Perception, 1995, volume 24, pages 1373-1382

Second-order texture contrast resolves ambiguous apparent motion

George Mather
Department of Experimental Psychology, University of Sussex, Brighton BN1 9QG, UK
Stuart Anstis
Department of Psychology, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0109, USA
Received 30 May 1995, in revised form 14 September 1995

Abstract. When a black and a white square on a grey surround exchange places, it was previously shown that on a dark surround it is the white square, and on a light surround it is the black square, that is seen in apparent motion (AM). Thus the higher-contrast square carries the AM. We now show that the same is true for second-order AM of texture-defined squares. Squares were defined by four different textures: by anisotropy (horizontal versus vertical random dashes), by alphanumeric letters, by hash marks, or by dot size. The result was that the square that differed more from the surround in texture properties carried the second-order AM. Judgments of texture salience revealed a high correlation between salience and apparent motion. In a third experiment, crossover AM between dissimilar textures was investigated, and it was found that the more salient textures carried the AM. Results cannot be explained by the concept of 'texture activity', but instead indicate that the system extracts a measure of 'texture contrast' prior to analysis of salience and apparent motion.

1 Introduction

If a black square on a grey surround suddenly disappears and an identical square appears a short distance away, one usually perceives apparent motion (AM) from one square to the other, rather than two unconnected perceptual events. The visual system detects correlated changes in image intensity over space and time because they signify image motion. An AM display mimics this spatiotemporal coincidence and therefore generates a response in motion-sensitive neurons. Previously we (Anstis and Mather 1985) considered the following questions. Imagine a black square and a white square side by side on a grey surround. Suddenly the black square becomes white and the white square becomes black. What does one perceive? Does one see two stationary squares, each changing simultaneously in brightness? Or does the white square appear to jump to the left, or the black square to the right? We found that the answer depends upon the luminance of the grey surround. On a light surround the dark square appears to jump, but on a dark surround the light square appears to jump. In other words the AM is assigned to the square that differs more from the surround. In our view, there is a motion signal in each direction, but the square with the higher contrast provides a stronger motion signal which prevails. Crossover motion can be compared to a counterphase flickering grating, which is mathematically identical to two superimposed gratings of the same spatial frequency moving in opposite directions. When the two opposed gratings have the same contrast then no net motion is seen. However, if the contrast of (say) the leftward moving grating is increased, then net motion to the left is perceived. In a similar way if the contrast of the square that is about to jump to the left is higher than that of the other square, then net apparent motion to the left is perceived.

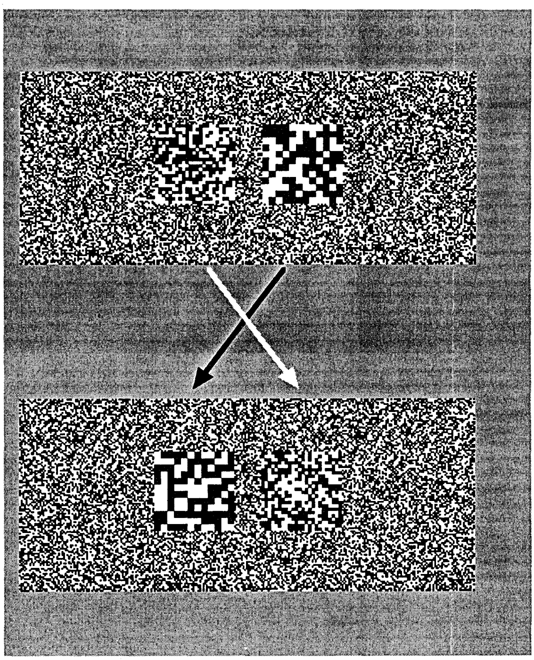
Our 1985 squares were defined by luminance, and therefore offered a signal to so-called short-range (Braddick 1974) or first-order (Cavanagh and Mather 1989) motion detectors whose response depends on Fourier energy (Adelson and Bergen 1985).

However, a 'second-order' square can be defined by texture, being plainly visible even if it has the same space-averaged luminance as the surround, provided that it differs in some other visual property such as colour, depth, or texture (Anstis 1989; Regan 1991). Effective AM displays can be constructed by using second-order squares, yet they offer no signal to Fourier-based detectors. A separate motion process has been invoked to account for AM in second-order displays (eg Chubb and Sperling 1988). Crossover displays can be used as a sensitive tool to examine the relative strengths of competing motion signals, as our experiments showed (Anstis and Mather 1985), so in this paper we shall use these displays to test whether second-order motion processes behave in the same lawful way as first-order processes.

2 Second-order crossover displays

In a second-order analogue of the earlier luminance-based crossover display, two textured squares are shown side-by-side against a textured background. The two squares suddenly exchange places. Which, if either, appears to move? An example is shown in figure 1. The background is made up of single-pixel random black-white texture. In the first frame, the left-hand square contains random texture made up of 2×2 pixel elements and the right-hand square contains texture made up of 4×4 pixel elements. In the second frame, the two squares exchange places (all textures are also rerandomised). Extrapolating from the earlier results, we reasoned that the square that differs most from the background in textural properties should appear to move. In the example of figure 1, we expected that the coarse texture in the right-hand square in the first frame would dominate AM when seen against a fine-texture background. To test the generality of such an effect, in experiment 1 we assessed four second-order texture properties for their ability to resolve ambiguities in crossover displays.

Frame 1



Frame 2

Figure 1. A second-order crossover display. In frame 1, two textured squares are displayed side by side, against a textured background. In frame 2, the squares exchange places (and all textures are rerandomised). Which square is seen to move in apparent motion, the coarse-textured square (black arrow) or the fine-textured square (white arrow)?

3 Experiment 1
3.1 *Method*
3.1.1 *Subjects.* Five subjects participated. All were experienced observers in motion experiments, but unaware of the purpose of the experiment.

3.1.2 *Apparatus.* A PC-compatible computer equipped with a high-resolution graphics board generated motion displays on a Hitachi 14MVX colour monitor (P22 phosphor). Responses were recorded by means of a standard two-button mouse.

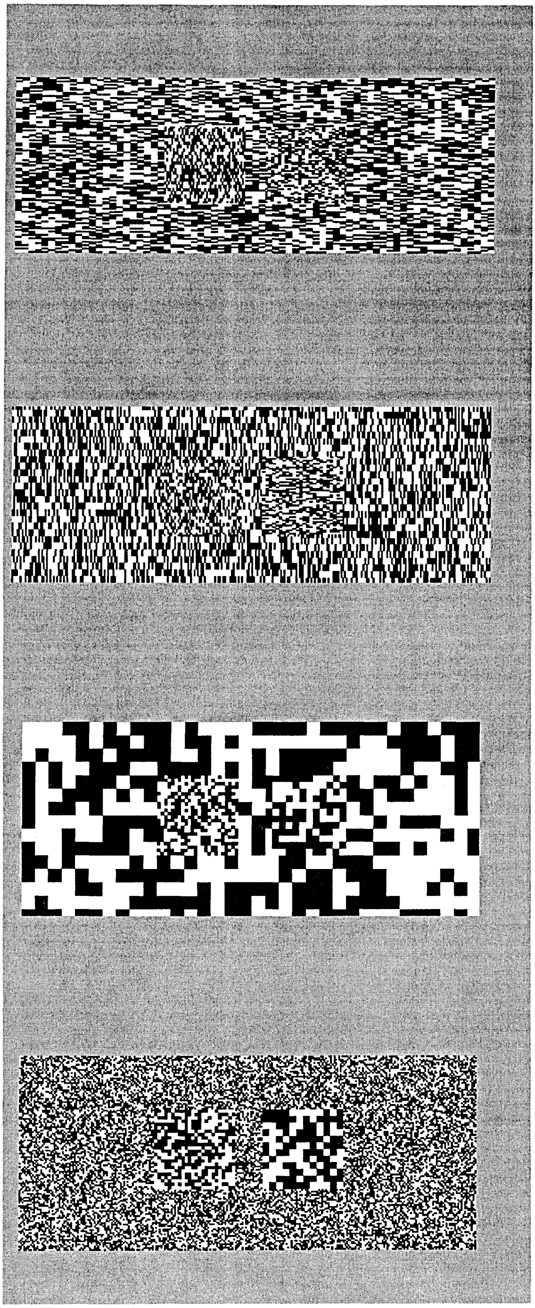
3.1.3 *Stimuli.* Eight two-frame AM displays were created, containing shapes defined by four different textural properties: anisotropy, size, letter shape, and 'hash' shape. Printouts of examples of actual experimental stimuli are shown in figure 2.

3.1.3.1 *Orientational anisotropy.* The base texture contained 50% black 50% white single-pixel random dots. This pattern was stretched to turn each dot into a dash and thereby generate four textures, either with short dashes (2×1 pixels) or with long dashes (4×1 pixels), and with the long axes of the dashes either horizontal or vertical. The two test patches in each AM display were composed of short dashes, one patch consisting of horizontal, the other of vertical short dashes. In one display the background was composed of long horizontal dashes, and in the other it was composed of long vertical dashes (see examples in figure 2). In both, the two test patches exchanged places to create a crossover display. Although the test patches were defined in the same way in both displays, we predicted that AM of the patch containing short horizontal dashes would predominate when the background contained long vertical dashes, and AM of the other patch (containing short vertical dashes) would predominate when the background contained long horizontal dashes.

3.1.3.2 *Dot size.* Four random-dot patterns were created, containing 50% black 50% white elements at different sizes. Dots were either 1×1, 2×2, 4×4, or 8×8 pixels square. Two AM displays were created from these textures. In both, one test patch contained 2×2 dots and the other patch contained 4×4 dots. In one display, the background contained 1×1 dots, and in the other display the background contained 8×8 dots (see examples in figures 1 and 2). We predicted that in the former, AM would be dominated by the 4×4 patch, and in the latter AM would be dominated by the 2×2 patch.

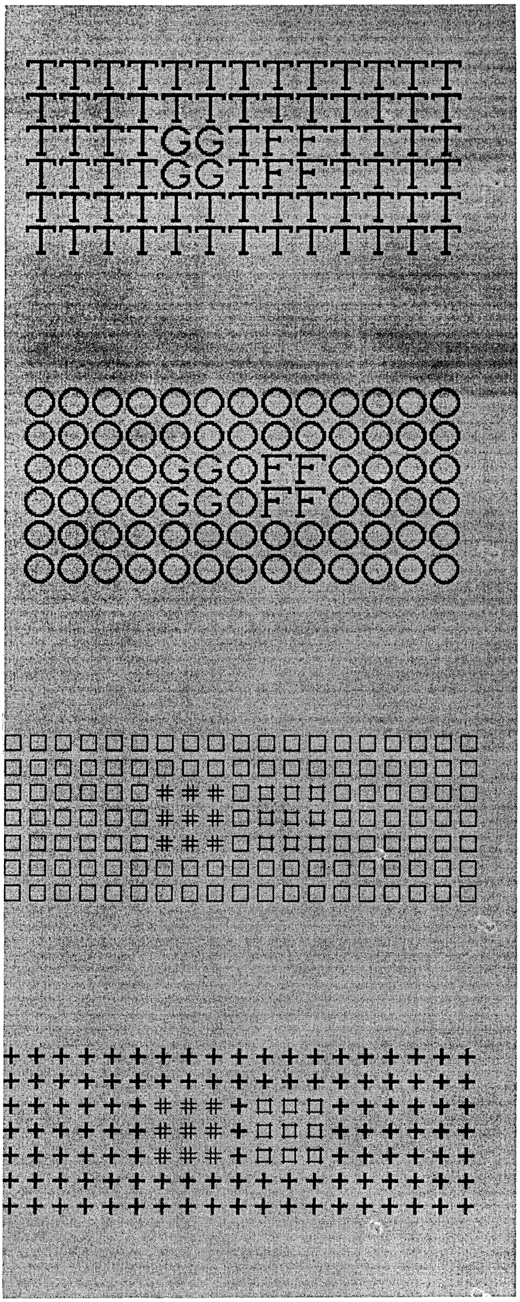
3.1.3.3 *Shape of letters.* Textures were made out of angular and rounded letters. The two patches in two AM displays were composed of capital F and G, which are, respectively, angular and rounded. These exchanged places on a surround texture which was made out of angular Ts in one display, and rounded Os in the other (see figure 2). Neisser (1963) showed that angular letters were more difficult to find than rounded letters in a context of other angular letters (and vice versa in a context of rounded letters). We reasoned that target letters which are easy to find against a set of distractor letters ought to make textures that are highly salient against background texture made from the same distractor letters. So we predicted that Gs would carry AM against a background of Ts, but Fs would dominate against a background of Os.

3.1.3.4 *Shape of hash marks.* We examined the properties of letter shape that underlay the motion competition by using an array of abstract elements. Each element contained two vertical and two horizontal line segments, which could be arranged to generate four different shapes. At one extreme, the four line segments formed a cross, and at the other extreme they formed a square. The two intermediate forms involved 'hash' shapes. Each hash shape contained two verticals, two horizontals, four intersections, and eight terminators. Two AM displays were constructed by using these four shapes. In one, the background contained a repetitive array of crosses, and in the other it contained a repetitive array of squares. The two test patches used in



Motion	Salience
88%	88%
21%	17.5%
90%	97.5%
14%	5.5%

Figure 2. Examples of stimulus frames used in all experiments. Each stimulus contains two textured squares against a textured background. In experiment 1, each frame was followed by a second frame in which the squares exchanged places (see figure 1) to create a crossover motion display. The column labelled 'motion' shows the mean percentage of trials in which observers reported rightward motion when the depicted frame was shown first, indicating that the left-hand square carried apparent motion (frame order was randomised during the experiment). In experiment 2, each frame was shown briefly by itself as a static display, and observers



Motion	Salience
82%	95.5%
3%	0%
89%	97.5%
53%	20.5%

Figure 2 (continued)
reported whether the left-hand or the right-hand square in each appeared the more salient (each square was actually shown an equal number of times on the left-hand and right-hand sides of the display). The column labelled 'salience' shows the mean percentage of trials in which the left-hand square was judged more salient. In experiment 3 each frame was again followed by a second frame to create a crossover display, but the two frames contained different textures (see figure 4).

both contained hash shapes. One patch contained an array of cross-like hashes and the other contained an array of square-like hashes (see figure 2). When the patches exchanged positions, we predicted that the square-like shapes should carry AM against a background of crosses, but the cross-like shapes should dominate against a background of squares.

All AM displays shared the same dimensions. The background was 20 deg wide and 8.5 deg tall. Each patch was 3.5 deg, and the distance between the nearest edges of the patches was 1.5 deg. Frame duration was fixed at 350 ms, with no interframe interval. As can be seen by inspection of figure 2, only three different intensities were used in all displays. They were 5.7, 47, and 116 cm m^{-2}.

3.1.4 *Procedure.* Subjects performed in a single experimental session, involving twenty presentations of each of the eight AM displays. Each presentation involved one two-frame exposure of an AM display, after which the subject pressed one of two mouse buttons to signify perceived direction. An interval of 1 s separated successive presentations, during which a uniform grey field was shown, containing a central red fixation cross. Order of presentation was randomised, and for each AM display the order of presentation of the two frames varied randomly.

3.2 *Results and discussion*

Results for each stimulus are presented in figure 2 adjacent to each example stimulus frame, in the column labelled 'motion'. Data are expressed in terms of mean percentage of trials in which rightward motion was reported when the frame depicted was shown first in the AM sequence. So high percentage values indicate that the left-hand patch in the examples dominated AM, and low values favour the right-hand patch. It is clear that in seven of the eight displays AM was seen in the patch that differed most from the background in terms of its predicted textural properties (the exception was one 'hash' display, which yielded ambiguous direction reports). For example, in the case of element size (shown in figure 1), the square that differed most in terms of element size relative to the background dominated AM reports.

We conclude that in second-order, as in first-order, crossover motion there is a competition between two opposed motion signals. The square that differs more from the surround, either in luminance or in texture, offers a stronger motion signal that determines the direction of the perceived AM. Perhaps the strength of motion signals generated by each square is related to the salience of the square against its background. In experiment 2, we tested this idea by obtaining judgments of texture salience in stationary texture-defined squares corresponding to those used in experiment 1. If crossover motion is related to texture salience, then there should be close agreement between the AM data of experiment 1 and salience judgments in experiment 2.

4 Experiment 2

4.1 *Method*

4.1.1 *Subjects.* Five observers participated. All also took part in experiment 1.

4.1.2 *Apparatus and stimuli.* The same equipment was used as in experiment 1. Sixteen different stationary displays were used, corresponding to the component frames of the eight AM displays used in experiment 1 (so eight pairs of frames were identical except for exchanging the content of the two stimulus patches). Each frame was presented for a fixed duration of 350 ms.

4.1.3 *Procedure.* Subjects performed in a single experimental session involving ten presentations of each of the sixteen frames, in random order. Each presentation involved a single exposure to one of the static frames, after which the subject was instructed to press one of two mouse buttons to signify whether the static square on the left or on the right side of the display appeared more salient. An interval of 1 s separated successive trials, as in experiment 1. To avoid effects due to the order in which subjects performed in all three experiments, session order varied between subjects.

4.2 *Results and discussion*

Means are shown in figure 2. Results are expressed in terms of reports that the left-hand square was more salient when the frame depicted was shown, to allow comparison with AM results from experiment 1. It is clear that there is indeed close agreement between salience responses and AM reports. This correspondence is shown graphically in the scatter plot in figure 3. Mean AM reports are plotted against mean salience responses for each of the eight stimuli. There is a high correlation between the results of the two experiments ($r = 0.96$). For example, in the case of element size (depicted in figure 1), the fine texture carried the motion on 90% of trials and was judged more salient on 97.5% of trials when viewed against a very coarse background. However, the scores of the fine texture on motion and salience fell to only 14% and 5.5% respectively when viewed against a very fine background.

Results therefore support the idea that the visual system assigns 'strength' values to texture boundaries, and can use them both to resolve ambiguous AM and to determine relative salience. Once assigned, are these 'strength' values independent of the texture properties that generated them? If so, we ought to be able to see AM in crossover displays in which the two frames contain dissimilar textures. The direction of AM seen in these displays should be predictable from the relative salience of the two squares in each frame. In experiment 3 we tested this idea.

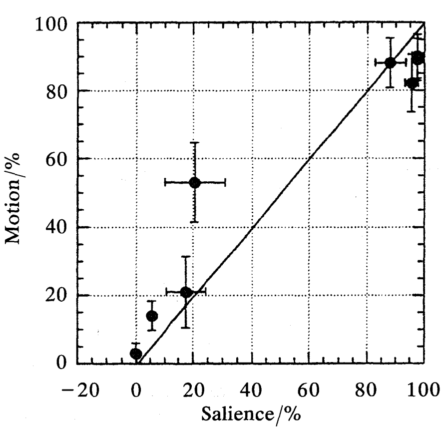


Figure 3. Scatter plot of mean motion judgments from experiment 1 against mean salience judgments from experiment 2. Vertical and horizontal bars show SE of each estimate. The solid line has a slope of unity.

5 Experiment 3

5.1 *Method*

5.1.1 *Subjects.* Five observers participated. All also took part in experiment 1.

5.1.2 *Apparatus and stimuli.* The same equipment was used as in experiment 1. Six different crossover AM displays were generated. The two frames in each display contained squares defined by different texture properties. The six displays comprised all combinations of the four texture properties used previously, as depicted in figure 4. The six displays are arranged in rows, and the figure depicts the texture elements used in the three regions of each frame (background, left-hand square, and right-hand square). Note that, on the basis of previous experiments, we predict AM from the left-hand square of frame 1 to the right-hand square of frame 2 in all six displays. Parameters defining actual stimulus frames were identical to those used in experiment 1.

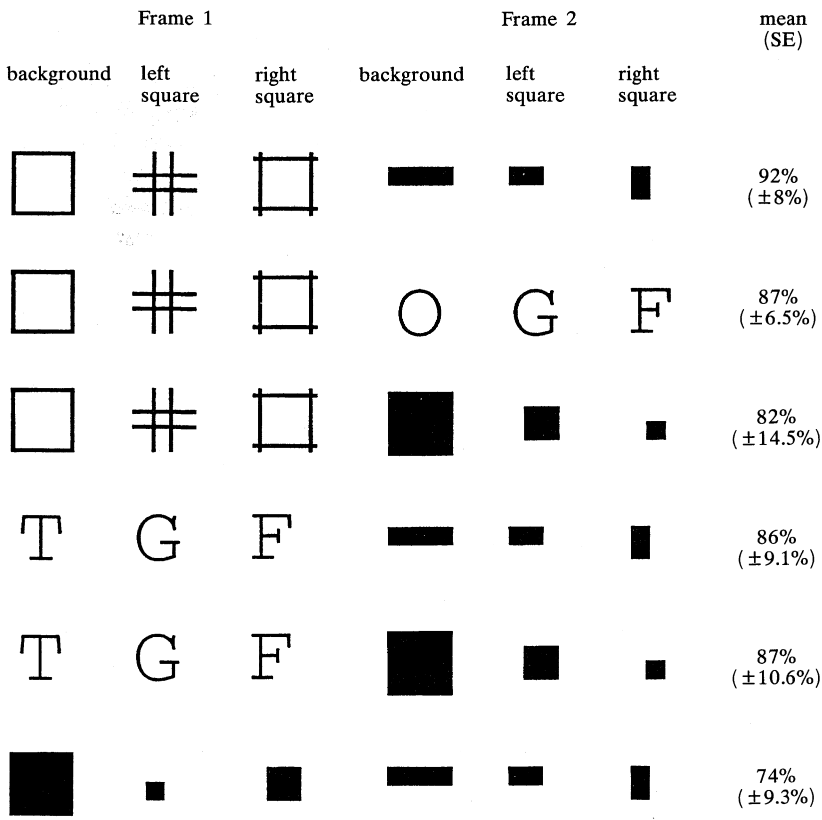


Figure 4. Stimuli and results of experiment 3. Each row depicts one of the six crossover displays used. The icons in each row represent the textures used in each region of each frame (background, left-hand square, and right-hand square). Actual stimuli corresponded to those shown in figure 2. Note that the two frames contained different texture, but on the basis of earlier experiments we predicted that motion should be seen from the left-hand square of the first frame to the right-hand square of the second frame (frame order was actually randomised during the experiment). The rightmost column shows mean and SE of subjects' direction reports, expressed in terms of reports of rightwards motion for the displays depicted (ie high percentages indicate that predictions were confirmed).

5.1.3 *Procedure.* Subjects performed in a single experimental session involving twenty presentations of each AM display, in random order. Each presentation involved a single exposure to one of the displays, after which the subject was instructed to press one of two mouse buttons to signify the direction of AM. An interval of 1 s separated successive trials, as in experiment 1. As mentioned previously, the order in which each subject performed in the three experiments was counterbalanced.

5.2 *Results and discussion*

Means and SEs of AM reports in the direction predicted (ie rightward for the sequences depicted in figure 4) are shown in the rightmost column of figure 4. In all cases, AM was seen in the direction predicted on the basis of previous experiments.

Even though textures in different frames were unrelated, consistent motion was seen on the basis of relative salience. This result supports the idea that the visual system attaches salience values to texture boundaries, and these values can mediate AM regardless of the textures that generated them.

6 General discussion

These experiments show that AM in second-order crossover displays is governed by the same rules as AM in first-order crossover displays. The displays set up a competition between two opposing motion signals, and the stronger signal prevails. The strength of the texture signal appears to determine the strength of the AM signal regardless of the particular second-order properties used in each animation frame.

Werkhoven et al (1993, 1994) also used a motion-competition paradigm, and found that motion could be seen between dissimilar stimulus patches even when the alternative motion path involved similar patches (their stimulus parameters were broadly similar to ours). Results were explained in terms of a 'single-channel motion computation', which involves a nonlinear transformation mapping texture onto a scalar value ('activity'). The strength of the motion signal between two texture patches is determined by the product of the 'activities' they generate. However, Werkhoven et al's concept of 'activity' cannot explain our results. According to their scheme, the activity generated by a texture patch depends only on how that patch is transmitted through a 'texture-grabbing' transformation involving linear filtering followed by rectification. Our experiments indicate that AM does not depend on the activity generated by a single texture patch, but depends instead on the texture 'contrast' between the patch and its background. In experiment 1 we found that the patch which dominates AM in crossover displays can be switched by changing only the background texture. Our results therefore indicate the need for a stage of differential filtering after the 'texture grabber' but before motion extraction, to extract texture contrast. Data support the conclusion that this texture-contrast signal can be used both in judgments of apparent motion and in judgments of perceptual salience.

Both physiological (Knierim and van Essen 1992) and psychophysical studies are consistent with the idea that the visual system extracts a measure of 'feature contrast' (Nothdurft 1991), or 'texture edge strength' (Landy and Bergen 1991) during texture segmentation.

Acknowledgements. This research was supported by a grant to George Mather administered by the Engineering and Physical Sciences Research Council. We are grateful to Linda Murdoch and Andrew Daniell for assistance in running the experiments.

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, 1989 "Kinetic edges become displaced, segregated, or invisible", in *Neural Mechanisms of Visual Perception* Eds D M-K Lam, C D Gilbert (Cambridge, MA: MIT Press) pp 247-260

Anstis S M, Mather G, 1985 "Effects of luminance and contrast on the direction of ambiguous apparent motion" *Perception* **14** 167-180

Braddick O J, 1974 "A short-range process in apparent motion" *Vision Research* **14** 519-527

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

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** 1986-2007

Knierim J J, Essen D C van, 1992 "Neuronal responses to static texture patterns in area V1 of the alert macaque monkey" *Journal of Neurophysiology* **67** 961-980

Landy M S, Bergen J R, 1991 "Texture segregation and orientation gradient" *Vision Research* **31** 679-691

Neisser U, 1963 "Decision-time without reaction-time: experiments in visual scanning" *American Journal of Psychology* **76** 376-385

Nothdurft H C, 1991 "Texture segmentation and pop-out from orientation contrast" *Vision Research* **31** 1073 - 1078

Regan D, 1991 "Spatial vision for objects defined by colour contrast, binocular disparity and motion parallax", in *Spatial Vision* Ed. D Regan (London: Macmillan) pp 135 - 178

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

Werkhoven P, Sperling G, Chubb C, 1994 "Perception of apparent motion between dissimilar gratings: spatiotemporal properties" *Vision Research* **34** 2741 - 2759