



d_{\max} for Stereopsis and Motion in Random Dot Displays

A. GLENNERSTER*

Received 3 December 1996; in revised form 19 June 1997

The upper displacement limit for motion was compared with the upper disparity limit for stereopsis using two-frame random dot kinematograms or briefly presented stereograms. d_{\max} (the disparity/displacement at which subjects make 20% errors in a forced-choice paradigm) was found to be very similar for motion and stereo at all dot densities, and to fall with increasing dot density (0.006% or two dots to 50%) according to a power law (exponent -0.2). If d_{\max} is limited by the spacing of false targets, this pattern of results suggests that the spatial primitives in the input to the correspondence process may be derived from multiple spatial scales. A model using MIRAGE centroids provides a good fit to the data. © 1998 Elsevier Science Ltd. All rights reserved.

d_{\max} Stereopsis Motion False targets MIRAGE

INTRODUCTION

Stereopsis and motion perception require the detection of correlation between images, either presented binocularly or at different times. There are close parallels between the mechanisms that have been proposed to carry out this process in each domain, for example, motion and disparity energy models (Adelson & Bergen, 1985; Ohzawa, DeAngelis & Freeman, 1990) and co-operative models of stereo and motion correspondence (Pollard, Mayhew & Frisby, 1985; Williams & Phillips, 1987). However, there is currently little agreement about what type of model is most appropriate to describe either correspondence process in human vision.

Broadly, two approaches have been used to investigate correspondence mechanisms. First, tolerance of the visual system to different kinds of decorrelation has been measured, for example by varying the ratio of correlated and uncorrelated dots in a motion or stereoscopic display (van Doorn & Koenderink, 1982; Tripathy & Barlow, 1996; Cormack, Stevenson & Schor, 1991) or by reducing the correlation between the disparity (Harris & Parker, 1992) or direction of motion (Williams & Sekuler, 1984) of neighbouring elements in a display. A

second method, which is used in the experiments described here, has been to apply a uniform disparity/displacement to all the elements in a pattern and measure the size of the shift at which stereopsis or motion perception fails.

For motion, there is a generally accepted definition of the upper displacement limit, ' d_{\max} ' (Braddick, 1974). No such consensus exists for stereo, where several subjective criteria have been described, such as the disparity at which diplopia occurs, the disparity at which maximum depth is perceived or at which no depth is perceived (e.g. Ogle, 1953; Richards & Kaye, 1974; Schor & Wood, 1983; see Tyler (1991) for a review). The term ' d_{\max} for stereopsis' has been attached to more than one of these definitions (Tyler & Julesz, 1980; Wilcox & Hess, 1995). An alternative, objective, definition of the upper disparity limit is the disparity at which subjects begin to make errors (above some criterion level) when asked to identify whether the stimulus is presented in front of or behind fixation. The method is directly analogous with that used by Braddick to determine d_{\max} for motion.

Using a forced-choice method such as this, the upper disparity limit for a single line target has been found to be many degrees (Westheimer & Tanzman, 1956; Blake-more, 1970), while for a 50% density random dot pattern it can be as small as 9 arcmin (Nielsen & Poggio, 1984). The difference in results for high and low density patterns may reflect properties of the stimulus and need not imply the existence of separate 'local' and 'global' stereoscopic processes (Richards & Kaye, 1974) just as, in the case of motion processing, the difference between d_{\max} for 50% and low density patterns does not necessarily reflect the action of different 'long range' and 'short range' mechanisms (e.g. Morgan, 1992; Eagle & Rogers, 1996).†

*University Laboratory of Physiology, Parks Road, Oxford OX1 3PT, U.K. [Fax: +44 1865 272469; Email: andrew.glennerster@physiol.ox.ac.uk].

†Element density is not the only important factor affecting performance. d_{\max} for both stereo (Tyler & Julesz, 1980) and motion (Lappin & Bell, 1976; Fig. 4 of Cleary & Braddick, 1990b) has been found to rise with the square root of stimulus width. The authors of the first two of these papers accounted for the square root relationship using a cross-correlation model operating at the scale of the dots, while Cleary and Braddick (1990b) attributed the same rise in d_{\max} to the increasing low-pass characteristics of eccentric retina stimulated by larger images.

The results presented in this paper show that d_{\max} for motion and stereo in random dot patterns are the same across a wide range of dot densities. It is argued that similar limitations must apply to the correspondence process in both domains, at least for briefly presented stereograms and two-frame kinematograms. In both cases, there is a gradual change in d_{\max} with dot density, as Eagle and Rogers (1997) have shown for motion. Differences in d_{\max} for these stimuli are likely to relate to the spatial or phase structure of the stimuli, since all the patterns have the same (flat) amplitude spectrum (see also Morgan & Fahle, 1992; Eagle & Rogers, 1996, Eagle & Rogers, 1997).

The model described in the Methods section ("The MIRAGE algorithm") is based on the spacing of spatial primitives in the stimuli, and is similar to the models proposed by Morgan (Morgan, 1992; Morgan & Fahle, 1992) and Eagle (Eagle, 1996; Eagle & Rogers, 1996). The details of the matching algorithm are not important, nor is it suggested that these accurately reflect the matching process in the visual system. Rather, the simulation provides one way to determine the relative spacing of spatial primitives in the input to the correspondence process which, in a pure false-targets model, is the limiting factor determining d_{\max} .

The spacing of image features, or the periodicity of the input, is also an issue for energy models, at least in the case of disparity detection and two-frame motion (Fleet, Wagner & Heeger, 1996). If an energy model is to account for the whole data set, similar issues about the statistical properties of the input to the motion or disparity detection system are likely to be important.

PSYCHOPHYSICS

Methods

Apparatus. Stimuli were generated on a Macintosh II computer and displayed on two monochrome monitors in a modified Wheatstone apparatus. Subjects sat with their head in a chin rest and viewed the monitors through two front-silvered mirrors, each set at 45 deg to the median plane. The viewing distance was 57 cm.

Stimuli. The stimuli consisted of random dot patterns in which the density of dots ranged from 50% (16 800 dots) down to 0.006% (2 dots). In Experiment 1 the dots were bright (32 cd/m²) on a dark background (0.12 cd/m²), in Experiment 2 the luminances of the dots and background were reversed. Pixel size was 2 arcmin, but dots were always 6 arcmin square (3 by 3 pixels) which matches the dot size used by Eagle and Rogers (1997). The stimuli subtended 21 deg (horizontally) by 16 deg (vertically), i.e. 630 by 480 pixels (the whole screen was 640 by 480

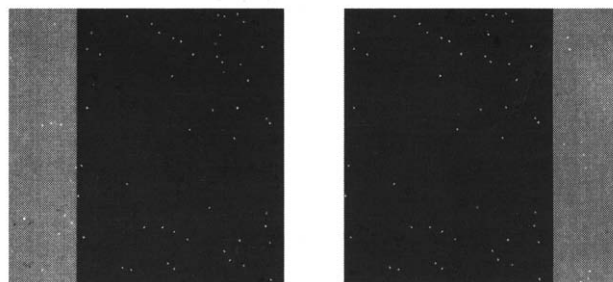


FIGURE 1. Pairs of random dot patterns were created, as illustrated here, and shown either as a two-frame kinematogram or as a stereo pair. The correlated dots were all given the same horizontal displacement. Uncorrelated dots filled the regions at the edges (shown, for illustration only, on a lighter background) so that the outline of the patch remained unchanged (i.e., stationary and at zero disparity).

pixels).^{*} The two eye's views were identical in the motion experiment (i.e. zero disparity).

For each trial, two images of random dots were created, one a displaced version of the other (as illustrated in Fig. 1). Patterns of 1% and greater were plotted probabilistically: each point at which a dot could appear had a given probability (equal to the dot density) of being bright (in Experiment 1) or dark (in Experiment 2). For patterns of lower densities (2–128 dots), the exact number of dots was plotted (at random x, y co-ordinates) so that random fluctuations in dot density were prevented. Dots from one image that were displaced outside the 'window' were re-plotted on the opposite side of the displaced image, i.e. dots 'wrapped round'. Although, in theory, the correlation of these dots across the two frames could be discovered, the displacement is so large that they are likely to be treated as uncorrelated. At low densities (2–128 dots) any dots that wrapped around were given a new vertical position so that the chance of spurious 'backward matches' was reduced.

Figure 1(a) shows in schematic form a pair of frames illustrating the displaced (correlated) dots and the strip of uncorrelated dots in each image (equal in width to the displacement applied). For the motion experiment, the images were displayed as a binocular two-frame apparent motion sequence, each frame lasting 150 msec with no inter-stimulus interval. For the stereo experiment, the two images were presented as a binocular pair (one to each eye). They were exposed simultaneously for 150 msec. In the case of an uncrossed disparity the stimulus has a simple planar interpretation (a surface of dots seen behind a dark window), but for a crossed disparity it does not. This asymmetry does not occur for motion (the surface is seen to move to the left or right behind a dark window). Given unlimited time to view such patterns, the crossed stimuli appear less 'solid' and the edge dots are seen as lustrous, which is not the case for the uncrossed stimuli. For the brief exposures used in this experiment, however, subjects could not detect a difference in the appearance of crossed and uncrossed stimuli and there was no obvious bias in subjects' responses. (The subject's responses were displayed at the end of each experimental

^{*}Eagle and Rogers (1997) investigated d_{\max} for motion at a range of stimulus sizes and found that d_{\max} did not rise above about 1/5 of the stimulus size, even at low dot densities. When the stimulus size was sufficiently large (25 deg), d_{\max} varied across the whole range of densities tested suggesting that, in this case, factors other than stimulus size were limiting d_{\max} .

run both as percent right button, from which any clear bias could be observed, and also as percent correct.)

After a trial the screen remained blank apart from a 15 by 15 arcmin fixation cross. In Experiment 1, the fixation cross was bright (32 cd/m^2) on a dark background (0.12 cd/m^2); in Experiment 2, it was dark on a bright background. The subject responded by pressing one of two switches, and this triggered the next display.

Psychometric procedure. The subject indicated by pressing one of two keys that the dots appeared to move left or right for the motion task or, for the stereo task, that they appeared in front of or behind the preceding fixation cross. In one run of trials all the stimuli were of a single density and all either motion or stereo. Five magnitudes

of displacement (each in two directions) were presented ten times, in random order, during a run of 100 trials. (The term 'displacement' is used here to refer both to the lateral displacement of the dots in the motion experiment and to the disparity added to the dots in the stereo experiment). The displacements used in any run were equally spaced (on a linear axis). Appropriate displacements (which would cover a range between a displacement at which a subject made 0% errors and one giving rise to 50% errors) were determined in a pilot run of 50 trials for each density. Larger displacements (and spacings) were used for the low density patterns.

Results from at least three runs were averaged (i.e. 60 trials per point). More data were gathered for low density

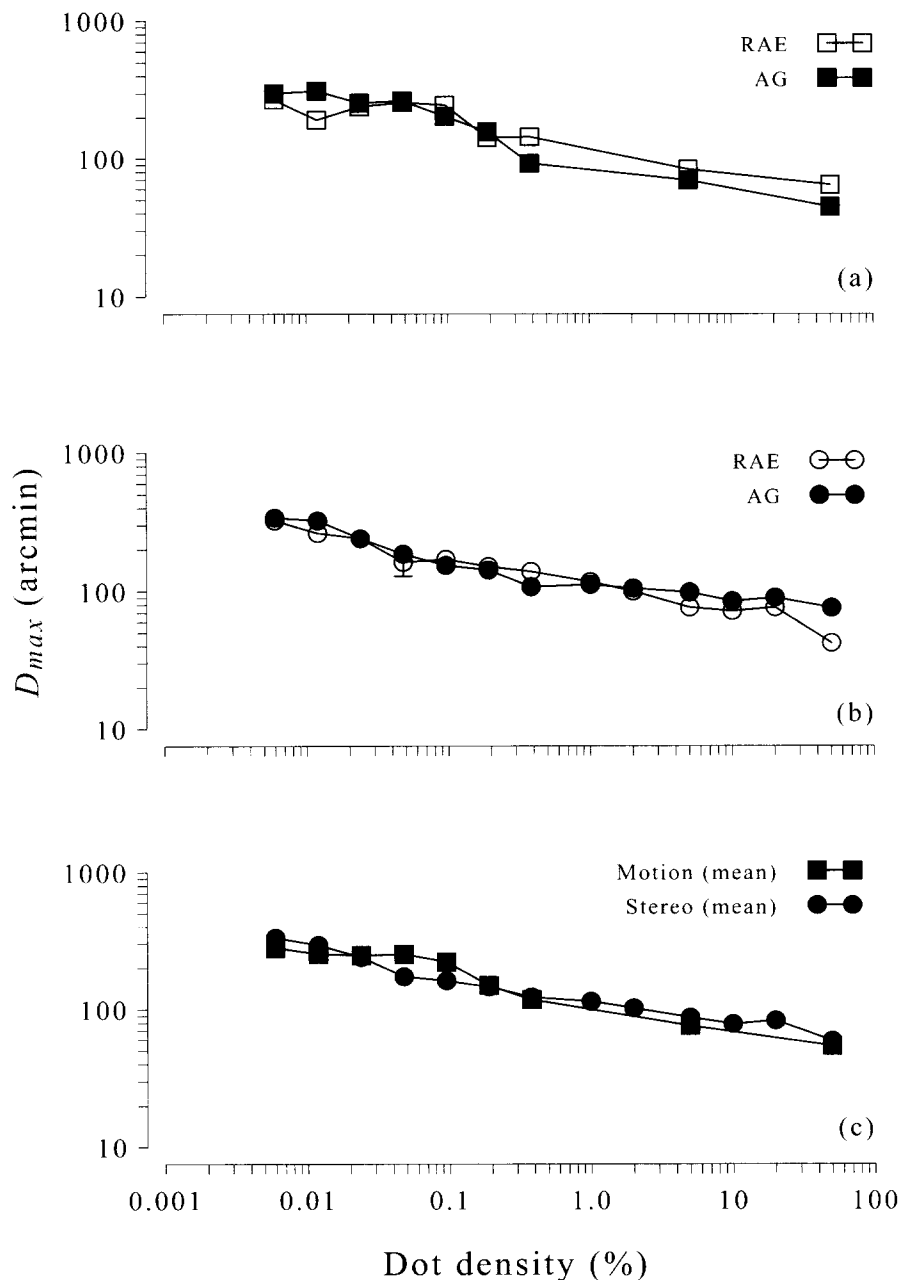


FIGURE 2. Results of Experiment 1. d_{\max} for (a) motion; and (b) stereo is shown for two observers, plotted against the density of dots in the stimulus, from 0.006% (i.e., two dots) to 50% density. The single error bar shows the largest SEM in the data. (c) Data for the two observers have been averaged so that results for the stereo and motion tasks can be compared directly.

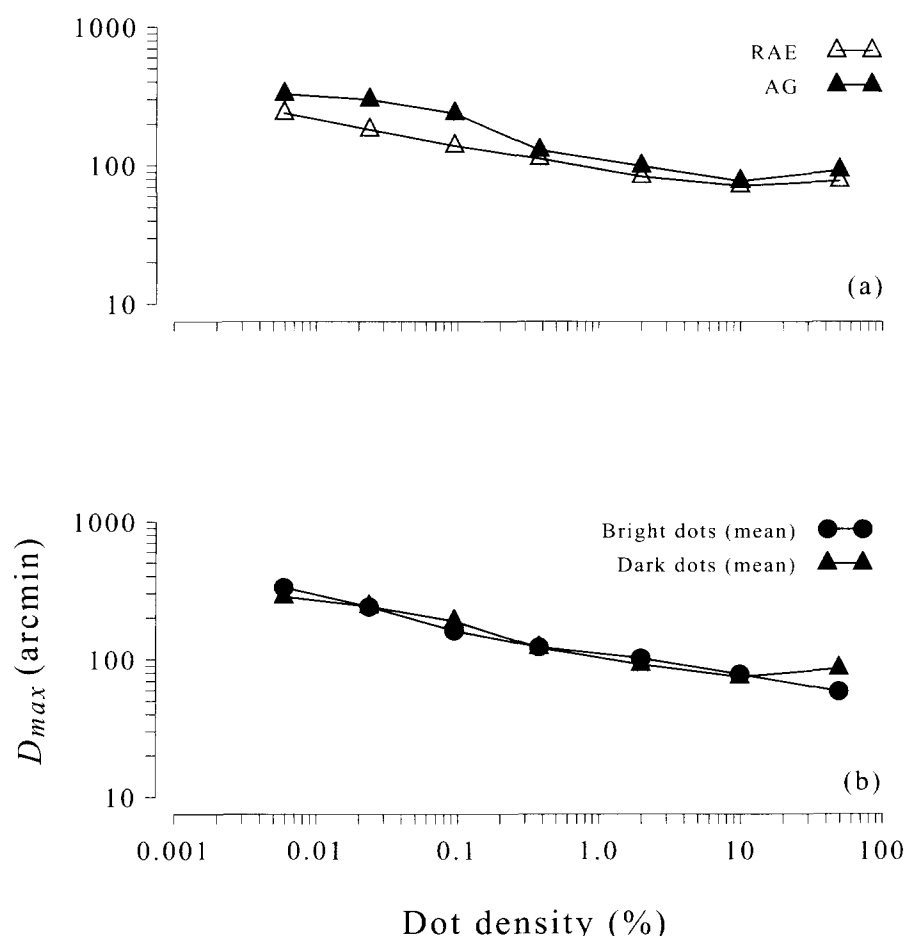


FIGURE 3. Results for Experiment 2 (dark dots on a bright background). (a) d_{max} for stereo for two observers. (b) Mean data for the two observers, plotted with data from the opposite contrast condition (see Fig. 2).

patterns (at least one extra run for stimuli containing less than 32 dots). Errors increase monotonically with increasing displacement up to 50% errors at large displacements. d_{max} , either for motion or stereo, was defined as the displacement that would give rise to 20% errors estimated by linear interpolation (Baker & Braddick, 1982).

Subjects. The author and one other experienced psychophysical observer acted as subjects. Both had 6/6 vision.

Experiment 1

In the first experiment all the patterns contained bright dots (32 cd/m²) on a dark background (0.12 cd/m²). Data were gathered for the motion and the stereo conditions.

Results. Figure 2(a) shows results for the motion task for two observers. The data are very similar to the results of Eagle and Rogers (1997): there is a smooth transition between d_{max} for a 50% pattern (about 50–60 arcmin) and d_{max} for the lowest density patterns of 5–6 deg. The

slope of the best fitting power function for the mean of the data for the two subjects is -0.18 . Results for the stereo task [Fig. 2(b)] are very similar. In this case, the mean slope is -0.20 .

Figure 2(c) shows data for both the motion and stereo experiments replotted (averaged across the two subjects) so that an explicit comparison can be made between performance in the two domains. Not only are the slopes of the functions similar, but also the absolute values of d_{max} for motion or stereo at each density.

Experiment 2

In Experiment 1, dot density varied over four log units and, hence, so did the mean luminance. This is potentially a reason for the observed fall in d_{max} with increasing dot density, since Dawson and DiLollo (1990) demonstrated a fall in d_{max} (for a single dot density*) when the mean luminance of the stimulus was increased. To control for this factor, in Experiment 2 the contrast of the dots was reversed (i.e., dark dots were shown on a bright background) so that in this case mean luminance fell as dot density was increased (albeit by a smaller magnitude when plotted on a log axis).

Results. Figure 3 shows d_{max} for stereo with patterns made up of dark dots on a bright background for two observers, for dot densities between 0.006 and 50%.

*Because dot size differed it is difficult to compare dot densities directly, but the patterns used by Dawson and DiLollo (1990) contained the same number of dots/degree as the 20% density pattern used in the experiments described here.

Although there is a small difference between the results of Experiments 1 and 2 for dot densities between 10 and 50%, over the majority of the range there is no significant difference between d_{\max} for stereo with the two opposite contrast patterns. This rules out the possibility that mean luminance is the only determinant of d_{\max} (since, as described above, it is changing in opposite directions in the two cases) and suggests that for densities below 10% mean luminance may have no effect at all on d_{\max} .

Above 10%, d_{\max} for stereo continues to fall for patterns of bright dots on a dark background but, for opposite contrast patterns, d_{\max} for stereo shows a small rise. A similar pattern of results has been observed in measurements of d_{\max} for motion with patterns of dark dots on a bright background (Eagle, 1992; Morgan & Fahle, 1992 for small dot sizes). To model the differences in d_{\max} over this range of densities, it is likely that effects of changes in mean luminance between the inter-trial interval and the stimulus will have to be taken into account (since the biggest difference between the results for the two conditions occurs at 50% density, where the only difference is the inter-trial screen luminance).

MODEL

Rationale

The close similarity between the results for d_{\max} for motion and stereo found in Experiment I suggests that, at least for these stimuli, similar limitations apply to the motion and stereo correspondence processes. In this section the possibility is explored that in both cases it is the density of false targets in the two images (left and right eye's images or first and second frames of the apparent motion sequence) that limits d_{\max} .

The basic principle is that each spatial primitive in one image is matched with only one primitive, its nearest neighbour, in the other image [just as Marr and Poggio (1979) described] and the direction of displacement assigned to the whole patch is determined by the proportion of matches made in a particular direction.

What constitutes a nearest neighbour depends on the algorithm used. Eagle and Rogers (1996) used points (2-D peaks in the luminance domain) as their spatial primitive and sought nearest neighbours within a sector (limited range of orientations) either side of the direction of displacement. Morgan (1992) and Eagle (1996), on the other hand, used zero-crossings as spatial primitives and restricted the search for nearest neighbours to one dimension (the direction of displacement).

False-target models account well for some d_{\max} results (e.g. Morgan, 1992; Eagle & Rogers, 1996) but provide a less convincing fit to data on random dot patterns of different densities (Morgan & Fahle, 1992, Eagle & Rogers, 1997). There are two possible reasons for this failure. One is that the spacing of primitives is not the only determinant of d_{\max} , i.e.:

$$d_{\max} = km, \quad (1)$$

where m is the mean spacing of elements and k is a scaling parameter whose value depends on various

factors such as the contrast or mean luminance of the pattern. The simplicity and intuitive appeal of a false-target model is lost by this modification. Nevertheless, it is the approach taken by Morgan and Fahle (1992) and Eagle and Rogers (1997) to explaining the discrepancy between their data and the spacing of the primitives they examined. Another possibility is that d_{\max} does always reflect the spacing of false targets. In this case, the aim is to find a primitive that will fit the data.

The spacing of primitives derived from the output of a single bandpass filter follow a characteristic pattern when plotted against dot density: a plateau at high densities and a rising portion at low densities (e.g. Figure 4 of Eagle & Rogers, 1996). The plateau is due to the fact that when the dot spacing is much smaller than the space constant of the filter, the spacing of false targets in the output is dependent only on the filter size (or spatial frequency tuning) and not at all on the density of the dots. The slope of the rising portion depends on the algorithm used. If nearest neighbours are sought within a 'sector' (Eagle & Rogers, 1996), mean distance to the nearest neighbour varies inversely with the square root of dot density. If, on the other hand, the search is 1-D in the direction of displacement (Morgan & Fahle, 1992), the rise in mean separation is even steeper (varying inversely with density).

The experimental data (Figs 2 and 3) do not fit any of these single-filter predictions. d_{\max} does not plateau at high densities nor does it rise steeply at low densities. The implication of a gradual, constant slope for a pure element-spacing model is that the density of false targets is changing gradually over the whole range of densities (4 log units). One way to achieve this pattern in a model is to include the output of a relatively coarse filter, so that some dots are blurred together at low densities, and also the output of a fine filter, which contributes extra primitives in a high density pattern.

The MIRAGE algorithm. There already exists at least one proposed method of combining the outputs of spatial-frequency tuned channels whose properties fit this description. The algorithm ('MIRAGE'), advocated by Watt and Morgan (1985) on the basis of a quite different set of experiments, involves three stages. First, the outputs from spatial filters at a range of different scales (spatial frequencies) are half-wave rectified. Second, the positive responses from all the scales are added together to form one signal ('S⁺') and the negative responses added to form a separate signal ('S⁻'). Third, the S⁺ and S⁻ signals are divided into zero bounded regions and described in terms of a few simple statistics. For a one-dimensional analysis, each zero-bounded region is described by its central moments, i.e. the area under the curve (mass), the mean (centroid), and the standard deviation of the distribution about that mean. The model described below uses 1-D centroids calculated in the direction of displacement (horizontal).*

*Morgan and Mather (1994) used a variant of the MIRAGE model, based on gaussian filters with the DC component removed.

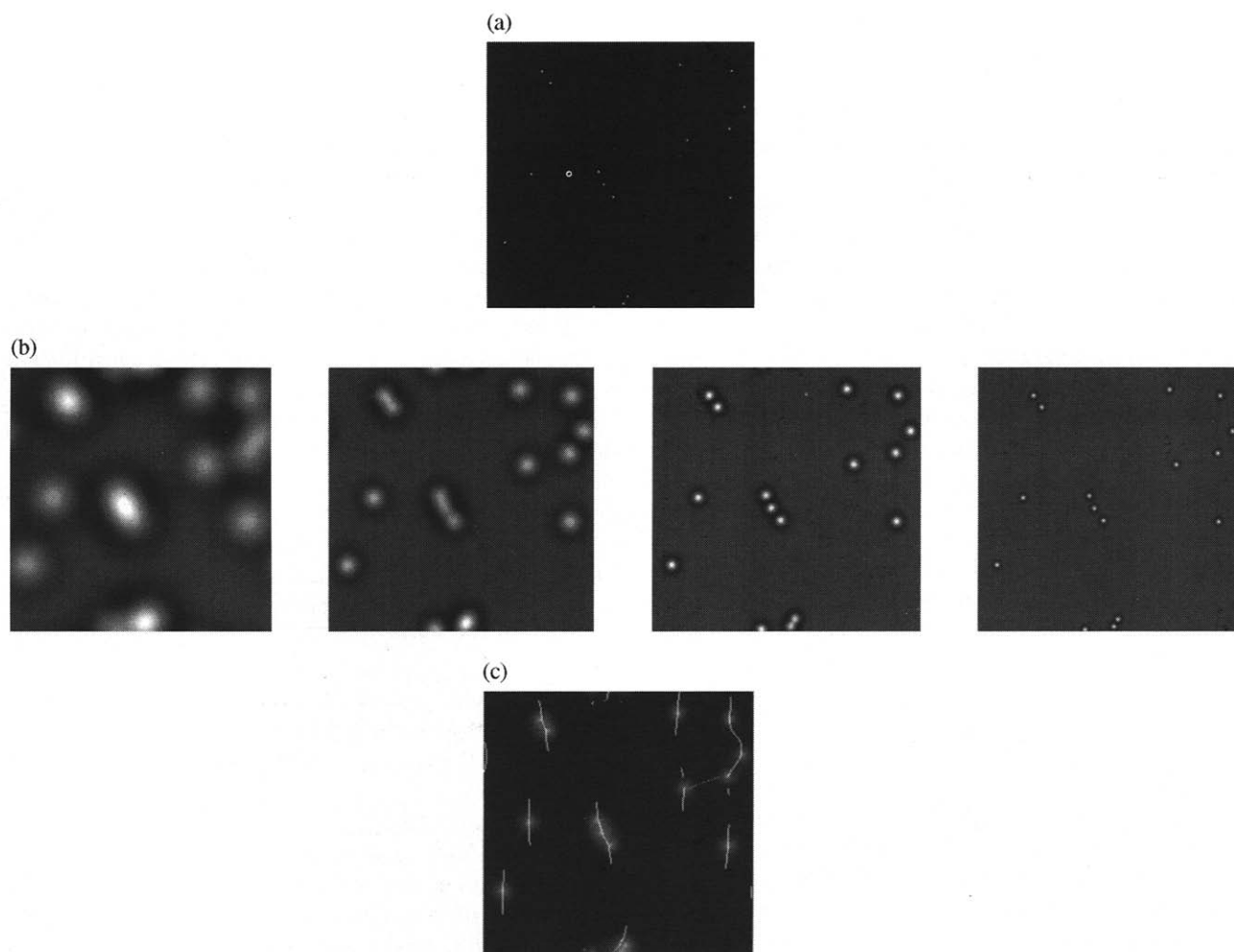


FIGURE 4. (a) A random dot pattern of 16 dots (0.024% density). (b) The output of four Laplacian of gaussian filters convolved with this pattern. σ is $1/16$, $1/32$, $1/64$ and $1/128$ of image size i.e. 16, 8, 4 and 2 pixels, respectively. In the psychophysical experiments dot size was 6 arcmin so these space constants correspond to 96, 48, 24 and 12 arcmin. (c) The images in (b) have been half-wave rectified and the positive responses at each scale summed together (MIRAGE S^+ response). Along each horizontal row, the centroids of zero-bounded regions are marked with a white pixel (see text for details).

Figure 4 shows the MIRAGE response to a low density pattern.* In the model described below, four Laplacian of gaussian filters with space constants of 12, 24, 48 and 96 arcmin make up the MIRAGE response. As will be seen later (Fig. 7), the precise sizes of the filters chosen are not critical, although the total span of spatial scales may be. The responses of the individual filters to the pattern are shown in Fig. 4(b).

The illustration in Fig. 4(c) shows that, for this low density pattern, the MIRAGE response is similar to the output of the largest (lowest frequency) filter. In both cases, when several dots lie close together they are blurred together and form one 'blob' or zero-bounded

region. For a high density pattern, on the other hand, the spatial primitives in the input to the correspondence process should be densely spaced, just as they are in the output of a relatively fine spatial filter. Figure 5 shows how this stipulation is also met by MIRAGE centroids.

Note that, for a 50% pattern, smaller zero-bounded regions fill in the gaps between the larger-mass zero-bounded regions. The smallest-mass zero-bounded regions are due entirely to the output of the finest filter. This is the reason why, in the model, d_{\max} for 50% density patterns is limited by the finest filter contributing to the MIRAGE output, as Cleary and Braddick (1990b) originally proposed. By comparing Figs 4 and 5 it is clear why the same would not be true of a low density pattern. The corresponding empirical prediction remains to be confirmed, i.e., that low-pass filtering will have a much smaller effect on d_{\max} for a low density pattern than it does for a 50% density pattern.

The next section describes a quantitative test of the MIRAGE centroid hypothesis. The simulation provides one way to determine the relative spacing of spatial primitives (in this case MIRAGE centroids) which, in a

*The MIRAGE S^- response for a 50% density pattern is the same as the S^+ response (statistically). For low density patterns of bright dots on a dark background the S^- response is a large 'sea' with a few 'holes' corresponding to each dot (see Watt, 1987, 1988 for examples and for a description of the psychophysical evidence and computational rationale for MIRAGE). For patterns consisting of dark dots on a bright background the characteristics of the S^+ and S^- signals are reversed.

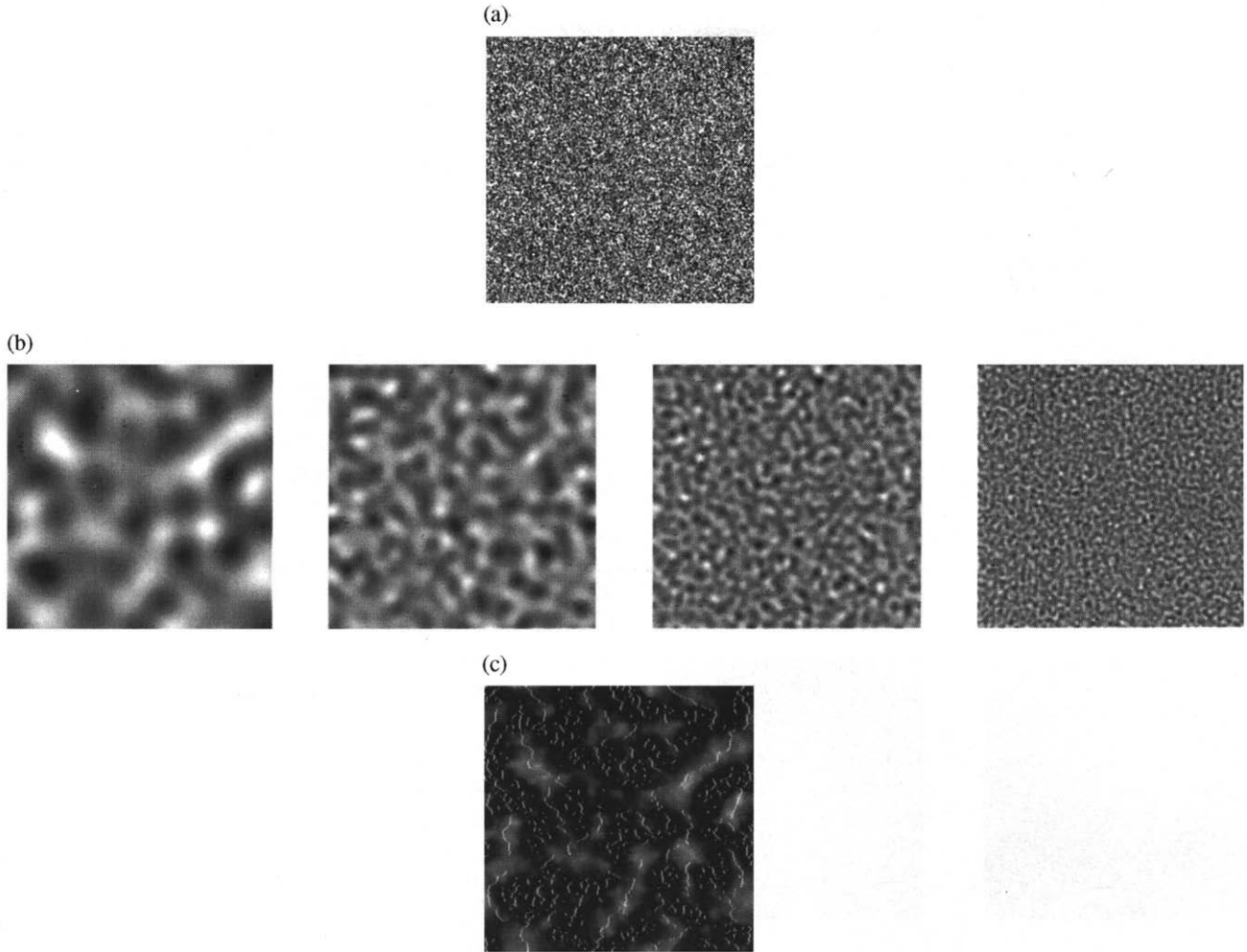


FIGURE 5. As for Fig. 4 but for a 50% density random dot pattern.

pure false-targets model, are the limiting factor in determining d_{\max} (see also Morgan, 1992; Eagle, 1996).

Methods

A pair of random dot patterns, each 256 by 256 pixels, was created by adding a given displacement to the dots in one image. Dots 'wrapped round' so that dots shifted out of the image were re-plotted on the opposite side of the image, as for the experimental stimuli. For low densities, as in the experiment, dots that wrapped round were given a new (random) vertical position. Each image was filtered with Laplacian of gaussian filters (space constants 2, 4, 8 and 16 pixels, i.e. modelling filters of 12, 24, 48 and 96 arcmin). The equation for a Laplacian of gaussian filter in the spatial domain is:

$$\nabla^2 G(r, \sigma) = \left(1 - \frac{r^2}{2\sigma^2}\right) e^{-\frac{r^2}{2\sigma^2}} \quad (2)$$

*If no threshold were put on either the level of response defining the limit of a 'zero'-bounded region or on the mass of centroids included in the analysis, then there would be a large number of very low-mass centroids in the S^+ and S^- responses which arise not from dots in the image but from 'ringing' caused by convolving with filters in the Fourier domain.

where r is the radial distance from the centre, σ is the space constant. The peak-to-trough amplitude of each filter in the spatial domain was equal, as in the original MIRAGE model (Watt & Morgan, 1985). The four filtered responses are:

$$R_i = F_i * I \quad (3)$$

where $i = 1$ to 4, $*$ refers to convolution and I is the input image. The output of each filter was half-wave rectified:

$$R_i^+ = R_i \quad \text{if } R > R_{\text{thresh}}, R_i^+ = 0 \quad \text{otherwise} \quad (4)$$

and

$$R_i^- = R_i \quad \text{if } R < -R_{\text{thresh}}, R_i^- = 0 \quad \text{otherwise.} \quad (5)$$

R_{thresh} was 0 in the Watt and Morgan algorithm; in the model used here it was set at 10^{-7} (where dot size = 1 pixel, dot luminance = 1 and background luminance 0), but as discussed below the exact value is not critical.* The positive responses are summed to give an S^+ signal:

$$S^+ = R_1^+ + R_2^+ + R_3^+ + R_4^+ \quad (6)$$

and similarly the negative responses are summed to give an S^- signal:

$$S^- = R_1^- + R_2^- + R_3^- + R_4^- \quad (7)$$

although only the S^+ signal was used in the modelling described here.

The filters are much larger than those described in Watt and Morgan (1985) but the patch size (and hence eccentricity) is greater and the exposure duration shorter (Cleary & Braddick, 1990b; Watt, 1987). The centroid of each zero-bounded distribution in the S^+ response was calculated in the direction of displacement, i.e., along horizontal raster lines. The 1-D centroid, P_i , is the position within a zero-bounded distribution about which the first order moment is zero:

$$P_i = \frac{\int_{Z_{c_i}}^{Z_{c_{i+1}}} x S^+(x) dx}{\int_{Z_{c_i}}^{Z_{c_{i+1}}} S^+(x) dx}, \quad (8)$$

where Z_{c_i} and $Z_{c_{i+1}}$ are the positions of adjacent zero-crossings and $S^+(x)$ is the S^+ response at point (x) along any particular raster line (Watt & Morgan, 1985). Only

centroids above a threshold mass (10^{-4}) were included in the input to the matching program but, as for R_{thresh} its exact value is not critical. An example of this input is illustrated in Fig. 6. For each centroid pixel derived from the left eye's image, the nearest centroid pixel from the right eye's image was found along the raster line. The proportion of matches made in the correct direction (i.e., in the direction of displacement of the dots) was recorded.

This was repeated for a range of dot displacements, covering the range from a displacement giving rise to no errors to one giving rise to 50% errors and the error rate for each displacement recorded. The whole process was repeated for at least 10 different random dot patterns at each density (more at low densities, when variability was greater just as it was in the psychophysical experiments). The mean proportion of centroid matches made in the correct direction was calculated for each displacement (averaging across the ten or more dot patterns). The displacement for which 20% of centroid matches were made in the wrong direction was defined as a theoretical d_{max} and is plotted in Fig. 7 showing the model results.*

Results

Results of the simulation are shown in Fig. 7(a). Also shown are the mean values of d_{max} for stereo and motion from Experiment 1 (dotted lines). The model fits the data well across the whole range of densities. The free parameters in the model have not been varied to maximize the fit. In fact, the pattern of centroids is relatively immune to large variations in most of the parameters, e.g. R_{thresh} and the threshold mass for centroids included in the analysis, which can both be increased or decreased by a factor of at least 10 without any appreciable effect. The same applies to the relative amplitude of the filters in the model: Fig. 7 shows results for MIRAGE in which the filters have an equal peak-to-trough height in the spatial domain, but if they have an equal peak amplitude in the Fourier domain the pattern of centroids and modelled d_{max} are barely affected.†

On the other hand, the range of filters contributing to

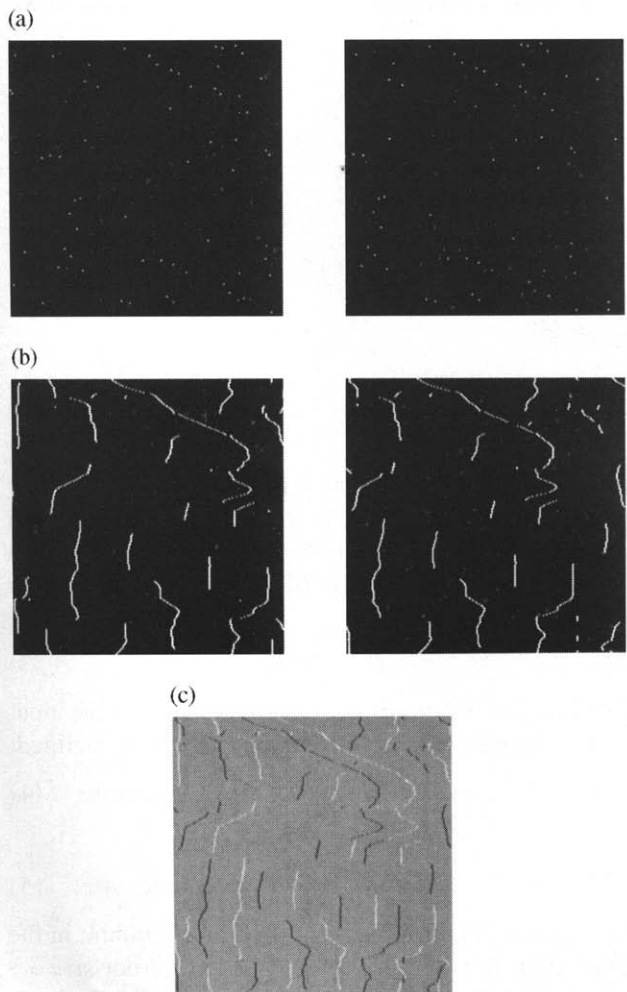


FIGURE 6. (a) A pair of random dot patterns in which the dots have been displaced (see Fig. 1). (b) The centroids of the MIRAGE S^+ response for these patterns (see Fig. 4). (c) A subtraction of the images in (b) to help illustrate the matching algorithm. For each centroid in one image (shown here as white pixels) the nearest centroid in the other image (i.e., a black pixel) is found and the direction of that match recorded. d_{max} for the model is the displacement at which 20% of nearest-neighbour matches are in the wrong direction (see Methods section for details).

*The mean and standard deviation of the distances to the nearest matches was also recorded. For small dot displacements, most matches are correct so the mean distance to the nearest match equals the dot displacement and the standard deviation of estimates is small. For very large displacements (or uncorrelated patterns) the mean is zero (there are an equal number of matches made in either direction) and the standard deviation is large. d_{max} lies between these extremes and could sensibly be defined as the displacement for which the standard deviation is equal to the mean of the distribution. In fact, this definition gives very similar values to those obtained using 20% errors to define d_{max} .

†At some point, depending on the level set for the threshold mass of zero-bounded distributions included in the input to the matching stage, reducing the relative amplitude of the input from fine spatial filters must reduce the number of primitives in a dense pattern and so raise modelled d_{max} . It would be relevant to explore the effects of varying these two parameters when modelling d_{max} results for flat and $1/f$ spectrum patterns (e.g. Bex, Brady, Fredericksen & Hess, 1995; Eagle, 1996).

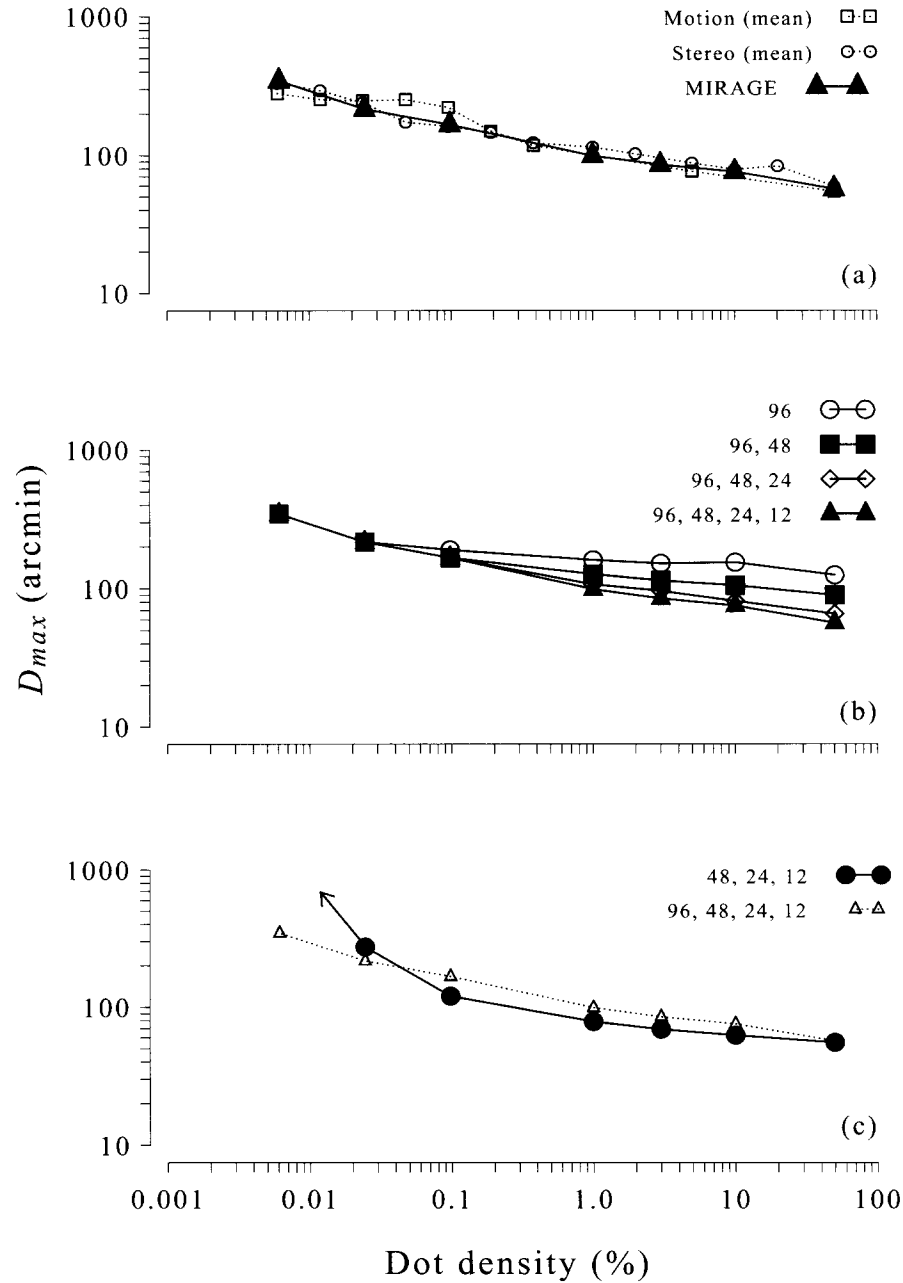


FIGURE 7. Results of the model. (a) The black triangles show simulated d_{\max} using MIRAGE centroids (see text for details). In this case, the space constants of the filters contributing to the MIRAGE signal were 96, 48, 24 and 12 arcmin. d_{\max} for the stereo and motion conditions are re-plotted from Fig. 2. (b) Simulated d_{\max} when the filters contributing to the MIRAGE signal are 96, 48, 24 and 12 [as in (a)]; 96, 48 and 24 arcmin; 96 and 48 arcmin; and 96 arcmin, i.e., the largest filter on its own. (c) The effect of omitting the largest filter (i.e., $a = 48, 24$ and 12 arcmin). d_{\max} could not be estimated by the model for the lowest density (0.006% or 4 dots) as in this case there is only one centroid on most horizontal rows.

the MIRAGE response does have a significant effect. Some of the possible variations are shown in Fig. 7(b) and (c). The effect of keeping the largest filter constant and progressively removing the smallest remaining filter [Fig. 7(b)] is to raise modelled d_{\max} for high density patterns but it has no effect on d_{\max} for the lowest density patterns. Removing the largest filter from the MIRAGE S^+ signal has quite a different effect [Fig. 7(c)]. At the lowest density, d_{\max} is unmeasurable as there are a large number of raster lines with no false match (for this image size). d_{\max} changes rapidly over the low density range

and much more slowly at high densities. In other words, theoretical d_{\max} becomes more like the predictions for a single, smaller filter, as expected.

Taken together, the variations on the model shown in Fig. 7(b) and (c) indicate that the spacing of MIRAGE centroids depends primarily on the size of the largest filter for sparse patterns and on the size of the finest filter for the highest density random dot patterns. This was the requirement, set out earlier ("Rationale") for any spatial primitive that will fit the experimental data using a pure false-target model.

DISCUSSION

The experiments described here show that d_{\max} for stereo in random dot patterns changes gradually with dot density, just as Eagle and Rogers (1997) found for motion using two-frame kinematograms. For both motion and stereo, d_{\max} is 5–6 deg at low densities but only 50–60 arcmin for a 50% density pattern (Experiment 1). The agreement between motion and stereo d_{\max} at all densities is close (see also Wattam-Bell, 1995 for data on 50% density patterns). The result strongly suggests that similar factors limit the correspondence process in motion and stereo, at least for briefly presented stereograms and two-frame kinematograms.

How can d_{\max} be explained?

Models of d_{\max} can be divided into two broad categories: those based on detectors with a fixed spatial limit (for either disparity or displacement), and false-target theories in which d_{\max} depends on the spacing of spatial primitives in the image. According to the fixed spatial limit theories, the optimum displacement for a detector is a fixed phase (i.e., it varies inversely with the peak spatial frequency the detector) and hence d_{\max} will be determined entirely by the spectral composition of the stimulus.

Currently, psychophysical evidence favours false-target theories. For dense patterns it is difficult to distinguish the two models: both predict that d_{\max} should rise as dot size is increased (e.g. Morgan, 1992) or as the pattern is bandpass filtered at progressively lower spatial frequencies (Cleary & Braddick, 1990a; Chang & Julesz, 1983; Bischoff & DiLollo, 1991; Smallman & MacLeod, 1994). For sparse patterns, the data follow the predictions of the false-target models (Eagle & Rogers, 1996; Morgan, Perry & Fahle, 1997). It might be argued that there are two mechanisms but, until compelling evidence can be brought against a false-target model, this dichotomy seems unnecessary.

Support for an element-spacing limit to d_{\max} does not restrict theories of motion or disparity detection to a particular correspondence algorithm. False-target models are, in general, agnostic about the nature of the matching process; they merely assert that d_{\max} is determined by the pattern statistics after a filtering stage.

One or many filters?

It has been suggested that the data on motion d_{\max} at different dot densities can be explained by considering the information in the stimulus at a single spatial scale (Morgan & Fahle, 1992; Eagle & Rogers, 1997). The two models are slightly different from each other and both differ significantly from the MIRAGE model presented here.

The model proposed by Morgan and Fahle (1992) predicts that there should be a range of densities over which d_{\max} does not change and, as density is reduced, a range over which d_{\max} rises inversely with dot density (a slope of -1 on a plot such as Fig. 2). The model put forward by Eagle and Rogers (1997) predicts that, over

the rising portion, d_{\max} should be proportional to the square root of density (a slope of -0.5). The difference is due to the way 1-D spacing is calculated, either as density along a single direction (Morgan and Fahle) or as the square root of the number of 2-D primitives within a 'sector' or range of orientations (Eagle and Rogers).

The data presented in this paper (Figs 2 and 3) and in the paper by Eagle and Rogers (1997) show, instead, a constant slope of about -0.2 over most of the 4 log unit change in dot density. Of the two single-filter models this is closer to the predictions of Eagle and Rogers (1997). They account for the discrepancies between their model and the data (Fig. 3 of their paper, as well as Figs 2 and 3 of the present paper) in terms of changes in the mean luminance of the pattern at high densities and changes in the r.m.s. contrast of the pattern at low densities. They show that, when the mean luminance of the stimulus and inter-trial interval is raised, the pattern of data is more like the predictions of a single-filter model (and also of the MIRAGE model with the largest filter removed [Fig. 7(c)]). An implication of the MIRAGE model described here is that the effects of changes in mean luminance and r.m.s. contrast with changes in dot density may not be as important in determining d_{\max} as Eagle and Rogers (1997) suggest, since most of the experimental data can already be accounted for by the spacing of false targets.

The other manipulation Morgan (1992) and Morgan and Fahle (1992) have carried out is to vary dot size. When dot size was large, they found a steep change in d_{\max} with changes in dot size (Morgan, 1992) or dot density (Morgan & Fahle, 1992), which fits the predictions of all three element-spacing models (Morgan & Fahle, 1992; Eagle & Rogers, 1997; and the MIRAGE model described here) when dots are larger than the filters in the model.

Conclusion

The fact that d_{\max} for stereo and motion is so similar across a wide range of dot densities strongly suggests that, at least for two-frame kinematograms and briefly presented stereograms, similar limitations apply to the correspondence process in both domains. One possibility is that both are limited by the density of spatial primitives in the representation preceding correspondence. If this is the only factor affecting d_{\max} then, because d_{\max} changes gradually over the whole range of densities tested, the primitive is likely to be derived from the outputs of filters tuned to a range of different spatial frequencies. One candidate, centroids of zero-bounded distributions in the output of the MIRAGE algorithm (Watt & Morgan, 1985) can account for the principal features of the data. Unlike similar false-target theories, the model does not incorporate an arbitrary multiplicative factor once an informational limit on d_{\max} has been computed.

REFERENCES

- Adelson, E. H. & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America*, A2, 284–299.

- Baker, C. L. & Braddick, O. J. (1982). The basis of area and dot number effects in random dot motion perception. *Vision Research*, 22, 1253–1259.
- Bex, P. J., Brady, N., Fredericksen, R. E. & Hess, R. F. (1995). Energetic motion detection. *Nature*, 378, 670–671.
- Bischoff, W. F. & DiLollo, V. (1991). On the half-cycle displacement limit of sampled directional motion. *Vision Research*, 31, 649–660.
- Blakemore, C. (1970). The range and scope of binocular depth discrimination in man. *Journal of Physiology (London)*, 211, 599–622.
- Braddick, O. J. (1974). A short-range process in apparent motion. *Vision Research*, 14, 519–527.
- Chang, J. J. & Julesz, B. (1983). Displacement limits for spatial frequency filtered random-dot cinematograms in apparent motion. *Vision Research*, 23, 1379–1385.
- Cleary, R. & Braddick, O. J. (1990a). Direction discrimination for band-pass filtered random dot kinematograms. *Vision Research*, 30, 303–316.
- Cleary, R. & Braddick, O. J. (1990b). Masking of low frequency information in short-range apparent motion. *Vision Research*, 30, 317–327.
- Cormack, L. K., Stevenson, S. B. & Schor, C. M. (1991). Interocular correlation, luminance contrast and cyclopean processing. *Vision Research*, 31, 2195–2207.
- Dawson, M. & DiLollo, V. (1990). Effects of adapting luminance and stimulus contrast on the temporal and spatial limits of short-range motion. *Vision Research*, 30, 415–429.
- van Doorn, A. J. & Koenderink, J. J. (1982). Spatial properties of the visual detectability of moving spatial white noise. *Experimental Brain Research*, 45, 189–195.
- Eagle, R. A. (1992). Spatial aspects of human visual motion detection. D. Phil. thesis, University of Oxford.
- Eagle, R. A. (1996). What determines the maximum displacement limit for spatially broadband kinematograms? *Journal of the Optical Society of America A*, 13, 408–418.
- Eagle, R. A. & Rogers, B. J. (1996). Motion detection is limited by element density not spatial frequency. *Vision Research*, 36, 545–558.
- Eagle, R. A. & Rogers, B. J. (1997). Effects of dot density, patch size and contrast on the upper spatial limit for direction discrimination in random-dot kinematograms. *Vision Research*, 37, 2091–2102.
- Fleet, D. J., Wagner, H. & Heeger, D. J. (1996). Neural encoding of binocular disparity: energy models, position shifts and phase shifts. *Vision Research*, 36, 1839–1857.
- Harris, J. M. & Parker, A. J. (1992). Efficiency of stereopsis in random-dot stereograms. *Journal of the Optical Society of America A*, 9, 14–24.
- Lappin, J. S. & Bell, H. H. (1976). The detection of coherence in moving random-dot patterns. *Vision Research*, 16, 161–168.
- Marr, D. & Poggio, T. (1979). A computational theory of human stereo vision. *Proceedings of the Royal Society of London (B)*, 204, 301–328.
- Morgan, M. J. (1992). Spatial filtering precedes motion detection. *Nature*, 335, 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.
- Morgan, M. J., Perry, R. & Fahle, M. (1997). The spatial limits for motion detection in noise depend on element size not on spatial frequency. *Vision Research*, 37, 729–736.
- Nielsen, K. R. K. & Poggio, T. (1984). Vertical image registration in stereopsis. *Vision Research*, 24, 1133–1140.
- Ogle, K. N. (1953). Precision and validity of stereoscopic depth perception from double images. *Journal of the Optical Society of America*, 43, 906–913.
- Ohzawa, I., DeAngelis, G. C. & Freeman, R. D. (1990). Stereoscopic depth perception in the visual cortex: neurons ideally suited as disparity detectors. *Science*, 249, 1037–1041.
- Pollard, S. B., Mayhew, J. E. W. & Frisby, J. P. (1985). PMF: A stereo correspondence algorithm using a disparity gradient limit. *Perception*, 14, 449–470.
- Richards, W. & Kaye, M. G. (1974). Local versus global stereopsis: two mechanisms? *Vision Research*, 14, 1345–1347.
- Schor, C. M. & Wood, I. (1983). Disparity range for local stereopsis as a function of luminance spatial frequency. *Vision Research*, 23, 1649–1654.
- Smallman, H. S. & MacLeod, D. I. A. (1994). Size disparity correlation in stereopsis at contrast threshold. *Journal of the Optical Society of America A*, 11, 2169–2183.
- Tripathy, S. P. & Barlow, H. B. (1996). The effect of dot number on correspondence noise in random dot kinematograms. *Investigative Ophthalmology and Visual Science*, 37, S745.
- Tyler, C. W. (1991). Cyclopean vision. In Regan, D. (Eds), *Binocular vision*. Basingstoke: Macmillan Press.
- Tyler, C. W. & Julesz, B. (1980). On the depth of the cyclopean retina. *Experimental Brain Research*, 40, 196–202.
- 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. (1988). *Visual processing: computational, psychophysical and cognitive research*. Hove: Lawrence Erlbaum.
- Watt, R. J. & Morgan, M. J. (1985). A theory of the primitive spatial code in human vision. *Vision Research*, 25, 1661–1678.
- Wattam-Bell, J. (1995). Stereoscopic and motion d_{\max} in adults and infants. *Investigative Ophthalmology and Visual Science*, 36, S910.
- Westheimer, G. & Tanzman, I. J. (1956). Qualitative depth localization with diplopic images. *Journal of the Optical Society of America*, 46, 116–117.
- Wilcox, L. M. & Hess, R. F. (1995). d_{\max} for stereopsis depends on size, not spatial frequency content. *Vision Research*, 35, 1061–1069.
- Williams, D. W. & Phillips, G. C. (1987). Cooperative phenomena in the perception of motion direction. *Journal of the Optical Society of America A*, 4, 878–885.
- Williams, D. W. & Sekuler, R. (1984). Coherent global motion percepts from stochastic local motions. *Vision Research*, 24, 55–62.

Acknowledgements—Supported by a SERC Image Interpretation Initiative studentship and MRC Career Development Award. I am grateful to Brian Rogers and Andrew Parker in whose laboratories the work was carried out. I would like to thank Richard Eagle for acting as an observer and for many interesting discussions, and Andrew Parker and Bruce Cumming for help in revising the manuscript.