

Correspondence Noise and Signal Pooling in the Detection of Coherent Visual Motion

Horace Barlow and Srimant P. Tripathy

Physiological Laboratory, Downing Site, Cambridge CB2 3EG, United Kingdom

In the random dot kinematograms used to analyze the detection of coherent motion in the middle temporal visual area (MT) and in psychophysical experiments the exact way that dots are paired between successive presentations is not known by the observer. We show how to calculate the limit to coherence threshold caused by this uncertainty, which we call “correspondence noise.” We compare ideal thresholds limited only by this noise with those of human observers when dot density, ratio of dot numbers in two fields, area of stimulus, number of fields, and method of generation of the coherent dots are varied. The observed thresholds vary in the same way as the ideal thresholds over wide ranges, but they are much higher. We think this difference is because the ideal detector takes advantage of the high precision with which dots are placed in the kinematograms, whereas the neural motion system can only operate with low precision. When kinematograms are generated with de-

creased precision of dot placement, the ideal detector no longer has this advantage, and the gap between ideal and actual performance is greatly reduced. Because the signals that result from objects moving in the real world are scattered over broad ranges of direction and velocity, high precision is not needed, and it is advantageous for the motion system to pool information over broad ranges. Other mismatches between kinematograms and the neural motion system, and internal noise, may also elevate human thresholds relative to the ideal detector. The importance of external noise suggests that the neurons of MT form a vast array of optimal filters, each matched to a different combination of parameters in the multidimensional space required to define motion in patches of the visual field.

Key words: correspondence noise; coherent motion; statistical efficiency; integration; matched filters; MT or V5; global motion

The motivation for the work to be described here was to find the natural difficulties and limiting factors for detecting motion in the random dot kinematograms that have been used so successfully to analyze the neuronal basis for the detection of coherent motion by monkeys (Newsome et al., 1989, 1990; Britten et al., 1992, 1995; Celebrini and Newsome, 1994). In this paradigm some of the dots are moved coherently in the same direction from field to field, whereas the remainder are replaced at random positions; the behavioral responses of the monkey, and the discharges of its cortical neurons, are tested for their ability to detect motions with varying percentages of coherence, and a fraction as low as 5% is often reliably detected both by the whole monkey and by single neurons in the middle temporal visual area (MT or V5). We thought that the value of the comparison between neurophysiology and behavior would be much increased if the limiting factors were better understood.

Figure 1, *top*, illustrates the correspondence problem, which arises whenever motion has to be detected and is specially important in random dot kinematograms, in which it has long been appreciated that it may be a limiting factor (Braddick, 1974; Morgan and Ward, 1980; van Doorn and Koenderinck, 1982a; Todd and Norman, 1995; Eagle and Rogers, 1996). But no one has shown how to calculate the magnitude of the noise that results from false correspondences, and this is the first problem we have tackled. We use the result to calculate ideal thresholds for detect-

ing coherent motion limited only by correspondence noise, and we compare these with thresholds measured in human subjects. The ideal thresholds are not based on any model of the motion-detecting mechanism but on knowledge of how random dot kinematograms are generated, because by definition ideal performance is limited by what is in the kinematograms and not by any properties of the visual system. Figure 1, *top*, shows dots from two successive fields, the ones from the first *filled* and those from the second *open*. At the *top left* all four dots have been coherently moved, but at the *top right* only one, marked by a *heavy arrow*, was moved in this direction; the four *light arrows* each show a spurious motion signal generated by pairing one of the first field dots with a second field dot; there are 15 such spurious arrows, because all the four filled dots can be paired with all the four open dots to form a total of 16 pairs, of which only one was generated deliberately. The spurious pairings are indistinguishable from the real one by an observer, so to decide whether there is coherent motion, all the pairs must be examined, and one must then find whether there is an excess over chance expectation in the number corresponding to a particular direction and velocity of motion.

We planned to vary the parameters of kinematograms and to compare their effects on the observed thresholds with the effects predicted by this strategy of examining all possible correspondences. The theoretical influences of the various stimulus parameters are set out below. Comparison of experimental and theoretical thresholds shows that some of the predicted relations hold over wide ranges, but even within these ranges the absolute level of performance achieved is not nearly as good as the theory allows, so other factors are important and need to be taken into account. We think the main one is the fact that the neural system pools motion signals over wide ranges of direction and velocity.

Received March 4, 1997; revised July 25, 1997; accepted July 25, 1997.

The work was supported by Grants from the Biotechnology and Biological Sciences Research Council and the Newton Trust. We thank Roland Baddeley for helping set up the early experiments and Valerie Bonnardel for her helpful comments as an observer for all of them.

Correspondence should be addressed to Horace Barlow at the above address.

Copyright © 1997 Society for Neuroscience 0270-6474/97/177954-13\$05.00/0

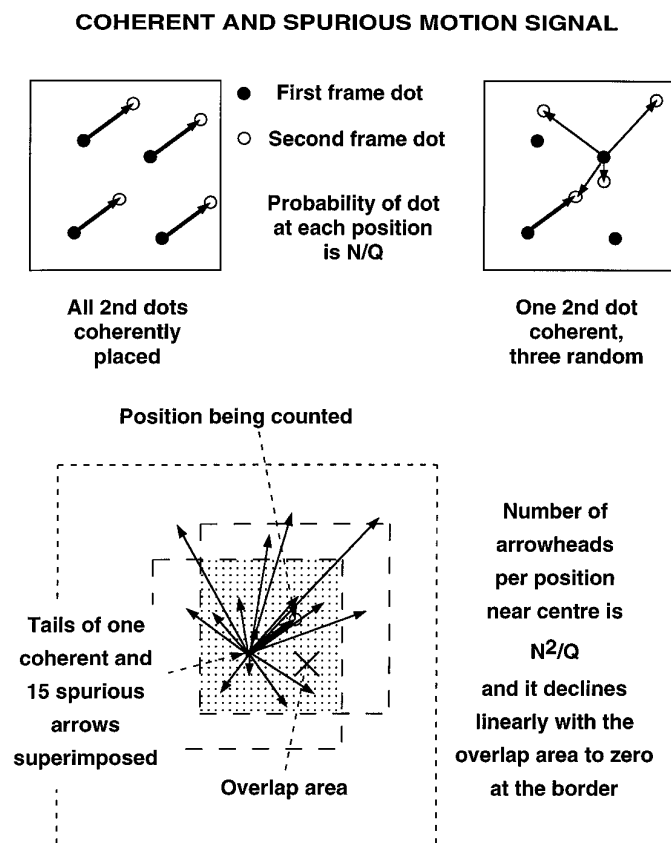


Figure 1. Correspondence noise. Spurious motion signals are generated when a dot in the first frame (filled circle) is incorrectly paired with a dot in the second frame (open circles), as shown at top right. For N dots in each frame there are N^2 possible pairings, of which CN are formed by coherent displacements and $N^2 - CN$ are spurious. Bottom, How the noise from spurious signals is calculated. The tails of all possible motion arrows are aligned, and the number of arrowheads is counted at the position corresponding to the motion that is to be detected. With N dots and Q possible positions, assuming a small movement and low coherence, the expected number of arrowheads at each position is nearly N^2/Q ; for larger movements, this figure decreases linearly to the edge of the overlap area. Assuming Poisson statistics, the noise is the square root of the expected number, approximately $(N^2/Q)^{1/2}$.

This is far from optimal for detecting coherence in kinematograms generated in the usual precise way, although it does appear to be well adapted to detecting motion in natural images. When the method of generating kinematograms is modified to require extensive pooling in the ideal detector, the ideal coherence thresholds are greatly elevated, and the difference between ideal and measured thresholds is correspondingly reduced.

To summarize our conclusion, we think that motion information in random dot kinematograms is pooled over wide ranges of direction and velocity as well as large areas of the visual field. Such pooling is desirable to capture all the motion signals in natural images, but it results in high levels of correspondence noise. This source of external noise is an important (although not necessarily the only) limiting factor in the task that has been such an effective tool in analyzing the neurophysiology of MT (V5). The fact that this example of cortical processing approaches a statistical limit inherent in the incoming signal has important implications for understanding how the cortex is organized to perform its sensory role.

THEORY

Definitions

N , total number of dots in a field; N_i , number for field i .

Q , total number of possible dot positions in a field; there are 1871 uniformly distributed dot locations per deg^2 in our conditions: note that usually $Q \gg N$.

$p = N/Q$, probability of a dot at a particular position.

A , stimulus area in deg^2 .

C , the proportion of dots coherently moved between fields.

C_θ , the threshold coherence, i.e., the coherence required for $d' = 1$.

$C_{\theta, \text{ideal}}$, the ideal threshold coherence.

$\langle S_C \rangle$, expected number of vectors for a particular coherence.

T , the number of displacements; the number of fields = $T + 1$.

α , the number of dot positions in which the head of a motion vector can fall and still be counted as coherent.

ϕ , the half-angle defining the sector within which a coherently moved dot is distributed in the randomization experiments.

In most of the experiments we have measured the proportion of dots that must be coherently moved for an observer to be able to discriminate between leftward and rightward motion with $d' = 1$ (see Materials and Methods for more details). We regard this two-alternative, forced choice direction discrimination (2AFC) task as a convenient way of estimating the detection threshold, that is, the proportion of dots that must be coherently moved to detect that motion is present with $d' = 1$, and the main theoretical exposition of the dependence on parameters of the stimulus is done for detection, because this is conceptually clearer and simpler. The theory for the 2AFC task is complicated by the change in SD of the decision variable when the coherence level rises, requiring a quadratic to be solved to predict threshold values of coherence. For simplicity we have skipped this stage in the exposition of the theory below, merely giving the expressions for d' in the 2AFC experiment. In checking the predicted performance of the ideal, correspondence noise-limited, thresholds, we have done extensive Monte Carlo simulations for which we have closely simulated the actual 2AFC experiments.

The ideal detector of coherent motion would base its decision on all the information present in the stimulus, so it would examine all possible correspondences, and count the number of vectors for motion with the particular direction and velocity of interest. Some vectors will result from the coherent displacement of first frame dots, but others will occur unintentionally as a result of a dot that was placed randomly in the second field occupying the position for the coherent motion of a first field dot. Note that some authors refer to the fraction of coherently moved dots C as the signal/noise ratio, but this is incorrect; it is random variation in the number of spurious motion vectors that sets the limit to the detection of coherent motion.

It may seem unrealistic to suppose that the real motion system counts vectors in this way and does it with the precision that is available in a typical screen display, but the object at this stage is to calculate ideal performance, irrespective of neural limitations. At a later stage we consider how limited precision of the motion detector system would influence the results.

Ultimately we predict how ideal thresholds would change when the following parameters of the random dot kinematograms are changed: dot density; ratio of dot numbers in the first and second fields; stimulus area; number of successive fields; method of dot generation; number of possible positions for the dots; and the number of dot positions over which the coherently displaced dots

are distributed. In the next section we show in detail how to predict the ideal coherence threshold as a function of the density of the dots in each field, using for clarity a slightly oversimplified theory; we neglect the border effects that result from either the first frame dot of a coherent pair or the second frame dot, lying beyond the edge of the stimulus; we assume that the threshold is at a low coherence and that the velocity and possible directions of the motion to be detected are known. Then in the following sections some of these complications are considered. The predictions for the other experimentally variable parameters of the stimulus are set out in conjunction with the methods and results for that particular experiment. Modified expressions giving d' for the 2AFC task as opposed to motion detection are included.

Dot density

In Figure 1, *bottom*, all possible dot pairs in two fields are represented by arrows with tails that have been superimposed; the limit imposed by correspondence noise is then brought out by examining the number of arrowheads at one particular position. With N dots in each field there are N^2 possible vectors. The optimum method for detecting movement of known direction and velocity is to count the vectors for that movement, because this measure includes all the signal dot pairs and does not include any unnecessary spurious pairs. For each dot in the first field the probability of a dot at the appropriate position in the second field is N/Q , and there are N dots in the first field. Hence when there are no coherently moved dots the expected number of arrowheads $\langle S_0 \rangle$ in the position corresponding to a particular motion, and its SD from binomial statistics, are:

$$\langle S_0 \rangle = \frac{N^2}{Q} \quad (1)$$

$$\sigma(S_0) = \sqrt{\frac{N^2}{Q} \left(1 - \frac{N}{Q}\right)}. \quad (2)$$

If a proportion C of the first frame dots is coherently moved, there will be CN additional arrowheads in the relevant positions, but the number expected there by chance is reduced, because the CN coherently moved ones are definitely there and, hence, removed from those available to occur there by chance. $\langle S_C \rangle$ is therefore given by Equation 3, and $\sigma(S_C)$ by Equation 4; note that the term CN in Equation 3 does not contribute to the variance of S_C . d' for detection is given by Equation 5 and for 2AFC discrimination by Equation 6:

$$\langle S_C \rangle = CN + (1 - C) \frac{N^2}{Q} \quad (3)$$

$$\sigma(S_C) = \sqrt{(1 - C) \frac{N^2}{Q} \left(1 - \frac{N}{Q}\right)} \quad (4)$$

$$d' \text{ (detection)} = \frac{\langle \Delta S \rangle}{\sigma(S_0)} = \frac{\langle S_C \rangle - \langle S_0 \rangle}{\sigma(S_0)} = C \sqrt{Q - N} \quad (5)$$

$$d' \text{ (2AFC)} = \frac{\langle \Delta S \rangle}{\sigma(\Delta S)} = \frac{\langle \Delta S \rangle}{\sqrt{\sigma^2(S_C) + \sigma^2(S_0)}} = C \sqrt{\frac{Q - N}{2 - C}}. \quad (6)$$

The threshold value of C_θ is given when $d' = 1$, so for detection:

$$C_\theta = \frac{1}{\sqrt{Q - N}}. \quad (7)$$

It will be seen by inspection that, if correspondence noise is the limiting factor, the number of dots N has very little effect on the ideal coherence threshold provided it is much less than Q , the number of available positions for dots. This somewhat counterintuitive prediction results from the fact that the number of spurious motion signals rises as the square of dot density, so its SD is directly proportional to dot density rather than to its square root, which is the more usual case (also see Laming, 1986; Maloney et al., 1987).

When Q is reduced in the quantization experiments to be described below, it becomes closer to N , and we no longer expect the coherence threshold to be uninfluenced by the number of dots.

Border effects

The distribution of *arrowheads* in Figure 1 is nonuniform, because a dot near an edge of the first field cannot be observed to move to a position outside the second field, and similarly, some dots in the second field cannot be observed to have moved from positions outside the first field. From the way the figure is constructed one can see that the density is actually proportional to the area of overlap of the two fields when one of them is displaced through a distance equal to the length of the vector for a given direction and velocity of motion, so the density is equal to N^2/Q at the center and declines linearly to the edge where there is no overlap between the two fields. Corrections can be calculated and are small for movements that are small compared with the width of the fields. The standard correction is not accurate when there is an additional random component to the displacement of the coherent dots (see Randomization in Results), and in these cases, as well as others, we have used Monte Carlo simulations.

Lack of independence of motion pairs

To calculate ideal performance in the experiments to be described in Randomization, the number of vectors had to be counted within a certain range of the vector corresponding to the mean coherent displacement. Under these conditions the same vector can contribute more than once to the total, and it is no longer possible to assume that the variance behaves according to binomial statistics. Again, this problem was handled by doing Monte Carlo simulations.

The dependence of ideal performance on changes in the other parameters we have varied are described with the experimental results.

MATERIALS AND METHODS

Equipment

The stimuli were generated using a Silicon Graphics Iris Indigo computer and displayed on a Silicon Graphics TFS6705KG-SG monitor with a frame rate of 67 Hz and a medium persistence P22 phosphor (the slowest phosphor decayed to <1% of initial luminance within 5 msec). Pixel separation was adjusted to be 0.23 mm in the horizontal and vertical directions. A computer mouse was used to input observer responses. A chin rest minimized head movements during the experiment. A black cardboard aperture was used to limit the visible area of the screen in early experiments and was replaced by a software aperture in later experiments.

Psychophysical procedure

Stimulus. As a result of preliminary experiments and a literature search (Morgan and Ward, 1980; van Doorn and Koenderinck, 1982a,b; Fredericksen, et al., 1993), we selected the following typical stimulus area, duration, and interstimulus interval. Most of our experiments have used

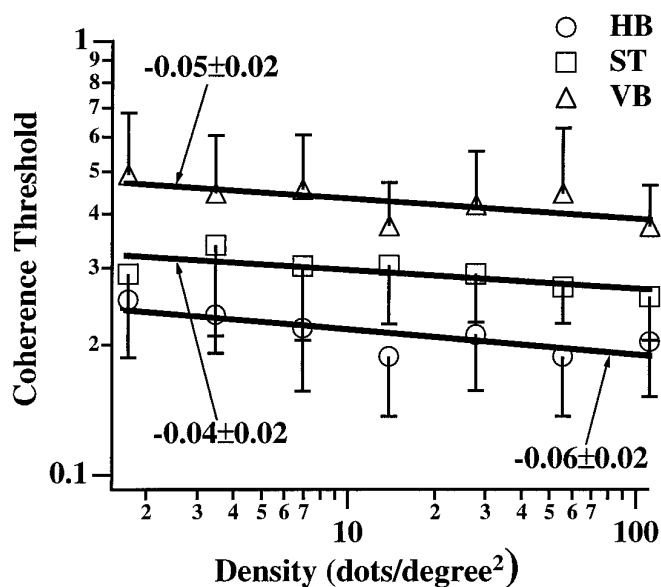


Figure 2. Effect of dot density. Coherence thresholds are plotted against dot density using logarithmic axes. Also shown are estimated SEs and regression lines with their slopes. In these three observers (*HB*, *ST*, *VB*), increasing the dot density 64-fold decreases the thresholds by 22, 15, and 19%.

only two sequentially presented fields, each having 100 dots on average. All experiments that used a software window had 100 dots exactly (i.e., all experiments except those depicted in Figs. 2, 4, 8, 9). Each field was displayed in the same position for 10 frames (150 msec), with no added interval between the fields. Each dot within a field was a square of size 2×2 pixels, and in most experiments a large area filled with such dots had a luminance of 78.2 cd/m^2 , the background luminance being 0.9 cd/m^2 ; the monitor had to be changed for a few of the later ones, and the replacement had a luminance of 46.3 cd/m^2 on a background of 0.3 cd/m^2 . The dots were randomly distributed over a circular aperture area of radius 2.15 deg when viewed from a distance of 114.6 cm. The area of such a field is 14.5 deg^2 , and the number of possible dot positions Q is 27,169. The maximum value of N in our experiments was 6400.

The motion signal on a trial was generated by displacing a proportion C of the dots by 16 pixels (11 arc-min) between the first and second fields and the remaining proportion $(1 - C)$ of the dots was randomly distributed within the aperture. A circular wraparound was used when the displaced dot moved out of the aperture. In most experiments the observers had to make a forced choice between rightward and leftward movement, but we have also done motion detection experiments in which the observers' task was to decide whether there was any coherent motion. The dot density, aperture size, and quantization experiments were conducted using both paradigms, but the direction discrimination paradigm gave less variability, and because the results were otherwise similar, only the direction discrimination experiments are described here.

Deviations from the typical stimulus are described below with the description of each particular experiment.

Procedure. A method of constant stimuli was used so that within a run the coherence level C took on one of nine predefined values, four leftward motion, four rightward, and one zero. The predefined values were selected so that the observer's responses covered a large proportion of the psychometric function. Four blocks of 180 observations were made, 20 at each coherence level. The first block was regarded as practice and was discarded. In addition, at the beginning of each block the observer could deliver sample displays by pressing a mouse button.

During testing the observer sat in the dark room viewing the display screen, with his or her chin on a chin rest. After the presentation of each stimulus the observer indicated, using appropriate buttons on the mouse, whether the motion was leftward or rightward. Observers also had the option of discarding trials (by pressing a third mouse button) in case of an attentional lapse. They were instructed not to use this option as a substitute for guessing when the motion stimulus was below threshold. Error feedback was provided when the observer reported the direction of

motion incorrectly. The experiments were self-paced, with each trial taking place only after the observer had responded to the previous trial.

Observers. The two authors and three naive observers participated in the various experiments. All observers had corrected to normal vision.

Data analysis. Probit analysis (based originally on the work of Finney, 1947) was used to evaluate the data. A cumulative normal Gaussian function was fitted to the data, giving percent of rightward responses versus percent coherence, which ranged from -100 (fully coherent leftward motion) to $+100$ (fully coherent rightward motion). The slope of the probit regression line corresponds to the SD of the best fitting cumulative Gaussian function, and this gives the coherence threshold for $d' = 1$. The calculation for each threshold was based on 540 observations (9 levels \times 60 observations).

Monte Carlo simulations

Theoretical predictions were backed up by Monte Carlo simulations when evaluating ideal performance in the quantization and randomization experiments (see Results). The positions of the dots in the simulated stimuli were generated using a procedure identical to that used for the psychophysical experiments. For each first field dot, two target zones were defined in the second field, one on either side of the first field dot. The number of left target zone dots was subtracted from the number of right target zone dots, yielding the signal for the residual rightward movement. On each trial this was summed over the second field target zones for all the first field dots to yield the net rightward signal. The ideal observer made a decision as to whether the motion was rightward or not from the value of this sum on each trial. The trial was repeated 300 times at a given coherence level to evaluate the proportion of rightward responses at that coherence level. The simulations were repeated at nine coherence levels, and probit analysis of the ideal observer's psychometric function was used to estimate the ideal coherence threshold, which was the change in coherence necessary to discriminate the direction of motion with a d' of 1.

Statistical efficiency

Statistical efficiency (Fisher, 1925; Swets, 1964) of the human observer was evaluated for the quantization and randomization experiments. In this case the evidence is not simply proportional to the number of dots in the stimulus, so the calculation of statistical efficiency η is based on the values of d' :

$$\eta = (d'_{\text{human}}/d'_{\text{ideal}})^2 = (C_{\theta, \text{ideal}}/C_{\theta, \text{human}})^2, \quad (8)$$

where the two discriminabilities are for stimuli of the same coherence level. $C_{\theta, \text{ideal}}$ can be evaluated from the Monte Carlo simulations described in the previous section.

Variable dot life kinematograms

In the majority of neurophysiological experiments kinematograms have been displayed point by point, and the coherence level has been varied by adjusting the probability of a given dot being coherently moved at each refresh cycle. Our kinematograms were generated and displayed field by field instead of point by point, and we usually selected the coherently moved dots from those that had *not* just been moved (the "different" method of Scase, et al., 1996). In variable dot life kinematograms, at low coherence levels the great majority of dots also move only once, but at high coherence levels a dot persists for more than the equivalent of two fields in our kinematograms. In a few experiments we used the same dots for each successive displacement, but confirming the results of Scase et al. (1996), this did not make much difference to the observed thresholds. We therefore do not think our different method of generation affects the comparison with neurophysiological experiments even when we were using multiple successive fields.

RESULTS

Variation of overall dot density

If correspondence noise limits performance, the prediction is that the coherence threshold will vary very little with dot density (see Theory), and Figure 2 shows the results for three observers on logarithmic axes. Within a block, the stimulus consisted of a fixed number of dots in each field. Between blocks the number was selected to be 25, 50, 100, 200, 400, 800, or 1600 dots, which correspond to densities from 1.7 to 111 dots/deg^2 . There was a

small but reliable decrease of coherence threshold with increasing dot density, the best-fitting regression lines having a mean slope of -0.05 ± 0.02 . Notice that this corresponds to a threshold drop of $<20\%$ for a 64-fold change of dot density.

We were unable to find the limits for this near invariance of coherence threshold with dot density. At the lowest density there were only 25 dots in the stimulus and only five coherent dots for the threshold coherence level of 20%. At high densities stimulus generation was becoming tediously slow, and the dot density was obviously far beyond a value at which individual dots were countable, so the neural system must already have been using an analog mechanism, presumably a correlation mechanism, a spatio-temporal filter, or some form of motion energy detector.

Separate variation of dots in first and second fields

With N_1 dots in the first field and N_2 dots in the second there are CN_1 coherently moved dots and N_1N_2 possible random pairings. The ideal coherence threshold $C_{\theta,ideal}$ can be derived:

$$\langle S_C \rangle = CN_1 + (1 - C) \frac{N_1N_2}{Q} \quad (9)$$

$$\sigma(S_0) = \sqrt{\frac{N_1N_2}{Q} \left(1 - \frac{N_2}{Q}\right)} \quad (10)$$

$$d' \text{ (detection)} = \frac{\langle S_C \rangle - \langle S_0 \rangle}{\sigma(S_0)} = C \sqrt{\frac{N_1(Q - N_2)}{N_2}} \quad (11)$$

$$d' \text{ (2AFC)} = C \sqrt{\frac{N_1(Q - N_2)}{N_2(2 - C)}} \quad (12)$$

$$C_{\theta,ideal} = \sqrt{\frac{N_2}{N_1(Q - N_2)}} \quad (13)$$

As before, Q is 27,169, and the maximum value of N is 6400, so the square root of the ratio N_2/N_1 dominates the relation if correspondence noise is the limiting factor. Note that coherence is expressed as a fraction of N_1 ; that is, the number of coherently moved dots is CN_1 .

The experimental results are shown in Figure 3. Within a block, the first field had a fixed value between 50 and 1600 for its N_1 dots, and the ratio N_2/N_1 was also at a fixed value between 0.5 and 4.0. Between blocks, N_1 and/or the ratio N_2/N_1 were varied. Measurements were not taken for the combination of $N_2/N_1 = 4$ with $N_1 > 100$, because of the high thresholds (approaching 100% coherence) and the limitation that the coherence level in the stimulus (C) could not physically exceed 100%. Also, measurements were not taken for values of $N_2/N_1 < 0.5$, because smaller values of N_2/N_1 meant smaller values for the range of C , because the proportion of coherent dots in the stimulus cannot exceed N_2/N_1 .

For each value of N_1 tested, Figure 3 plots threshold coherence against the ratio of the number of dots in the second field to the number of dots in the first field on logarithmic axes. Figure 3A shows results for three observers for a displacement of 16 pixels (11 arc-min), and Figure 3B shows results for two observers at a displacement of 8 pixels (5.5 arc-min). The *solid lines* represent straight-line fits to the data for $0.5 < N_2/N_1 < 2.0$. The data for $N_2/N_1 = 4$ were excluded from the fit because observers experienced difficulty in making the judgments, and the points are obviously above the line passing through the other data. Possible reasons for this are (1) the difference in mean luminance of the two fields makes the matching task very difficult; and/or (2)

backward masking from the second frame affects the visibility of the first frame dots.

The results up to $N_2/N_1 = 2$ fall along lines having slopes ranging from 0.52 to 0.65; the SEs in the estimates of the slopes are ± 0.05 . Thus the observed slopes, although reliably greater, are close to the slope of 0.5 predicted from the correspondence noise limit.

Variation of stimulus area

From the expression (Eq. 7) derived in Theory, it will be seen that ideal threshold should be proportional to $(Q - N)^{-1/2}$, where Q is the number of possible positions for a dot in the stimulus, and for the current experiments $N \gg Q$. Because Q is proportional to stimulus area A , $C_{\theta,ideal}$ should therefore also be closely proportional to $A^{-1/2}$.

For the results shown in Figure 4 the dot density was 6.9 dots/deg² as in the typical stimulus, but now the area was varied using five different circular apertures in cardboard sheets varying from 3.6 to 57.8 deg². In a sixth condition, the cardboard sheet was removed, and the stimulus consisted of the entire rectangular screen of area 171.6 deg². The threshold coherence is plotted as a function of effective aperture area for two observers on logarithmic axes. The effective aperture area is the area of the stimulus that contributes to the motion signal after the correction (derived geometrically; see Fig. 1) for dots moving out of or into the stimulus region. The *solid line* shown has a slope of 0.5.

For effective areas below ~ 3 deg² the data definitely have a slope >0.5 , and again when the area exceeds ~ 12 deg² the data have a slope <0.5 . There is a transitional region of two octaves in area where the square root law predicted from the correspondence noise limit holds approximately. Other factors must be sought to explain the deviations at smaller and larger areas.

Variation of number of displacements: "different" generation

The way that ideal performance depends on the number of fields varies according to the way that the displays are generated. In the "different" method of generating coherent motion (as defined by Scase et al., 1996) the CN dots in a field that have coherently moved partners in the next field are selected at random from dots that were not coherently moved from the previous field; in the "same" method (see below), the same dots are coherently moved between each successive pair of fields. To detect coherence optimally in "different" kinematograms, each successive pair of fields must be treated independently, because this corresponds to the way they are generated. In these kinematograms there will be no coherent signal from nonconsecutive frames, but the neural system may well be sensitive to such correlations, and spurious pairs in nonconsecutive frames could contribute to noise. These possibilities would need to be considered in a fuller treatment.

If the T displacements between the $T + 1$ fields are independent, the optimal treatment is simply to add the number of dots at the predicted positions over all successive field pairs:

$$\langle S_C \rangle = CTN + (1 - C) \frac{TN^2}{Q} \quad (14)$$

$$\sigma(S_0) = \sqrt{\frac{TN^2}{Q} \left(1 - \frac{N}{Q}\right)} \quad (15)$$

$$d' \text{ (detection)} = \frac{\langle \Delta S \rangle}{\sigma(S_0)} = C \sqrt{T(Q - N)} \quad (16)$$

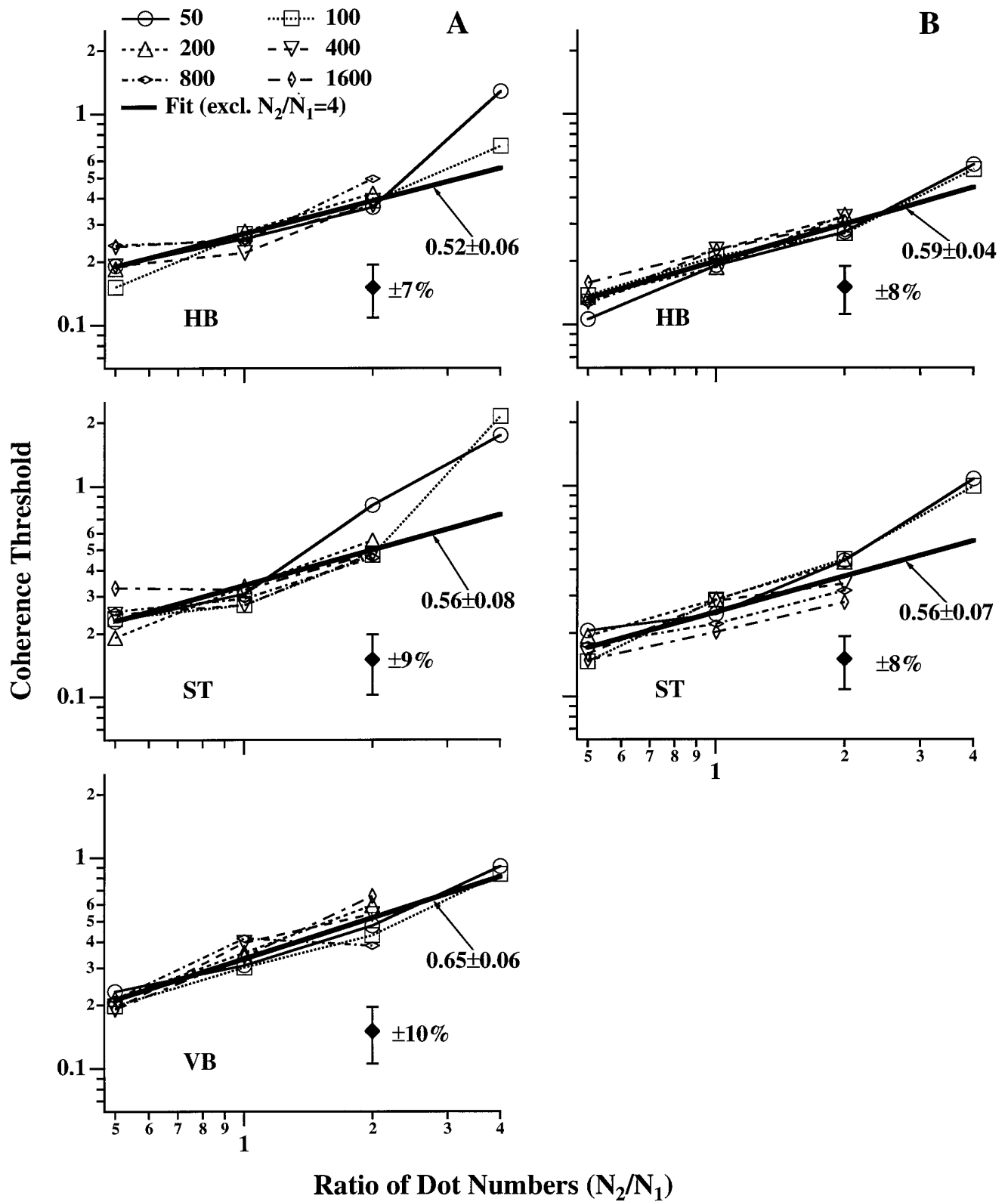


Figure 3. Effect of ratio N_2/N_1 . Coherence thresholds are plotted against the ratio of number of dots in the second frame to number of dots in the first frame using logarithmic axes. Thresholds are shown for six different first frame dot numbers for a displacement of 11 arc-min and three observers (A), and a displacement of 5.5 arc-min and two observers (B). The prediction from the correspondence noise limit is a line of slope 0.5. The thick lines are regressions excluding (excl.) the data for 4:1 ratio, in which the task was made difficult by the second frame being much brighter than the first.

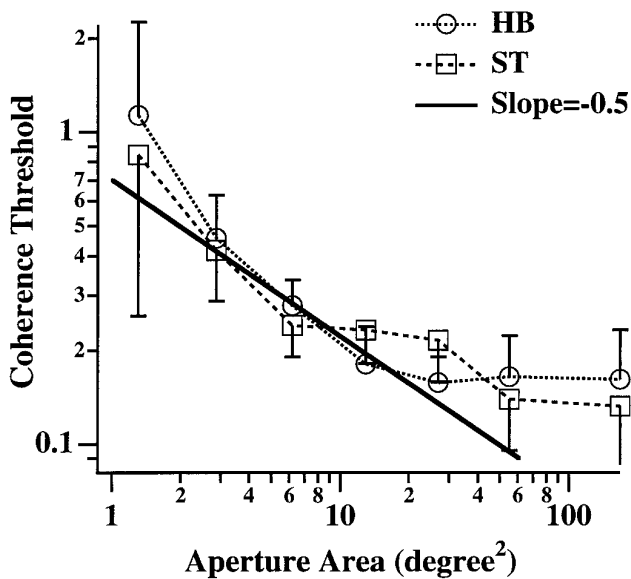


Figure 4. Effect of aperture area. Coherence thresholds are plotted against stimulus area for two observers (HB, ST) using logarithmic axes. The area of the stimulus was corrected for the border effect using the geometric principle illustrated in Figure 1. The line has a slope of -0.5 , the value predicted from the correspondence noise limit. Deviations from this prediction are evident at $<3 \text{ deg}^2$ and $>12 \text{ deg}^2$.

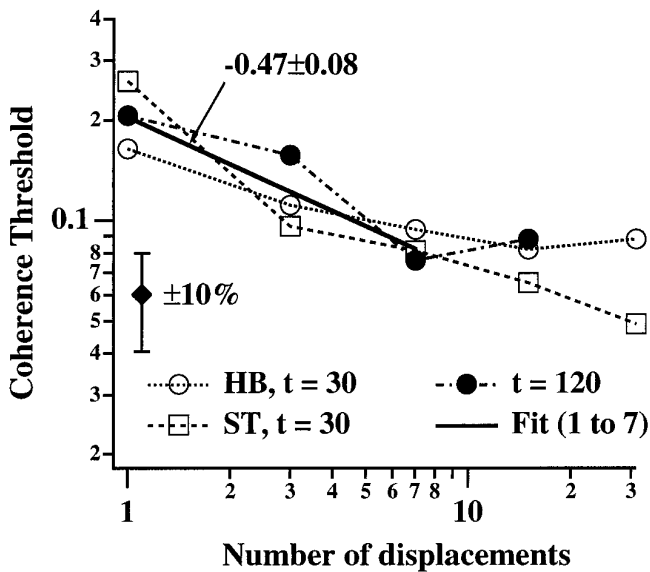


Figure 5. Effect of multiple fields for different-generated kinematograms. Using logarithmic axes, coherence thresholds are plotted against the number of displacements for two observers at a field duration of 30 msec (frames repeated twice) and for one observer at a field duration of 120 msec (frames repeated 8 times). The correspondence noise limit predicts a slope of 0.5 . The thick line with a slope of -0.47 ± 0.08 is the best fit to the data up to 7 displacements (8 fields).

$$d' (2AFC) = C \sqrt{\frac{T(Q - N)}{2 - C}} \quad (17)$$

$$C_{\theta, \text{ideal}} = \frac{1}{\sqrt{T(Q - N)}} \quad (18)$$

The ideal coherence threshold will therefore be proportional to $T^{-1/2}$ if correspondence noise is the limiting factor.

For the results shown in Figure 5 each stimulus consisted of either 2, 4, 8, 16, or 32 fields, the number being fixed within a block. Between fields n and $n + 1$, a proportion C of the dots in field n was displaced using the “different” method of coherent dot generation described above, whereas the rest were randomly replaced. Measurements were made when each field in the stimulus was presented for durations of either 30 msec (two frames) or 120 msec (eight frames), although the combination of 32 fields with 120 msec field duration was not used because of the tediously long duration of each stimulus.

In Figure 5 coherence thresholds for two observers at a field duration of 30 msec and one observer at a field duration of 120 msec are plotted on logarithmic axes as a function of the number of displacements. The theory predicts that thresholds will fall along a line with a negative slope of 0.5 . Deviations from this prediction appear to set in at about seven displacements, although the threshold goes on dropping out to 31 displacements. The thick line shows the best fit to the data when the number of displacements ranged from one to seven and has a slope of -0.47 ± 0.08 , which is reasonably close to the theoretical prediction.

Variation of number of displacements: same generation

In real life, moving objects can often be followed for considerable periods. This can be imitated in a random dot kinematogram by moving the same dots from field to field, rather than making the coherently moved pairs different, as was done above. This is the “same” method of generating kinematograms defined by Scase et al. (1996), and Watamaniuk et al., (1995) have shown that we are extremely sensitive to trajectories generated by this method; a single dot tracing a trajectory among dots in Brownian motion can be reliably detected.

If a kinematogram has been generated by the “same” method, for a fraction of the first frame dots there will be a dot at the expected position in every subsequent frame. The optimum strategy is therefore to inspect the string of positions in subsequent fields for all the first field dots and to count the number of strings for which all positions are occupied. Such a string may have been caused by a coherently moved dot, or it might have arisen from chance occupancy in each successive field. For the first field, N positions are occupied. For the second field, the number expected by chance in the selected positions is Np , where $p = N/Q$ as before. After T displacements the expected number of strings in which all positions are occupied is Np^T .

$$\langle S_C \rangle = CN + (1 - C)N\left(\frac{N}{Q}\right)^T \quad (19)$$

$$\sigma(S_0) = \sqrt{N\left(\frac{N}{Q}\right)^T \left(1 - \frac{N^T}{Q^T}\right)} \quad (20)$$

$$d' (\text{detection}) = \frac{\langle S_C \rangle - \langle S_0 \rangle}{\sigma(S_0)} = C \sqrt{\frac{Q^T - N^T}{N^{T-1}}} \quad (21)$$

$$d' (2AFC) = C \sqrt{\frac{(Q^T - N^T)}{N^{T-1}(2 - C)}} \quad (22)$$

$$C_{\theta, \text{ideal}} = \sqrt{\frac{N^{T-1}}{Q^T - N^T}} \quad (23)$$

Note again that $Q \gg N$.

A multiple-coincidence detecting system would definitely help distinguish from correspondence noise an object moving contin-

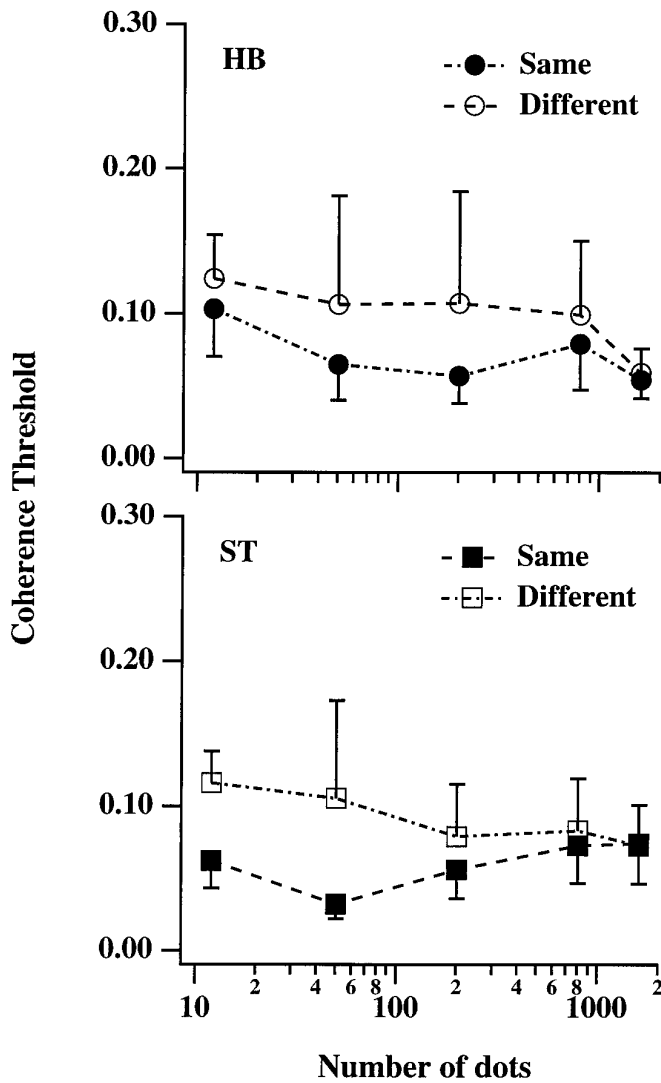


Figure 6. Effect of dot density with multiple fields. Coherence thresholds are plotted against dot density for two observers for same-generated kinematograms of 5 fields. Results with different-generated kinematograms are included as a control. The thresholds do not rise with the dot density raised to the power of 1.5, as predicted by the correspondence noise limit under the assumption that advantage is taken of the same dots being moved coherently in the same condition (see Eq. 23).

uously across the field of view, so it is interesting to look for evidence for its presence. Two quantitative predictions can be tested: (1) for “same” generated kinematograms with fixed T , the coherence threshold should be strongly dependent on dot density, unlike the case with “different” generation of the kinematogram; and (2) the square of the coherence threshold should decline exponentially with the number of fields.

Figure 6 tests the first prediction. Four displacements (five fields) were used, so the coherence threshold should be proportional to $N^{3/2}$. In fact there was very little if any dependence on N . For comparison, results using the different method of generation are also shown; these are higher than those for the same generated kinematograms and show the small decline with N that was previously found (Fig. 2).

Figure 7 tests the second prediction. The stimulus was similar to the one used for Figure 5, with the coherently moved dots

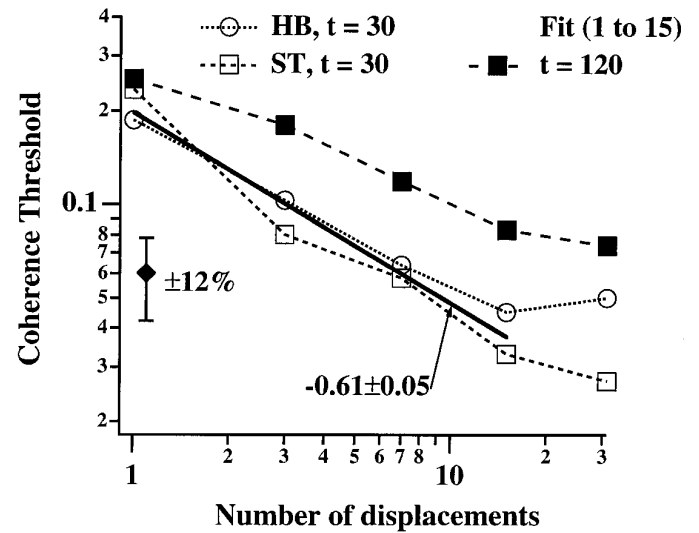


Figure 7. Effect of multiple fields for same-generated kinematograms. Coherence thresholds are plotted against the number of displacements for two observers (*HB*, *ST*) at a field duration of 30 msec and one observer at a field duration of 120 msec. There is no evidence for the threshold dropping exponentially with the number of displacements, as it would if there were a mechanism taking advantage of the same dots being moved from field to field, and if correspondence noise were limiting (see Eq. 23). The *thick line* shows the best fit to the 30 msec field duration data for up to 15 displacements.

displaced using the “same” method of generation. The coherence thresholds for the two observers at a field duration of 30 msec and one observer at a duration of 120 msec are plotted as a function of the number of displacements. Coherence thresholds were again lower than those found with the “different” method of generation, but they showed no sign of dropping off exponentially, as predicted by the theory that multiple coincidences are used. The *thick line* is the best fit to the 30 msec field duration data for up to 15 displacements and has a slope of -0.61 ± 0.05 . Thresholds measured with a 120 msec field duration were slightly higher and were excluded from the fit. The thresholds for the “same” generation of kinematograms fall off rather more steeply, and they continue to fall over a larger number of displacements than was the case for the different generation.

The fact that thresholds are lower with “same” generation remains to be discussed, but there is no evidence in these results of a mechanism that would optimally detect same-generated kinematograms up to the correspondence noise limit, even when only four displacements are used.

Absolute efficiencies

The theory has so far shown five ways in which measured coherence thresholds should change with the parameters of the stimulus if correspondence noise is the limiting factor, and experimental results have shown the conditions in which these predictions are followed and not followed. The theory also predicts the absolute performance, and under the conditions in which the variations with a stimulus parameter indicate that correspondence noise is limiting, one might expect human performance to approach the theoretical limit. In fact, the theoretical limit is enormously better than human performance under all the conditions so far described, so much better that the theoretical limit has not even been indicated in the figures. To understand the motivation for the next experiments, a possible reason for this discrepancy must be explained.

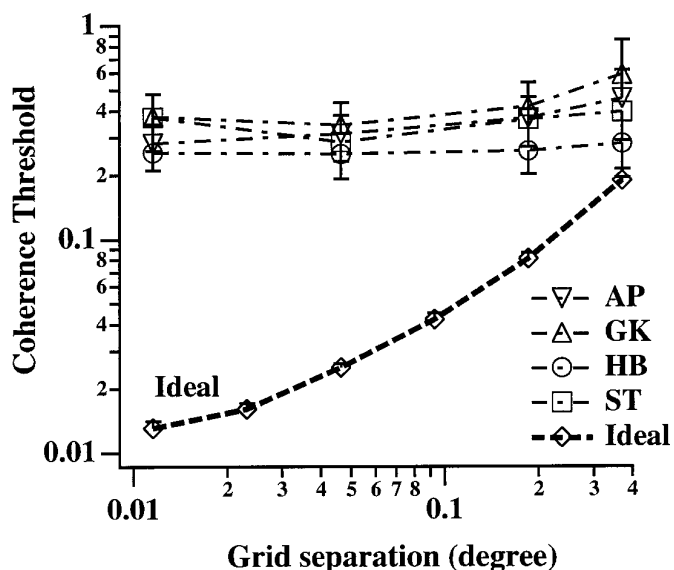


Figure 8. Effect of quantization on coherence thresholds. The coherence thresholds when dots are confined to lattice points with varying grid separations are shown for four observers (*AP*, *GK*, *HB*, *ST*) on logarithmic axes. Also shown are thresholds for an ideal detector. As expected, coarse quantization impairs ideal performance greatly. It has little effect on human coherence thresholds, presumably because the neural motion system is insensitive to precise dot positioning.

Ideal thresholds have been calculated on the assumption that the position of every dot is known to the system with the precision with which it is displayed, but it is unreasonable to assume that this is true for the neural mechanism, which is likely to treat as coherent any vector with a head that lies close to the expected position. Suppose that α such positions are accepted in an otherwise ideal detector. Then one can recalculate Equations 1–7 on this basis and reach the conclusion:

$$C_{\theta, \text{ideal}} = \frac{1}{\sqrt{Q/\alpha - N}} \approx \sqrt{\alpha/Q} \text{ if } N \ll Q/\alpha. \quad (24)$$

Because the efficiency is the square of the ratio of ideal to actual C_{θ} , a detector with a raised value of α will be very inefficient in comparison with the ideal detector, and we think this may be a major factor hampering the performance of the human motion-detecting system. Furthermore, Equation 24 shows that the ratio α/Q is the important factor, so the ability of the ideal detector to exploit the high precision of dot placement would be affected by changing either of them.

In the next experiments the precision of dot placement in the kinematograms was reduced to test whether this reduced the discrepancy between ideal and actual performance to reasonable values. In the quantization experiment this was done by reducing Q , and in the randomization experiment it was done by forcing the ideal detector to use a high value of α . Forcing the ideal detector to use a high value of α has been shown to increase the absolute efficiency of symmetry detection, which is in some ways a comparable problem (Barlow and Reeves, 1979).

Quantization

Figure 8 shows an experiment in which the separation of lattice points was varied in steps between 1 and 32 pixels, decreasing Q to 1/1024 of its usual value by coarsening the grid on which the dots were constrained to fall. With coarser quantization the

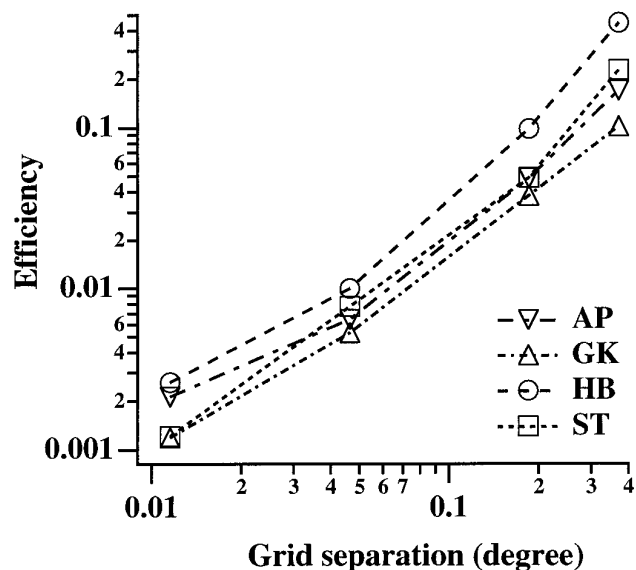


Figure 9. Effect of quantization on efficiency. Using log axes the estimated statistical efficiency is plotted as a function of grid separation for four observers (*AP*, *GK*, *HB*, *ST*). Efficiency rises when kinematograms are generated in a way that matches the coarse resolution of the motion-detecting system, so that the ideal observer cannot gain an advantage from its greater precision.

probability of occupancy of a position rose so that N was no longer much smaller than Q , as it has been in the other experiments so far described. When the grid separation was 1, dots occasionally partially overlapped, but controls showed that this had little effect on the results. In the condition in which the grid separation was 32 pixels (22 arc-min), the second field had its entire grid displaced laterally by 16 pixels (11 arc-min) to accommodate a 16 pixel (11 arc-min) displacement to the right or left.

In Figure 8 the coherence thresholds for four observers are plotted against the grid separation on logarithmic axes, together with the ideal coherence thresholds obtained from simulations. The following observations can be made about the data: (1) for the human observer the coherence threshold changes only slightly over the range of grid separations tested; (2) for the ideal observer the coherence threshold rises dramatically over the same range; and (3) for a finely quantized stimulus (stimulus with small grid separation) the threshold for the ideal observer is much lower than that of the human observer, whereas for a coarsely quantized stimulus (stimulus with large grid separation) the threshold for the ideal observer approaches that of the human.

Figure 9 shows the effect of grid separation on statistical efficiency calculated from the data shown in Figure 8. As the grid separation (i.e., the coarseness of quantization) is increased, efficiencies increase to values ranging from 10 to 44%. The interobserver variability is exaggerated, because coherence thresholds are squared when calculating efficiency.

In the coarsely quantized stimulus with 100 dots, a large proportion of the grid locations were occupied by dots. To see whether the high probability of occupancy was important, we repeated the experiment using the largest grid separation but with 25 or 50 dots. The efficiencies found were comparable with those shown for the same grid separation in Figure 9.

Randomization

If coherently moved dots are scattered over a range of positions instead of being placed at their precisely correct positions, then it

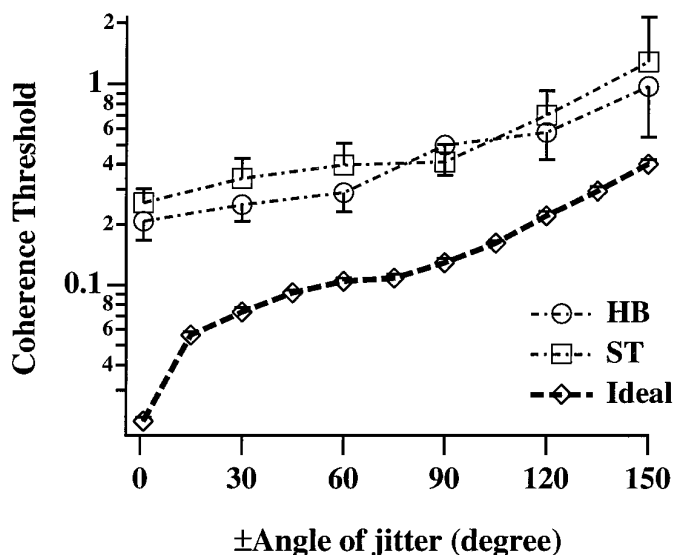


Figure 10. Effect of randomization on coherence thresholds. The positions of the coherently moved dots were randomized to varying extents. The dots were displaced to random positions within a sector of radius 8 ± 7.5 arc-min and \pm angle as shown on the *abscissa*. The corresponding coherence thresholds of two observers (*HB*, *ST*) are shown on log-linear axes. As expected, if the neural motion system is insensitive to precise dot positioning, coarse randomization impairs ideal performance to a greater extent than it impairs human coherence thresholds.

will be necessary for the ideal detector to pick up signals from this range of scattered positions; if it does not, it will fail to count some coherently displaced dots and will perform nonoptimally. Accordingly, in these experiments the signal dots were displaced randomly and uniformly in a sector that was $\pm\phi^\circ$ of horizontal and 12 ± 11 pixels (8 ± 7.5 arc-min) to the left or right. ϕ took the values 1, 30, 60, 90, 120, and 150°. Note that this procedure forces the ideal detector to use the appropriate value of α in Equation 24, and for the largest sector this has a value of 1364 pixels.

For each value of ϕ the human observer's coherence thresholds were measured, and the corresponding theoretical limits were determined by simulations as described before. Simulations were also performed at other intermediate values of ϕ . The variation of coherence thresholds with ϕ for two observers, along with the corresponding theoretical limits, is shown using log-linear axes in Figure 10. Thresholds, both human and ideal, increase with increasing angle of jitter over the range shown. As ϕ is increased, the human thresholds approach the ideal threshold but do not get as close as was observed in the case of quantization in Figure 8. Measurements were also made for $\phi = 180^\circ$, but performance was at chance level.

Absolute efficiencies were determined as before, from the square of the ratio of ideal and human coherence thresholds, for the data shown in Figure 10. Figure 11 plots these efficiencies against ϕ for two observers using log-linear axes. The highest efficiencies found were 17% and 10% for H.B. and S.T., respectively, and were approximately half those observed for quantization in Figure 9.

Neighboring values for displacement and the amount of random jitter were explored, keeping ϕ fixed at 90°. The highest efficiency found for observer S.T. was 18% for a displacement of 24 ± 16 pixels (17 ± 11 arc-min) and for observer H.B. was 14% for a displacement of 16 ± 15 pixels (11 ± 10 arc-min). We have

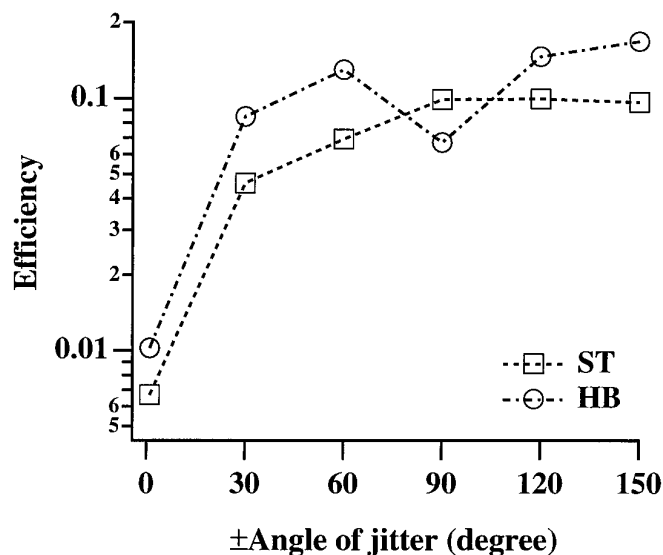


Figure 11. Effect of randomization on efficiency. The coherently moved dots were displaced to random positions within a sector of radius 8 ± 7.5 arc-min and \pm angle as shown on the *abscissa*. The corresponding statistical efficiencies of two observers (*ST*, *HB*) are shown on log-linear axes. As with quantization, randomization increases efficiency by matching the way kinematograms are generated to the coarse resolution of the motion-detecting system, so that the ideal observer cannot gain an advantage from its greater precision.

also done experiments in which the randomized dots were scattered over square regions instead of circular sectors; similar efficiencies were obtained for equal scatter areas.

The conclusion from the randomization experiments is that this procedure decreases the difference between ideal and actual performance by forcing the ideal detector to use less precision in counting the coherently moved dots. It thus supports the view that one cause of the low absolute efficiency of the neural motion system is that it cannot make use of the high precision with which dots are placed in normally generated kinematograms.

DISCUSSION

Our conclusions do not depend on any specific model of how MT works or how coherent motion is detected in the brain but are obtained by comparing ideal and actual performance. Knowledge of how the stimuli are generated enables the ideal performance to be calculated with certainty, and there are no assumptions about the mechanism involved in the measurements of actual performance.

The points that need discussion are (1) the relation of this work to previous work, (2) the ranges over which correspondence noise is an important factor, (3) the causes of lost efficiency, (4) comparisons with neurophysiological results, (5) the implications for the psychology of perception, and (6) the implications for the way the cortex performs its work.

Relation to previous psychophysical work

Many of the results of this paper confirm previous ones. Williams and Sekuler (1984) and Downing and Movshon (1989) reported the lack of influence of dot density on coherence threshold, but the effect of varying the density independently in two fields has not been explored previously. The effect of stimulus area (Baker and Braddick, 1982; van Doorn and Koenderinck, 1982b; Fredericksen et al., 1993; Eagle and Rogers, 1997) and the number of

successive fields (van Doorn and Koenderinck, 1982a; Frederickson et al., 1993, 1994; Festa and Welch, 1997) have been studied previously, and our results do not conflict in important respects with these. Watamaniuk (1993) considered a fine direction discrimination task from a signal/noise perspective but did not take the correspondence problem into account. The possible importance of correspondence noise has been recognized in many studies, and its importance in limiting D_{\max} has recently been clearly demonstrated both by Todd and Norman (1995) and Eagle and Rogers (1996); what we have done in this paper is to show how to calculate correspondence noise under various conditions, to extend the range of the experimental observations, and to compare them systematically with predictions from the theory that correspondence noise limits the detection of coherent motion in random dot kinematograms.

Ranges over which correspondence noise predictions hold

The hypothesis correctly predicts the changes in threshold with changes in parameters over the following ranges: (1) dot density from 1.7 to 111 dots/deg², (2) ratio of dot numbers in the two fields from 0.5 to 2, (3) area of the fields over a narrow range from 3 to 12 deg², and (4) number of consecutive fields from two to approximately eight.

The absolute performance compared with the ideal is obviously of crucial importance. Initially the estimated coherence thresholds for the ideal observer were much lower than the observed thresholds, indicating very low statistical efficiencies, but ideal coherence thresholds were much elevated when the kinematograms were generated in such a way that it was necessary for the ideal observer to pool motion information over broad ranges of direction and velocity, in the way we suspect the human system does. Thus changes in the following parameters also fit the correspondence noise hypothesis, with the supplementary hypothesis that the coherent motion system is coarsely tuned for direction and velocity: (1) the number of possible dot positions, and (2) the area over which coherently moved dots are randomly scattered.

Causes of lost efficiency

We think these results taken together provide good evidence for the importance of correspondence noise. It would be even better established if we could show that the statistical efficiency approaches 100%, because there would then be no room for other factors having an important influence under optimal conditions, and one would simply have to explain why performance declined under nonoptimal conditions. But the highest efficiencies we have consistently found are in the neighborhood of 30% for coarse quantization of dot positions and 15% for randomization. There are many detailed ways in which the generation of the kinematograms might be modified to match the detector mechanisms better. For instance, graded onsets and offsets of the fields might be better than the square wave stimuli we have used, and it is most unlikely that the rectangular grid for quantization and the sharply defined sectors used for the randomization experiments are optimally matched to the neural system. Higher efficiencies could probably be achieved by changes along these lines, but the figures are already high enough to establish that correspondence noise is important.

The fact that the efficiency is still rising as quantization becomes coarser in Figure 9 prompts the obvious suggestion that the observations be extended to coarser quantization, but this would

be difficult, because the dot positions were already very sparse, and only about 110 grid locations lay in the viewing area. Although the impressions of motion were still genuine under these conditions, we feared they would become become intellectual assessments rather than sensory judgments under even more extreme conditions.

With regard to area, the predicted relation does not hold beyond ~ 12 deg², or a field diameter of 4°. This presumably reflects the size of the receptive fields of neurons in MT (V5) in macaque (Raiguel et al., 1995); the average area within 15° eccentricity is between 10 and 20 deg², but there is much variability (Gattass and Gross, 1981; Maunsell and Van Essen, 1983b; Albright, 1984; Felleman and Kaas, 1984; Desimone and Ungerleider, 1986; Snowden et al., 1992), and there certainly are some much larger fields, especially among those that do not have inhibitory surrounds (Born and Tootell, 1992). There is a loss of efficiency for very small areas, presumably because the receptive fields of V5, being larger than the stimulus, then collect more noise than necessary. The effect of eccentricity has not been investigated, and some caution is also needed in interpreting these experiments because border effects become very important when the field is comparable in size with the displacement.

Predictions were formulated on the hypothesis that there is a mechanism for exploiting the continuous motion of dots in same-generated kinematograms, and that correspondence noise limits this mechanism. These were not fulfilled; our results confirm those of Scase et al. (1996) in showing surprisingly small differences in performance for “same” and “different” generated kinematograms. This implies that there is no mechanism that can take full advantage of the additional information available in the “same” kinematograms, at least for the range of conditions we have tested. But it must be appreciated that the predictions were made for an extreme form of detector that only counted the occasions when there was a dot in every expected position in the successive frames, and it is easy to imagine less extreme forms. Furthermore, the results of Watamaniuk et al. (1995) show that continuously moving objects can be successfully tracked, and Mikami (1992) found that 22% of cells in MT required more than one displacement to yield directionally selective responses. The problem then is to define the conditions when such tracking occurs and to formulate the properties of a detector that could account for such performance.

These experiments did show that coherence thresholds were usually lower with “same” than “different” generated kinematograms, but there is a simple possible explanation for this. Our predictions for “different” kinematograms do not take into account signal or noise resulting from dots lying close together in space but occurring in nonconsecutive fields. Considering that there is considerable temporal integration, there is likely to be substantial additional signal available from nonconsecutive frames in the “same” generated kinematograms but not in the “different” generated ones.

Thus the limits of predicted performance that we find suggest the following additional limiting factors: (1) the collecting fields do not match test stimuli effectively outside the range from 3 to 12 deg²; (2) motion information cannot be effectively summated beyond 0.5–1 sec; and (3) although coherence thresholds drop when the target persists for more than two fields, the full advantage theoretically available is not obtained.

Neurophysiological evidence for coarse tuning

There is ample neurophysiological evidence for the supplementary hypothesis that the direction and velocity tunings of the mechanism are very coarse. Single-unit recordings from MT (V5) indicate an average width of tuning of $\pm 45^\circ$ at half-height, although the range of widths is very large (Maunsell and Van Essen, 1983b; Albright, 1984; Felleman and Kaas, 1984; Rodman and Albright, 1987; Snowden et al., 1992). Velocity tuning is also broad in V5 (Maunsell and Van Essen, 1983b; Felleman and Kaas, 1984; Rodman and Albright, 1987), and the variations appear to be related to varying distances over which correlations are used rather than varying temporal intervals (Mikami et al., 1986; Newsome et al., 1986). The broad tuning of V5 neurons is likely to be advantageous in improving the signal/noise ratio for the detection of moving objects in natural images because of the spread of motion energy away from the true direction of motion. This coarse tuning appears to result partly from V5 neurons receiving inputs with a range of different preferred directions, because the V1 neurons that project to V5 are more narrowly tuned than most V5 neurons (Movshon and Newsome 1996).

Implications for psychology of motion perception

The reviews of Attneave (1974) and Dawson (1991) show that the psychology of motion perception and interpretation are not well understood, but neither of them consider the problem as a signal/noise discrimination. Some of the puzzling features may be the consequence of mechanisms that pool motion information to combat noise and clutter. The fact that some of the noise is external and enters the brain inextricably mixed with the signal makes a large difference to how we must interpret the organization of motion-detecting mechanisms, because it means that good performance can only be obtained by combining signals over large ranges of their parameters. The emphasis has often been on improving signal/noise ratios by combining signals of different neurons responding to the same spatio-temporal region, as in the approach of Zohary et al. (1994). This can only be effective when the noise in different neurons is independent; as Zohary et al. (1994) found, combining the signals from several neurons shows only limited improvement of signal/noise ratios when the noise is correlated, as it will be when the noise is external and enters with the sensory signals. This demonstrates why it is so important to know the extent to which correspondence noise limits the task of detecting coherent motion.

Implications for cortical organization

The connections to neurons in MT (V5) seem well designed to provide a sensitive, rapidly available, but rather coarse-grained map of the motions occurring all over the visual field, such as is needed for representing optic flow from self-motion. We think the fact that some of the noise is external gives new insight into the organization of the cerebral cortex for carrying out this work. When the limiting noise is external the engineering rule is to collect together as much as possible of the appropriate information by matching the collecting range to the ranges available in the signal and doing this for all the parameters of the stimulus. This is precisely what is done by the functional and anatomical arrangements for sorting and pooling motion information revealed by the anatomical connections (Zeki, 1975; Maunsell and Van Essen, 1983a). The parameters required to characterize a patch of motion in the visual field are position, size, direction, velocity, and depth. These are also the variable parameters of neurons in MT, so they evidently form a vast array of filters, each with a

different combination of these parameters, between them capable of providing a near optimal match for a huge range of motion stimuli at all positions in the visual field.

MT has an area of $\sim 30\text{--}40\text{ mm}^2$ in the macaque (Van Essen et al., 1981), and with $200,000\text{ neurons/mm}^2$ the total number of neurons is about 7×10^6 . The possible parameters of motion in small patches of the visual field define a multidimensional space, and the number of MT neurons required to sample this space adequately is very large. To illustrate this, suppose the receptive fields are centered 2 deg apart and cover a hemisphere uniformly, so more than 5000 would be needed. Then suppose their preferred directions are 30 deg apart, so 12 are required at each location, and that they have four different preferred velocities and come in four different sizes and three different disparities. With these numbers almost 3×10^6 neurons would be required, and because they do not all project to the same destination, we must allow for some reduplication. It is often assumed that the number of cortical neurons is vastly greater than the number required to sample the image adequately, but the above figures suggest that this is probably not the case in MT and that each neuron has a distinct job to perform.

If MT neurons form the suggested array of matched filters, almost all the information would be carried by the small number of neurons with parameters that best match those of any particular patch of movement. A motion field can be represented with a far smaller number of elements (e.g., the quadrature pairs or quadruples of Adelson and Bergen, 1985), but an array of matching filters is very efficient if the main problem is to pick out the coherent signal from the other disturbing signals that arrive with it.

If this interpretation of MT (V5) is correct, then one begins to see that the principles on which information is collected together may be equally important for the tasks performed in the primary visual cortex, in other extrastriate areas, in other sensory areas, and for that matter in many areas of the cortex that are not directly concerned with sensory information. When external noise is the limiting factor, collecting relevant information is the really important operation, and it would be well performed by a cortex with cells that constitute a vast array of filters, each filter matched to one of the myriad of possible combinations of features that we need both to detect and to discriminate from each other. Such a system would enable speedy and appropriate responses to be made as soon as crucial events occur in the cluttered and noisy world that surrounds us.

REFERENCES

- Adelson EA, Bergen JR (1985) Spatio-temporal energy models for the perception of motion. *J Opt Soc Am A* 2:284–299.
- Albright TD (1984) Direction and orientation selectivity of neurons in visual area MT of the macaque. *J Neurophysiol*, 52:1106–1130.
- Attneave F (1974) Apparent movement and the what–where connection. *Psychologia* 17:108–120.
- Baker CL, Braddick OJ (1982) The basis of area and dot number effects in random dot motion perception. *Vision Res*, 22:1253–1259.
- Barlow HB, Reeves HB (1979) The versatility and absolute efficiency of detecting mirror symmetry in random dot displays. *Vision Res* 19:783–793.
- Born RT, Tootell RBH (1992) Segregation of global and local motion processing in primate middle temporal visual area. *Nature* 357:497–499.
- Braddick O (1974) A short-range process in apparent motion. *Vision Res* 14:519–527.
- Britten KH, Shadlen MN, Newsome WT, Movshon JA (1992) The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J Neurosci* 12:4745–4765.

- Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA (1995) A relationship between behavioural choice and the visual responses of neurons in macaque MT. *Vis Neurosci* 13:87–100.
- Celebrini S, Newsome WT (1994) Neuronal and psychophysical sensitivity to motion signals in extrastriate area MST of the Macaque monkey. *J Neurosci* 14:4109–4124.
- Dawson MRW (1991) The how and why of what went where in apparent motion: modeling solutions to the motion correspondence problem. *Psychol Rev* 98:569–603.
- Desimone R, Ungerleider LG (1986) Multiple visual areas in the caudal superior sulcus of the macaque. *J Comp Neurol* 248:164–189.
- Downing C, Movshon JA (1989) Spatial and temporal summation in stochastic random dot displays. *Invest Ophthalmol Vis Sci [Suppl 30]*:72.
- Eagle RA, Rogers BJ (1996) Motion detection is limited by element density not spatial frequency. *Vision Res* 36:545–558.
- Eagle RA, Rogers BJ (1997) Effect of dot density, patch size and contrast on the upper spatial limit for direction discrimination in random-dot kinematograms. *Vision Res* 37:2091–2101.
- Felleman DJ, Kaas JH (1984) Receptive field properties of neurons in middle temporal visual area (MT) of owl monkeys. *J Neurophysiol* 52:488–513.
- Festa EK, Welch L (1997) Recruitment mechanisms in speed and fine-direction discrimination tasks. *Vision Res*, in press.
- Finney DJ (1947) Probit analysis. Cambridge: Cambridge UP.
- Fisher RA (1925) Statistical methods for research workers. Edinburgh: Oliver and Boyd.
- Fredericksen RE, Verstraten FAJ, van de Grind WA (1993) Spatio-temporal characteristics of human motion perception. *Vision Res* 33:1193–1205.
- Fredericksen RE, Verstraten FAJ, van de Grind WA (1994) An analysis of the temporal integration mechanism in human motion perception. *Vision Res* 34:3153–3170.
- Gattass R, Gross CG (1981) Visual topography of striate projection zone (MT) in posterior superior temporal sulcus of the macaque. *J Neurophysiol* 46:621–638.
- Laming D (1986) Sensory analysis. London: Academic.
- Maloney RK, Mitchison GJ, Barlow HB (1987) The limit to the detection of Glass patterns in the presence of noise. *J Opt Soc Am A* 4:2336–2341.
- Maunsell JHR, Van Essen DC (1983a) The connections of the middle temporal visual area (MT) and their relation to a cortical hierarchy in the macaque monkey. *J Neurosci* 3:2563–2586.
- Maunsell JHR, Van Essen DC (1983b) Functional properties of neurons in the middle temporal area of the macaque monkey. *J Neurophysiol* 49:1127–1147.
- Mikami A (1992) Spatiotemporal characteristics of direction-selective neurons in the middle temporal area of the macaque monkeys. *Exp Brain Res* 90:40–46.
- Mikami A, Newsome WT, Wurtz RH (1986) Motion selectivity in macaque visual cortex. II. Spatiotemporal range of directional interactions in MT and V1. *J Neurophysiol* 55:1328–1339.
- Morgan MJ, Ward R (1980) Conditions for motion flow in dynamic visual noise. *Vision Res* 20:431–435.
- Movshon JA, Newsome WT (1996) Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *J Neurosci* 16:7733–7741.
- Newsome WT, Mikami A, Wurtz RH (1986) Motion selectivity in macaque visual cortex. III. Psychophysics and physiology of apparent motion. *J Neurophysiol* 55:1340–1351.
- Newsome WT, Britten KH, Movshon JA (1989) Neuronal correlates of a perceptual decision. *Nature* 341:52–54.
- Newsome WT, Britten KH, Salzman CD, Movshon JA (1990) Neuronal mechanisms of motion perception. *Cold Spring Harb Symp Quant Biol* 55:697–705.
- Raiguel S, Van Hulle MM, Xiao D-K, Marcar VL, Orban GA (1995) Shape and spatial distribution of receptive fields and antagonistic motion surrounds in the middle temporal area (V5) of the macaque. *Eur J Neurosci* 7:2064–2082.
- Rodman HR, Albright TD (1987) Coding of visual stimulus velocity in area MT of the macaque. *Vision Res* 27:2035–2048.
- Scase MO, Braddick OJ, Raymond JE (1996) What is noise for the motion system? *Vision Res* 36:2579–2586.
- Snowden RJ, Treue S, Andersen RA (1992) The response of neurons in areas V1 and MT of the alert rhesus monkey to moving patterns of random dots. *Exp Brain Res* 88:389–400.
- Swets JA, ed (1964) Signal detection theory by human observers. New York: Wiley.
- Todd JT, Norman JF (1995) The effect of spatio-temporal integration on maximum displacement thresholds for the detection of coherent motion. *Vision Res* 35:2287–2303.
- van Doorn AJ, Koenderinck JJ (1982a) Temporal properties of the visual detectability of moving spatial white noise. *Exp Brain Res* 45:179–188.
- van Doorn AJ, Koenderinck JJ (1982b) Spatial properties of the visual detectability of moving spatial white noise. *Exp Brain Res* 45:189–195.
- Van Essen DC, Maunsell JHR, Bixby JL (1981) The middle temporal visual area in macaque: myeloarchitecture, connections, functional properties and topographic representation. *J Comp Neurol* 199:293–326.
- Watamaniuk SNJ (1993) Ideal observer for discrimination of the global direction of dynamic random-dot stimuli. *J Opt Soc Am A* 10:16–28.
- Watamaniuk SNJ, McKee SP, Gryzywacz N (1995) Detecting a trajectory embedded in random-direction motion noise. *Vision Res* 35:65–77.
- Williams DW, Sekuler R (1984) Coherent global motion percepts from stochastic local motions. *Vision Res* 24:55–62.
- Zeki SM (1975) The functional organization of projections from striate to pre-striate visual cortex in the rhesus monkey. *Cold Spring Harb Symp Quant Biol* 40:591–600.
- Zohary E, Shadlen MN, Newsome WT (1994) Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* 370:140–143.