Vision Res. Vol. 31, No. 1, pp. 161–167, 1991
Printed in Great Britain.

0042-6989/91 $3.00 + 0.00

SHORT COMMUNICATION

FIRST-ORDER AND SECOND-ORDER VISUAL PROCESSES IN THE PERCEPTION OF MOTION AND TILT

GEORGE MATHER
University of Sussex, Experimental Psychology, Brighton BN1 9QG, U.K.

(*Received* 7 *March* 1990)

It has been known for many years that the mammalian visual system contains neurones which respond selectively to the orientation of contours in the image, and to the direction of their movement. Numerous perceptual studies have also found evidence for the existence of orientation and direction specific channels in the human visual system. A new theoretical framework is currently emerging, which sub-divides direction selective channels into first-order and second-order classes. Data from two experiments reported here show that each class of mechanism can be activated in isolation from the other, using appropriate motion stimuli. Both can mediate direction discrimination, and both give rise to after-effects. Further, analogous illusions of tilt are presented which demonstrate that the first vs second-order distinction applies to orientation coding processes as well. The dichotomy between first and second-order processes therefore may reflect a more general property of the visual system's organisation.

A recent evaluation of psychophysical data on motion perception proposed that motion detectors in the human visual system can be divided into two broad classes: first-order detectors, and second-order detectors (Cavanagh & Mather, 1989). Broadly, first-order detectors respond to the movement of contours defined by intensity gradients in the image (i.e. first-order differences in the intensity distribution), whereas second-order detectors respond to the movement of contours defined in more abstract terms (texture density, flicker, binocular disparity, motion parallax, etc.). Chubb and Sperling (1988) have developed a range of second-order stimuli, which are invisible to first-order detectors ("drift-balanced"). The two classes of detector may correspond to, or at least overlap with the two processes previously identified as "short-range" and "long-range" (Braddick, 1980). Figure 1 illustrates motion stimuli designed to activate each class in isolation from the other. The leftmost panel is an *xt* plot of a standard apparent motion stimulus—space (x-position) is represented on the horizontal axis, and time on the vertical axis. A dark bar is shown occupying a series of static positions at successive time intervals. If the spatial and temporal steps are chosen appropriately, an illusion of rightward apparent motion will be seen by an observer. One can represent a simple first-order detector which responds to this rightward motion as having a receptive field in the *xt* plot with the shape shown in the figure, so that it is tuned to a particular "spatiotemporal orientation". The tilt of the receptive field's long axis specifies its preferred velocity (distance travelled per unit of time). Detectors tuned to other velocities, either rightward or leftward, would have receptive fields with different spatiotemporal orientations. There is good evidence that the visual system does possess motion detectors tuned to spatiotemporal orientation, and several models have been proposed (Adelson & Bergen, 1985; Ross & Morrone, 1986; Emerson, Citron, Vaughn & Klein, 1987; Watson & Ahumada, 1985).

Consider the middle panel of Fig. 1. Here, a bar is shown alternating between two positions over successive time intervals. Contrast polarity is maintained for rightward shifts, but reverses for leftwards shifts. It was first reported by Anstis in 1970 that shifts accompanied by contrast reversals produce reversals in apparent direction, and recent perceptual, computational and physiological data confirm that directional signals are simply inverted by contrast reversal

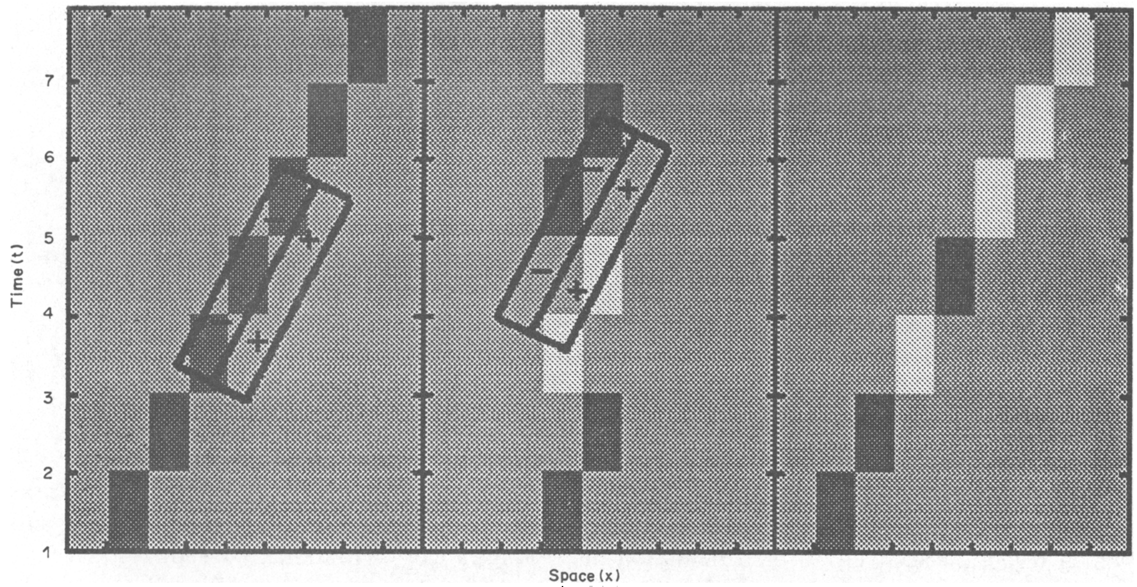

Fig. 1. xt-Plots of three different motion stimuli. Abscissa represents x-position, ordinate represents time. Lefthand panel: a dark bar is shown occupying a position to left of centre at time 1 (if bar is assumed to be vertical there is no need to add the third, y-dimension because it remains fixed). At time 2 the bar suddenly shifts position to the right by one bar width, and repeated shifts occur at subsequent time intervals. These shifts create a staircase across the xt-plot describing a conventional apparent motion sequence. Receptive field shown, tuned to a specific spatiotemporal orientation (velocity), would signal rightward motion. Middle panel: the dark bar shifts rightward between time 1 and time 2, but then shifts leftward again (3), at the same time reversing contrast to become bright. At time 4 bar (now bright) shifts to the right, and then reverses contrast again on its leftward shift (5). The sequence constititues a basic "four-stroke cycle" of alternating position and contrast over time, arranged to create an impression of unidirectional rightward motion. Only detectors with receptive fields tuned to rightward motion (as shown) will respond. Righthand panel: bar shifts consistently to the right over time, but its contrast polarity at each position varies randomly between light and dark. Thus half of the shifts will activate detectors for rightwards motion, with receptive fields such as those shown in other panels, and half will activate leftwards detectors. Any apparent unidirectional motion would have to be mediated by detectors which disregard contrast polarity.

(Emerson et al., 1987; Mather, 1991; Sato, 1989). Thus the sequence depicted should create an illusion of consistent rightward motion (cf. Anstis & Rogers, 1986). The illusion could be explained straightforwardly in terms of first-order detectors tuned to spatiotemporal orientation: contrast-reversal activates not detectors tuned to the leftward orientation of the position shift, but detectors tuned to rightward orientation, as shown.

In the rightmost panel of Fig. 1, a bar consistently shifts in a rightward direction. However, its contrast polarity at each position is selected at random, being negative (dark) at half of the positions and positive (light) at the rest. First-order detectors will supply no consistent unidirectional signal over the whole sequence, because half of the shifts (those not involving contrast reversal) will create rightwards signals and the rest (those involving reversals) will create leftwards signals (i.e. it is drift-balanced; Chubb & Sperling, 1988, 1989). Any rightwards motion reported by observers could be mediated by second-order mechanisms which respond to the presence of a bar but disregard the sign of its contrast. Note that such second-order detectors could not give a consistent response to the stimulus in the middle panel, because there the bar simply alternates in position. Thus a detector which could respond consistently to one of the stimuli shown in the two righthand panels could not respond consistently to the other. Either kind of detector would respond to the conventional stimulus shown in the lefthand panel.

Two experiments were conducted using stimuli based on the sequences shown in Fig. 1, firstly to determine whether subjects actually perceive unidirectional apparent motion in all three, and secondly to compare the stimuli for their ability to generate motion after-effects (MAEs—illusory opposite movement in a stationary stimulus after exposure to a moving stimulus). The experimental stimulus was a circular patch (dia. 3.4 deg arc) of vertical bars against a darker background. Each bar was 2.8 min arc wide and was separated from its neighbour by 22.4 min arc (8 bar widths). Luminances used were: bright bars 60 cd/m^2; dark bars 20 cd/m^2; gaps between bars 40 cd/m^2; background 15 cd/m^2. The bars could be displaced as a group either to the right (as in Fig. 1) or to the left. For the stimulus in the lefthand panel of Fig. 1, the contrast polarity of the whole group of bars (either dark or light) was selected randomly from trial to trial. For the stimulus in the righthand panel, each bar's contrast polarity in each frame of the motion sequence was set independently of the other bars, and varied randomly from frame to frame. In the first experiment on direction discrimination the three different stimulus conditions were performed in separate sessions, order randomised across five subjects. Eighty trials were presented in each session, 20 trials at each of four frame durations (order randomised): 1, 3, 5 and 7 TV frames (each 1/60 sec). A single trial consisted of a 1-sec presentation of the stimulus, during which the subject fixated on a small red spot located at the centre of the stimulus patch. Stimulus direction (left vs right) was selected randomly from trial to trial, and the subject indicated perceived direction by pressing one of two response keys. A short interval (stimulus field blank) separated successive trials. A MicroPDP11/73 computer equipped with a high-resolution raster graphics system generated the stimuli and recorded responses. Results of the experiment are shown in Fig. 2a. Almost perfect direction discrimination was achieved in all conditions except those involving the first-order stimulus (middle panel of Fig. 1) at shorter frame durations. The latter effect can be explained by the temporal properties of first-order processes, which sample moving stimuli at intervals of about 40 msec (Baker & Braddick, 1985), or once every two TV frames, making the judgement reliable only at longer durations.

In each trial of the second experiment, the adapting stimulus (specifications as in the first experiment, using only 50 msec frame durations) was presented for 60 sec. After an audible warning the stimulus stopped moving and the subjects pressed a response key when it no longer appeared to move in the opposite direction (or reported to the experimenter if no after-effect at all was visible). Subjects fixated on the central spot through adaptation and testing.

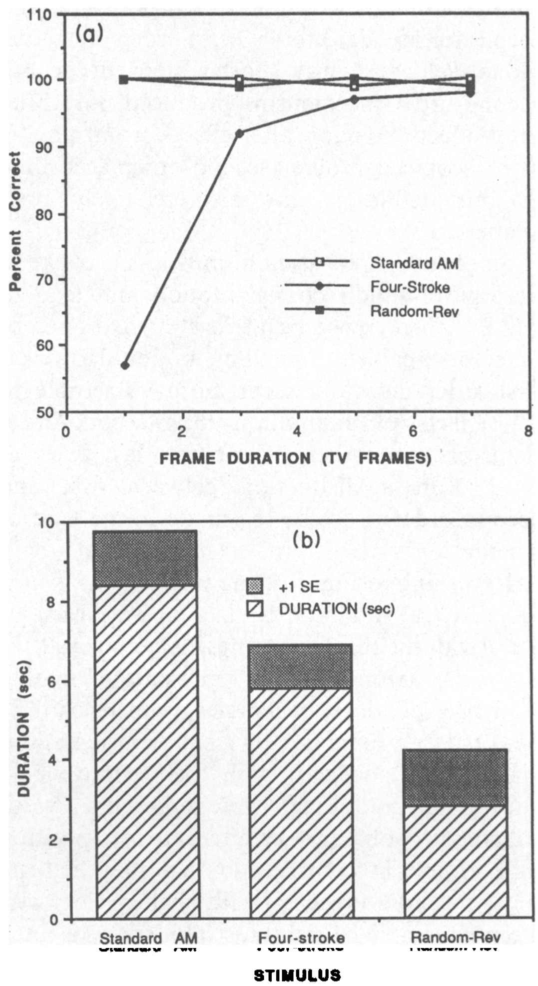

Fig. 2. Results of two experiments employing motion stimuli based on the sequences depicted in Fig. 1. (a) Mean results (five subjects) in a direction discrimination task. (□)—leftmost panel in Fig. 1 (standard apparent motion); (◆)—middle panel (four-stroke cycle); (■)—rightmost panel (random contrast reversal). Standard errors (SE) are omitted for clarity, but were below 1% for standard apparent motion and random reversals, and 5% on average for the four-stroke cycle. Discrimination performance was very high in all but two conditions. (b) Mean durations (7 subjects) of MAEs generated by the three motion stimuli. Shaded zones represent +1 SE. Duration was longest for the stimulus tapping both first and second-order processes. First-order processes alone gave slightly shorter after-effects, but second-order processes gave relatively weak after-effects.

Four after-effect durations were obtained for each of the three stimulus types, from each of seven subjects. A separate session was used for each stimulus type (order randomised across subjects), and an interval of 4 min separated successive duration measurements within a session. Mean motion after-effect duration after adaptation to each of the three types of stimulus are shown in Fig. 2(b). Durations were longest for the stimulus which activated both first and second-order detectors. First-order detectors alone yielded slightly shorter after-effects, but second-order mechanisms produced only brief after-effects. Motion after-effects from moving stereoscopic contours (second-order according to our definition) are also relatively brief (Papert, 1964).

Since all three stimuli in Fig. 1 evoke a percept of unidirectional motion and lead to MAEs, there must be at least two classes of detector capable of signalling motion. However, first-order detectors seem more vulnerable to the effects of adaptation than second-order detectors, since their after-effects last twice as long. Can a distinction between first and second-order processes be made in the case of orientation?

Returning to Fig. 1, if the vertical axis is now labelled "y" and the horizontal axis remains "x" (with the third constant axis becoming "t") then the leftmost plot represents a spatial staircase of dark rectangles, which can be detected as a tilted line by a simple orientation-specific (OS) detector with the receptive field shown. The middle plot now represents a vertical column of bars which alternate in x-position and reverse in contrast, so that they activate only detectors tuned to a tilt slightly clockwise from vertical. The rightmost plot shows a tilted column of rectangles with random contrast polarities, so that there will be no consistent tilt over the whole column signalled by OS detectors, and any perceived tilt must be mediated by second-order orientation detectors which disregard contrast polarity. Complete stimuli corresponding to those in Fig. 1 are shown in Fig. 3. In the top panel, the word LIFE is drawn using vertical and horizontal lines. In the middle and lower panels, LIFE is defined using first and second-order tilts respectively (cf. Fig. 1). Tilted lines can be perceived in both the first and second-order stimuli, confirming that the distinction also applies to the coding of orientation. Once created, the first-order stimulus was revealed as a minimal rendering of Fraser's twist-cord illusion (Fraser, 1908; hence the adoption of the LIFE configuration), which now can be viewed as a stimulus arranged selectively to activate first-order tilt mechanisms. Its appearance is paradoxical in the same way that the equivalent in the spatiotemporal domain is paradoxical—the lines appear tilted (or seem to move) but never actually change position over space (or time, much like a motion after-effect), perhaps as a result of the conflicting signals from first and second-order processes, one signalling tilt in Fig. 2, and the other signalling no consistent tilt. Variants of Fraser's tilt illusions can be decomposed into elements which basically create a spatial "four-stroke cycle" of alternating contrast and position. Morgan and Moulden (1986) demonstrated that other tilt illusions are essentially variants of Fraser's twisted cord.

Thus, the distinction between first and second-order detectors applies to both motion and tilt, since analogous illusions exist in the two domains. Individual detectors are probably tuned jointly to spatial orientation and to spatiotemporal orientation i.e. motion direction (Heeger, 1987; Henry, Bishop & Dreher, 1974; Schiller, Finlay & Volman, 1976; Wenderoth, Bray & Johnstone, 1988). There are, however, asymmetries between the spatial and spatio-temporal dimensions which limit the degree to which one can consider them to be interchangeable. For example, detectors can respond to both left and right inputs in space, but only to past inputs in time. Functionally, there appear to be many size (spatial frequency) tuned detectors, but only a few flicker (temporal frequency) tuned detectors, creating as asymmetry in the sampling density between the spatial and temporal dimensions.

The distinction between first and second-order processes may apply to other stimulus dimensions as well, such as size (spatial frequency) and depth, reflecting a general property of the visual system's organisation. Orientation and direction-specific processes are also known to be spatial frequency selective, which leads to the prediction that first and second-order effects will be found in size perception too. For example, a drifting grating consisting of bars defined by second-order textural differences not only supports motion perception (Cavanagh & Mather, 1989), but also has an apparent orientation and spatial frequency. To date, perceptual research has concentrated on studying first-order stimuli defined by intensity differences, and the processes which detect them, but the emerging class of second-order visual processes raises new questions for future research: how do the two processes differ in their response properties, what are their functional roles, and how do they interact to determine perceptual experience?

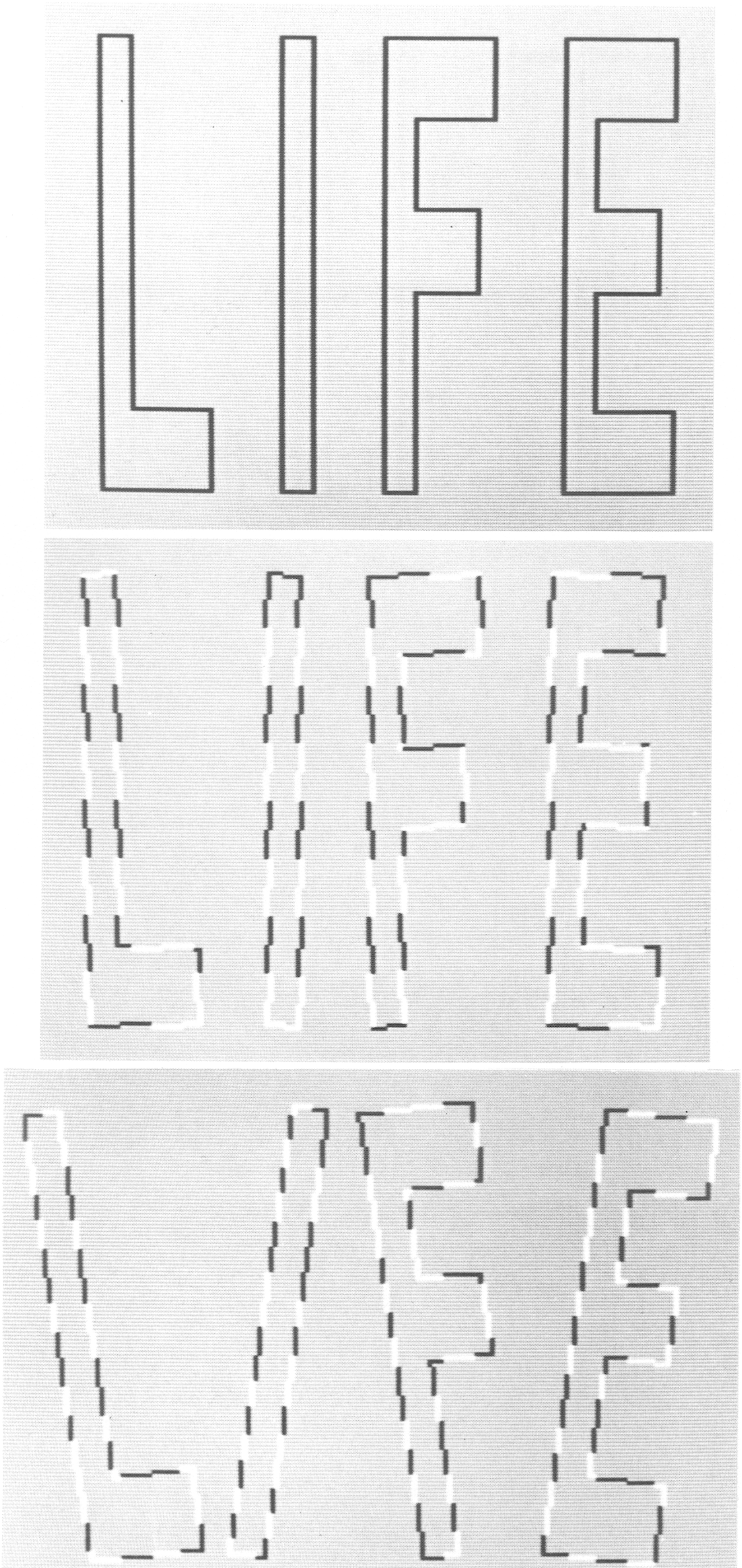

Fig. 3. Stimuli designed to activate first and second-order orientation coding processes, based on sequences shown in Fig. 1 (assuming that the time axis is re-labelled as y): top—vertical letters signalled by both classes of detector; middle—tilted letters signalled only by first-order processes; bottom—tilted letters signalled only by second-order processes.

Acknowledgements—This research was supported by the Science and Engineering Research Council, under the Image Interpretation Initiative. I am grateful to Alexandra O'Halloran for assisting in the execution of 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. (1970). Phi movement as a subtraction process. *Vision Research, 10,* 1411–1430.

Anstis, S. M. & Rogers, B. J. (1986). Illusory continuous motion from oscillating positive-negative patterns: Implications for motion perception. *Perception, 15,* 627–640.

Baker, C. L. & Braddick, O. J. (1985). Temporal properties of the short-range process in apparent motion. *Perception, 14,* 181–192.

Braddick, O. J. (1980). Low-level and high-level processes in apparent motion. *Philosophical Transactions of the Royal Society, London B, 290,* 137–151.

Burr, D. C., Ross, J. & Morrone, M. C. (1986). Seeing objects in motion. *Proceedings of the Royal Society, London B, 227,* 249–265.

Cavanagh, P. & Mather, G. (1989). Motion: The long and the 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,* 1986–2006.

Chubb, C. & Sperling, G. (1989). Two motion perception mechanisms revealed through distance-driven reversal of apparent motion. *Proceedings of the National Academy of Science of the U.S.A., 86,* 2985–2989.

Emerson, R., Citron, M., Vaughn, W. & Klein, S. (1987). Nonlinear directionally selective subunits in complex cells of cat striate cortex. *Journal of Neurophysiology, 58,* 33–65.

Fraser, J. (1908). A new illusion of visual direction. *British Journal of Psychology, 2,* 307–320.

Heeger, D. (1987). Model for the extraction of image flow. *Journal of the Optical Society of America A, 4,* 1455–1470.

Henry, G. H., Bishop, P. O. & Dreher, B. (1974). Orientation, axis and direction as stimulus parameters of striate cells. *Vision Research, 14,* 767–777.

Mather, G. (1991). Computational modelling of motion detectors: Responses to two-frame displays. *Spatial Vision,* in press.

Morgan, M. J. & Moulden, B. (1986). The Münsterberg figure and twisted cords. *Vision Research, 26,* 1793–1800.

Papert, S. (1964). Regularities in the time courses of some visual processes. *MIT Quarterly Progress Report, 73,* 239–244.

Sato, T. (1989). Reversed apparent motion with random dot patterns. *Vision Research, 29,* 1749–1758.

Schiller, P. H., Finlay, B. L. & Volman, S. F. (1976). Quantitative studies of single-cell properties in monkey striate cortrex. II. Orientation specificity and ocular dominance. *Journal of Neurophysiology, 39,* 1320–1333.

Watson, A. B. & Ahumada, A. J. (1985), Model of human visual-motion sensing. *Journal of the Optical Society of America A, 2,* 322–342.

Wenderoth, P., Bray, R. & Johnstone, S. (1988). Psychophysical evidence for an extrastriate contribution to pattern-selective motion after-effect. *Perception, 17,* 81–91.