Parallel search for conjunctions with stimuli in apparent motion

Clara Casco
Department of General Psychology, University of Padova, via Venezia 8, 35131 Padova, Italy;
e-mail: casco@psico.unipd.it
Giorgio Ganis
Department of Psychology, Harvard University, 33 Krikland Street, Cambridge, MA 02138, USA
Received 21 May 1996, in revised form 17 July 1998

Abstract. A series of experiments was conducted to determine whether apparent motion tends to follow the similarity rule (ie is attribute-specific) and to investigate the underlying mechanism. Stimulus duration thresholds were measured during a two-alternative forced-choice task in which observers detected either the location or the motion direction of target groups defined by the conjunction of size and orientation. Target element positions were randomly chosen within a nominally defined rectangular subregion of the display (target region). The target region was presented either statically (followed by a 250 ms duration mask) or dynamically, displaced by a small distance (18 min of arc) from frame to frame. In the motion display, the position of both target and background elements was changed randomly from frame to frame within the respective areas to abolish spatial correspondence over time. Stimulus duration thresholds were lower in the motion than in the static task, indicating that target detection in the dynamic condition does not rely on the explicit identification of target elements in each static frame. Increasing the distractor-to-target ratio was found to reduce detectability in the static, but not in the motion task. This indicates that the perceptual segregation of the target is effortless and parallel with motion but not with static displays. The pattern of results holds regardless of the task or search paradigm employed. The detectability in the motion condition can be improved by increasing the number of frames and/or by reducing the width of the target area. Furthermore, parallel search in the dynamic condition can be conducted with both short-range and long-range motion stimuli. Finally, apparent motion of conjunctions is insufficient on its own to support location decision and is disrupted by random visual noise. Overall, these findings show that (i) the mechanism underlying apparent motion is attribute-specific; (ii) the motion system mediates temporal integration of feature conjunctions before they are identified by the static system; and (iii) target detectability in these stimuli relies upon a nonattentive, cooperative, directionally selective motion mechanism that responds to high-level attributes (conjunction of size and orientation).

1 Introduction

Apparent motion is usually experienced when an object is displaced in discrete steps in successive presentations with appropriate temporal and spatial parameters (Anstis 1970; Kolers 1972; Korte 1915; Ullman 1979; Wertheimer 1912).

Apparent motion is perceived with a variety of moving images. The most common type of object motion is produced by moving objects that are lighter or darker than their background. In this case, the retinal image contains luminance discontinuities, that is luminance steps at the edge of such objects, that change position over time. With these moving stimuli, apparent motion depends on the activation of either a low-level luminance correlator that computes the spatiotemporal correspondence of local luminance discontinuities or a high-level motion system for which the extraction of shape information in each static frame precedes motion computation (Braddick 1980).

Another kind of object motion is produced by moving objects or subregions which differ from their background only in colour, texture, or flicker rate. Recent work has shown that form cues such as chromatic and textural discontinuities are in no way inferior to luminance discontinuities in their potential for establishing spatiotemporal correspondence. Indeed, the matching between elements that are similar in colour, orientation,

2-D shape, and spatial frequency tends to be preferred to the matching between dissimilar elements (Casco 1990; Green 1986; Green and Odom 1986; Kolers 1972; Navon 1983; Watson 1986). Apparent motion in these conditions has been interpreted as being due to *intra-attribute* matching, that is matching of similar elements (as opposed to *interattribute* matching). In this case, the matching of elements between frames follows what is referred to as the *similarity rule* (Ullman 1979); recent support for it has been provided by Gorea and co-workers (Gorea et al 1992; Gorea and Papathomas 1989, 1991, 1993). Intraattribute matching has been thought to be based on a high-level motion mechanism that computes similarity and matches similar elements in different positions. The most detailed model of intra-attribute matching is a mechanism that tracks feature position over time (Braddick 1980). This mechanism has been described by Ullman (1979), who developed a 'minimal mapping theory' that defines a set of algorithms for computing the likelihood of the correspondence between low-level tokens (eg edges and corners) detected at different times. Recently, the dichotomy between low-level and high-level motion has been reinterpreted by Cavanagh (1991, 1992) to incorporate the distinction between active and passive motion detection. According to this author, the high-level mechanism is active, ie *attentive*, and can operate by tracking of visible objects or features in the image irrespective of how they are defined. Thus, this mechanism is engaged whenever there are visible features in the image regardless of whether luminance discontinuities are present or not.

Compelling demonstrations of intra-attribute matching are obtained with a kinematogram (Cavanagh et al 1985; Simpson 1990) that consists of successive texture images each containing a central subregion of elements that differ from those in the background along a single dimension such as hue, phase, or orientation. It has been suggested that a luminance correlator (low-level mechanism) may be involved in detecting moving contours defined by texture (Mather and West 1993). According to Cavanagh's dichotomy, this mechanism is passive, ie *nonattentive*, and can mediate detection of contours defined by both luminance (by extracting invariants of spatial structure over time) and texture. Thus, according to this view, some form processing can be carried out directly by the motion system on the basis of a low-level, nonattentive mechanism. A considerable amount of theoretical work has been conducted to determine how the low-level passive mechanism could respond to the motion of contours defined by texture. For example, Chubb and Sperling (1988, 1989) have shown theoretically that in many circumstances a simple nonlinear transformation (eg rectification) of the luminance profile of the image can be used to transform the image from a feature domain to a luminance domain, so that the final operation is essentially edge detection in the luminance domain.

The goal of the present study was to determine whether apparent motion of contours defined by texture is mediated by a high-level attentive mechanism or by a low-level nonattentive one. This was achieved by investigating whether the motion system can extract motion of contours defined by texture when their detection in static viewing requires visual selective attention. It is known (Treisman and Gelade 1980) that, in static viewing, detecting and identification of a target in a background of distractors is serial when the target is defined by a conjunction of features. Though there are many exceptions (Houck and Hoffman 1986; McLeod et al 1988; Sagi 1988) most interpretations of serial search link integration of features in a conjunction task to visuo-spatial attention (eg Treisman and Sato 1990; Wolfe et al 1989). Thus, we predicted that if motion of contours defined by texture is carried out by a low-level nonattentive motion mechanism that is attribute-specific, then apparent motion of conjunction of features should be detected preattentively, even though the search for the same conjunction of features would require visual selective attention in static viewing.

To test this prediction, we devised a new stimulus—a variant of the kinematogram— that could drive such a low-level motion mechanism. In its classical version, the

kinematogram consists of two or more frames containing uncorrelated sparse random elements. A central region of elements in frame i is displaced in frame $i + 1$ and, within this region, the elements are spatially correspondent from frame to frame. Because of this spatial correspondence, in the classic kinematogram the motion of a target region can be detected by a luminance correlator which uses, as information invariant, aspects of spatial structure over time. In our stimulus, spatial correspondence was abolished by randomising target position in each frame independently so that motion could not be detected by means of a luminance correlator. This way, motion of the target area could only be detected by a motion mechanism sensitive to texture discontinuities between the target and background regions, ie an attribute-specific motion mechanism.

To summarise, the stimulus employed differed from the standard kinematogram in three main ways: (i) the position of the background elements changed randomly from frame to frame; (ii) the position of the target elements within the target area also changed randomly, so as to abolish correspondence between frames; and (iii) the target elements were defined by the conjunction of orientation and size.

We show that apparent motion of the target area can still be perceived in this variant of the kinematogram. To assess whether motion can be perceived between elements defined by the conjunction of features before these elements are identified in each static frame, we compared temporal thresholds for direction of motion (ie frame duration to detect the moving target) with thresholds for location (ie frame duration to detect target location). Lower thresholds with moving than with static targets were taken as an indication that the identification of target elements mediated by the motion system preceded shape identification by the static system.

Since the position of background elements changes randomly from frame to frame, on average a random subset of elements is seen as moving left, and the remainder as moving right. In other words, there is not a global bias in the direction of apparent motion of the background elements. On the contrary, the position of target elements changes randomly from frame to frame *within* the target area. Since this area is shifted leftwards or rightwards from frame to frame, there is a global bias in the direction of target area displacement. Such a global bias underlies apparent motion of the target area in this type of stimulus.

Overall, the ten experiments presented in this paper show that (i) the motion system has access to the information conveyed by texture discontinuities; (ii) the underlying mechanism cannot be a high-level motion system that first identifies the targets and then matches them across frames; and (iii) the underlying mechanism is a low-level motion system whose directionally selective units are tuned to feature conjunctions.

2 General method
2.1 Subjects
Eight psychophysical observers, only one of whom was experienced (the first author), from the Psychology Department of the University of Padova participated in the experiments. Their average age was 28 years. All subjects had normal or corrected-to-normal vision.

2.2 Apparatus and stimulus
The stimuli were presented on the monitor of a Macintosh SE/30. The stimuli were black (1.78 cd m^{-2}) on a white background (71.9 cd m^{-2}) and were viewed under dim fluorescent room lighting. Responses were made orally after each trial and were recorded by the experimenter. Viewing distance was 228 cm. Display size was 3.6 deg $\times$ 1.6 deg. Each stimulus area contained a total of 60 elements 10 min of arc long; half of the background elements were wide (5 min of arc) horizontal bars, the other half were narrow (2.5 min of arc) vertical bars. Mean interelement separation was 18 min of arc. 1, 4, 6, or 8 target elements were used in independent conditions. Target elements were

wide vertical bars and narrow horizontal bars in experiment 2, and wide vertical bars in all the other experiments. This way, target elements were defined by conjunction of orientation and size. The theoretical number of serials scans, S_S, was defined as (Johnson and Kotz 1977):

$$S_S = \frac{gN + 1}{t + 1},$$

where N is the total number of display items ($N = 60$ in this experiment), t is the number of targets in the display ($t = 1$, 4, 6, or 8), and g is the minimum number of targets that must be detected to make a (positive) response ($g = 1$). The corresponding four S_S values were: 30.5, 12.2, 8.71, and 6.78. Note that S_S increases as the number of targets decreases. We kept the number of distractors fixed and varied the number of targets to avoid stimulus area and interelement distance covarying with S_S. Indeed, these spatial factors are known to affect performance in both visual search (Sagi and Julesz 1987; Toet and Levi 1992) and motion perception (Snowden and Braddick 1989b).

The stimulus was created with a three-step algorithm (figures 1 and 2). In the first step, 60 locations within the stimulus area were chosen by means of a pseudo-random number generator. In the second step, 1, 4, 6, or 8 of these locations were selected randomly within a nominally defined subregion of the display (target area). In the third step, target elements were placed at these selected locations while distractors were placed at the remaining locations. Elements were not allowed to be in contact or to overlap within a frame. This procedure was repeated for every frame in the stimulus sequence with a new seed for the pseudo-random generator. The size of the rectangular target area was 0.9 deg × 1.6 deg in all experiments except experiment 4, where a condition with a subregion of 0.45 deg × 1.6 deg was used.

2.2.1 *Static stimuli.* In the static viewing conditions (figure 1) only one frame was presented, immediately followed by a 250 ms mask. Static stimulus duration corresponded to frame duration in the dynamic stimuli. The mask was used to terminate stimulus processing and to ensure that participants could use only information from the first fixation (Breitmeyer 1984). The centre of the target region was 1 deg left or right relative to the centre of the stimulus. The observer's task was to say "left" or "right" according to whether the target elements were perceived in the left or in the right side of the display in all experiments except experiments 7 and 8 in which a present versus absent task was used.

2.2.2 *Dynamic stimuli.* In the dynamic viewing conditions (figure 2), apparent motion was generated by presenting sequences of frames with no interframe interval (IFI). In experiments 1, 2, 4, 5, 7, 8, 9, and 10 the stimulus consisted of three frames. The number of frames was a factor with three levels in experiment 3 (two, three, and six frames) and two levels in experiment 6 (three and six frames). With the centre of the screen at the centre of the coordinate system, the horizontal position over time of a target area moving rightwards was:

$$X_i = -d\left[\tfrac{1}{2}(n+1) - i\right], \tag{1}$$

where X_i is the horizontal position of the centre of the target region in frame i, d is the interframe distance, n is the number of frames in the sequence, and i is frame number ($1 \leqslant i \leqslant n$). Note that locations in the left half of the screen have negative coordinates. Similarly, for a target area moving leftwards:

$$X_i = d\left[\tfrac{1}{2}(n+1) - i\right]. \tag{2}$$

The target area was displaced by a fixed amount from frame to frame, from left to right in half of the trials and from right to left in the other half [d in equations (1) and (2)]. It may be worth reiterating that target element position within the target
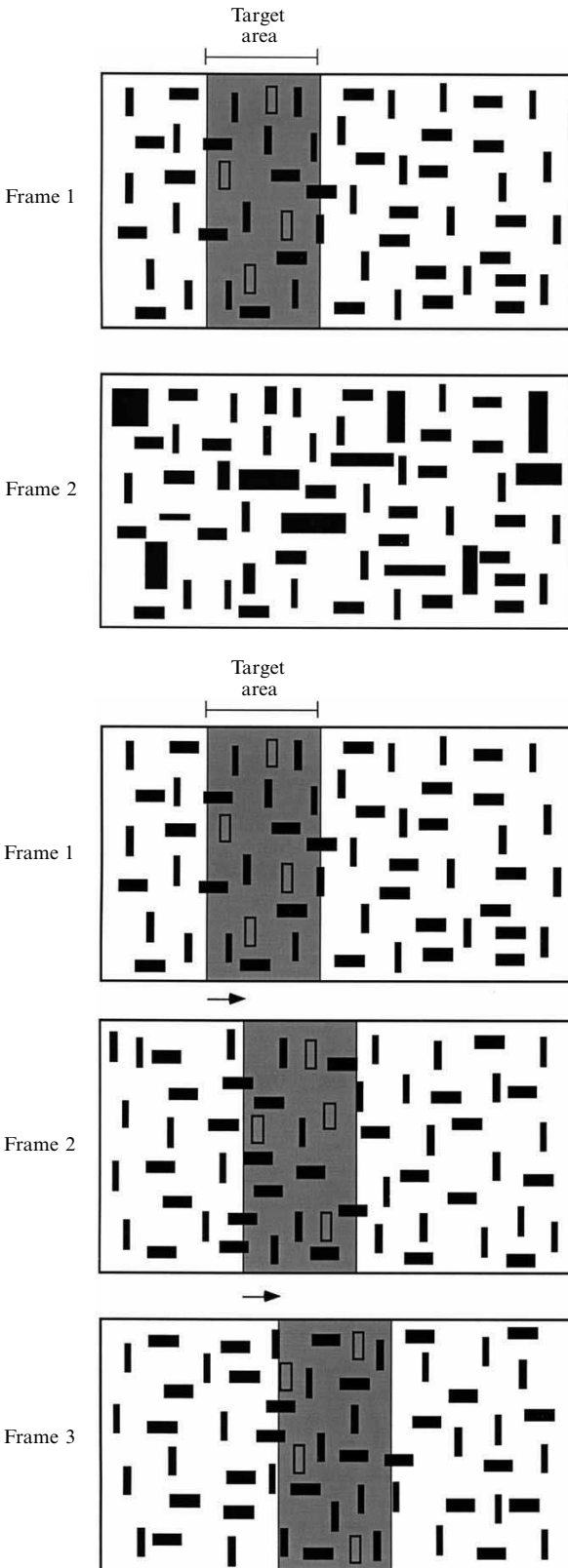
Figure 1. Static display. Distractors (wide horizontal bars and narrow vertical bars) are placed at random locations within the stimulus area. Target elements (wide vertical bars in all experiments with the exception of experiment 2 where they are both wide vertical bars and narrow horizontal bars) are placed at random location within the target area (1 deg off the centre of the stimulus). Target elements are outlined for illustrative purposes. The stimulus is followed with no interval by a 250 ms mask (frame 2).

Figure 2. Dynamic display. The stimulus consists of multiple frames presented in sequence with no interval between them. Each frame is generated with the same algorithm as that used for the static display, with the constraint that the target area is shifted leftwards or rightwards from frame to frame (see text for details).

area was recomputed by a pseudo-random algorithm in each frame. This was the case also for the background elements within the entire stimulus area. The displacement d was equal to 18 min of arc in all experiments (with the exception of experiment 9 in which d was 0). This way, the *average* position of target elements was shifted by 18 min of arc to the right or to the left from frame to frame. In experiment 5, in addition to 18 min of arc, two additional values of d were used: 36 and 54 min of arc. Note that, since 18 min of arc is about one third of the target area width, target areas in successive frames largely overlap. In the standard random-dot kinematogram, target elements are correspondent as their spatial configuration remains fixed from frame to frame. In the present studies, this kind of correspondence between stimulus elements in successive frames was abolished because their position changed randomly. Thus, in the dynamic stimulus, detection of the displacement direction of target elements cannot rely upon correspondence based on spatial configuration matching between frames; rather, it must rely upon a motion mechanism sensitive to feature conjunctions.

The observer's task was to respond "left" or "right" depending on whether the target elements moved left or right in all experiments except experiments 7 and 8, in which a present versus absent task was used. There was no need for a mask in the dynamic condition because (i) frame $i + 1$ is a mask for frame i, and (ii) observers could not rely on the last frame to infer the direction of motion since in the final frame the target area was only slightly off the stimulus centre. However, to assess the effect of a mask more directly we conducted a control experiment where a mask followed the motion sequence (experiment 6).

2.3 Psychophysical procedure

The psychophysical procedure used in these experiments was a binary choice in which the observer had to detect either the position (static conditions) or the direction of movement of target elements (motion conditions). Observer thresholds were defined as the frame duration for 75% correct performance (halfway between chance and perfect performance). After a warning sound the stimulus would appear from a gray background. In both static and dynamic stimulus conditions, frame duration varied between 17 and 250 ms in independent trials. Trials were presented in blocks of twelve stimuli (four with target elements displaced right, four displaced left, and four catch trials in which the target region was at the centre of the screen and did not move between frames). In experiments 7 and 8, the catch trials did not contain targets. In half of the blocks, frame duration was decreased progressively from trial to trial; in the other half, frame duration was increased. Twenty-four blocks (288 trials) were used for each S_S condition. Within a block, two trials were devoted to each duration level. Blocks were presented in random order.

2.4 Statistical analyses

Analyses were performed with an ANOVA on experiments with four or more subjects and with a nonparametric bootstrap test (Efron 1993) in the other cases. The statistic used in the bootstrap test was the difference between the means.

3 Experiment 1

In the first experiment, thresholds for detecting the position of targets in the static display were compared with those for detecting the direction of motion of similar targets in the dynamic display.

3.1 Results and discussion

The results are shown in figure 3, where thresholds obtained in a basic condition of search for a single target ($S_S = 30.5$) were compared with those obtained in three levels of S_S (12.2, 8.3, 6.8) independently for the two stimulus conditions (static versus dynamic)
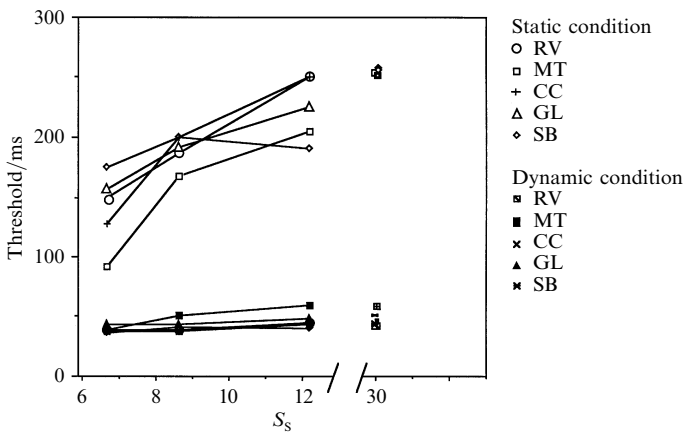
Figure 3. Results of experiment 1. Thresholds (defined as stimulus duration required for 75% correct responses) plotted as a function of the theoretical number of serial scans, S_S, independently for five observers and two viewing conditions (static: unfilled symbols; dynamic: filled symbols).

and five observers. An ANOVA was conducted on the raw data, with stimulus condition and S_S as factors. The results show that thresholds in the static condition were larger than those in the dynamic one (187.7 versus 42.4; $F_{1,4} = 180.3$, $p < 0.001$). Pairwise comparisons (Newman – Keuls) indicate that the difference between thresholds for the two stimuli is significant (all $ps < 0.01$) at each of the S_S levels (140 versus 38.4, 199.1 versus 41.8, and 224 versus 47.0 for S_S equal to 6.8, 8.3, and 12.2, respectively). Pairwise comparisons also indicate that thresholds increase significantly ($p < 0.01$) with S_S in the static, but not in the dynamic condition. Significance in this case refers to the comparison between thresholds for the smallest and largest values of S_S (dynamic, 38.4 versus 41.8; static, 47 versus 224). The higher thresholds in the static relative to the dynamic condition indicate that the perceptual segregation of target elements in each static frame is not necessary for detecting the direction of their displacement in successive frames.

An ANOVA was conducted on the slopes of the linear regression lines fitted to individual data, which represent the mean processing time per S_S. The only factor used in this ANOVA was condition (static versus dynamic). The slope of the regression lines was larger in the static than in the dynamic condition (15.42 versus 1.65; $F_{1,4} = 20.38$, $p < 0.05$). The difference in the slope of the regression line obtained in the two tasks suggests that the search for conjunction is serial and involves focal attention in the static task, whereas it is parallel and preattentive in the motion task.

These results show that visual search for conjunction is slower in the static than in the motion task. This suggests that target detection in the dynamic condition is not performed by a process that explicitly identifies target elements in each static frame and then matches them across frames. We propose that the underlying mechanism may be a Reichardt-like detector which operates on the basis of the output provided by direction-selective units (Barlow and Levick 1965; Reichardt 1961; Van Santen and Sperling 1985). The existence of motion detectors tuned to different resolution scales is supported by the effects of spatial frequency on the strength of apparent motion (Casco 1990; Green 1986; Watson 1986). Moreover, van den Berg and van de Grind (1990) demonstrated that motion detectors exhibit orientation selectivity. We suggest that at least some motion detector subunits are tuned for specific combinations of size and orientation and that these may constitute a subset of units specialised for the extraction of contours defined by textures.

4 Experiment 2

An alternative explanation of the results of experiment 1 is that the visual system detects the difference in the relative density of bars with different orientation. Remember that in experiment 1 the target elements were wide vertical bars. Therefore, the density of vertical elements within the target area was higher than that in the background area. In principle, this information could be used by an apparent-motion mechanism that detects the displacement direction of an area in which the relative density of horizontal and vertical bars differs from that of the background. This possibility was investigated in experiment 2.

4.1 Method

In this experiment, target elements were both wide vertical bars and narrow horizontal bars. Therefore in this stimulus the density of vertical and horizontal bars was the same, locally as well as across the entire display. S_S levels were the same as in experiment 1.

4.2 Results and discussion

The results are shown in figure 4, independently for three observers. The results are very similar to those obtained in experiment 1 in that (i) the slope of the search function was higher in the static than in the motion condition (15.97 versus 0.56, $p = 0.008$); and (ii) thresholds increased with S_S in the static (76 versus 139; $p = 0.01$) but not in the dynamic condition (36 versus 39; $p = 0.1$). Significance in this case refers to the comparison between thresholds for the smallest and largest values of S_S. These results show that the relative density of vertical bars was not the cue used during the detection task.
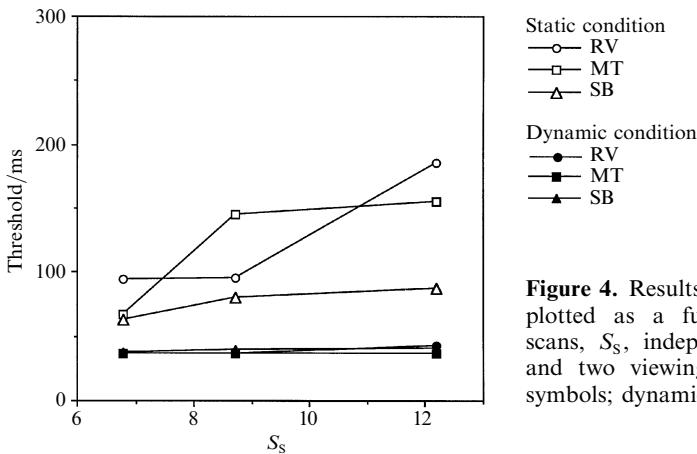

Figure 4. Results of experiment 2. Thresholds plotted as a function of theoretical serial scans, S_S, independently for three observers and two viewing conditions (static: unfilled symbols; dynamic: filled symbols).

Overall, the results of these experiments suggest that motion perception of target elements defined by the conjunction of features is mediated by a motion system similar to that described by Barlow and Levick (1965) or Reichardt (1961) but tuned to high-level attributes (in this case, a combination of width and orientation). The alternative explanation that motion perception in our conditions may be based on a high-level attentive system that first identifies forms and then tracks them (Cavanagh 1992; Ullman 1979) is not supported by our data.

Since the displacement of the target area is one third its width, the overlap between the target area in successive frames may produce some pairings of target elements in the direction opposite to that of overall displacement, ie false pairings (Casco and Morgan 1987). How can the direction of overall displacement be detected under these conditions? We suggest that motion signals generated by false pairings are inhibited by a cooperative process based on inhibitory and excitatory connections between motion detectors: units tuned to the same direction excite each other, while units tuned to opposite directions

inhibit each other. It has been suggested that such a cooperative process can solve the correspondence problem in the standard random-dot kinematogram, where the motion signals in the same direction of overall displacement are thought to reinforce each other while inhibiting a minority of signals in the opposite direction due to false pairings (Chang and Julesz 1984). Such a cooperative process can effectively amplify small biases present in the individual motion signals over a population of motion detectors. In our stimulus, such bias is produced by the displacement of the target area, if it is assumed that the underlying motion mechanism is tuned to the conjunction of size and orientation.

5 Experiment 3
The cooperative process mentioned above may combine signals over space and time (Chang and Julesz 1984). In the random-dot kinematogram the combination of motion signals over time has been shown to occur up to six 45-ms-duration frames. This is known as 'temporal recruitment', and can improve detection in a unit that summates signals from subunits in a chain. In the next experiments, we investigated whether temporal recruitment was also present in our stimulus by measuring thresholds as a function of number of frames.

5.1 Method
Thresholds for $S_S = 12.2$ (four target elements) were measured as a function of the number of frames. Two, three, and six frames were used. All other parameters were the same as in experiment 1.

5.2 Results
Figure 5 shows thresholds as a function of the number of frames for three observers.

Thresholds in the motion task in the two-frame condition are not different from those obtained in the static task in experiment 1 (220 ms versus 224 ms; $p = 0.46$). Note that in the two-frame condition the *total* duration of the dynamic stimulus is twice that of the static stimulus. If the motion task was carried out by identifying target position in each frame one would expect to observe an advantage in the dynamic versus the static stimulus condition. The lack of such an advantage suggests that the observers were truly performing a direction-of-motion discrimination task.
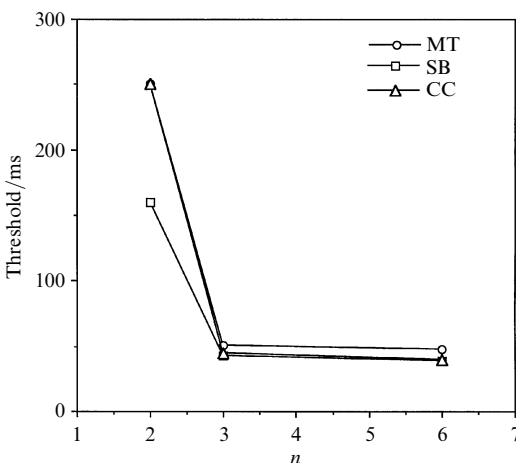

Figure 5. Results of experiment 3. Thresholds in the dynamic task plotted as a function of the number of frames, n, independently for three observers ($S_S = 12.2$).

Thresholds in the three-frame and six-frame conditions were lower than those in the static condition (48 ms versus 224 ms and 42 ms versus 224 ms; $p = 0.01$ in both cases). The threshold in the six-frame condition was lower than that in the three-frame one (42 ms versus 48 ms; $p = 0.01$). It has been suggested that performance increases with the number of frames because of temporal recruitment (McKee and Welch 1985; Nakayama

and Silverman 1984) by an integrative cooperative process that combines information from motion detectors over space and time (Snowden and Braddick 1989a, 1989b, 1991). If the underlying motion mechanism responds to high-level attributes, the cooperative integrative property may account for the detectability of a small bias of individual target element motion in the direction of overall displacement. As described above, this bias exists because of target area displacement. However, this bias is not easy to detect with two frames because displacement is one third the width of the target area and thus successive target areas largely overlap producing false pairings. The results of the current experiment suggest that the increase in performance with the number of frames is due to temporal recruitment, that is integration of motion signals across frames. The result of such temporal recruitment is an increase of the strength of motion signals in the direction of overall target displacement relative to the opposite one.

6 Experiment 4

If the interpretation of the previous experiment is correct, then the facilitatory effect of increasing the number of frames should be produced also by decreasing the width of the target area in the two-frame condition. Indeed, reduction of the width of the target area results in a smaller overlap between successive frames (ie fewer false pairings), thus increasing the size of the motion bias in the direction of target motion.

6.1 Method

The two-frame condition of experiment 3 was employed with a target area that was half its original width (0.45 deg). All other parameters were the same as in the previous experiment.

6.2 Results

The results are shown in figure 6 in which thresholds are shown independently for two observers. Thresholds drop drastically in the small-target-area condition (250 ms versus 70 ms, $p = 0.04$). This what one would expect if detection relied on a cooperative process in which detectability increases with the proportion of motion units being stimulated in the direction of motion of the target area.

The results of the last two experiments suggest that integration occurs in the direction of motion, provided a bias in this direction physically exists. We have shown that detection improves when such a bias is increased either by temporal recruiting (by using more than two frames) or by increasing the local motion bias in the direction of target motion (by reducing the width of the target area).
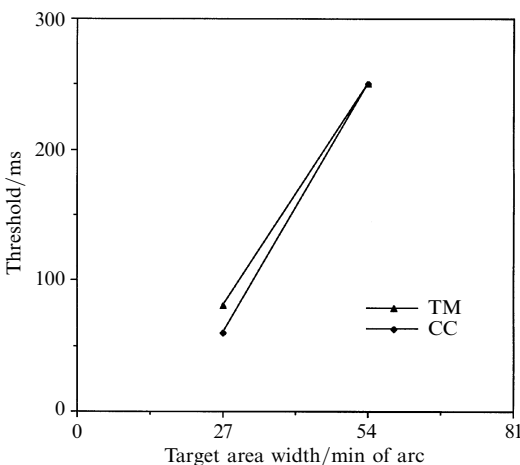

Figure 6. Results of experiment 4. Thresholds obtained in two target-area-width conditions (0.45 and 0.9 deg) plotted independently for two observers ($S_S = 12.2$).

Control experiments
7 Experiment 5
In this experiment we investigated whether preattentive search of conjunctions in dynamic conditions also occurs for long-range motion. Dick et al (1987) showed that direction of motion is detected in parallel (preattentively) only with small displacements (less than 17.5 min of arc). Our previous results were obtained with small displacements (short-range motion conditions). The rationale for this experiment is that, if parallel search in the dynamic stimuli depended upon short-range motion, then it would not take place in a long-range motion condition.

7.1 *Method*
All parameters in this experiment were the same as in the dynamic condition in experiment 1, with the only exception that two different displacement conditions were tested (d of 36 and 54 min of arc, instead of 18 min of arc).

7.2 *Results*
In figure 7, thresholds obtained in experiment 1 (18 min of arc displacement condition) are compared with two long-range displacement conditions: 36 and 54 min of arc, independently for three observers.
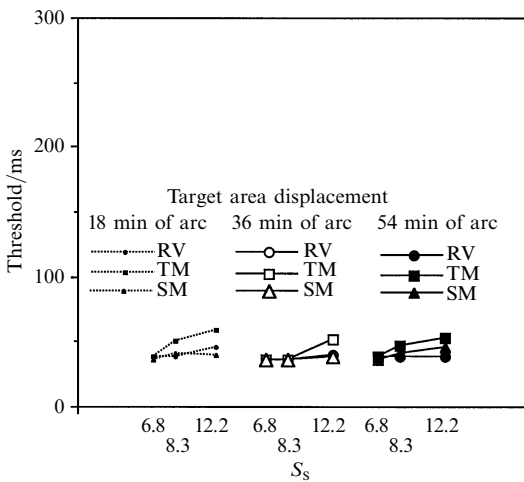

Figure 7. Results of experiment 5. Thresholds in the dynamic condition plotted as a function of S_S (6.8, 8.3, and 12.2) for three levels of target area displacement (18, 36, and 54 min of arc).

The slope of the search function was close to 0 in all cases. Slopes tended to decrease with displacement: 2.5, 1.45, and 0.8 for displacements of 18, 36, and 54 min of arc, respectively. The small effect could be due, at least in part, to a practice factor as this experiment was conducted after experiments 1 and 2 on a subset of the same subjects.

These results indicate that, regardless of displacement size of the target area, visual search for these conjunctions involves parallel preattentive processes. This interpretation seems at odds with the suggestion that long-range motion requires serial search (Dick et al 1987). Remember, however, that in Dick et al's experiment motion was one of the features that defined the target, while in our experiment it was not. In addition, Snowden and Braddick (1989b) and Nakayama and Silverman (1984) have shown that the displacement limit for perceiving motion in the random-dot kinematogram with three frames is greater than twice that observed with kinematograms (the ones used by Dick et al 1987), consistent with the hypothesis of a process that combines motion information over time.

8 Experiment 6

A potential confound is the use of a mask in the static but not in the dynamic stimulus. In principle, the motion task could be carried out by detecting the target area in the last frame, not masked by a subsequent image, and by inferring the direction of motion (eg if the target area in the last frame is on the right, motion must have been rightwards). This control experiment assessed this issue by using the dynamic stimuli employed in experiment 1, with the only difference that the stimulus was followed by a mask.

8.1 *Method*

Only the four targets were used ($S_S = 12.2$); thresholds were measured in both three-frame and six-frame conditions.

8.2 *Results and discussion*

In figure 8 the results for two observers are compared with those obtained in experiment 1 (no mask). It is clear that thresholds are virtually unaffected by the mask (48 ms versus 45 ms in the mask and no-mask conditions, respectively; $p = 0.2$). Moreover, the mask had no effect regardless of the number of frames (53 ms versus 48 ms in the three-frame condition, $p = 0.25$; 45 ms versus 44 ms in the six-frame condition, $p = 0.27$). This rules out the possibility that the motion task is executed by detecting target position in the last frame and inferring the direction of motion.
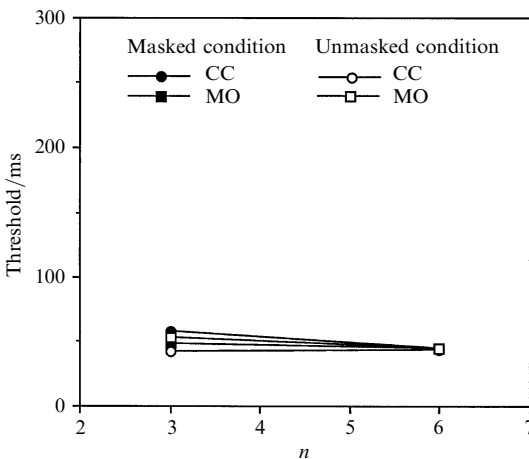

Figure 8. Results of experiment 6. Thresholds plotted as a function of the number of frames, *n*, for the masked (filled symbols) and unmasked (open symbols) conditions independently for two observers.

9 Experiment 7

Another potential confounding factor is the difference in the tasks used in the static and dynamic conditions, a left versus right *location* decision in one case and a left versus right *motion direction* decision in the other. At least part of the effect may be attributable to this task difference because factors of difficulty and eccentricity could vary in these two tasks. For example, one important difference is that the location decision becomes easier the closer the stimulus is to the fovea. To rule out this possibility, in this experiment observers were asked to report whether the targets were present or absent in both static and dynamic conditions.

9.1 *Method*

The stimulus was the same as in experiment 1, except that only one target was used ($S_S = 30.5$) and that in each of twenty-four blocks the target was absent in four out of twelve trials.

9.2 *Results and discussion*

The percentages of correct responses in present and absent conditions are shown in figure 9 as a function of frame duration levels in both static and dynamic condition for two observers. The percentage of correct responses is much higher in the dynamic than in the static task (92% versus 52%; $p = 0.03$); pairwise comparisons show that this is the case for every duration level (all $ps < 0.05$). These results confirm that the facilitation effect in the motion task is not a task-related artifact.
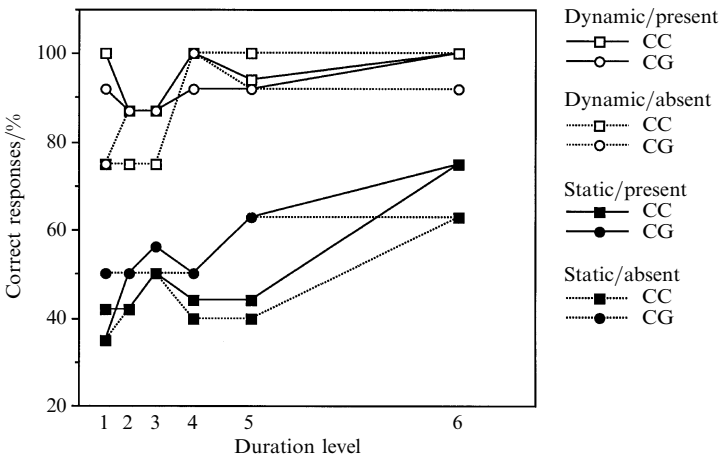

Figure 9. Results of experiment 7. The percentage of correct responses in a present versus absent task ($S_S = 30.5$) plotted for two observers as a function of six frame duration levels (average frame duration was 250, 113, 75, 50, 33, and 17 ms in levels 6, 5, 4, 3, 2, and 1, respectively). Present trials are shown with continuous lines and absent trials with dotted lines independently for the static (filled symbols) and dynamic (empty symbols) stimulus conditions.

10 Experiment 8

The main result in this study is that the motion and location tasks differ not only in absolute level of duration thresholds but also in the rate of change with the number of targets. The difference in the slope of the regression line obtained in the two tasks suggests that the search for conjunction is serial in the static task but parallel and preattentive in the motion task. However, Sagi (1990) has shown that search for conjunction of spatial frequency and orientation in static stimuli can be parallel. Although we varied size rather than spatial frequency, search in our static stimulus could be parallel as suggested by the rate of change (about 25 ms per item) which is at the fast end of serial search (Treisman and Sato 1990). To test this possibility, standard search functions were measured with both static and dynamic stimuli.

10.1 *Method*

Only one target was presented with a variable number of distractors. Observer thresholds were defined as in the previous experiments (ie frame duration for 75% correct responses). The number of distractors varied in independent trials according to three levels (20, 40, and 60 in the motion task; and 10, 20, and 30 in the static task) giving S_S values of 10.5, 20.5, and 30.5 versus 5.5, 10.5, and 15.5, respectively. Stimulus duration varied between blocks according to five levels. To obtain a psychometric function, the temporal distance between successive duration levels was chosen according to the difficulty of the task: 66 ms and 17 ms for the static and dynamic tasks, respectively. Trials were presented in blocks of twelve stimuli (four with target elements displaced right, four with target elements displaced left, and four with no target). The observer's task was to report whether the target was present or absent.

10.2 *Results and discussion*

The results for three observers are shown in figure 10. Thresholds are plotted as a function of S_S independently for the two stimulus conditions.

Linear regression lines were fitted to individual data in the two tasks. The average slope of the regression line (corresponding to the mean processing time per S_S) was 34 ms in the static task and 0.35 ms in the dynamic task. Although experimental conditions can be found (Sagi 1990) in which search for spatial frequency and orientation is effortless, the result of this experiment shows that search for conjunction of size and orientation is serial with the static stimuli employed here. Perhaps this discrepancy is not surprising, because the stimulus used in Sagi's demonstration differs in two important aspects from other conjunction stimuli used in classical search tasks (Treisman 1986). First, both target and distractors contain a component of horizontal and vertical orientation as well as of low and high spatial frequency (the difference is solely in the way orientation and spatial frequency are combined). Second, in Sagi's demonstration distractors are arranged in a regular grid, while in our stimulus they are not.
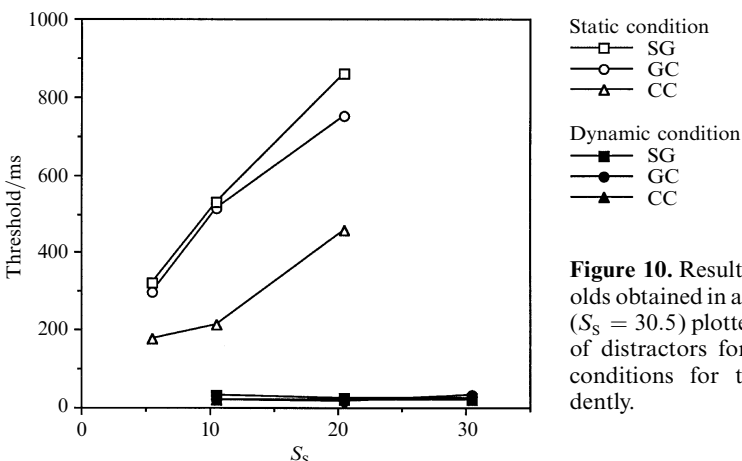

Figure 10. Results of experiment 8. Thresholds obtained in a present versus absent task ($S_S = 30.5$) plotted as a function of number of distractors for the static and dynamic conditions for three observers independently.

11 Experiment 9

One possible artifact in our experiment is that the motion response could be based on location because it is correlated with motion direction. For instance, if the target is on the left at the beginning of a stimulus sequence then it moves rightwards during the rest of the sequence. The observer could learn this pairing, especially in the longer-duration trials. In other words, the motion response could be inferred from target location, which in turn may be easier to see in the multiple frame display (though the last frame conveys no particular advantage as shown in experiment 6). To rule out this possibility we used the standard three-frame stimulus but the target area was not displaced between frames [ie d in equations (1) and (2) was set to 0]. If the motion response was based on location, then observers should be able to perform the location task (left versus right) in these stimulus conditions.

11.1 *Method*

The stimulus was the same as in experiment 7 but the target area was not displaced from frame to frame [$d = 0$ in equations (1) and (2)].

11.2 *Results and discussion*

The percentage of correct responses as a function of frame duration levels is shown in figure 11. The results obtained in the static and dynamic conditions of experiment 1 are also reported for comparison.
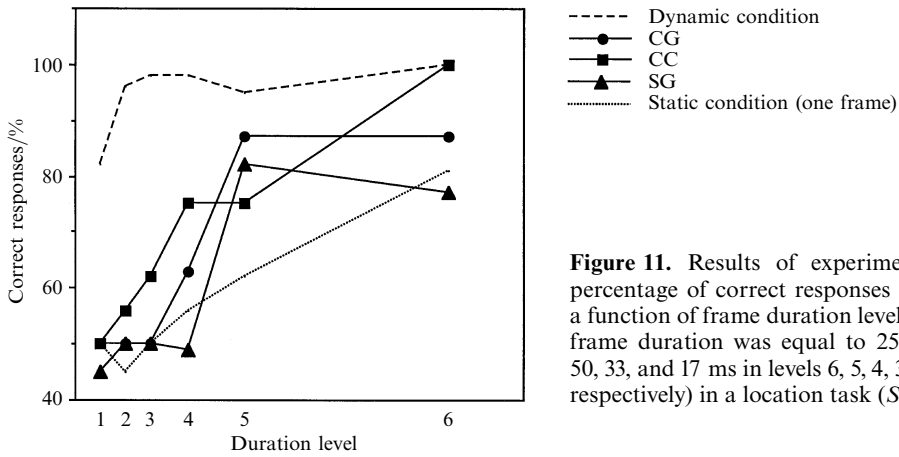
Figure 11. Results of experiment 9. The percentage of correct responses plotted as a function of frame duration levels (average frame duration was equal to 250, 113, 75, 50, 33, and 17 ms in levels 6, 5, 4, 3, 2, and 1, respectively) in a location task ($S_S = 30.5$).

At short frame durations (levels 1 and 2), the results in the two static conditions are similar (all ps > 0.2) while at longer frame durations the performance in the three frame conditions is in between that obtained in the static and the dynamic conditions in experiment 1 (all ps < 0.05). This is not surprising because location judgment in our stimulus could still be mediated by independent target motion, although without any net direction.

The main finding of this experiment is that observers' response in the dynamic stimulus is not based (or at least not entirely) on location information. Indeed, it is difficult to detect the location of the target area in the absence of a global motion bias.

12 Experiment 10
If the advantage in our dynamic displays is due to the activation of a low-level motion mechanism tuned to feature conjunctions, then such advantage should be reduced or abolished by an experimental manipulation that interferes with the response of this motion mechanism. The last experiment was conducted to test this prediction.

12.1 *Method*
The stimulus was the same as that used in experiment 7 with the only difference that successive target frames were interleaved with dynamic visual noise. Dynamic visual noise consisted of a two-frame sequence in which each frame contained randomly placed nontarget elements. Dynamic visual noise was added to the three-frame sequence used in experiment 7 (F1, F2, F3) resulting in the new nine-frame sequence (Ft1, Fn1, Fn2, Ft2, Fn1, Fn2, Ft3, Fn1, Fn2), where Ft1, Ft2, and Ft3 are the frames containing the targets while Fn1 and Fn2 are the visual noise frames. The observer task was to detect the direction of motion (left versus right) of the target area.

12.2 *Results and discussion*
Thresholds are shown independently for two observers in figure 12 together with those obtained in the motion condition without interference tested in experiment 7. Thresholds are much higher in the presence of dynamic visual noise (250 ms versus 50 ms; $p = 0.04$). These results show that the advantage for dynamic displays disappears when the response of the motion mechanism is disrupted by dynamic visual noise. This provides further support for the claim that the matching of conjunctions in the dynamic displays is carried out by a motion mechanism.
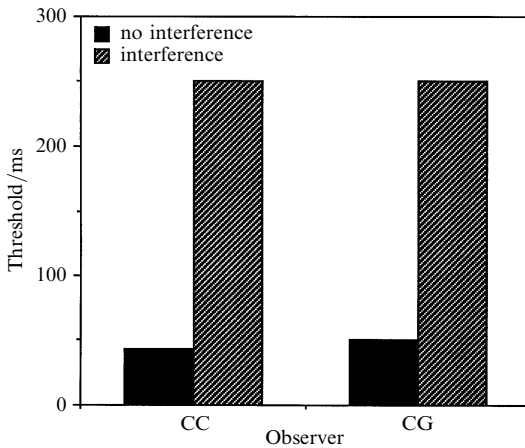
Figure 12. Thresholds obtained in a direction-of-motion task ($S_S = 30.5$) plotted independently for two observers. Dynamic visual noise was added to the three-frame sequence (F1, F2, F3) resulting in a new frame sequence (Ft1, Fn1, Fn2, Ft2, Fn1, Fn2, Ft3, Fn1, Fn2) where Ft1, Ft2, and Ft3 are frames containing the target, while Fn1 and Fn2 are the visual noise frames. Thresholds are compared with those obtained in experiment 7 in the present versus absent task without interference.

13 Discussion and conclusions

In the literature there is disagreement on (i) whether apparent motion tends to follow the similarity rule (ie is attribute-specific) and, if so, (ii) what the responsible mechanism might be. A number of studies based on apparent-motion techniques have provided mixed evidence on these issues (eg van den Berg and van de Grind 1990; Casco 1990; Cavanagh et al 1989; Green 1986; Green and Odom 1986; Ramachandran et al 1983; Simpson 1989; Ullman 1979; Victor and Conte 1990; Watson 1986).

With regard to the first issue, by using a stimulus in which moving texture discontinuities are defined by the conjunction of size and orientation, we provide evidence that the mechanism underlying apparent motion is attribute-specific.

With regard to the second issue, we provide evidence that the responsible mechanism is a low-level nonattentive motion mechanism. Two findings support this conclusion. First, the search for conjunctions is much slower with static than with dynamic displays; second, the search for conjunctions is serial with static but not with dynamic displays, suggesting that visual attention is not required for the detection of moving targets defined by the conjunction of size and orientation.

Previous work had already reported that feature conjunctions can be detected in parallel (eg McLeod et al 1988; Nakayama and Silverman 1986; Wolfe et al 1989), consistent with neurophysiological findings indicating that feature conjunctions are directly encoded in the visual system of many organisms. For example, in primate, most cells in V1 are tuned for spatial frequency and orientation (De Valois et al 1982); many cells in extrastriate areas (such as V2, V4, TEO, and TE) are tuned to increasingly complex conjunctions of features (Desimone et al 1985; Maunsell and Van Essen 1983; Tanaka 1993; Thorell et al 1984); cells in area MT seem to encode not only motion information but also information about the orientation and the shape of the moving stimuli (Malonek et al 1994). However, to our knowledge this is the first study to show that targets defined by the conjunction of certain visual features can be detected preattentively when they are in motion but not when they are stationary.

The result that the motion system can identify targets defined by the conjunction of features before they are identified by the static system has important implications for theories of motion perception. Indeed it addresses the issue of the interactions between motion and shape processing. Our results do not support modular theories that postulate independent processing of shape and motion (eg Braddick 1974, 1980; Maunsell and Van Essen 1983). Instead, they suggest that the motion system can access shape information directly by responding to complex attributes (eg van den Berg and van de Grind 1990).

In recent literature many hypotheses have been advanced to account for the apparent motion of texture discontinuities. The simplest and more parsimonious explanation is that they are detected by a correspondence-based mechanism. The most-articulated correspondence-based mechanism is a feature-tracking one, also referred to as long-range mechanism (Cavanagh 1992). Our results rule out the possibility that apparent motion of texture discontinuities requires attentive tracking of features over time. Indeed, they show that the detection of conjunction of features in motion is faster than when they are static and does not involve attention as would be required by an attentive tracking motion system.

Our results do not support the hypothesis that apparent motion of texture discontinuities is mediated by a low-level nonattentive mechanism operating directly on image intensities. Indeed, this low-level correspondence mechanism would be based on the extraction of invariant aspects of spatial structure over time which, by construction, are not available in our stimulus.

Our results suggest that the mechanism underlying apparent motion of texture discontinuities is a low-level mechanism that is sensitive to the similarity of local features in different frames. We provide evidence that such a mechanism might be based on the output of a population of motion detectors (Barlow and Levick 1965; Reichardt 1961) tuned to the conjunction of high-level attributes (Mather and West 1993). This motion mechanism can be a Reichardt-like 'delay and compare' detector that compares the outputs of two spatial filters (subunits) characterised by (i) spatially offset receptive fields, and (ii) similar spatial frequency and orientation tuning. However, the Reichardt model has difficulty in predicting complex motion in multielement images like those we used, in which motion velocity varies across the image (Adelson and Bergen 1985). More appropriate to do the task would be a mechanism specific for extracting motion of texture discontinuities (Mather and West 1993; Wilson et al 1992; Wilson and Mast 1993). In this model the stimulus is filtered out at the first stage by a mechanism selective for orientation or the combination of size and orientation (in our stimulus). The second stage involves rectification followed by a low-spatial-frequency filtering (to extract texture discontinuities) and extraction of motion energy. Motion energy is extracted by means of a spatiotemporally oriented filter (Adelson and Bergen 1985). Filtering is a continuous operation and leads to a continuously varying output, whereas matching is discrete, taking place between images samples at two particular moments in times.

In principle, this second model may account for our findings providing that two conditions are satisfied: (a) The first stage filtering is carried out by mechanism responding to the conjunction of orientation and spatial frequency. Our results show that the motion system possesses higher sensitivity to these conjunctions relative to the static system. (b) Since target motion is perceived with a single target element, motion energy can also be extracted at the fine scale level before the second filtering operation (which allows the extraction of texture discontinuities at the coarse level). This operation is not necessary to do the task in our stimulus, at least when only one target is present. Many of our results, like the temporal recruitment and the target-area-width effect, suggest that encoding of moving textures discontinuities occurs at a fine scale level, that is at the level of local motion signals resulting from displacement of element defined by the conjunction of features.

We show that performance increases either by increasing the number of frames (resulting in temporal recruitment, as in experiment 3) or by reducing the width of the target area (resulting in fewer false matchings, as in experiment 4). These effects may be accounted for by a mechanism characterised by integrative and cooperative properties that is capable of combining information over space and time. A mechanism operating this way would increase the bias of individual motion signals in the direction of target displacement (Casco and Morgan 1987). This bias physically exists with three frames

and only one target, and our results show that it can be detected. Finally, experiment 5 showed that our results with moving stimuli are not confined to short-range motion parameters, suggesting that the mechanism underlying motion perception of contours defined by texture can operate over long distances (Cavanagh and Mather 1989).

A number of control experiments were conducted to rule out the possibility that the differences between the static and dynamic conditions were artifactual. First, the static display was masked, whereas the dynamic display was not. Experiment 6 showed that our results are not an artifact of the mask. Second, the tasks employed in the static and motion conditions were different. Experiment 7 showed that the same differences between the static and dynamic conditions are present when the same task is used in the two conditions. Third, the motion task could in principle have been performed on the basis of location information. Experiments 6 and 9 showed that observers were using motion information to perform the motion task. Finally, an additional demonstration that motion information was used to perform the motion task was provided in experiment 10. There we show that performance in the motion task is disrupted by adding random visual noise to the dynamic stimulus which interferes with the extraction of motion information.

Our results have implications also for theories of visual search. Indeed, it is difficult to interpret our findings in the context of the classical feature integration theory, which predicts that features are analysed separately (Treisman 1986). The results in the dynamic condition seem to favour network theories of visual search which predict that conjunctions can be detected in parallel by the motion system (Green 1991). These theories assume a cooperative process amongst neighbouring units in which the final response is enhanced when they signal similar features. We have shown that a cooperative operation may account for the results obtained with our moving stimuli.

In conclusion, the present studies show that neither a high-level attentive mechanism nor a low-level passive one based on luminance correlation underlies the detection of moving texture discontinuities (Mather and West 1993). Rather, the responsible mechanism is a low-level, passive, and attribute-specific one. Mather and West (1993) have suggested that this attribute-specific mechanism is selectively activated by the motion of second-order contours which define image regions that have the same mean intensity but differ in texture, contrast, motion etc. Since in our displays the target and the background region have the same luminance, target motion can be regarded as motion of a second-order discontinuity. Our data suggest that these texture discontinuities can be segregated from the background by means of a low-level motion mechanism, presumably a motion-energy mechanism (Chubb and Sperling 1989; Werkhoven et al 1993; Wilson et al 1992) tuned to feature conjunctions.

References
Adelson E H, Bergen J R, 1985 "Spatiotemporal energy models for the perception of motion" *Journal of the Optical Society of America A* **2** 284 – 299
Anstis S M, 1970 "Phi movement as a subtraction process" *Vision Research* **10** 1411 – 1430
Barlow H B, Levick W R, 1965 "The mechanism of directionally selective units in rabbit's retina" *Journal of Physiology (London)* **178** 477 – 504
Berg A van den, Grind W van de, 1990 "Motion detection in the presence of local orientation changes" *Journal of the Optical Society of America A* **7** 933 – 939
Braddick O J, 1974 "A short range process in apparent motion" *Vision Research* **14** 519 – 529
Braddick O J, 1980 "Low-level and high-level processes in apparent motion" *Philosophical Transactions of the Royal Society of London, Series B* **290** 137 – 151
Breitmeyer B G, 1984 *Visual Masking: An Integrative Approach* (New York: Oxford University Press)
Casco C, 1990 "The relationship between visual persistence and event perception in bistable motion display" *Perception* **19** 437 – 445

Casco C, Morgan M J, 1987 "Detection of moving local density differences in dynamic random patterns by human observers" *Perception* **16** 711 – 717

Cavanagh P, 1991 "Short-range vs long-range motion: not a valid distinction" *Spatial Vision* **5** 303 – 309

Cavanagh P, 1992 "Attention based motion perception" *Science* **257** 1563 – 1565

Cavanagh P, Arguin M, Grünau M von, 1989 "Interattribute apparent motion" *Vision Research* **29** 1197 – 1204

Cavanagh P, Boeglin J, Favreau O, 1985 "Perception of motion in equiluminous kinematograms" *Perception* **14** 151 – 162

Cavanagh P, Mather G, 1989 "Motion: The long and short of it" *Spatial Vision* **4** 103 – 129

Chang J J, Julesz B, 1984 "Cooperative phenomena in apparent motion perception of random-dot cinematograms" *Vision Research* **24** 1781 – 1788

Chubb C, Sperling G, 1988 "Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception" *Journal of the Optical Society of America A* **5** 1986 – 2006

Chubb C, Sperling G (Eds), 1989 *Second-order Motion Perception: Space/Time Separable Mechanisms"* (Washington DC: IEEE Computer Society Press) pp 126 – 138

De Valois R L, Yund E W, Hepler N, 1982 "The orientation and direction selectivity of cells in macaque visual cortex" *Vision Research* **22** 531 – 544

Desimone R, Schein S J, Moran J, Ungerleider L G, 1985 "Contour, color and shape analysis beyond the striate cortex" *Vision Research* **25** 441 – 452

Dick M, Ullman S, Sagi D, 1987 "Parallel and serial processes in motion detection" *Science* **237** 400 – 402

Efron B, Tibshirani R J, 1993 *An Introduction to the Bootstrap* (New York: Chapman and Hall)

Gorea A, Lorenceau J, Bagot J D, Papathomas T V, 1992 "Sensitivity to colour- and to orientation-carried motion respectively improves and deteriorates under equiluminant background condition" *Spatial Vision* **6** 285 – 302

Gorea A, Papathomas T V, 1989 "Motion processing by chromatic and achromatic pathways" *Journal of Optical Society of America A* **6** 590 – 602

Gorea A, Papathomas T V, 1991 "Texture segregation by chromatic and achromatic visual pathways: An analogy with motion processing" *Journal of the Optical Society of America A* **8** 386 – 393

Gorea A, Papathomas T V, 1993 "Double opponency as a generalized concept in texture segregation illustrated with stimuli defined by color, luminance, and orientation" *Journal of the Optical Society of America A* **10** 1450 – 1462

Green M, 1986 "What determines correspondence strength in apparent motion?" *Vision Research* **26** 599 – 607

Green M, Odom J V, 1986 "Correspondence matching in apparent motion: Evidence for three-dimensional spatial representation" *Science* **233** 1427 – 1429

Green M, 1991 "Visual search, visual streams and visual architectures" *Perception & Psychophysics* **50** 388 – 403

Houck M R, Hoffman J E, 1986 "Conjunction of color and form without attention: Evidence from an orientation contingent color after-effect" *Journal of Experimental Psychology: Human Perception and Performance* **12** 186 – 199

Johnson N L, Kotz S, 1977 *Urn Models and Their Applications: An Approach to Modern Discrete Probability Theory* (New York: John Wiley)

Kolers P A, 1972 *Aspects of Motion Perception* (New York: Pergamon Press)

Korte A, 1915 "Kinematoskopische Untersuchungen" *Zeitschrift für Psychologie* **72** 194 – 296

McKee S P, Welch L, 1985 "Sequential recruitment in the discrimination of velocity" *Journal of the Optical Society of America A* **2** 243 – 251

McLeod P, Driver J, Crisp L, 1988 "Visual search for a conjunction of movement and form is parallel" *Nature (London)* **332** 154 – 155

Malonek D, Tootell R B, Grinvald A, 1994 "Optical imaging reveals the functional architecture of neurons processing shape and motion in owl monkey area MT" *Proceedings of the Royal Society of London, Series B* **258** 109 – 119

Mather G, West S, 1993 "Evidence for second-order motion detectors" *Vision Research* **33** 1109 – 1112

Maunsell J H R, Van Essen D C, 1983 "Functional properties of neurones in middle temporal visual area of the macaque monkey: I. Selectivity for stimulus directions, speed and orientation" *Journal of Neurophysiology* **49** 1127 – 1147

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

Nakayama K, Silverman G H, 1986 "Serial and parallel processing of visual feature conjunction" *Nature (London)* **320** 264 – 265

Navon D, 1983 "Preservation and change of hue, brightness, and form in apparent motion" *Bulletin of the Psychonomic Society* **21** 131 – 134

Ramachandran V, Ginsburg A, Anstis S, 1983 "Low spatial frequencies dominate apparent motion" *Perception* **12** 457 – 461

Reichardt W, 1961 "Autocorrelation as a principle of the evaluation of sensory information by the central nervous system", in *Sensory Communication* Ed. W Rosenblith (Boston, MA: MIT Press, and New York: John Wiley)

Sagi D, 1988 "The combination of spatial frequency and orientation is effortlessly perceived" *Perception & Psychophysics* **43** 601 – 603

Sagi D, 1990 "Detection of an orientation singularity in Gabor textures: Effect of signal density and spatial frequency" *Vision Research* **30** 1377 – 1388

Sagi D, Julesz B, 1987 "Short range limitations on detection of features differences" *Spatial Vision* **2** 39 – 49

Simpson W A, 1990 "The use of different features by the matching process in short-range motion" *Vision Research* **30** 1421 – 1428

Snowden R J, Braddick O J, 1989a "The combination of motion signals over time" *Vision Research* **29** 1621 – 1630

Snowden R J, Braddick O J, 1989b "Extension of displacement limits in multiple-exposure sequence of apparent motion" *Vision Research* **29** 1777 – 1787

Snowden R J, Braddick O J, 1991 "The temporal integration and resolution of velocity signals" *Vision Research* **31** 907 – 914

Tanaka K, 1993 "Neural mechanisms of object recognition" *Science* **262** 685 – 688

Thorell L G, De Valois R, Albrecht D G, 1984 "Spatial mapping of monkey V1 cells with pure colour and luminance stimuli" *Vision Research* **24** 751 – 769

Toet A, Levi D M, 1992 "The two-dimensional shape of spatial interactions zones in the parafovea" *Vision Research* **32** 1349 – 1357

Treisman A, 1986 "Features and objects in visual processing" *Scientific American* **255** 114 – 125

Treisman A, Gelade G, 1980 "A feature integration theory of attention" *Cognitive Psychology* **12** 97 – 136

Treisman A, Sato S, 1990 "Conjunction search revised" *Journal of Experimental Psychology: Human Perception and Performance* **16** 459 – 478

Ullman A, 1979 *The Interpretation of Visual Motion* (Cambridge, MA: MIT Press)

Van Santen J P H, Sperling G, 1985 "Elaborated Reichardt detectors" *Journal of the Optical Society of America A* **2** 300 – 321

Victor J D, Conte M M, 1990 "Motion mechanisms have only limited access to form information" *Vision Research* **30** 289 – 301

Watson A B, 1986 "Apparent motion occurs only between similar spatial frequencies" *Vision Research* **26** 1727 – 1730

Werkhoven P, Sperling G, Chubb C, 1993 "The dimensionality of texture-defined motion: a single channel theory *Vision Research* **33** 463 – 485

Wertheimer M, 1912 "Experimentelle Studien über das Sehen von Bewegung" *Zeitschrift für Psychologie und Physiologie der Sinnesorgane* **61** 161 – 265 [see also Sekuler R, 1996 "Motion perception: A modern view of Wertheimer's 1912 monograph" *Perception* **250** 1243 – 1258]

Wilson H R, Ferrera V P, Yo C, 1992 "A psychophysically motivated model for two-dimensional motion perception" *Visual Neuroscience* **9** 79 – 97

Wilson H R, Mast R, 1993 "Illusory motion of texture boundaries" *Vision Research* **33** 1437 – 1446

Wolfe J M, Cave K R, Franzel S L, 1989 "Guided search: An alternative to the feature integration model for visual search" *Journal of Experimental Psychology: Human Perception and Performance* **15** 419 – 433