Human Sensitivity to Expanding and Rotating Motion: Effects of Complementary Masking and Directional Structure

TOM C. A. FREEMAN,* MIKE G. HARRIS*

Received 30 November 1989; in revised form 17 April 1991

A comparison of sensitivity to expanding, rotating, translating and random motion suggests the existence of specialised mechanism for the detection of expansion and rotation. Complementary masking shows that the detection of expansion is unaffected by the presence of rotation, and vice versa. These results are interpreted in terms of a Relative Motion System, which combines the outputs of localised motion detectors in a variety of ways, and which functions to analyse complex image motion into simpler, more useful, components.

INTRODUCTION

As an observer moves about the world, the retinal image changes smoothly over time. This smooth change, commonly termed retinal flow contains a great deal of information about the observer's movement and about the 3-dimensional layout of the world (e.g. Clocksin, 1980; Cutting, 1986; Gibson, 1950, 1979; Koenderink, 1985, 1986; Koenderink & van Doorn, 1975, 1976, 1981; Longuet-Higgins & Prazdny, 1980; Prazdny, 1983; Rieger & Lawton, 1985; Waxman & Wohrn, 1988).

How might this information be extracted and used by the human visual system? The informative aspects of retinal flow lie not in the way that individual points change their positions in the image, but rather in how they change their relationships to each other. For example, if the observer looks more or less where he or she is going, the central region of the image expands and the distance between individual features gradually increases. A system capable of extracting useful information from retinal flow ought to contain mechanisms which are sensitive to such changing relationships, so we might expect the visual system to contain (at least) two stages of motion analysis. The first, which we term the Classical Motion System (CMS), should be capable of signalling the change in position (i.e. speed and direction) of individual image features. The second, which we term the Relative Motion System (RMS), should combine CMS outputs to produce mechanisms sensitive to the relative motion of groups, or constellations, of features.

What types of relative motion should the RMS encode? Relative image motion can be broken down into just three components—expansion, rotation and deformation (e.g. Koenderink & van Doorn, 1976). The first two of these are self-explanatory, the last is a form of shear, involving an expansion along one axis and a complementary compression along the axis at right angles to it, which results in a change in shape without a change in area. If the RMS contained mechanisms that were sensitive to just these three types of relative motion, then it would be able conveniently to describe the useful aspects of any retinal flow pattern. Of course, to be useful, these mechanisms would have to function independently of each other so that, for example, the expansion-sensitive mechanism should be completely insensitive to rotation, and vice versa.

The experiments reported here investigated whether the visual system contains RMS mechanisms that are selectively sensitive to different types of relative motion, and whether these mechanisms function independently of each other. For simplicity, we considered only rotation and expansion and not deformation because there is already considerable psychophysical and neurophysiological interest in separate expansion- and rotation-sensitive mechanisms. We aimed to provide further psychophysical evidence for their existence and to extend the previous work by showing that they are involved in the analysis of complex flow patterns.

In previous psychophysical studies of image expansion and rotation, Regan and Beverley have shown that: (a) adaptation to a rotating stimulus decreases sensitivity to rotation more than does adaptation to a stimulus containing the same amount of linear motion but no rotation (Regan & Beverley, 1985); (b) adaptation to an expanding square decreases sensitivity to expansion...
more than sensitivity to linear motion along the same axes (Regan & Beverley, 1978b); (c) adaptation to a square whose size expands with a ramp waveform produces an after-effect with a time-course different from that found with classical motion displays (Regan & Beverley, 1978a). These results suggest the existence of separate RMS mechanisms, selectively sensitive to expansion or rotation and quite distinct from the simple mechanisms of the CMS. Similar RMS mechanisms have been proposed to explain the spiral after-effect (Cavanagh & Favreau, 1980; Hershenson, 1984, 1987).

There is also good neurophysiological evidence that the visual system contains mechanisms selective for expanding or rotating retinal flow. Regan and Cynader (1979) described units in area 18 of the cat which were sensitive to small-field expansions but the most relevant studies have concentrated upon later stages of the cortical pathway. Directionally-selective units (D-cells) have frequently been encountered in MT and in MST, to which the pathways of RMS mechanisms that are specifically sensitive to expansion or rotation, these two components are detected independently of each other.

**EXPERIMENT 1**

Experiment 1 used random dot flow patterns to compare sensitivity to expanding, rotating and "destructured" motion. Destructured stimuli contained the same distribution of speeds across the retina as their structured counterparts, and differed only in their overall directional organisation. If the visual system contains RMS mechanisms that are specifically sensitive to expansion or rotation, then these types of structured motion should be more easily detected than destructured motion.

**Method**

**Subjects.** One of the authors and one other experienced psychophysical observer, who did not know the purpose of this experiment, acted as subjects.

**Stimuli.** Stimuli were generated by DEC LSI 11/23 computer and displayed upon a Hewlett Packard 7161 oscilloscope (P31 phosphor) through a CED 502 laboratory interface at a frame rate of 77 Hz. All stimuli consisted of 128 dots in a random annular pattern with an inner radius of 0.6 deg and an outer radius of 2.5 deg. A stationery dot at the centre of the annulus served as a fixation point. Individual dots subtended approx. 1.5 min of arc, and were bright (120 cd/m²) on a dark (15 cd/m²) background. Viewing was binocular from a distance of 114 cm. Viewing distance was maintained by a backrest and was frequently checked.

In the expanding stimulus, each dot moved radially away from the central fixation point. In the rotating stimulus, each dot moved around a circular path centred upon the fixation point. In both stimuli, the speed of each dot was proportional to its radial distance from the fixation point, so that the resulting pattern of flow was equivalent to that produced by the vertical surface either rotating or changing its distance from the viewer. The destructured equivalents of these stimuli were produced from them by rotating the trajectory of each dot through a different random angle. In effect the two types of destructured stimulus were identical except that individual dot trajectories were curved in the rotating condition and straight in the expanding condition.

In order to minimise any motion after-effects, the flow was modulated sinusoidally over time so that it completed a single cycle during the display period. Thus, in the expanding stimulus, each dot moved first outward and then inward before returning outward to its original position. In the rotating stimulus, each dot moved clockwise and then anti-clockwise before returning clockwise to its original position. The amplitude of the temporal modulation determined the amount of expanding or rotating motion.

Using 12-bit DACS, the minimum possible motion between frames of any dot was a little over 2 sec of arc. Even with this fairly high resolution, at the very low rates of movement sometimes required in these experiments, dot movement sometimes became probabilistic: not all the dots moved on every frame. Nonetheless, careful observation at short viewing distances confirmed that all the stimuli appeared to move smoothly and coherently, suggesting that the mechanisms with which we are concerned are insensitive to small, local, instantaneous distortions of the overall pattern.

The stimuli included one further necessary refinement. Structured expansion is accompanied by a change in the overall size of the stimulus, whereas destructured expansion is not. This cue can be eliminated simply by fixing the overall size of the annular display window, so that dots moving outside its boundaries are deleted. In order to prevent the instantaneous appearance and disappearance of individual dots, which are perceptually salient
events and which differ in number between structured and destructured conditions, the edges of the window were made gradual, rather than abrupt, by decreasing luminance linearly to 0 over a 0.25 deg margin. Dots approaching the boundaries then increased or decreased smoothly in luminance over the course of several frames. Experiment 1 investigated the importance of the overall size change of expanding stimuli by including both “size change” and “no size change” conditions.

Procedure. Each trial consisted of two 650 msec stimulus presentations separated by 500 msec and each accompanied by a high-pitched auditory tone. One presentation contained the appropriate moving stimulus and the other contained a stationary random dot pattern. The subject’s task was to indicate by a button-press which interval contained the moving stimulus. This response initiated the next trial after a delay of 1 sec. Correct responses were signalled by a brief low-pitched tone.

Each session investigated one of six stimulus conditions (structured or destructured rotation; structured or destructured expansion, with or without size change), and consisted of 20 practice followed by 100 experimental trials. The experimental trials consisted of 20 replicates of each of 5 motion amplitudes, in random order and with the moving stimulus equally likely to appear in the first or second interval. Each subject undertook at least 10 sessions for each of the 6 stimulus conditions, with the order of conditions randomised across sessions.

Analysis. GLIM was used to fit a psychometric function (specifically, the logit approximation to the Normal ogive) to the data from each individual session. A session was repeated if the chi-squared value for the fit was significant at the 5% level. Less than 5% of sessions needed repeating. The fitted curve for each session was used to derive the Maximum Likelihood Estimate of the target amplitude yielding 75% correct detection. This point was used as the estimate of detection threshold.

Results

Figure 1 shows the mean log detection thresholds obtained for each of the 6 motion conditions. For both subjects, structured expansion was detected more easily than destructured expansion [Fig 1(a): $F(1,14) = 688.86$, $P < 0.01$, TCAF; $F(1,9) = 110.68$, $P < 0.01$, GHW]. Similarly, structured rotation was detected more easily than destructured rotation [Fig. 1(b): $F(1,9) = 376.41$, $P < 0.01$, TCAF; $F(1,9) = 62.92$, $P < 0.01$, GHW]. Size change had no significant overall effect upon the detection of expansion [$F(1,14) = 1.52$, TCAF; $F(1,9) < 1$, GHW], and there was no significant interaction between size change and structure type [$F(1,14) = 4.39$, TCAF; $F(1,9) < 1$, GHW].

Conclusions

The significant effects of structure provide support for the existence of RMS mechanisms that combine the motions of several individual dots. The finding that overall size change has no effect upon the detection of expanding stimuli is important in demonstrating that for the expanding stimulus, as well as the rotating stimulus, enhanced sensitivity depends upon the internal structure of the stimulus, rather than upon grosser features such as overall size.

We cannot yet conclude, however, that the pooling is specifically selective for either expansion or rotation. Within each small region of the structured stimuli, the dots all move in roughly the same direction, whereas in the destructured stimuli they do not. Any mechanism which pools motion in a single direction over a moderate retinal area could account for the enhanced sensitivity demonstrated in Expt 1. Experiment 2 demonstrates that unidirectional pooling is not, in fact, an adequate explanation of enhanced sensitivity to expansion and rotation.

EXPERIMENT 2

Experiment 2 compared sensitivities to simplified expanding, rotating, translating and destructured stimuli. In the translating condition, all the dots moved in the same direction. If enhanced sensitivity to structured motion is due to simple unidirectional pooling, then sensitivity to this stimulus should be at least as high as that to expanding or rotating stimuli.

Method

Stimuli. The display was that used in the previous experiment, consisting of an annular pattern of 128 randomly-placed dots. Dot luminance at the inner and outer edges of the annulus was smoothly windowed, as in Expt 1, to prevent overall size and position change. As before, dot motion was sinusoidally modulated over time. For the translating stimulus all the dots moved horizontally at the same speed. The other three stimulus types were derived from this basic pattern simply by rotating the trajectories of individual dots so that in the expanding condition all the dots moved radially, in the rotating condition they all moved tangentially, and in
the destructured condition they all moved in different random directions.

The probabilistic nature of dot movement at low speeds might disadvantage translating motion because it only has a horizontal component, whereas the other conditions have both horizontal and vertical components. To counter this, translating motion was actually generated in a diagonal direction and the screen was rotated to make the display horizontal.

Procedure. As in Expt 1, each trial consisted of a moving stimulus in one presentation and a stationary pattern in the other. The subject's task was to indicate which interval contained the moving stimulus. The two authors each undertook 5 replications of each of the 4 motion conditions.

Results

Figure 2 shows the mean detection thresholds in each motion condition. Contrary to the prediction of unidirectional pooling, sensitivity to translating motion is well below sensitivity to either expansion or rotation. Perhaps more surprisingly, sensitivity to translation is no better than sensitivity to destructured motion. The Tukey Honestly Significant Difference (HSD) test confirms that all differences are significant except those between destructured motion and translation (HSD = 0.203, mean difference (MD) = 0.087, TCAF; HSD = 0.141, MD = 0.042, MGH), and between expansion and rotation (HSD = 0.203, MD = 0.112, TCAF; HSD = 0.141 MD = 0.015, MGH).

Conclusions

These results cannot be explained by mechanisms which simply pool unidirectional motion across retinal space. They suggest instead RMS mechanisms which are selective for the particular relationships between individual dots that characterise rotating or expanding stimuli. The important factor is presumably the systematic arrangement of motion directions in the stimulus. Enhanced sensitivity does not seem to require the smooth variations in speed or, for rotating patterns, the smoothly curved trajectories which would normally be encountered in the natural world.

EXPERIMENT 3

To test the notion that image expansion and rotation are analysed independently, Expt 3 measured sensitivity to expansion in the presence of rotating masks, and to rotation in the presence of expanding masks.

Method

Subjects. One of the authors (TCAF) took part in the main experiment. Twenty volunteers (11 male, 9 female) took part in a smaller-scale experiment to confirm the main results. None of the subjects, except TCAF, knew the aim of the work. All had normal, or corrected-to-normal vision, and worked, or were students, in the School of Psychology, Birmingham University.

Stimuli. The display was the same as that used in the previous experiments, consisting of an annular pattern of 128 randomly-positioned dots. In the main "orthogonal" masking condition, two components of motion were assigned to each dot on each frame: a radial ("expansion") component, directed along the line connecting the fixation point to the dot's current position, and a tangential ("rotation") component at right angles to this. As in Expt 1, the magnitudes of both components were proportional to radial distance, so that the resulting pattern of image motion resembled that produced by movement relative to a vertical surface. The flow was sinusoidally modulated over time and the amplitudes of the two temporal sinusoids could be independently varied to determine the amounts of expanding and rotating motion.

We also included a "non-orthogonal" masking condition to check that the masking levels used were effective stimuli. For these stimuli, both components of motion were of the same type (i.e. both expanding or both rotating).

Procedure. The procedure was as before except that, on each trial, one presentation contained the mask alone (e.g. rotation) and the other contained target plus mask (e.g. expansion plus rotation in the orthogonal condition, or rotation plus rotation in the non-orthogonal condition). The subjects task was to indicate by a button-press which presentation contained the target. Subject TCAF undertook 6 replications at each of 5 mask amplitudes in the orthogonal condition, and 5 replications at the same mask amplitudes in the non-orthogonal condition. Ten paid volunteers undertook 1 replication at each of 3 mask amplitudes in the orthogonal condition, while ten others undertook 1 replication at each of 5 mask amplitudes in the non-orthogonal condition. The mask amplitudes used were 0 (static), 1, 2, 4, and 8 times the typical thresholds of experienced observers, determined in a preliminary experiment. They were selected in a different random order for each subject. All subjects undertook four practice sessions before any data were collected.

Results

The mean effect of mask amplitude upon target detection threshold is shown for subject TCAF in Fig 3.
The data clearly indicate that the detection of expansion or rotation is unaffected by the presence of an orthogonal mask (Fig. 3: solid symbols). This interpretation is supported by 2 separate 1-factor analyses of variance on the data for expanding and rotating targets. In neither case did mask amplitude have a significant effect upon target detectability [expanding target: $F(4,20) = 1.21$; rotating target: $F(4,20) < 1$. The lack of masking effect cannot be attributed to any general ineffectiveness of the mask, since the same values produced significant effects upon non-orthogonal targets [Fig. 3: open symbols; expanding target: $F(4,16) = 7.07$, $P < 0.005$; rotating target: $F(4,16) = 25.53$, $P < 0.001$].

Mean detection thresholds for the volunteers are shown in Fig. 4. Again, orthogonal masks had no significant effect upon target detection threshold [expanding target: $F(2,18) = 1.85$; rotating target: $F(2,18) < 1$], while non-orthogonal masks did have a significant effect [expanding target: $F(4,36) = 11.17$, $P < 0.001$; rotating target: $F(4,36) = 3.86$, $P < 0.025$].

**Conclusions**

Experiment 3 demonstrates that rotation and expansion are detected independently of each other: the detection of rotation is unaffected by the presence of an expanding mask, and the detection of expansion is unaffected by the presence of a rotating mask. It is tempting to suggest that the two types of motion are detected by separate mechanisms, and that these mechanisms form part of the proposed Relative Motion System. It remains possible, of course, that the necessary motion discrimination is performed entirely at a local level, and reflects the functioning of the Classical Motion System alone. After all, rotation and expansion are orthogonal to each other, so the two motion components of each dot in a target-plus-mask stimulus should be detected by different local (CMS) mechanisms tuned to different directions of motion. However, in view of the results of Expts 1 and 2, and those of other studies (e.g. Regan & Beverley, 1978a,b, 1985) it seems more likely that the two components of motion are detected independently by specialised RMS mechanisms sensitive to the overall directional organisation of the stimulus.

The results of Expt 3 are rather different from those of Simpson (1988), who found that the ability to discriminate between the times-to-collision of two expanding targets was reduced by the addition of rotation, and improved after adaptation to rotation. It therefore remains possible that the ability to detect expansion and rotation independently is not completely preserved or made use of in more complex tasks.

**DISCUSSION**

Our results are difficult to explain in terms only of localised motion detectors (the Classical Motion System, or CMS). Taken together, they provide further support for a Relative Motion System (RMS) that combines the output of localised motion detectors to produce separate mechanisms capable of analysing complex motion fields into their important parts. Firstly, simple, coherent flow patterns are detected by specialised mechanisms which seem specifically sensitive to the relative directions of individual display elements, rather than to grosser features of the stimulus such as the distribution of speeds across the retina or, in the case of expansion, overall change in size (Expt 1 and 2). Secondly, these mechanisms may function to analyse complex flow patterns into their simpler directional components (Expt 3).
A number of important questions about the detailed properties of the putative RMS mechanisms are yet to be answered before their role can be properly understood. Firstly, the results of Expt 2 suggest that although expansion and rotation mechanisms are sensitive to directional relationships, they do not require spatial speed gradients. This finding fits well with the report that S and R cells in MST respond just as well to expansions or rotations from which all speed gradients have been removed (Tanaka et al., 1989). The flow patterns produced by movements of the observer about the world typically contain smooth gradients of speed both along and at right angles to the direction of flow. Whilst the direction of flow is determined by the observer’s locomotory heading and by rotations of the eyes and head, the speed gradients are determined by the 3-dimensional layout of the world. Thus, mechanisms that are sensitive to directional relationships but not speed gradients might be adequate for signalling locomotory heading (Warren & Hannon, 1988) and for sensing eye and head rotations (Simpson, Graf & Leonard, 1981), but they would be quite inadequate for encoding useful information about surface layout. It is particularly interesting that studies of perceived shape from motion (e.g. Braunstein, 1968; Braunstein & Andersen, 1984; Braunstein & Tittle, 1988; Rogers & Graham, 1979) and more recent work on the perception of surface rigidity (De Bruyn & Orban, 1990) suggest that the human visual system can indeed make use of speed gradients in random dot flow patterns.

A second, closely related, question concerns the receptive field size of expansion and rotation mechanisms. Although our findings show that these mechanisms are sensitive to the internal directional pattern of the flow rather than to coarser features such as overall change in size, they do not bear directly upon receptive field size. On the other hand, the neurophysiological data suggest that S- and R- cells have very large receptive fields. Again, such mechanisms might be adequate for sensing locomotory heading and eye/head rotations, since these types of movement affect the whole retinal flow field. But they would be less suitable for signalling surface layout because flow patterns typically consist of a number of distinct local regions, one for each of the surfaces in the imaged world. Although the directional pattern of motion is coherent throughout the whole flow field, the pattern of speed gradients varies from region to region and, within each region, carries the necessary information about the layout of the corresponding surface.

In the limit, if the receptive fields of expansion and rotation mechanisms were very small indeed, then they would provide a reasonable approximation to the differential operators \( \text{div} \) and \( \text{curl} \) which, roughly speaking, respectively measure the speed gradients along and at right angles to the direction of flow. This might be very useful because the measurement of \( \text{div} \) and \( \text{curl} \) over the whole flow field would provide a convenient way to segment the flow field and to recover information about surface layout (e.g. Koenderink & van Doorn, 1975, 1976; Koenderink, 1985; Harris & Freeman, 1990; see also Nakayama & Loomis, 1974; Frost & Nakayama, 1983). Mechanisms with moderate receptive field sizes, though providing worse approximations to \( \text{div} \) and \( \text{curl} \), might still be adequate in a world consisting of relatively few, fairly flat surfaces, whilst mechanisms with receptive fields covering the whole retinal field could provide little or no useful information about surface layout.

Although there is increasing support for the existence of a Relative Motion System within human vision containing, at least, mechanisms sensitive to expansion or rotation, it is not yet possible to distinguish between two very different functional roles. If the proposed mechanisms have large receptive fields and are insensitive to speed gradients, then their role may be confined to signalling locomotory heading and, possibly, to decomposing complex flow fields into the relatively uninteresting components produced by rotations of the eyes and head, and the more interesting components produced by translations of the observer. This latter aspect would be equivalent to recovering optic flow from retinal flow. On the other hand, if the mechanisms have small receptive fields and can signal information about speed gradients along and at right angles to the direction of flow, then they may be directly concerned with perception of 3-dimensional surface layout. Further psychophysical and neurophysiological work is urgently needed to elaborate the properties of RMS mechanisms, and to establish their relationship to such mathematical descriptors as \( \text{div} \) and \( \text{curl} \).

**REFERENCES**


Acknowledgements—This work was supported by a SERC Research Studentship to TCAF. We should like to thank the anonymous referee who provided helpful comments on an earlier version of this paper.