

## THE COMBINATION OF MOTION SIGNALS OVER TIME

R. J. SNOWDEN\* and O. J. BRADDICK

Kenneth Craik Laboratory, Department of Experimental Psychology, University of Cambridge,  
Cambridge, England

(Received 21 November 1988; in revised form 13 February 1989)

**Abstract**—The improvement in performance with increasing number of frames in a random-dot kinematogram (temporal recruitment) was assessed by measuring threshold signal-to-noise ratios for direction discrimination. At fast frame presentation rates (50 Hz) thresholds fell sharply as the number of frames in the sequence increased, whereas at slow frame presentation rates (20 and 10 Hz) there was a less dramatic fall in thresholds. The similarity between the results at 20 and 10 Hz suggests that the mechanism of this less dramatic rise is relatively independent of temporal factors. The recruitment effect also does not appear to be limited by a maximum spatial range. We propose that temporal recruitment may occur via two mechanisms. One involves stimulating motion detectors with greater spans and delays, whilst the other involves the co-operative interaction of signals from units tuned to similar directions and having similar spans and delays. This distinction is supported by a further experiment which eliminates the first of these recruitment mechanisms by destroying possible correlations between non-adjacent frames.

Motion    Co-operatively    Temporal recruitment    Bilocal detectors

### INTRODUCTION

As we move around our environment the objects we see are generally in motion due to our own head, body and eye movements as well as their own motion through space. There have been several attempts recently to model a visual system's motion sensors (Reichardt, 1961; van Santen & Sperling, 1985; Watson & Ahumada, 1985; Adelson & Bergen, 1985). All these models are conceptually similar in that they sample the image at two points in space at slightly different times. A similar model was also suggested after exploration of directionally selective cells of the rabbit retina (Barlow & Levick, 1965). They suggest that cells achieve directional selectivity by passing inhibitory signals to units in the non-preferred direction. Such a claim has received recent support (Mikami, Newsome & Wurtz, 1986; Poggio & Koch, 1987). Whatever the nature of the process by which the two signals are compared this general class of sensor can be thought of in a simplistic form as "bilocal" detectors (Koenderink, Van Doorn & Van de Grind, 1985). Such a detector has a characteristic spatial span,  $ds$ , between the subunits that sample the image and a characteristic time delay,  $dt$ ,

which is the difference in time that impulses would arrive at the comparator unit if both subunits were stimulated simultaneously.

A sensor of this kind inevitably has ambiguities in its output. If a moving texture passes over the receptive field it is possible that one texture element could stimulate subunit *A* at time *T* and that a different texture element could stimulate subunit *B* at time *T* + *dt* causing the sensor to signal motion in its preferred direction when the motion was in fact in another direction. We therefore expect that when any complex image moves there will be a problem of "false" motion signals being present in the output of these sensors. This is analogous to the "correspondence problem" (Marr, 1983) of stroboscopic motion—which element of frame 1 matches which element of frame 2. Clearly if the true direction of motion is to be obtained there needs to some process that combines these motion signals over space and/or time. Such processes operating over a range of detectors are termed 'global' opposed to the 'local' process of the motion detector. This paper aims to explore the nature of the combination of motion signals over time.

There is already considerable evidence for the improvement of performance as the temporal extent of a motion sequence is increased. This is true for both continuous and stroboscopic motion. For instance Sperling (1976) reports that

\*Present address: Department of Brain and Cognitive Sciences, E25-236, Massachusetts Institute of Technology, MA 02139, U.S.A.

subjects rated a prolonged sequence of a dot moving stroboscopically to 'more closely approximate real motion' than a simple two frame version. Van Doorn and Koenderink (1983) report that signal-to-noise ratios for movement detection fall with increasing viewing duration. Lappin and Fuqua (1982) report a similar phenomenon with increasing number of frames of stroboscopic motion. McKee and Welch (1985) report that Weber fractions for velocity discrimination fall rapidly as the motion sequence is extended to 100 msec. Interestingly they report that this time period is independent of the number of frames appearing within it. Nakayama and Silverman (1984) also report that the maximum displacement limit (Braddick, 1974) for perceived motion in random dot kinematograms of 3 frames is greater than twice of a 2 frame kinematogram, suggesting a process which combines motion information over time. This process has been further investigated by Snowden and Braddick (1989). They report that motion information may be combined for up to 6 displacements. Interestingly they report that this figure of 6 displacements did not change even when the frame rate underwent a five fold change. In other words the combination was not limited by a temporal integration time. So there appears to be some discrepancy in the literature. McKee and Welch found that "sequential recruitment" of velocity information is dependent of a temporal interval (i.e. it was independent of the number of displacements). Snowden and Braddick found that "temporal recruitment" of direction information is independent of the temporal interval, and dependent upon the number of frames. Any theory of the combination of motion signals over time will have to explain this apparent discrepancy.

If temporal recruitment is not limited by an integration time, as claimed by Snowden and Braddick (1987), what are the limits of this process? One possibility is a limited area over which signals may be combined: a spatial limit. This could explain both the results of Snowden and Braddick and of McKee and Welch. McKee and Welch found a constant time limit for recruitment as the frame rate was varied. However, as actual velocity was held constant, the constant time limit was also a constant distance traversed. Snowden and Braddick found that recruitment was limited by a constant number of frames even though frame rate varied. However, in their experiment changing the frame rate also changed the velocity but not the dis-

tance traversed, as displacement per frame was held constant. In order to assess the possibility of a limited area of spatial summation we decided to take measurements of performance for displacements smaller than the maximum displacement limit ( $d$ -max) which was used in the previous study. If the hypothesis of a constant spatial area of summation is correct, we should expect to see that many more displacements may be summated as the actual size of that displacement is decreased.

If we use displacements much smaller than the maximum displacement limit what indicator of performance can be taken? We adopted the approach of Van Doorn and her co-workers (Van Doorn and Koenderink, 1982a, b, 1983; Van de Grind, Van Doorn & Koenderink, 1983, 1986) in using a signal-to-noise ratio paradigm. However, the way in which our equipment generated the dot patterns required a somewhat different way of setting the signal-to-noise ratio. We used sparser dot patterns than Van Doorn et al., and manipulated the detectability of coherent motion by varying the percentage of dots that underwent that motion. All dots that did not undergo this motion were randomly redistributed on the screen. We define those that were coherently displaced as the "signal", and those that were randomly redistributed as the "noise". Thus the signal-to-noise ratio is the percentage signal divided by the percentage noise. For example if 80 dots were signal whilst 320 were noise the signal-to-noise ratio would be 0.25. Using this measure as an indicator of performance we measured signal-to-noise ratios required to discriminate motion to the left from motion to the right for sequences of random dot apparent motion in which the displacement size, frame rate and number of frames per sequence could all be manipulated.

## METHODS

### *Stimuli*

Random-dot kinematogram frames were produced by software running on a PDP 11/10 minicomputer which generated random X- and Y-coordinates. The dots were plotted on a HP 1319 c.r.t. display (P31 phosphor) controlled by a Sigma QVEC display processor. 400 dots (dia. 0.5 mm) were displayed in an area 19.5 cm square. When viewed from 3 m this subtended a 3.7 deg arc. Dot density was therefore 29 dot/deg<sup>2</sup>, making dot centre separation an average of 11 min arc. The role of dot density

upon  $d$ -max and  $d$ -min has been previously assessed (Baker & Braddick, 1982, 1985a). They show that large changes in dot densities have remarkably little effect upon these thresholds. We therefore suspect that the results reported here would also remain constant for quite large changes in dot density.

To produce coherent displacement of the dots the X-coordinate of all dots was increased or decreased by a set amount. Any dot that would now be plotted outside the original 19.5 cm square was "wrapped around" to the opposite side of the display.\* The shifted pattern was then displayed. This process could be repeated for up to 10 frames with the memory available. The effect was of uniform displacement within a window. The screen was refreshed at a rate of 100 Hz, hence stimulus onset asynchronies (the time between the onset of successive frames) are always multiples of 10 msec.

The above technique produces patterns of infinite signal-to-noise ratio. For smaller signal-to-noise ratios, a subset of the dots were chosen and shifted whilst the rest were given new random X and Y coordinates. The selection of these shifted dots could be governed by either of two rules. Firstly, as used in expt 1, the dots that were chosen (randomly) for the first displacement were then also shifted for all subsequent displacements, this is the "same" rule. Alternatively, as used in expt 2, the dots to be shifted were only picked from those that had been given new coordinates on the previous frame, this is known as the "different" condition. Note that if this technique was employed all signal-to-noise ratios must remain below 1.0 (i.e. less than 50% of the dots coherently displaced).

#### *Procedure*

In each trial a fixation point was displaced for 0.5 sec, followed by the kinematogram whose direction of displacement (left or right) was random from trial to trial. The subject reported as to the perceived direction by pressing one of two buttons. No feedback was provided. Subjects viewed the display in a dimly lit room.

\*It is worth noting that these dots which were wrapped around to the opposite side of the display could not contribute to the "signal", hence true signal to noise ratios are somewhat smaller than the results presented here. If all dots were displaced, on each displacement around 14 dots would be wrapped around. At lower signal to noise ratios fewer dots would be wrapped around (as fewer are displaced beyond the edge of the display), such that at a signal to noise ratio of 0.1 an average of 1.3 dots would be wrapped around.

In all experiments a staircase procedure was employed. Two successive correct responses for a particular staircase caused the value of the next signal-to-noise ratio of that staircase to be decremented. A single incorrect response produced the corresponding increment; this procedure tracks the 71% correct level. Each staircase consisted of 11 reversals (here defined as two or more successive correct responses followed by an incorrect response). Up to 6 staircases were run in one experimental block, with the selection of the staircase to be presented chosen randomly from trial to trial. All staircases began with a signal-to-noise ratio of 1.0. In order to reach the region of threshold performance quickly the increment/decrement value was 0.2 log units before any reversal occurred. This dropped to 0.1 log units between reversal 1 and reversal 2, and to 0.05 log units for all subsequent trials. Threshold was taken as the mean of all signal-to-noise ratios after the third reversal.

Our subjects had normal or corrected to normal vision. Both were experienced psychophysical observers. I.H. was not informed as to the exact purpose of the experiments.

#### EXPERIMENT 1. THE EFFECT OF MULTIPLE-DISPLACEMENTS ON SIGNAL-TO-NOISE RATIOS

In this first experiment we were interested in the number of successive displacements over which motion information could be recruited. We measured the signal-to-noise ratio required for the direction of a motion sequence to be reliably discriminated. All displacements were 7 min arc and sequences could vary in the number of frames presented. If recruitment is limited by a spatial area of integration we should expect two results:

(1) recruitment effects should be observed over a similar number of displacements irrespective of stimulus onset asynchrony. Snowden and Braddick (1989) show this to be true over a wide range of stimulus onset asynchronies when the threshold  $d$ -max was measured as a function of the number of displacements;

(2) when the displacement is small (i.e. 7 min arc) recruitment effects should be observed over a greater number of displacements than when the displacement is large (i.e. near  $d$ -max). Snowden and Braddick (1989) found that recruitment took place over 5–6 displacements each of about 45 min arc. If this is due to a spatial limit it implies a region of integration

about 230 min arc across (3.8 deg arc), which is similar to the dimensions of the display. Within this region of integration over 30 displacements of 7 min arc could be recruited. This, however, may be an unrealistic prediction as it would be most surprising if there was no temporal limit whatsoever upon the recruitment effect. Therefore, to test for any possible restriction of recruitment by temporal limits, the experiment was performed with three values of stimulus onset asynchrony: 20, 50 and 100 msec. It is important to note that in this experiment the "signal" dots were selected according to the same rule.

### Results

The results of this experiment are displayed for two subjects in Fig. 1. The data are plotted with decreasing signal-to-noise ratio increasing

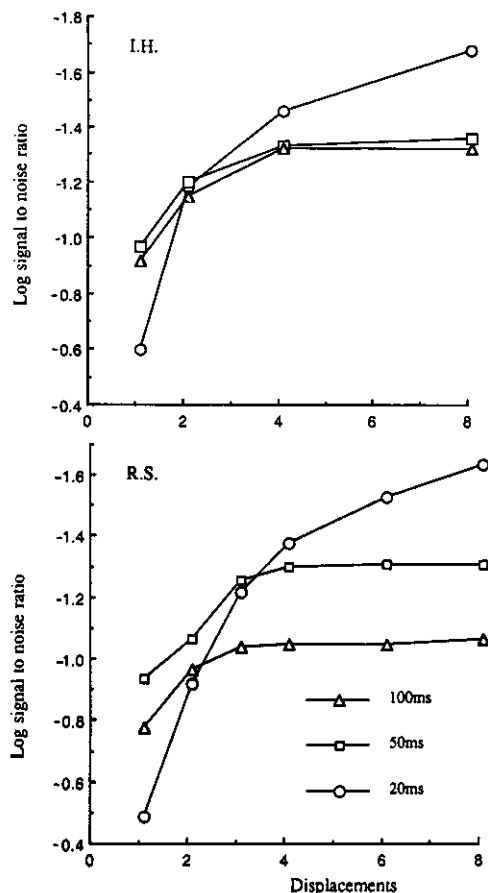


Fig. 1. The results of experiment 1 for two subjects. The logarithm of the threshold signal-to-noise ratio is plotted against the number of displacements in the sequence. Note the steep rise in the graph (fall in signal-to-noise ratio) for a stimulus onset asynchrony of 20 msec, but the less dramatic rise of stimulus onset asynchronies of 50 and 100 msec.

up the ordinate (i.e. performance is better for points at the top of a graph). There are 5 notable points about this data. (1) At all stimulus onset asynchronies, threshold signal-to-noise ratios decrease as the number of displacements increases, i.e. conspicuous temporal recruitment effects are present. (2) For the medium and large stimulus onset asynchronies signal-to-noise ratios asymptote after approx. 5–6 displacements. The shape of this curve is very similar to the data pertaining to the maximum displacement limit and number of displacements (Snowden & Braddick, 1989). (3) The data for a small stimulus onset asynchrony (20 msec) follow a much steeper function which starts at a higher signal-to-noise ratio and shows no sign if an asymptote even when there are 8 displacements per sequence. (4) The required signal-to-noise ratio for the 20 msec stimulus onset asynchrony after 8 displacements is only 0.02, this corresponds to only 8 dots being coherently displaced whilst the other 392 are randomly repositioned. Clearly the discrimination of direction of motion can be supported by very sparse and noisy information. (5) Though the data for stimulus onset asynchronies of 50 and 100 msec are of the same shape for both subjects, R.S. has better performance for a stimulus onset asynchrony of 50 msec than for one of 100 msec, whereas there is little difference for I.H. This is somewhat puzzling, but probably reflects individual differences. Such differences have been noted before in studies of lowest thresholds of motion (Baker & Braddick, 1985a) and are apparent (though small) in studies of the upper threshold of motion (see Baker & Braddick, 1985b, Fig. 4).

### Discussion

From these results we conclude firstly that the "temporal recruitment" reported by Nakayama and Silverman (1984) who investigated maximum displacement thresholds, also appears to operate for displacements less than the upper limit ( $d_{\max}$ ). Secondly, the data reported here are very similar to those of Lappin and Fuqua (1982) who find a rapid increase in detectability as the number of frames is increased for high frame rates, and a more moderate increase for low frame rates. Our data extend these results to moderate frame rates and therefore allow a more rigorous testing of several hypotheses.

Can the hypothesis of a limited spatial area of integration explain the present data? First, the prediction that the temporal recruitment effect

is independent of temporal factors does not entirely hold. A stimulus onset asynchrony of 20 msec gives a function of different shape from that with 50 or 100 msec stimulus onset asynchrony. Note that the curves for stimulus onset asynchrony of 50 and 100 msec both asymptote after 5–6 displacements, as was found for all stimulus onset asynchrony at much greater displacements (Snowden & Braddick, 1989). We suggested above that there might also be effects due to temporal constraints. However, temporal constraints can not explain our data. If, for example, there were a upper temporal limit on recruitment of 200 msec this could explain why performance for 20 msec stimulus onset asynchrony improved over many more displacements than the 50 or 100 msec stimulus onset asynchrony. However, the similarity between the data for the larger stimulus onset asynchrony argues against such a notion. There is no single temporal constraint which could explain the current pattern of results. Secondly, the data for the longer stimulus onset asynchrony reach an asymptote for a total path length of about 35 min arc (5 by 7 min arc displacements), much smaller than the apparent spatial range of integration in our earlier experiments measuring *d*-max. The 20 msec stimulus onset asynchrony yields integration over a longer range, but we have seen that the variation with stimulus onset asynchrony is not readily explained by a temporal limit. We conclude that the results depicted in Fig. 1 are not readily explained by any combination of limits on integration in time and space.

The hypothesis of a limited spatial integration region was proposed to explain the apparent discrepancy between data on the maximum displacement threshold and on velocity discrimination. Is there an alternative explanation? Lappin and Fuqua (1982) suggest that for frames of long duration (120 msec) displacements add independently to detectability, as in a linear system, whereas for short frame durations (12 msec) the rapid increase on detectability reflect the operation of a non-linear co-operative process involving recruitment among successive displacements. We would like to dispute this interruption and suggest one of our own.

Co-operative processes have already been demonstrated in human motion perception, such as ambiguous motion (Chang & Julesz, 1984) and hysteresis (Williams, Phillips & Sekuler, 1986). We suggest that the recruitment

effects demonstrated here, and in previous papers (Nakayama & Silverman, 1984; Snowden & Braddick, 1989) are manifestations of such a process. The co-operative process consists of inhibitory and excitatory interconnections between motion detectors such that those within a local area that are tuned to the same direction of motion send facilitatory signals to one another, whereas if they are tuned to different directions of motion they are mutually inhibitory. The resulting competition between directions of motion when a textured pattern is moved should result in a single correct motion being signalled and the "false" signals being suppressed (though there are exceptions in this simple case; for example two random dot sheets can be seen to move simultaneously through one another; Clark, 1977; Snowden, 1989). By this process the "correspondence problem" may be solved (Marr & Poggio, 1976). The phenomenon of temporal recruitment may also be explained by co-operative processes. When a displacement of random texture occurs local motion measurements will take place and a co-operative/competitive struggle will ensue. This process will leave the network with a bias for this particular direction of motion which will last for a short (as yet unspecified) time. If a second displacement takes place in the same direction within this time it will be affected by this bias and may be detected at a lower signal-to-noise ratio than if the previous displacement had not taken place; this is temporal recruitment. At this point the system will have an even greater bias, hence if a third displacement takes place it will require an even lower signal-to-noise ratio to be detected. Note that this recruitment process is limited by the ability of the system to extract some motion information from the first displacement so as to bias the system to some degree.

#### *Two forms of recruitment*

How does this co-operative system explain the sharper rise in performance for the 20 msec stimulus onset asynchrony over the 50 or 100 msec stimulus onset asynchrony? The simple answer is that by itself it does not. It does explain the similarity of the 50 and 100 msec data, and of the invariance of stimulus onset asynchrony found for the maximum displacement limit in our previous study (Snowden & Braddick, 1989). To explain the 20 msec data we have to postulate a second type of recruitment based upon the known properties of the short-

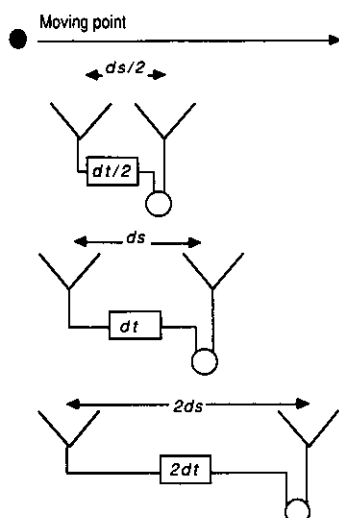


Fig. 2. A moving dot is portrayed moving over a range of possible bilocal detectors. Note that whilst each of these detectors has a different span and delay, they are tuned to the same velocity. If the dot is of infinite duration it will excite all detectors. However, if it is sufficiently truncated in time it will not excite the largest of these units. In other words as the duration of movement increases units of greater span may be excited; this is hetero recruitment.

range process (Braddick, 1980). Imagine a dot moving at velocity  $V$  for time  $T$ . It will excite units tuned to that velocity only if their span is less than  $VT$ , or put another way if their time delay (see Fig. 2) is less than  $T$ . We should therefore expect that as  $T$  is increased it will be able to stimulate (recruit) units of greater and greater span and time delay (see Fig. 2). That the human motion system detects a single velocity with units that have varying spans (and therefore time delays) has been well demonstrated by Van Doorn and Koenderink (1983). Therefore in order to stimulate all this range of detectors the stimulus must be of sufficient time duration. Braddick (1973) and Morgan and Ward (1980) have demonstrated that the upper time limit over which correlations may be made is no more than 80 msec. We therefore predict that there will be improvement due to "recruitment of units with greater spans" over this time period. The same argument may be applied to the apparent motion of our displays if, and only if, it can be demonstrated that correlations can be made between temporally non-adjacent frames e.g. between frames 1 and 3. Such correlations have been unequivocally demonstrated by Van Doorn and Koenderink (1982a). So it is now possible that in 3 frames of 20 msec that, not only are there correlations made by detec-

tors tuned to 7 min arc, but also by detectors tuned to 14 min arc. Hence the steep rise of the 20 msec condition. Why is there not such a steep rise for the 50 and 100 msec conditions? This is accounted for by the fact that temporally non-adjacent frames are now sufficiently separated in time to be beyond the upper temporal limit of the short-range process (i.e. the largest delay of any relevant detector) and therefore motion signals only arise from detectors tuned to 7 min arc. Hence we claim that there are two types of temporal recruitment: one due to the excitation of units with greater spans and time delays (which I shall term "hetero recruitment"), and one due to the bias in a co-operative system caused by previous stimulation (which I shall term "homo recruitment"). Of course for most "real" motion both of these recruitments will be in operation.

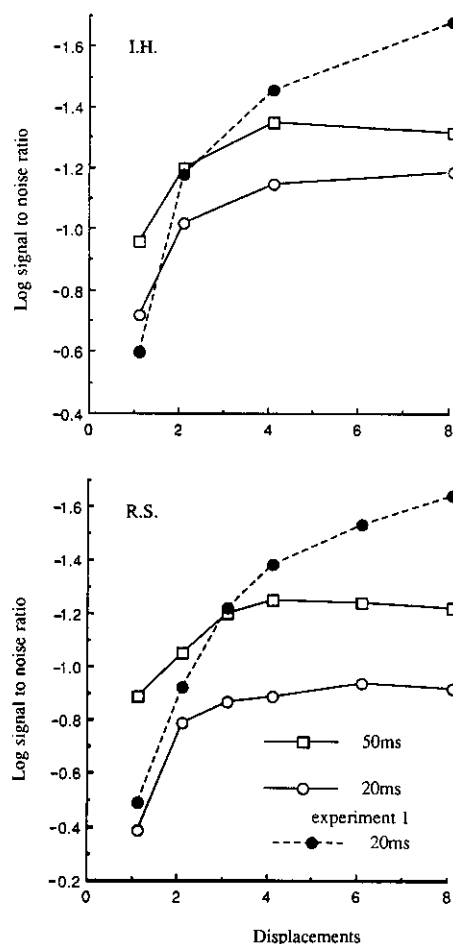


Fig. 3. The results of experiment 2 for two subjects, the results are plotted as in Fig. 1. Note now that the curves for the stimulus onset asynchronies of 20 and 50 msec are of similar shape. In addition the results for a 20 msec stimulus onset asynchrony from experiment 1 are reproduced for comparison.

In order to test this hypothesis we designed a stimulus in which all correlations between temporally non-adjacent frames were destroyed (see Methods). We predict that under these conditions the results from the 20 msec stimulus onset asynchrony will resemble in shape those of the 50 and 100 msec stimulus onset asynchrony of the previous experiment.

## EXPERIMENT 2. THE ROLE OF DOT PATH IN RECRUITMENT EFFECTS

In this experiment we repeated the conditions of expt 1, except that the "signal" dots on any displacement were always picked from those that were "noise" dots on the previous displacement (the "different rule"); hence no intentional correlations existed between frames 1 and 3, 2 and 4 etc. In doing this we hoped to isolate the homo recruitment by eliminating the possibility of hetero recruitment. From previous experimentation (Snowden & Braddick, 1987; 1989) we know that homo recruitment is unaffected by the path of individual dots and that this recruitment appears to generalize to other dots undergoing a similar motion.

The results for stimulus onset asynchronies of 20 and 50 msec are displayed in Fig. 3. As predicted the curve for a stimulus onset asynchrony of 20 msec is now the same shape (though shifted down the ordinate) as the 50 msec condition in both this and the previous experiment. Threshold signal-to-noise ratios decrease with increasing length of sequence up to 5–6 displacements. The data for a stimulus onset asynchrony of 20 msec from expt 1 is also reproduced to aid comparison between the two experiments.

## GENERAL DISCUSSION

Let us now consider some previous research in the light of these speculations of two types of recruitment. The data of Lappin and Fuqua (1982) are very similar to those of expt 1, and therefore our explanation of the present data can be applied to their's.

McKee and Welch (1985), report that velocity discrimination depends upon time duration rather than number of displacements. Their critical time is around 80–100 msec, and we note the similarity of this figure to the upper temporal limit of the short-range process (Braddick, 1980; Morgan & Ward, 1980). Note this fixed temporal limit (which in their experiment is a

variable number of displacements) is in contrast to the variable time but fixed number of displacements which seem to limit the extent of homo recruitment. If a temporal delay occurred between displacements their recruitment effect was greatly diminished (see their Fig. 8) unlike the findings of our experiments and those of Nakayama and Silverman (1984). The idea that two different processes are at work is further supported by McKee and Welch's finding that improvement in velocity discrimination is limited to the specific motion path of the line, unlike the process which continued to operate in the "different" condition in our experiments. We therefore suggest that the recruitment that McKee and Welch observe is governed by the hetero process, that of stimulating units with larger spans and time delays as the duration of the stimulus increases. Note that we predict that homo recruitment should also occur in their paradigm. Its effect, however, may be disguised by a ceiling effect in some of their data. We note that when they used a stimulus onset asynchrony of 100 msec there was indeed some improvement with increasing number of displacements. Large stimulus onset asynchrony will eliminate the effects of hetero recruitment, hence this improvement might well be attributable to homo recruitment.

More recently de Bruyn and Orban (1988) have confirmed many of the major findings of McKee and Welch using random dot patterns which eliminate any use of "long-range" processes (Braddick, 1980) to aid velocity discrimination. Further, the use of random dot patterns allowed them to measure Weber fractions for direction discrimination. This threshold was also found to decrease with the increasing duration of the stimulus.

The work of Van Doorn et al. has been previously mentioned and we would like to consider some of their data in the light of the present discussion. However, it should be noted that there are some important differences between their studies and the present data. For instance, they employed patterns in which each pixel was either black or white, and defined threshold by means of varying the contrast of a noise pattern relative to a superimposed signal pattern. Van Doorn and Koenderink (1983) report that threshold signal-to-noise ratios for detecting motion fall with increasing stimulus duration (see their Fig. 10). They also demonstrate that there is "spatial recruitment", in that limiting the viewing aperture also has a detri-

mental effect upon performance (though the possible role of a greater intrusion of the peripheral retina must also be considered). They too interpret their results in terms of a large spectrum of detectors for a given velocity, in which limiting time or space prevents the largest of these units from being activated. Whilst our results partially support their interpretation, this interpretation does not take into account any distinction between hetero and homo recruitment. Van Doorn et al.'s account would imply that we are recruiting (in the hetero sense) units with delays of 1000 msec, a value which does not correspond well with electrophysiological data (e.g. Mikami et al., 1986), and is at least ten times the upper temporal limit of the short-range process! If we allow for homo recruitment these units with very large delays are not required. It is of interest to calculate over what sort of time period we predict recruitment to occur. From the data of Nakayama and Silverman (1984, Fig. 7) it appears that the recruitment effect between two displacements may exist for up to 300 msec. From the present data and that of our previous study (Snowden & Braddick, 1989) recruitment may occur for up to 6 displacements. Hence we predict that recruitment may occur for up to around 1800 msec. This data of Van Doorn and Koenderink can therefore be explained without having to postulate units with time delays greater than the postulated figure of the short-range process.

These two forms of recruitment can also be considered in analyzing the results of Van de Grind et al. (1983). This study systematically measured the maximum velocity for motion detection for various field sizes and locations on the retina. Performance was invariant throughout the temporal visual field, provided all stimuli were scaled according to the cortical magnification factor to obtain equivalent cortical sizes and velocities at all eccentricities. The upper velocity limit increased linearly with the width of the square stimulus. Although this stimulus was of unlimited duration maximum velocity was always such that any pixel crossed the field in 5–9 steps (depending upon subject), or a crossing time of 50–90 msec as the frame rate was 100 Hz. We can assume hetero recruitment does not play an important role as near maximum velocity single steps will be near the maximum displacement limit and double steps well beyond it. So it appears that maximum velocity occurs when the stimulus is configured

in such a way that 5–9 steps may be recruited. This looks very similar to the apparent motion stimulus of our experiments (Snowden & Braddick, 1987) whose own maximum velocity (maximum displacement/stimulus onset asynchrony) also peaked when a point travelled 5 steps (approximately 220 min arc (field size) divided by 45 min arc (asymptotic maximum displacement)). It is worth noting that the very high frame rate used in Van de Grind et al.'s experiment means that their display was equivalent for human vision to real continuous motion (Watson & Ahumada, 1983; Burr, Ross & Morrone, 1986). Thus the agreement with our own study gives us confidence that the processes we propose are not a special or artefactual result of discontinuous or stroboscopic motion.

#### *Estimate of span and delay of bilocal units*

The present data also help solve discrepancies in the literature over the spans and delays of the proposed bilocal units underlying the perception of a particular velocity. Much work has been performed by Koenderink and co-workers (Van Doorn & Koenderink, 1982a, b; Koenderink et al., 1985; Van de Grind et al., 1986). However, whilst the overall pattern of results is consistent between studies (larger velocities being subserved by greater spans and smaller delays) the actual estimation of those spans varies widely between studies. For instance Van Doorn and Koenderink (1982b) propose the formula:

$$S = 4.2 V^{0.6},$$

( $S$  in min arc,  $V$  in deg arc/sec) to estimate the detector span from the velocity. Hence a velocity of 10 deg/sec would be subserved by units with a mean span of 17 min arc. However, inspection of Fig. 3 of Van de Grind et al. (1986) estimates spans for 10 deg arc/sec from around 20–80 min arc. Why is there such a discrepancy? We suggest that the effects of recruitment have been overlooked in the second study hence these figures overestimate the true detector values.

To obtain these estimations Van de Grind et al. (1986) reduced the size of a strip of coherently moving dots, which was surrounded by incoherent movement, until this perceived coherency disappeared. They argue that coherency disappears because the incoherent movement is now encroaching onto the outer spatial limits of the detector. In other words the span (here meaning the maximum distance from the outer



edges of the two receptive fields) of the detector subserving this motion is on the order of the width of this strip. We have previously shown that this cannot be true. Snowden and Braddick (1989) demonstrate that asymptotic  $d$ -max occurs after around 5 displacements and that this asymptotic  $d$ -max is around 1/5 of the display area. Most importantly they demonstrate that the individual dot paths are unimportant to this. This is one again demonstrated by comparing the data for a stimulus onset asynchrony of 50 msec in expts 1 and 2. The results are essentially identical, showing individual dot paths are unimportant to homo recruitment (Williams & Sekuler, 1984, demonstrate a similar phenomena).

Our finding that maximal total displacement coincides approximately with the display width is not due to individual dots transversing the whole width, but to the display providing the available region for "homo" recruitment, plus the known variation of  $d$ -max with eccentricity (Baker & Braddick, 1985a). With this in mind it is clear that the units detecting each individual displacement are not on the order of the same size as the size of the display. They must be at least 5 times smaller than the display size. We therefore suggest that the study of Van de Grind et al. (1986) gives us an indication of the spatio-temporal recruitment area rather than the spans and delays of individual detectors. Any quantitative studies of maximum displacement or velocity must take into account both these effects as well as the hetero recruitment of detectors with different spans and delays.

**Acknowledgements**—I thank Ian Holliday for acting as a subject during these experiments. The research was supported by a S.E.R.C. grant to R.J.S. and a M.R.C. grant to O.J.B. A brief preliminary report of these experiments was presented to the 11th European Conference on Visual Perception, Bristol, U.K. September 1988.

## REFERENCES

- Adelson, E. H. & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America*, *A2*, 281–299.
- Baker, C. L. Jr & Braddick, O. J. (1982). The basis of area and dot number effects in random dot motion perception. *Vision Research*, *22*, 1253–1260.
- Baker, C. L. Jr & Braddick, O. J. (1985a). Eccentricity-dependent scaling of the limits for short range apparent motion. *Vision Research*, *25*, 803–812.
- Baker, C. L. Jr & Braddick, O. J. (1985b). Temporal properties of the short-range process in apparent motion. *Perception*, *14*, 181–192.
- Barlow, H. B. & Levick, W. R. (1965). The mechanism of directionally selective units in rabbit's retina. *Journal of Physiology, London*, *178*, 477–504.
- Braddick, O. J. (1973). The masking of apparent motion in random dot patterns. *Vision Research*, *13*, 355–369.
- Braddick, O. J. (1974). A short-range process in apparent motion. *Vision Research*, *14*, 519–527.
- Braddick, O. J. (1980). Low-level and high-level processes in apparent motion. *Philosophical Transactions of the Royal Society, London*, *B290*, 137–151.
- Burr, D. C., Morrone, C. & Maffei, L. (1981). Intra-cortical inhibition prevents simple cells from responding to textured visual patterns. *Experimental Brain Research*, *43*, 455–458.
- Burr, D. C., Ross, J. & Morrone, M. C. (1986). Smooth and sampled motion. *Vision Research*, *26*, 643–652.
- Chang, J. J. & Julesz, B. (1984). Cooperative phenomena in apparent motion perception of random-dot cinematograms. *Vision Research*, *24*, 1781–1788.
- Clark, P. G. H. (1977). Subjective standstill caused by interaction of moving patterns. *Vision Research*, *17*, 1243.
- De Bruyn, B. & Orban, G. A. (1988). Human velocity and direction discrimination measured random dot patterns. *Vision Research*, *28*, 1323–1336.
- Koenderink, J. J., Van Doorn, A. J. & Van de Grind, W. A. (1985). Spatial and temporal parameters of motion detection in the peripheral visual field. *Journal of the Optical Society of America*, *A2*, 252–259.
- Lappin, J. S. & Fuqua, M. (1982). Nonlinear recruitment in the detection of moving patterns. *Arvo Abstracts*, 123.
- Marr, D. (1983). *Vision*. San Francisco, CA: Freeman.
- Marr, D. & Poggio, T. (1976). Cooperative computation of stereo disparity. *Science, New York*, *194*, 283–287.
- McKee, S. P. & Welch, L. (1985). Sequential recruitment in the discrimination of velocity. *Journal of the Optical Society of America*, *A2*, 243–251.
- Mikami, A., Newsome, W. T. & Wurtz, R. H. (1986). Motion selectivity in Macaque visual cortex. I. Mechanisms of direction and speed selectivity in extrastriate area MT. *Journal of Neurophysiology*, *55*, 1308–1327.
- Morgan, M. J. & Ward, R. (1980). Conditions for motion flow in dynamic visual noise. *Vision Research*, *20*, 431–435.
- Nakayama, K. & Silverman, G. H. (1984). Temporal and spatial characteristics of the upper displacement limit for motion in random dots. *Vision Research*, *24*, 293–299.
- Poggio, T. & Koch, C. (1987). Synapses that compute motion. *Scientific American*, *256*, 42–48.
- Reichardt, W. (1961). Autocorrelation; a principle of the evaluation of sensory information by the central nervous system. In Rosenblith, W. A. (Ed.) *Sensory communication*. New York: Wiley.
- Snowden, R. J. (1989). Motions in orthogonal directions are mutually suppressive. *Journal of the Optical Society of America*, *A*, *6*, 1096–1101.
- Snowden, R. J. & Braddick, O. J. (1987). Extension of displacement limits in multiple-exposure sequences of apparent motion. *Perception*, *16*, A39.
- Snowden, R. J. & Braddick, O. J. (1989). Extension of displacement limits in multiple-exposure sequences of apparent motion. *Vision Research*, *29*, (in press).
- Sperling, G. (1976). Movement perception in computer-driven displays. *Behaviour Research and Instrumentation*, *8*, 144–151.
- Van Doorn, A. J. & Koenderink, J. J. (1982a) The temporal properties of the visual detectability of moving spatial white noise. *Experimental Brain Research*, *45*, 179–188.

- Van Doorn, A. J. & Koenderink, J. J. (1982b). The spatial properties of the visual detectability of moving spatial white noise. *Experimental Brain Research*, *45*, 189–195.
- Van Doorn, A. J. & Koenderink, J. J. (1983). Spatiotemporal integration in the detectability of motion. *Vision Research*, *23*, 47–56.
- Van de Grind, W. A., Van Doorn, A. J. & Koenderink, J. J. (1983). Detection of coherent movement in peripherally viewed random-dot patterns. *Journal of the Optical Society of America*, *73*, 1674–1683.
- Van de Grind, W. A., Koenderink, J. J. & Van Doorn, A. J. (1986). The distribution of human motion detector properties in the monocular visual field. *Vision Research*, *26*, 797–810.
- Van Santen, J. P. H. & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America*, *A2*, 300–321.
- Watson, A. B. & Ahumada, A. J. (1983). A look at motion in the frequency domain. NASA Technical Memorandum, TM-84352.
- Watson, A. B. & Ahumada, A. J. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America*, *A2*, 322–341.
- Williams, D. W. & Sekuler, R. (1984). Coherent global motion percepts from stochastic local motions. *Vision Research*, *24*, 55–62.
- Williams, D. W., Phillips, G. & Sekuler, R. (1986). Hysteresis in the perception of motion direction as evidence of neural cooperativity. *Nature, London*, *324*, 253–255.