functions were used to find the approx. 70% correct point, and the resulting predicted D_{max} values are plotted in Fig. 8 along with the experimental values for each observer.

Qualitatively, the data and predictions in Fig. 9 show similar trends, in particular: (1) as the size of the extrinsic blur is increased, performance initially shows little change, then D_{max} starts to rise in the low-broad case while remaining the same in the broad-low case; (2)

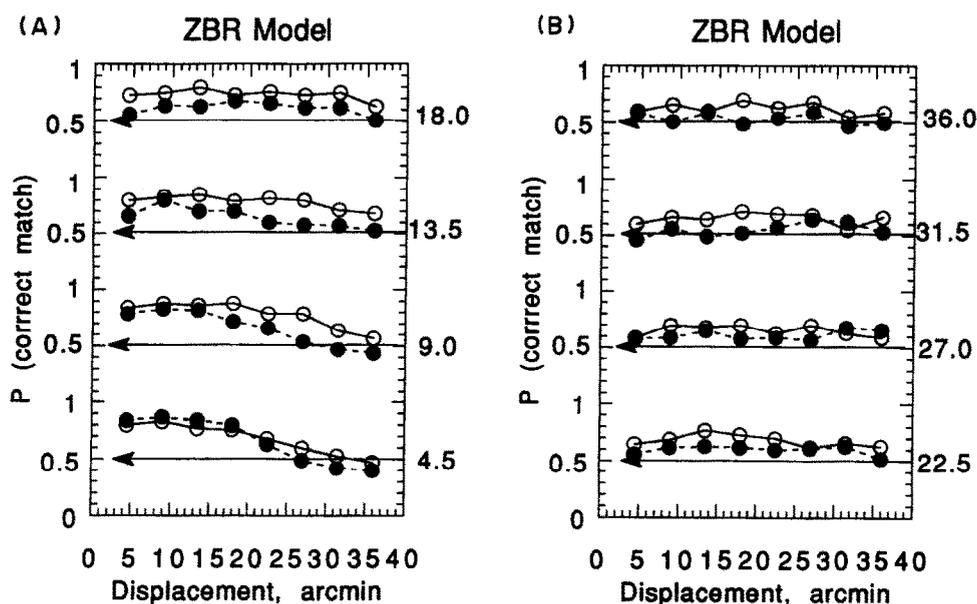


FIGURE 8. The figure shows computed psychometric functions for the ZBR model in the conditions of Expt 1, with a broadband pattern in one frame and, in the other frame, the pattern filtered with a Gaussian having the space constant shown to the right of each pair of curves. Open symbols: the low-pass filtered pattern is subjected to an intrinsic filter that is a DoG and the broadband pattern is subjected to an intrinsic Gaussian filter. Solid symbols: the low-pass filtered pattern is subjected to an intrinsic filter that is a Gaussian and the broadband pattern is subjected to an intrinsic DoG filter. These conditions are intended to correspond to the different temporal orders: low-broad and broad-low respectively, on the assumption that the first frame is subjected to DoG and the second to Gaussian filtering.

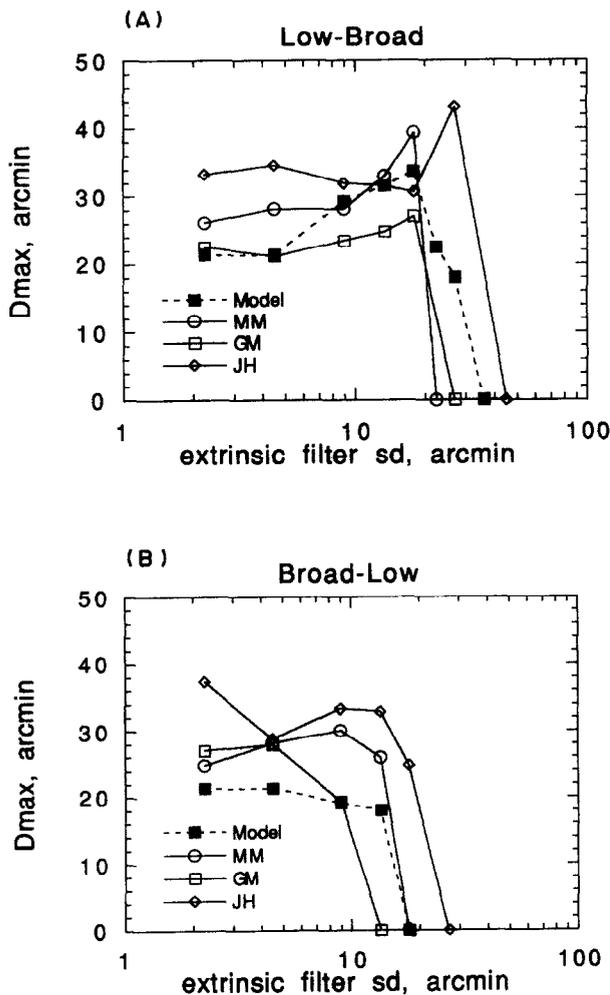


FIGURE 9. The figure shows D_{\max} values computed from the psychometric functions in Fig. 8 (0.70 correct point) along with data from three observers in Expt 1.

beyond a limiting value of extrinsic blur, which is smaller in the broad-low case than in the low-broad, performance collapses abruptly and completely. It will be obvious that we could have obtained closer fits by tailoring the space constants of the intrinsic filter to the individual observers, and also by varying the ratio of space constants in the DoG filter to make it more or less like the low-pass filter. However, with a set of minimal assumptions, the way in which the model accounts for the main features of the data is suggestive.

A delay of about 20–50 msec in the surround contribution relative to the direct cone contribution has been reported in the retina of the mudpuppy (Werblin, 1977; see review by Richter & Ullman, 1982). However, the lag in the indirect input to retinal ganglion cells in the cat and in primates is about a factor of 10 times smaller than in the mudpuppy and turtle (see, again, the review by Richter & Ullman, 1982) so it is not likely that the effects we are observing have their origin in the retina.

GENERAL DISCUSSION

The modelling presented here relies heavily upon the concept of a single prefilter before motion detection,

and we attempt first to respond to possible criticisms of this notion. The putative filter was identified from data showing that D_{\max} is insensitive to variation in pattern element size in the range 1–10 min arc, but thereafter rises to become asymptotically linear with element size.

The alternative interpretation of the dependence of D_{\max} on element size is that the span of bi-local motion detectors is inversely proportional to spatial frequency (van de Grind, Koenderink & van Doorn, 1992). To explain why D_{\max} does not vary with element size in the range 1–10 min arc van de Grind *et al.* suggest an ingenious possibility. They note that when a constant field size is used, reducing the element size causes an increase in element number, and they suggest that this increase in number exactly compensates for the expected decrease in D_{\max} . Using displays where small pixel sizes are combined with small field sizes, they find astonishingly small D_{\max} values, as little as 2 min arc for element sizes of about 15 sec arc.

The problem we have with D_{\max} values obtained with small angular field sizes is that values may be depressed because of the artifact noted by Eagle and Rogers (1991). As field size is reduced the number of elements that stay within the field when displaced rather than moving outside it is reduced, thus diminishing the effective signal-to-noise ratio. We also find the idea of a fortuitous cancellation between the effects of element number and element size conceptually less simple than that of a single prefilter. Nevertheless, van de Grind *et al.* may be correct, and it would be interesting to see how well their own “stack of stacks” model could account for the results of the present experiment.

A second criticism, also arising from the work of the Utrecht group, is that D_{\max} changes with eccentricity (e.g. van de Grind *et al.*, 1983) and so it is unlikely that our 5×5 deg field can be adequately described by a single filter. However, if we are correct in supposing that the receptive field size of cells in the magnocellular pathway is a major determinant of D_{\max} , it is relevant to note that these may scale less with eccentricity than traditional measures of cortical magnification would suggest. A recent anatomical study (Dacey & Petersen, 1992) of man has shown that the difference in dendritic field size between the parasol and midget cells increases from a ratio of $\sim 3:1$ in the retinal periphery to $\sim 10:1$ at 3 deg eccentricity. Another possibility for explaining our single prefilter is that it represents the largest filter size available within the stimulus area (cf. Eagle & Rogers, 1991).

We now turn to the implications of our mixed-frame data. Values of D_{\max} in the mixed-frame sequences were clearly more similar to those of the broadband patterns than to those of the low-pass, at least until the point where motion discrimination disappeared altogether. Our view that different spatial frequencies interact within a single prefilter appears to account for the data quite economically. Motion discrimination is possible with moderate amounts of blur of one stimulus because there is a close spatial correspondence between features (ZBRs) in the blurred and in the broadband pattern.

D_{\max} , however, will continue to be determined by the mean spacing of features in the broadband pattern, because these will produce false matches with the sparser set of features in the more blurred pattern. Finally, the collapse of motion perception when blur reaches a critical value is accounted for by the fact that spatial correspondence between the two patterns has broken down.

The concept of a common motion processor across frequency must be modified, however, by the finding of a temporal order asymmetry of the broadband and low-pass patterns. We have argued that this may arise because of a delay in the transmission of high spatial frequencies to the central motion processor. A similar suggestion has recently been made by Anderson (1993) to account for distortions in the perceived appearance of drifting complex gratings. In our model the high and low spatial frequency signals are pooled *before* motion analysis takes place. This is an alternative to the usual idea that motion analysis is carried out initially in independent frequency channels, and frequency tuned responses are pooled *after* motion analysis to form the basis for a final decision about motion direction. According to this conventional view, the deleterious effects of high frequencies on D_{\max} in broadband patterns arises because their motion responses mask or inhibit those from low frequencies. We argue against this conventional view because it cannot explain our basic result that D_{\max} in mixed broadband-low-pass sequences remained constant at the value obtained for broadband patterns. Removal of high-frequency motion signals from the pooled signal should have allowed D_{\max} to rise in the mixed sequences as it did in the low-pass-low-pass sequences. The data are consistent with the alternative explanation that spatial frequency information is pooled prior to motion analysis. It is interesting to note that this is the conclusion reached about the processing of spatial frequencies in static spatial vision by Watt and Morgan (1985) in their "MIRAGE" model.

We have suggested that the delayed high-frequency signal arrives at the CMP via an indirect route involving the parvocellular pathway, but this is purely speculative. It is possible that *within* the magno-pathway, there are different spatial frequency channels with different temporal impulse response characteristics. Or, as we have shown, the temporal order asymmetry may arise from a temporal transition from low-pass to band-pass filtering, caused by a delay in surround inhibition. Alternatively, a motion signal from the parvo-pathway may arise from a sub-system with low-pass spatial frequency characteristics, such as the parvo-interblob system (Livingstone & Hubel, 1988; De Yeo & Van Essen, 1988). If this view is correct, the motion detection system would make use of at least three major pathways, just as Tyler (1990) has suggested for stereopsis. This could be tested by seeing whether the delayed signal involved in the temporal order asymmetry is sensitive to colour, using the method described by Morgan and Cleary (1992).

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. (1993). Visual processing delays alter the perceived spatial form of moving gratings. *Vision Research*, *33*, 2733–2746.
- Anderson, S. J. & Burr, D. C. (1985). Spatial and temporal selectivity of the human motion detection system. *Vision Research*, *25*, 1147–1154.
- Anstis, S. M. (1970). Phi movement as a subtraction process. *Vision Research*, *10*, 1411–1430.
- van den Berg, A. V. & van de Grind, W. A. (1991). Conditions for the detection of coherent motion. *Vision Research*, *31*, 1039–1051.
- Bischof, W. F. & Di Lollo, V. (1990). Perception of directional sampled motion in relation to displacement and spatial frequency: Evidence for a unitary motion system. *Vision Research*, *9*, 1341–1362.
- Braddick, O. (1974). A short-range process in apparent motion. *Vision Research*, *14*, 519–527.
- Burr, D. C., Ross, J. & Morrone, M. C. (1986). Seeing objects in motion. *Proceedings of the Royal Society of London B*, *227*, 249–265.
- Chang, J. J. & Julesz, B. (1983). Displacement limits, directional anisotropy and direction versus form discrimination in random-dot cinematograms. *Vision Research*, *23*, 639–646.
- Cleary, R. F. & Braddick, O. J. (1990). Direction discrimination for band-pass filtered random dot kinematograms. *Vision Research*, *30*, 303–316.
- Dacey, D. M. & Petersen, M. R. (1992). Dendritic field size and morphology of midgital and parasol ganglion cells of the human retina. *Proceedings of the National Academy of Sciences, U.S.A.*, *89*, 9666–9670.
- De Yeo, E. A. & Van Essen, D. C. (1988). Concurrent processing streams in the monkey visual cortex. *Trends In Neuroscience*, *11*, 219–226.
- Eagle, R. A. & Rogers, B. J. (1991). Maximum displacement (d_{\max}) as a function of density, patch size, and spatial filtering in random-dot kinematograms. *Investigative Ophthalmology and Visual Science (Suppl.)*, *32*, 893.
- Fahle, M. & Poggio, T. (1981). Visual hyperacuity: Spatiotemporal interpolation in human vision. *Proceedings of the Royal Society of London B*, *213*, 451–477.
- van de Grind, W. A., Koenderink, J. J. & van Doorn, A. J. (1992). Viewing-distance invariance of movement detection. *Experimental Brain Research*, *91*, 135–150.
- van de Grind, W. A., van Doorn, A. J. & Koenderink, J. J. (1983). Detection of coherent movement in peripherally-viewed random-dot patterns. *Journal of the Optical Society of America*, *73*, 1674–1683.
- Lappin, J. S. & Bell, H. H. (1976). The detection of coherence in moving random-dot patterns. *Vision Research*, *16*, 161–168.
- Livingstone, M. S. & Hubel, D. H. (1988). Segregation of form, color, movement and depth: Anatomy, physiology and perception. *Science*, *240*, 740–749.
- Marr, D. & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London B*, *207*, 187–217.
- Merigan, W. H. (1991). Does primate motion perception depend on the Magnocellular pathway? *Journal of Neuroscience*, *11*, 3422.
- Morgan, M. J. (1980a). Analogue models of motion perception. *Philosophical Transactions of the Royal Society of London B*, *290*, 117–135.
- Morgan, M. J. (1980b). Spatiotemporal filtering and the interpolation effect in apparent motion. *Perception*, *9*, 161–174.
- Morgan, M. J. (1992). Spatial filtering precedes motion detection. *Nature (London)*, *365*, 344–346.
- Morgan, M. J. & Cleary, R. F. (1992). Effects of colour substitution upon motion detection in spatially random patterns. *Vision Research*, *32*, 815–821.
- 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 B*, *248*, 189–198.

- Morgan, M. J. & Ingle, G. (1994). Effects of colour and luminance substitutions upon motion discrimination. *Vision Research*. Submitted.
- Richter, J. & Ullman, S. (1982). A model for the temporal organization of X- and Y-type receptive fields in the primate retina. *Biological Cybernetics*, *43*, 127-145.
- Robson, J. G. (1966). Spatial and temporal contrast sensitivity functions of the visual system. *Journal of the Optical Society of America*, *56*, 1141-1142.
- Ross, J. & Burr, D. (1983). The psychophysics of motion. In Arbib, M. & Hanson, A. R. (Eds), *Vision, brain & cooperative processes*. Amherst, Mass.: Bradford.
- van Stanten, J. P. H. & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A*, *2*, 300-321.
- Tolhurst, D. J. (1975). Reaction times in the detection of gratings by human observers: A probabilistic mechanism. *Vision Research*, *15*, 1143-1149.
- Tyler, C. W. (1990). A stereoscopic view of visual motion processing. *Vision Research*, *30*, 1877-1895.
- Watson, A. B. (1990). Optimal displacement in apparent motion and quadrature models of motion sensing. *Vision Research*, *30*, 1389-1393.
- Watson, A. B. & Ahumada, A. J. Jr (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A*, *2*, 322-342.
- Watt, R. J. (1987). Scanning from coarse to fine spatial scales in the human visual system after the onset of a stimulus. *Journal of the Optical Society of America A*, *4*, 2006-2021.
- Watt, R. J. & Morgan, M. J. (1985). A theory of the primitive spatial code in human vision. *Vision Research*, *25*, 1661-1674.
- Werblin, F. S. (1977). Regenerative amacrine cell depolarization and formation of on-off ganglion cell response. *Journal of Physiology, London*, *264*, 767-785.
- Wilson, H. R. & Bergen, J. R. (1979). A four mechanism model for threshold spatial vision. *Vision Research*, *19*, 19-33.

Acknowledgements—This work was supported by a grant from the Image Interpretation Initiative of the SERC to MM and by an MRC grant to GM.