



Temporal Filtering Enhances Direction Discrimination in Random-dot Patterns

GEORGE MATHER,*† HILARY TUNLEY*

Received 28 March 1994; in revised form 27 June 1994

In conventional presentations of random-dot kinematograms, two frames of random dots are presented in temporal sequence, separated by a blank inter-stimulus interval, and a coherent offset in spatial position is added to dots in one frame relative to dots in the other frame. Direction discrimination performance is limited temporally to inter-stimulus intervals below about 100 msec (T_{\max}). Experiments are described in which temporal smoothing was applied to the onset and offset of each frame in the kinematogram. T_{\max} was found to increase in proportion with the time constant of the temporal smoothing function. An explanation based on contrast-dependent responses in simple motion detectors cannot accommodate the results. Instead, the increase in T_{\max} with temporal smoothing, and analogous increase in spatial limit (D_{\max}) with spatial blurring, can be related to the spatiotemporal frequency content of the stimulus. Random-dot kinematograms can be viewed as continuously drifting patterns that have been discretely sampled at regular spatiotemporal intervals. Sampling introduces artefacts (alias signals), which become more intrusive as sampling rate declines (i.e. inter-stimulus interval or spatial displacement increases) and consequently limit discrimination performance. Temporal smoothing or spatial blurring extends performance because it removes alias signals generated by high spatiotemporal frequencies in the pattern. Computational modelling to estimate the Fourier energy available in random-dot kinematograms confirmed that the sampling account can predict the proportional increase in T_{\max} and D_{\max} limits as filter time or space constant increases.

Motion Random-dot kinematograms Sampling Temporal frequency

INTRODUCTION

Random-dot kinematograms (RDKs) have played a central role in the development of theories relating to early motion analysis in the visual system. In a simple RDK, two frames containing random black–white dots are presented in temporal sequence at the same spatial location. The only difference between the two frames is that a coherent shift in position is added to some or all of the dots in the second frame relative to the first, either in one direction or its opposite. The observer's task is to identify the direction of shift (if all dots shift position), or identify the shape of the displaced region (if only some dots shift position). Early findings were interpreted as evidence for a “short-range” motion process. Braddick (1973, 1974) found that shape discrimination was reliable for short spatial displacements (below about 0.25 arc deg), but once an upper limiting displacement was exceeded (the so-called D_{\max} limit) discrimination was no longer possible. Similarly, good discrimination performance was limited temporally to inter-stimulus intervals below about 100 msec (T_{\max}). These performance

limits have been replicated in numerous experiments, and are usually attributed to physiological properties of simple motion detectors. Each detector consists of two input receptive fields positioned at adjacent retinal locations, whose responses are multiplied together at a comparator neuron. If a temporal delay (dt) is imposed on the signal arriving from one of the two inputs, then the comparator's response is maximal only when the input fields are activated sequentially in the appropriate order (not in the reverse order). D_{\max} in RDKs was taken as an estimate of the spatial offset between the detector's input receptive fields, and T_{\max} was taken as an estimate of dt (e.g. Baker & Braddick, 1985).

Low-pass spatial filtering of RDKs has been found to extend the D_{\max} limit, with D_{\max} increasing in proportion with the space constant of the filter [at least above a certain minimum filter space constant (e.g. Morgan, 1992)]. According to the physiological account of detection limits, the increase in D_{\max} with spatial blurring can be related to the variation in receptive field size of motion sensors, under the assumption that sensors tuned to lower frequencies can detect greater spatial displacements (Cleary & Braddick, 1990; Burr, Ross & Morrone, 1986). Of course broadband patterns contain low as well as high spatial frequencies, but Cleary and Braddick (1990) argue that performance with these patterns is

*Experimental Psychology, University of Sussex, Brighton BN1 9QG, England [Email georgem@epvax.sussex.ac.uk].

†To whom all correspondence should be addressed.

limited by incoherent responses generated in high-frequency detectors by large displacements. Alternatively, the D_{\max} spatial limit can be viewed as representing an informational limit rather than a physiological limit on discrimination. The arguments put forward by Bischof and DiLollo (1990), Eagle and Rogers (1991), and Morgan and Mather (1994) can be summarized as follows. The probability of a correct response at a certain displacement depends on how great that displacement is relative to the density of features in the "neural image" of the dot pattern (i.e. after spatial filtering by early receptive fields). Beyond a certain limiting displacement (which sets D_{\max}), the ambiguity caused by the intrusion of false matches between features in the two stimulus frames is too great to allow reliable discrimination of direction. Thus low-pass spatial filtering extends D_{\max} because it decreases the density of features in the image and ameliorates the false matching problem. The informational theory can be re-cast in the Fourier domain by first considering RDKs as spatiotemporal samples of drifting patterns. Sampling introduces artefacts (alias signals) which become more intrusive as sampling rate declines (i.e. spatial displacement increases). Beyond a limiting displacement, corresponding to D_{\max} , alias signals prevent reliable direction discrimination. In this version of the theory, low-pass filtering extends D_{\max} because it removes the high frequencies that require high sampling rates to avoid aliasing. The space domain and frequency domain versions of the informational theory are clearly homologous, but we shall argue below that the frequency domain version is more general because it can be applied to the effects of temporal filtering as well as to the effects of spatial filtering.

In this paper we examine whether *temporal* filtering extends the temporal limit T_{\max} in the same way that spatial filtering extends D_{\max} . The physiological account of spatial limits can be extended to predict that low-pass temporal filtering will extend T_{\max} limits because it will isolate responses in low temporal frequency tuned sensors that can cope with longer inter-stimulus intervals. However, temporal filtering should be less effective than spatial filtering, because there are relatively few channels (two or three) tuned to different temporal frequencies but at least twice as many channels tuned to different spatial frequencies (Anderson & Burr, 1985; Burr *et al.*, 1986; Hess & Snowden, 1992). On the other hand, the informational account predicts the temporal filtering will produce equivalent effects to spatial filtering—both remove high spatiotemporal frequencies responsible for aliasing, so both should extend detection limits.

EXPERIMENT 1

In conventional presentations of RDKs, the onset and offset of each frame is abrupt, involving stepwise increases and decreases in dot contrast (upper traces in Fig. 1). We applied temporal smoothing to the onset and offset of each frame, using a cumulative Gaussian smoothing function, so that dots appeared and disappeared gradually (lower traces in Fig. 1). To assess the

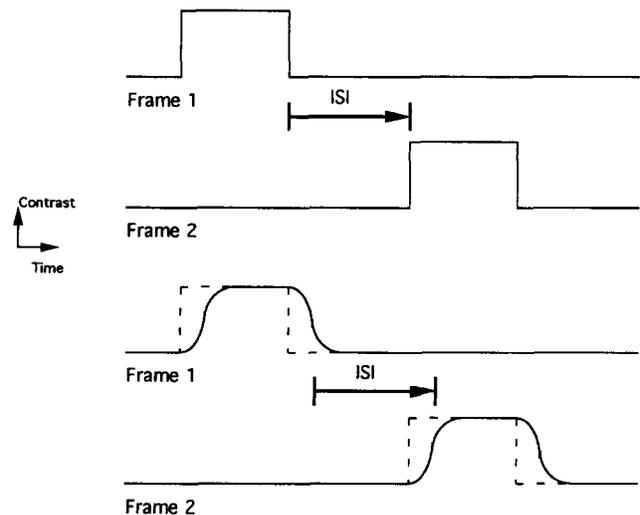


FIGURE 1. Time-course of stimulation in RDK displays. The top two traces depict the abrupt onset and offset of each frame in conventional displays. The bottom two traces depict the original abrupt onsets and offsets, and corresponding temporally smoothed changes used in the present experiments.

effect of temporal smoothing, we measured T_{\max} for discriminating the direction of a fixed small displacement, as a function of the time constant of the smoothing function. A fixed frame duration was used, long enough for maximum dot intensity to be reached even for the longest filter time constants.

Method

Subjects. Six observers participated, the two authors and four others who were unaware of the purpose of the experiment.

Apparatus and stimuli. Visual displays were generated by a PC-compatible computer equipped with a high-resolution raster graphics sub-system, and displayed on a Hitachi 14MVX monitor (P22 phosphor) at a frame rate of 83 Hz (non-interlaced). In between trials, the screen was uniformly illuminated at $30 \text{ cd} \cdot \text{m}^{-2}$, except for a small red fixation cross at its centre. 625 pairs of dots (each subtending 2 arc min) were plotted at random positions in a central 4×4 arc deg screen area (viewing distance 114 cm), in two different look-up table (LUT) numbers and offset spatially by a fixed horizontal distance. The random locations of dot pairs were constrained to avoid overlaps between pairs, so on average 13% of the allowable dot positions were filled in each frame. By means of LUT manipulation, the intensity of each dot could be varied independently between background level ($30 \text{ cd} \cdot \text{m}^{-2}$) and maximum intensity ($70 \text{ cd} \cdot \text{m}^{-2}$). When one dot was made to appear and disappear before the other, an impression of apparent motion was seen. To create temporal smoothing at the onset and offset of each "frame", dot intensity increased gradually from grey to white over a series of TV frames, and then decreased gradually back to background level. The (γ -corrected) changes in grey level conformed to a cumulative Gaussian profile. The following Gaussian time constants were used in different presentations:

0 (i.e. conventional onset/offset), 12, 24, 48 and 96 msec. At a time constant of 0 msec, each dot changed from background intensity to maximum intensity in successive TV frames. At a time constant of 12 msec, each dot faded up from background to maximum intensity in four TV frames, while the longest time constant of 96 msec involved fading over 32 TV frames (in all cases the transition from background intensity to maximum intensity occurred over a time period spanning four time constants of the Gaussian function). Frame duration can be defined as the interval between the onset of each frame (time at which the dots begin to increase in grey level), and its offset (time at which the dots begin to decrease in grey level), as depicted in Fig. 1. Equivalently, it can be defined as the time interval separating the mid-points of the onset and offset smoothing functions. Inter-stimulus interval (ISI) can be defined as the time period separating the offset of the first frame and the onset of the second frame (see Fig. 1). Frame duration was fixed at 384 msec (32 TV frames), just sufficient for dots to reach maximum intensity at the longest time constant used. ISI was varied in different presentations, to permit estimation of T_{max} .

Design and procedure. The design involved 50 factorial combinations of five ISIs (24, 48, 96, 192 and 384 msec), five filter time constants (0, 12, 24, 48 and 96 msec), and two displacements (2 and 16 arc min). Data were accumulated for each subject over 10 experimental sessions, each involving randomly ordered presentations of stimuli from all conditions, until 40 trials had been presented for each stimulus. Each trial involved a single two-frame presentation of the stimulus, direction selected at random, following which the observer pressed one of two buttons to indicate perceived direction. The spatial arrangement of dots varied randomly from trial to trial.

Results and discussion

All subjects showed near-perfect direction discrimination at small ISIs, with performance declining as ISI increased. Figure 2 plots mean discrimination performance as a function of ISI, with smoothing function time constant as the parameter. Figure 2(a) depicts results with the small displacement, and Fig. 2(b) depicts results with the large displacement (SEs have been omitted for clarity, but were reliably below 5%). In a three-factor ANOVA, the effects of ISI and filter time constant were highly significant (ISI, $F_{4,20} = 105.51, P = 0.0001$; time constant, $F_{4,20} = 10.55, P = 0.0011$), as was their interaction ($F_{16,80} = 3.86, P = 0.0001$). Thus, temporal smoothing enhanced performance at longer ISIs. To quantify this effect, T_{max} was calculated for each filter time constant, defined as the 75% point of the psychometric function relating direction discrimination to ISI. Figure 3 shows T_{max} as a function of filter time constant.

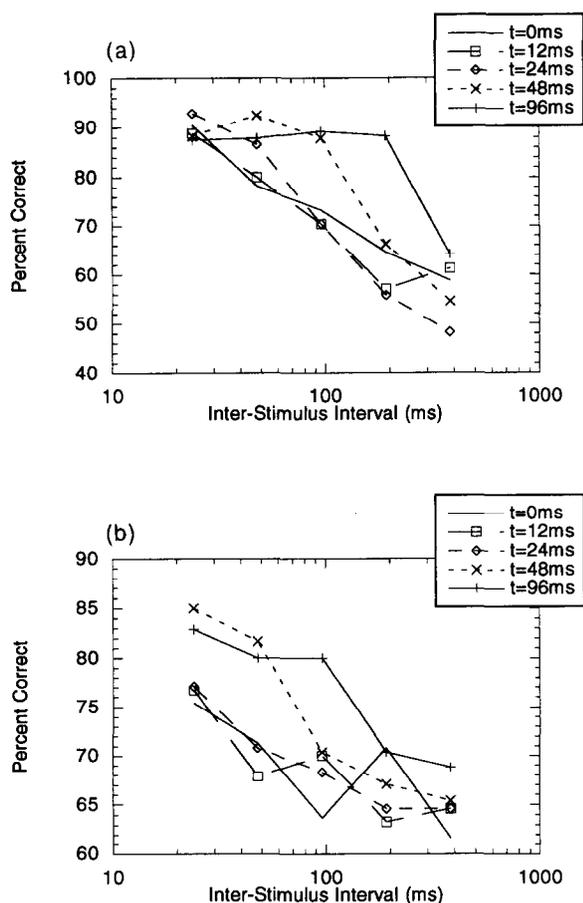


FIGURE 2. Results of Expt 1, showing mean percentage correct in a direction discrimination task as a function of ISI. (a) Results obtained using a displacement of 2 arc min; (b) results obtained using a displacement of 16 arc min. Different curves in each figure represent results when the onset and offset of each frame was temporally smoothed using Gaussian filters with different time constants (specified in the inset).

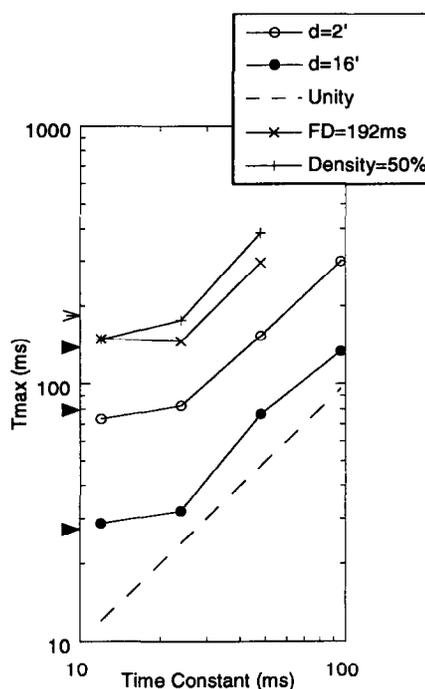


FIGURE 3. T_{max} values calculated from the psychophysical data obtained in Expts 1, 2 and 3. T_{max} is defined as the ISI yielding 75% correct in the direction discrimination task, calculated by linear interpolation. Circles—results from Expt 1 at two displacements; \times —results from Expt 2 involving a shorter frame duration; $+$ —results from Expt 3 involving 50% black-white patterns.

Open circles depict results at the short displacement, and solid circles depict results at the long displacement. The arrows on the ordinate adjacent to each data set give performance using unfiltered patterns (i.e. time constant = 0 msec). It is clear that, beyond a time constant of 24 msec, T_{\max} increases in proportion with the time constant of the smoothing filter.

The effect of temporal filtering is strikingly similar to the effect of spatial filtering, in extending performance limits in proportion to the degree of filtering. Before discussing possible explanations for this effect, supplementary experiments will be described which attempted to replicate the result under different stimulus conditions.

EXPERIMENT 2

The data of Expt 1 were collected using a single, rather long, frame duration (384 msec). To assess whether the effects generalize to shorter frame durations, Expt 1 was repeated using a frame duration of 192 msec. Recall that the longer frame duration was chosen to allow temporally filtered patterns to reach maximum contrast at the longest filter time constant. We applied the same restriction in the second experiment, so the longest temporal filter time constant used was 48 msec.

Method

Subjects. Four subjects took part. All had performed in Expt 1.

Apparatus, stimuli and procedure. All details were identical to those given above for Expt 1, with the following exceptions. Only one displacement was used (2 arc min); frame duration was fixed at 192 msec; and only four temporal filter time constants were used (0, 12, 24 and 48 msec).

Results and discussion

T_{\max} values were calculated from the mean data relating percent correct discrimination to ISI, and are plotted in Fig. 3 (\times). As in Expt 1, temporal smoothing enhanced performance at longer ISIs. In a two-factor ANOVA, the effects of ISI and filter time constant were significant (ISI, $F_{4,12} = 22.32$, $P = 0.0001$; time constant, $F_{3,9} = 5.72$, $P < 0.018$), as was their interaction ($F_{12,36} = 2.05$, $P < 0.049$). The results of Expt 2 replicate those of Expt 1, though it seems that a shorter frame duration improves the overall level of performance. Experiment 3 tested whether the results of the first two experiments can be replicated using high-density dot patterns.

EXPERIMENT 3

In Expts 1 and 2, only 625 visible dots were visible in a 120×120 dot area of the display. Dot density is known to have a major effect on performance (e.g. Morgan & Fahle, 1992), so it is important to establish whether

results generalize to dense dot patterns. Experiment 3 employed 50% black-white random dot patterns.

Method

Subjects. Five observers took part, three of whom had served in Expt 1.

Apparatus and stimuli. Equipment was identical to that used in previous experiments. However, a different animation technique was employed, which allowed presentation of dense random-dot patterns. The two frames of the random-dot pattern were presented using an interleaving technique, so that frame 1 was displayed in even-numbered TV frames, and frame 2 was displayed in odd-numbered TV frames. The TV refresh rate was 83 Hz. Each RDK frame contained 50% black-white random dots and, as usual, dots in frame 2 were given a coherent spatial displacement relative to dots in frame 1 (fixed at 4 arc min, either leftward or rightward). If dots in frame 2 were set to zero contrast, while dots in frame 1 were set to maximum contrast, then only frame 1 was visible in the interlaced display. If, during the flyback period following a view of dot frame 1, dots in frame 2 were switched to maximum contrast and dots in frame 1 were switched to zero contrast, apparent motion was seen between frames 1 and 2 in the usual manner. To create temporal smoothing at the onset and offset of each frame, contrast was increased and decreased gradually over a series of TV frames. Unlike the technique used in Expts 1 and 2, this technique places no constraints on the spatial properties of the dot pattern. However, the interlacing procedure halves the effective refresh rate of each frame of dots from 83 to 46.5 Hz, resulting in coarser control of temporal modulation.

Temporal modulation of dot contrast conformed to a cumulative Gaussian profile. The following Gaussian time constants were used: 0, 12, 24 and 48 msec. The dot pattern consisted of an array of 64×64 dots, with each dot subtending 4 arc min on a side. The mean luminance of all patterns was $60 \text{ cd} \cdot \text{m}^{-2}$. Dot contrast in each RDK frame was 0.48 (allowing for the attenuation caused by interlacing). The duration of each RDK frame was fixed at 384 msec. As before, ISI was varied in different presentations, to allow estimation of T_{\max} .

Design and procedure. Stimuli were presented 20 times in random order to each subject in a single session. As in previous experiments, each trial consisted of a single two-frame exposure of the stimulus, direction selected at random, following which the observer pressed one of two buttons to indicate perceived direction.

Result and discussion

As in previous experiments, T_{\max} values were calculated from the data, and are plotted in Fig. 3 (+). T_{\max} increased in proportion with smoothing filter time constant for the two longer time constants, replicating earlier results. Having established that temporal smoothing extends T_{\max} in both low density and high density patterns, we now consider possible explanations for the effect.

EXPERIMENT 4

Do results reflect the functional properties of motion detectors, or the varying information content of filtered kinematograms? The proportional increase in T_{\max} with filter time constant, equivalent to previously reported effects of spatial filtering on D_{\max} , is inconsistent with the physiological explanation based on selective activation of low-frequency channels. There should be less scope for improved detection in the case of temporal filtering compared to spatial filtering, because there are fewer temporally tuned channels. A possible alternative functional explanation runs as follows. Earlier, we defined ISI as the time interval between the offset of frame 1 (or the time at which the intensity of its dots falls through the midpoint of the temporal smoothing function) and the onset of frame 2 (or the time at which the intensity of its dots rises through the midpoint of the smoothing function). Defined in these terms, the maximum ISI supporting direction discrimination increases in proportion with the time constant of the smoothing function, as shown in Fig. 3. However, it may be that as far as the visual system is concerned the *effective* temporal interval is the time period between the dots in frame 1 falling below some minimum contrast level, and the dots in frame 2 rising above this threshold. If the effective interval supporting motion detection (d_t) was fixed, as described in the Introduction for simple motion detectors, then our results can be explained, at least qualitatively. At any one ISI, temporal smoothing brings the threshold increments and decrements from successive frames closer together in time, reducing the *effective* interval between frames. A fixed *effective* interval (corresponding to d_t) would occur at progressively longer ISIs as filter time constant increases. Clearly, this 'threshold' account predicts that an increase in dot contrast should produce similar effects to an increase in filter time constant. If pattern contrast is increased while all other parameters remain fixed, the *effective* interval between frames will decrease, because dots in the first frame will fall below threshold later after they being fading off, and dots in the second frame will rise above threshold sooner after they begin fading on. To test the prediction of contrast dependency, we measured T_{\max} in dense, temporally filtered patterns as a function of contrast.

Method

Subjects. Four subjects took part, three of whom had served in previous experiments.

Apparatus and stimuli. The same equipment and stimulus generation techniques were used as in the previous experiment. Pattern contrast in each trial was selected from the following four values: 0.05, 0.1, 0.2 and 0.4 (values adjusted for the effect of interleaving). Temporal filter time constant was fixed at 48 msec, and frame duration was fixed at 384 msec. ISI was manipulated to allow estimation of T_{\max} .

Design and procedure. Stimuli at different ISIs were presented 40 times each in random order to each subject, spread over four sessions. As in previous experiments,

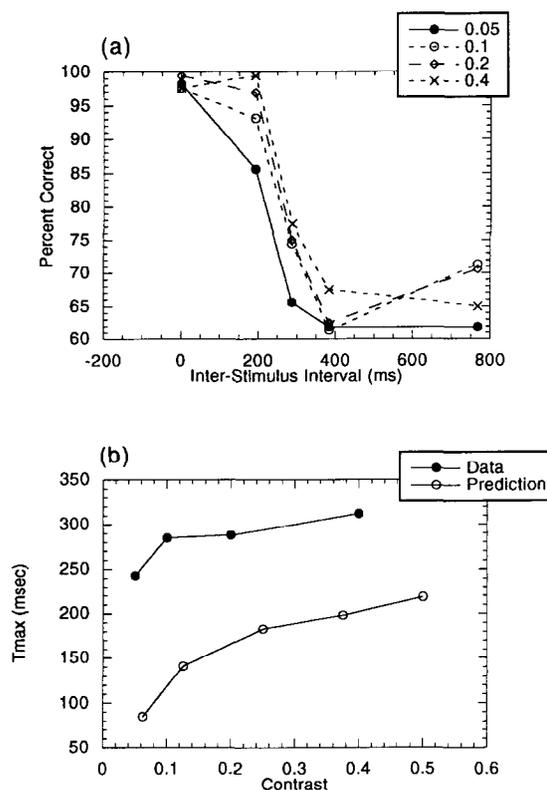


FIGURE 4. Results of Expt 4. (a) Mean percentage correct as a function of ISI, with different curves depicting results at different contrasts. (b) T_{\max} as a function of contrast (●), calculated from the psychophysical data as described in Fig. 3. ○ show the contrast effect predicted by the threshold explanation (see text for details).

each trial consisted of a single two-frame exposure of the stimulus, direction selected at random, following which the observer pressed one of two buttons to indicate perceived direction.

Results and discussion

Figure 4(a) plots mean percentage correct as a function of ISI and contrast, and Fig. 4(b) plots T_{\max} as a function of contrast (solid symbols), calculated from the data in Fig. 4(a). There is only a relatively small effect of contrast. To assess the plausibility of the threshold explanation, we derived quantitative predictions for the effect of contrast as follows. Assume that the threshold level of contrast necessary to attain 75% correct in the task is fixed at 2.5%, that the dot patterns have a maximum contrast of 100%, and that the effective temporal interval supporting motion detection (d_t) is fixed at 60 msec. If a temporal smoothing filter with a time constant of 48 msec is applied to each frame (as in the experiment), then T_{\max} will be reached at an ISI of 248 msec, calculated as $(60 + 94 + 94)$ msec. The 60 msec corresponds to the fixed interval d_t ; 94 msec is the time interval required after the offset of frame 1 for its dots to reach 2.5% contrast (1.96 SDs after passing the midpoint of the smoothing function). A further 94 msec is added because frame 2 will reach 2.5% contrast 1.96 SDs before its dots reach the midpoint of the smoothing function (recall that ISI is defined as the interval separating the mid-points). If contrast is reduced to 50%,

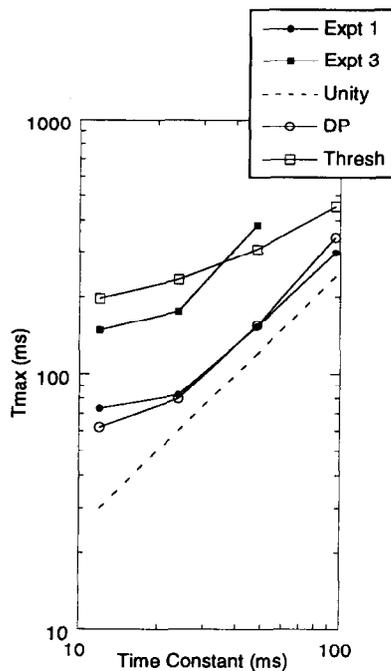


FIGURE 5. Comparison of psychophysical T_{\max} values obtained in Expts 1 and 3 (solid symbols) with values predicted by the threshold model (\square) and the sampling model (\circ). Psychophysical data is replotted from Fig. 3, and derivation of predictions is described in the text. Note that DP predictions have been scaled to msec by equating one time-frame in Fig. 9 to 100 msec (i.e. re-scaling the abscissa in Fig. 9 so that the predicted curves superimpose on the data curves in Fig. 2).

T_{\max} falls to 218 msec ($60 + 79 + 79$), because the 2.5% threshold is reached 1.65 SDs from the mid-point of the function. A further reduction in contrast to 25% shortens T_{\max} to 182 msec ($60 + 61 + 61$), because the 2.5% threshold is reached 1.28 SDs away from the mid-point of the function. The predicted decline in T_{\max} with contrast is plotted in Fig. 4(b) as open symbols. The relatively small effect of contrast evident in the data is well predicted by the threshold model. However, predicted T_{\max} values are clearly much lower than those obtained. Predicted values can be brought into alignment with the psychophysical data by increasing the value of dt from 60 to 160 msec, which shifts the predicted curve up the ordinate in Fig. 4 (changes to the threshold value of 2.5% only serve to shift the curve sideways on the abscissa, re-scaling the contrast values associated with particular levels of performance). Such a large dt value seems implausible, since it implies rather large T_{\max} values even in unfiltered patterns. We derived further predictions from the threshold model for the variation in T_{\max} with filter time constant, using the same procedure as above. Threshold was fixed at 2.5%, and dt was set at 160 msec, for consistency with the results of Expt 4. Predictions are plotted against data from Expts 1 and 3 in Fig. 5 (recall that the stimulus used in Expt 3 was identical to that used in Expt 4). The threshold model (\square) predicts a more gradual decline in T_{\max} than is

evident in the psychophysical data (solid symbols). Predicted T_{\max} approaches an asymptotic value of 160 msec (the fixed value of dt) as filter time constant approaches zero. The predicted function can be steepened by decreasing dt , but only approaches the proportionality shown in the data when dt approaches zero. So the threshold explanation requires large dt values to accommodate the contrast data, but small dt values to accommodate the filtering data. It cannot offer a good fit for both sets of data with a single dt value.

However, the picture is complicated by the fact that motion response is likely to be a nonlinear function of contrast. So, as a further test of the threshold explanation that does not require assumptions about contrast response, we conducted a final experiment to compare direction discrimination performance for three stimuli:

Step—a two-frame RDK without temporal filtering (rectangular onset and offset);

Gaussian—a two-frame RDK with a Gaussian smoothing filter applied to the onset and offset of each frame (time constant 96 msec);

Tailed—identical to Gaussian, except that only the offset tail of the first frame and the onset tail of the second frame were visible.*

It is straightforward to predict that performance with the Gaussian stimulus will be better than performance with the Step stimulus, as found in previous experiments. The threshold explanation of temporal filtering effects predicts that performance with the Tailed stimulus will be comparable to performance with the Gaussian stimulus, because the two stimuli have the same “tails” which mediate extended detection levels at large ISIs. The informational account, which explains filtering effects in terms of removing sampling artefacts generated by high frequencies, predicts that performance with the Tailed stimulus will be worse than performance with the Gaussian stimulus because the Tailed stimulus contains high frequencies introduced by the sudden onset and offset of each tail.

EXPERIMENT 5

Method

Subjects. Four subjects took part, both authors and two naive observers who had participated in previous experiments.

Apparatus, stimuli and procedure. All equipment and stimulus details are identical to those given for Expt 1. Only one displacement was used (2 arc min), and frame duration was fixed at 384 msec. The Step and Gaussian stimuli were identical to corresponding stimuli in Expt 1 (i.e. time constants of 0 and 96 msec respectively). The Tailed stimulus was identical to the Gaussian stimulus except that only the descending portion of frame one and the ascending portion of frame two were visible. Five different ISI values were used in different trials (24, 48, 96, 192 and 384 msec). A total of 600 trials (3 stimuli \times 5 ISIs \times 40 trials) were presented in random order over two experimental sessions.

*We are grateful to an anonymous referee for suggesting this stimulus.

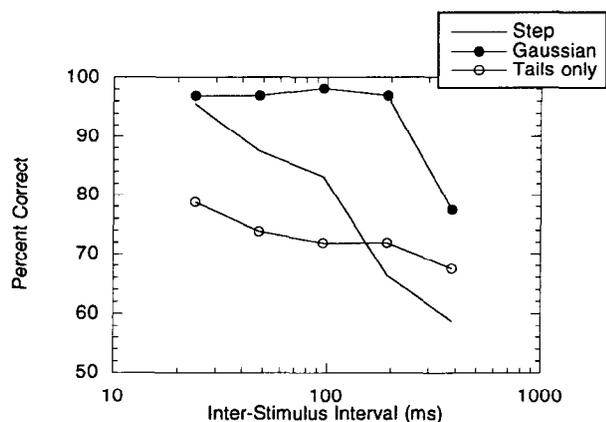


FIGURE 6. Results of Expt 5, show mean percentage correct in a direction discrimination task as a function of ISI, for three different stimuli (see key).

Results and discussion

Figure 6 shows mean direction discrimination performance as a function of ISI, for each of the three stimuli. SEs have been omitted for clarity, and were on average 4.3%. Results for the Step and Gaussian stimuli are very similar to those in Fig. 2—discrimination performance is much better for the temporally filtered stimulus than for the unfiltered stimulus. Crucially, performance with the Tailed stimulus is much worse than performance with either of the other stimuli at most ISIs, indicating that discrimination is not mediated by responses to the tails of the two frames.

Taking the results of Expts 4 and 5 together, we conclude that the threshold model does not provide a satisfactory explanation for the effects of temporal smoothing on detection limits. An alternative approach will now be proposed, which can embrace both temporal filtering effects and spatial filtering effects.

A SAMPLING MODEL OF FILTERING EFFECTS IN RDKS

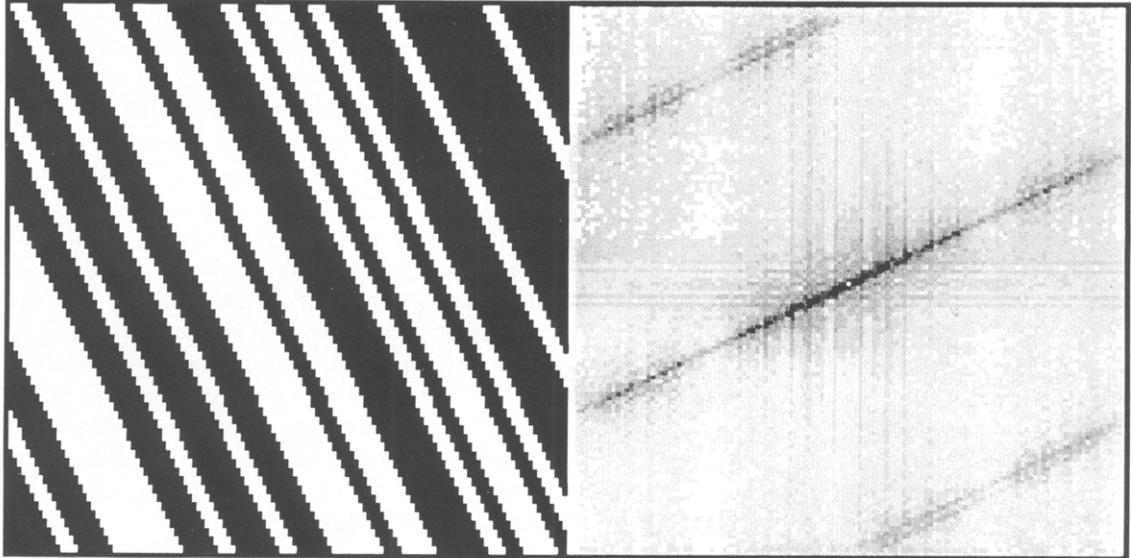
As a starting point, it is useful to consider the RDK as a directly sampled version of a continuously drifting random-dot pattern. The left-hand side of Fig. 7 shows xt plots of a row of random black-white elements (horizontal axis) at different instants in time (vertical axis). Figure 7(a) depicts a continuously drifting pattern (within the resolution limits of the pixel-based data), and Fig. 7(b) depicts a sampled version in which the row of elements undergoes discrete rightward displacements at successive time intervals (separated by uniform grey ISIs). The right-hand side of Fig. 7 shows the spatio-temporal Fourier transform of each xt plot. In the case of continuous drift (upper spectrum), Fourier energy is confined to a tilted line in frequency space. The angle of tilt specifies velocity, and in crude terms energy in the top-right and bottom-left quadrants signifies rightward motion, and energy in the top-left and bottom-right quadrants signifies leftward motion (see Watson, Ahumada & Farrell, 1986). In the sampled pattern

(lower spectrum), energy corresponding to the signal again falls along a line passing through the origin, but the spectrum contains repeating replicates of the signal. The distance between replicates depends on sampling rate, in accordance with the Sampling Theorem. Thus, the horizontal separation between replicates (i.e. on the spatial frequency axis) depends on the size of the spatial displacement separating samples in the xt plot, and the vertical separation between replicates (i.e. on the temporal frequency axis) depends on the duration of the blank ISI in the xt plot. A number of current models of low-level motion detection in the visual system assume that detectors sample small regions in spatiotemporal frequency space: detectors tuned to rightward motion have receptive fields symmetrically placed in the top-right and bottom-left quadrants, and detectors tuned to leftward motion have receptive fields in the top-left and bottom-right quadrants (see Adelson & Bergen, 1985; Watson & Ahumada, 1985). In the case of continuous drift, the stimulus clearly offers a strong signal for detectors tuned to rightward motion, since energy is confined to the “rightward” quadrants. In the case of sampled motion, detector response will be contaminated by energy from replicates spilling over into the “leftward” quadrants. The contamination becomes more severe as the spatial and/or temporal sampling interval increases, and the replicates move closer in frequency space to the signal. Watson *et al.* (1986) and Burr *et al.* (1986) demonstrated that subjects’ ability to discriminate between continuous motion and sampled motion can be explained by the intrusion of alias signals falling inside the visual system’s spatiotemporal “window of visibility”. In the case of direction discrimination in unfiltered RDKs, alias signals falling inside the window of visibility should impair performance at longer ISIs and/or larger spatial displacements, in agreement with the standard psychophysical results obtained from these stimuli.

Turning to filtering effects, why does temporal smoothing improve performance at longer ISIs? Temporal smoothing is a form of low-pass filtering, which effectively removes high spatiotemporal frequencies. Since alias signals are generated by high frequencies (which require higher sampling rates), low-pass filtering reduces the aliasing problem, allowing high levels of performance to extend to lower sampling rates (i.e. longer ISIs). Low-pass filtering is a standard technique in signal processing to remove sampling artefacts from sampled signals. Figure 8 shows examples of xt plots and their Fourier spectra before and after temporal smoothing. The top xt plot shows a two-frame RDK of the kind used in our experiments. Its spatiotemporal Fourier spectrum is severely contaminated by alias signals. The bottom xt plot is identical to the top plot, except that it has been smoothed or blurred along the vertical axis, to simulate the effects of temporal smoothing in our experiments. The frequency spectrum of the smoothed pattern contains no energy in the high spatiotemporal frequencies which tend to carry alias signals in the unfiltered pattern, and should therefore offer a more coherent signal for detectors tuned to rightward motion.



(a)



(b)

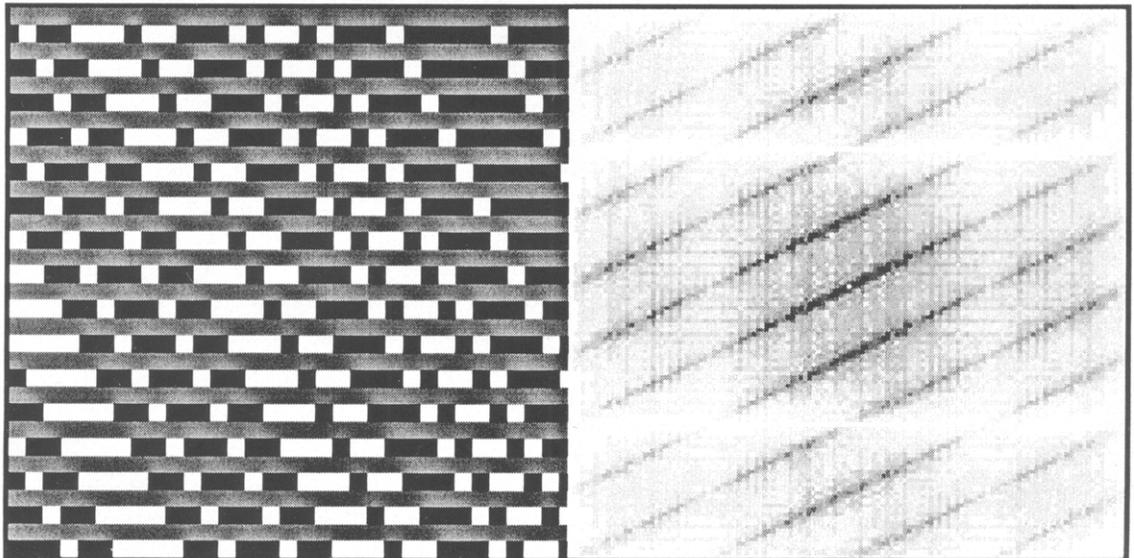
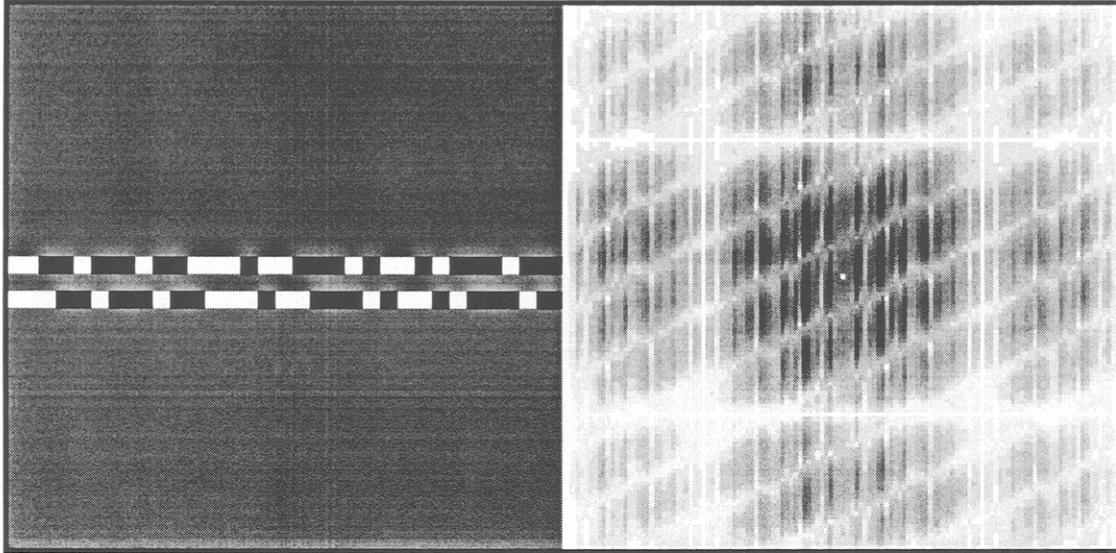


FIGURE 7. xt plots of random element patterns, and their spatiotemporal Fourier transforms. (a) Continuous motion; (b) sampled motion, as used in RDKs. Each xt plot was held as a 128×128 element array, so each transform represents 64 spatial frequencies horizontally, and 64 temporal frequencies vertically. Each pixel in the transform represents the Fourier amplitude at that frequency, $|F(u,v)|$, scaled conventionally as follows: $\log[1 + |F(u,v)|]$. Scaled amplitudes were quantised to 256 grey levels for display, with darker pixels representing higher amplitudes.



(a)



(b)

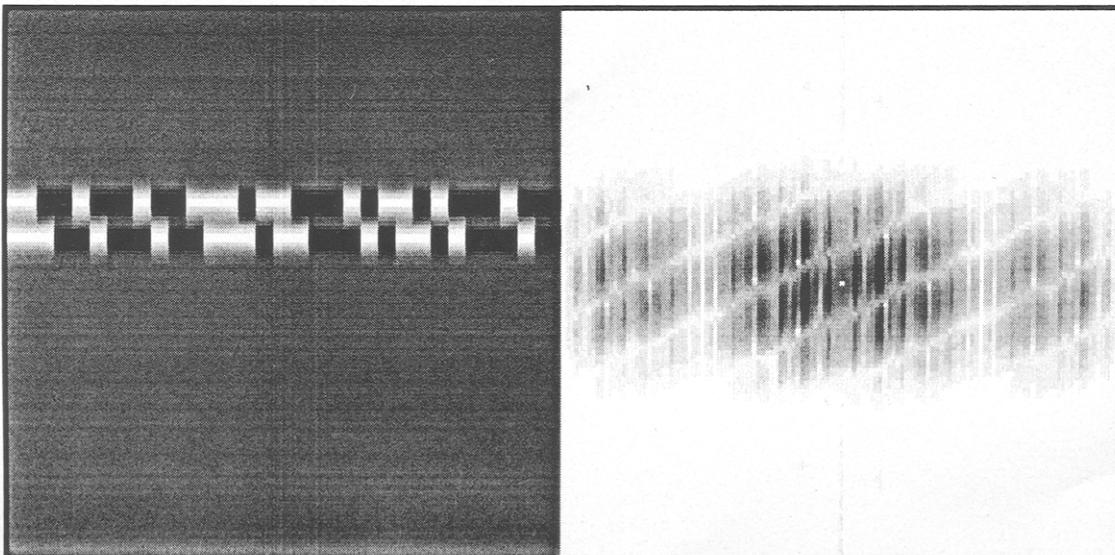
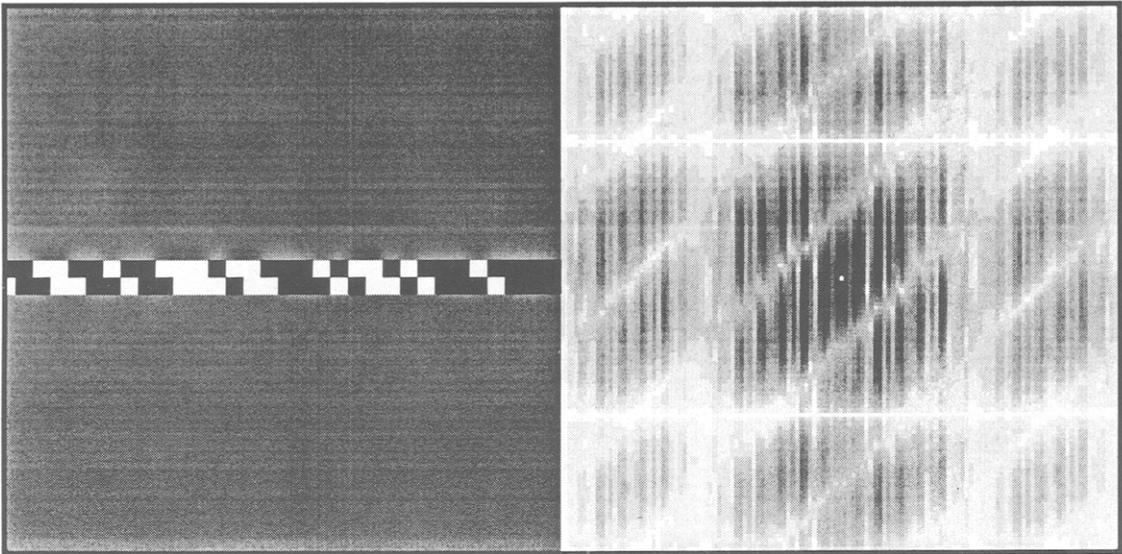


FIGURE 8. xt plots and Fourier transforms of a two-frame random-element kinematogram before (a) and after (b) application of a temporal smoothing filter (i.e. along the vertical axis). Conventions as in Fig. 6.



(a)



(b)

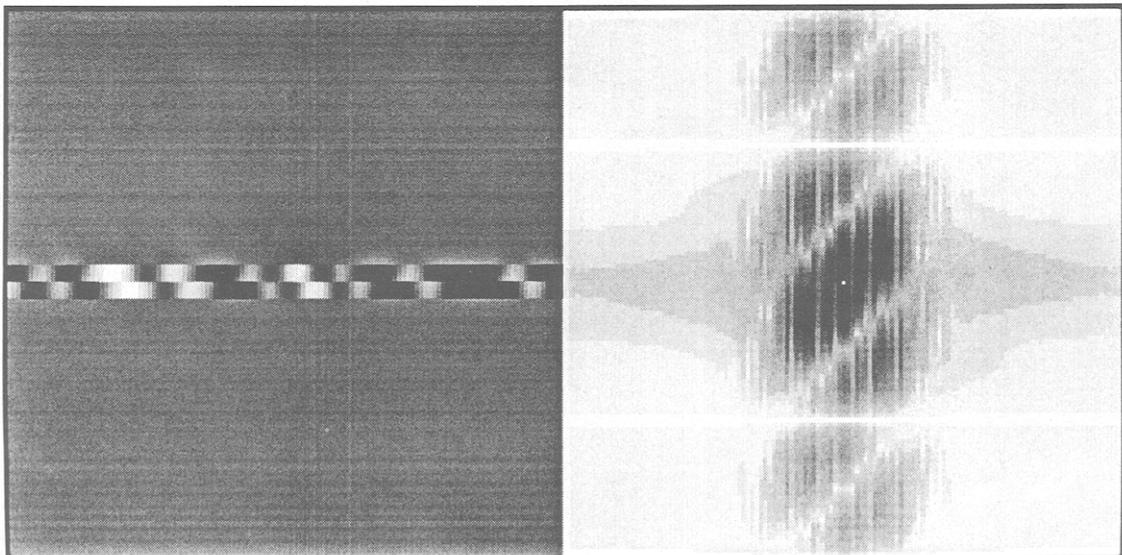


FIGURE 9. xt plots and Fourier transforms of a two-frame random-element kinematogram before (a) and after (b) application of a spatial smoothing filter (i.e. along the horizontal axis). Conventions as in Fig. 6.

The same ‘‘sampling’’ explanation can obviously be applied to the improvement in displacement limits resulting from low-pass *spatial* filtering of RDKs. Figure 9 shows *xt* plots and Fourier transforms for a two-frame RDK before and after blurring along the horizontal (spatial) axis. Again, blurring removes alias signals at high frequencies.

Why should Gaussian blurring result in a *proportional* increase in T_{max} and D_{max} limits? Informally, we can attribute the proportional relationship to the fact that the Fourier transform of a Gaussian function is itself a Gaussian function, and the SD of the latter is inversely proportional to the SD of the former. So increases in filter time (or space) constant correspond to proportional decreases in frequency response. We attempted a more rigorous assessment of the sampling explanation by deriving quantitative predictions for the effects of spatial and temporal filtering, based on the information available in Fourier transforms of the patterns. First, a set of *xt* plots was created, each containing a row of 32 random black–white elements in 32 time-frames. Each *xt* plot was similar to those shown in Figs 8 and 9 (i.e. only two time-frames contained dots and the remainder were set to grey). Plots differed in terms of the frame-to-frame displacement depicted (expressed as multiples of dot width), and the ISI (expressed as multiples of time-frames). Further *xt* plots were generated from this basic set by filtering each *xt* plot with a one-dimensional temporal Gaussian filter [i.e. on the vertical axis only, as in Fig. 8(b)], or with a one-dimensional spatial Gaussian filter [i.e. on the horizontal axis only, as in Fig. 9(b)]. The space constant or time constant of the filter was varied parametrically in different *xt* plots. We took the Fourier transform of each *xt* plot and, as a simple estimate of the information available for direction discrimination, computed directional power (DP), defined as the ratio of summed rightward power to summed leftward power in the transform (i.e. ratio of summed power in top-right and bottom-left quadrants to summed power in top-left and bottom-right quadrants). A display containing no net motion signal would yield a DP ratio of 1.0. Ratios above 1.0 indicate net rightward energy, and ratios

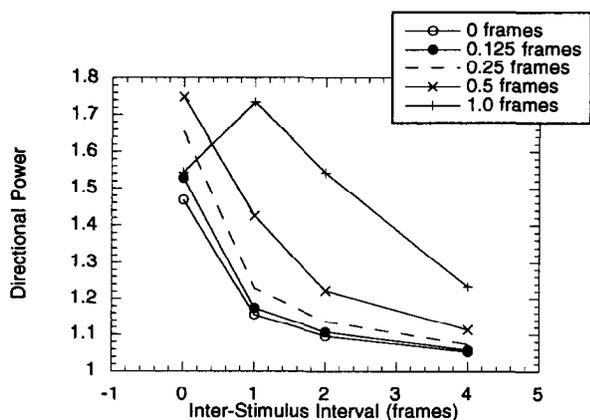


FIGURE 10. DP as a function of ISI (abscissa) and filter time constant (different curves), computed from spatiotemporal Fourier spectra such as that shown in Fig. 7. Computations are described in detail in the text.

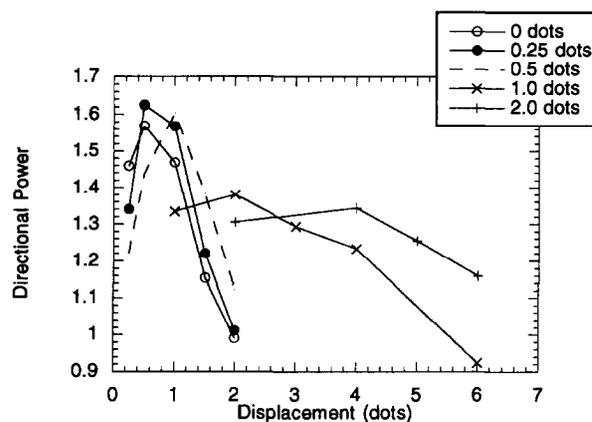


FIGURE 11. DP as a function of displacement (abscissa) and filter space constant (different curves, specified in inset), computed from spatiotemporal Fourier spectra such as that shown in Fig. 8. Computations are described in detail in the text.

below 1.0 indicate net leftward energy. A number of recent papers have used variants of this DP measure to estimate the information available for Fourier-based motion detection in simple displays (e.g. Doshier, Landy & Sperling, 1989; Nishida & Sato, 1992; Boulton & Baker, 1993). Figure 10 plots DP as a function of ISI for a fixed frame-to-frame displacement of one dot width, with temporal filter time constant as the parameter. DP declines as ISI increases, consistent with standard psychophysical results. However, temporal filtering permits higher DP values at long ISIs, again consistent with the psychophysical data in Fig. 2. Assuming that discrimination performance depends directly on DP, an arbitrary value of DP was chosen to represent a specific level of performance (equivalent to the arbitrary percent correct

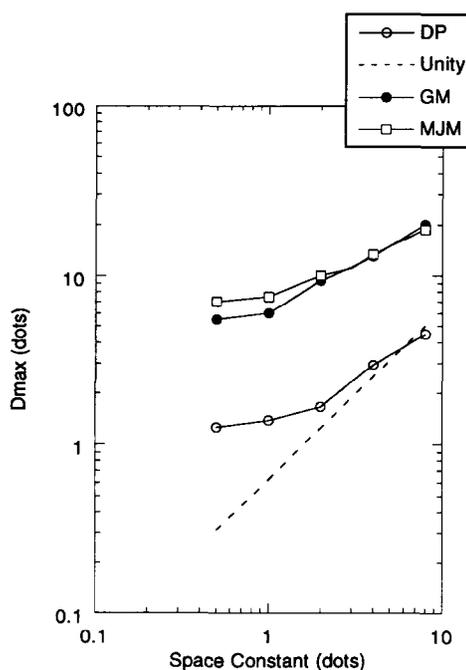


FIGURE 12. Comparison of D_{max} computed from DP values in Fig. 11 with psychophysical data for two subjects (GM and MJM) replotted from Morgan and Mather (1994).

used to define T_{\max}), and T_{\max} values were calculated as the ISIs yielding this criterion DP level at different filter time constants. Figure 5 plots computed T_{\max} against filter time constant, along with psychophysical data from two experiments. Note that the computed T_{\max} values increase in proportion with the time constant of the filter, just as the psychophysical T_{\max} values do. Turning to effects of spatial filtering on D_{\max} values, Fig. 11 plots DP as a function of displacement (zero ISI), with filter space constant as the parameter. Again, DP levels decline with displacement, but are higher for low-pass filtered patterns. Using the same criterion level of DP as before, D_{\max} values were calculated from the data, and are plotted against some psychophysical data from Morgan and Mather (1994) in Fig. 12. The DP measure correctly predicts the proportional increase in D_{\max} with filter space constant. It is important to note that the only free parameter in the computations is the criterion level of DP, and the result of choosing different criterion values would be to shift the DP curves in Figs 5 and 12 along the ordinate. The computations weighted all frequency components equally, without regard to a "window of visibility". Despite this simplicity, the sampling account can explain the proportional effects of filtering on performance limits in RDKs, at least at larger filter time constants and space constants. It may be necessary to introduce a "window of visibility" to maximize the fit of the DP data with the psychophysics at very small filter time and/or space constants (where filtering has little effect because it removes frequency components which fall outside the window).

As mentioned in the Introduction, spatial filtering effects have been explained in the literature in terms of changes in the mean separation of features in the "neural" image of blurred dot patterns (Bischof & DiLollo, 1990; Morgan, 1992; Eagle & Rogers, 1991). We prefer to cast the explanation in the spatiotemporal frequency domain because it provides an integrated framework within which to consider both spatial and temporal filtering effects. Of course the sampling explanation falls well short of providing a comprehensive theory of direction discrimination in RDKs, since it includes no statements about the visual processes which extract Fourier energy, or how their responses determine discrimination performance. However, the sampling approach does make it clear that important features of the psychophysical data can be related simply to the physical information content of random-dot patterns, and it is not necessary to attribute those features to physiological properties of the visual system.

SUMMARY AND CONCLUSIONS

The upper ISI supporting direction discrimination (T_{\max}) was found to increase when each frame of the random-dot pattern was temporally smoothed at its onset and offset. An explanation based on contrast-dependent responses in simple motion detectors could accommodate the psychophysical data satisfactorily. Computational modelling results show that the effect

of temporal filtering, and previously published effects of spatial filtering, can be related to the sampling properties of RDKs. The lower sampling rates resulting from long ISIs and large displacements introduce spurious alias signals which contaminate the spatiotemporal frequency spectrum of the stimulus. Temporal or spatial blurring extends performance limits (T_{\max} or D_{\max}) because it removes the high spatiotemporal frequencies which generate alias signals.

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*, 25, 1147–1154.
- Baker, C. L. & Braddick, O. J. (1985). Temporal properties of the short-range process in apparent motion. *Perception*, 14, 181–192.
- 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*, 30, 1341–1362.
- Boulton, J. C. & Baker, C. L. (1993). Different parameters control motion perception above and below a critical density. *Vision Research*, 33, 1803–1811.
- Braddick, O. J. (1973). The masking of apparent motion in random-dot patterns. *Vision Research*, 13, 355–369.
- Braddick, O. J. (1974). A short-range process in apparent motion. *Vision Research*, 14, 519–527.
- Burr, D. C., Ross, J. & Morrone, M. C. (1986). Smooth and sampled motion. *Vision Research*, 26, 643–652.
- Cleary, R. & Braddick, O. J. (1990). Masking of low frequency information in short-range apparent motion. *Vision Research*, 30, 317–327.
- Dosher, B. A., Landy, M. S. & Sperling, G. (1989). Kinetic depth effect from optic flow—I. 3D shape from Fourier motion. *Vision Research*, 29, 1789–1813.
- 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.
- Hess, R. F. & Snowden, R. J. (1992). Temporal properties of human visual filters: Number, shapes and spatial covariation. *Vision Research*, 32, 47–59.
- Morgan, M. J. (1992). Spatial filtering precedes motion detection. *Nature*, 355, 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 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.
- Nishida, S. & Sato, T. (1992). Positive motion after-effect induced by bandpass-filtered random-dot kinematograms. *Vision Research*, 32, 1635–1646.
- Watson, A. B. & Ahumada, A. J. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A*, 2, 322–342.
- Watson, A. B., Ahumada, A. J. & Farrell, J. E. (1986). Window of visibility: A psychophysical theory of fidelity in time-sampled visual motion displays. *Journal of the Optical Society of America A*, 3, 300–307.

Acknowledgement—This work was supported by a grant from the MRC.