



# The Perceived Speed of Second-order Motion and its Dependence on Stimulus Contrast

TIMOTHY LEDGEWAY,\* ANDREW T. SMITH\*

Received 13 April 1994; in revised form 25 July 1994

Speed matches were obtained, using a spatial two-alternative forced-choice task, between a second-order motion stimulus and a first-order motion stimulus. The second-order motion stimulus was composed of contrast-modulated noise [produced by multiplying two-dimensional (2-d), static noise by a drifting, one-dimensional (1-d) sinusoid]. The first-order motion stimulus was composed of luminance-modulated noise (produced by summing, rather than multiplying, 2-d noise and a drifting sine grating). In Expt 1, the relationship between the perceived speed of first- and second-order motion was examined. The motion stimuli had the same spatial frequency (1 or 3 c/deg) and were equated for visibility by presenting them at the same multiple of direction-identification threshold. Over a range of physical speeds, the perceived speeds of the first-order and second-order motion stimuli were identical when their physical speeds were the same. In Expt 2, the effect of varying stimulus "contrast" (contrast modulation depth) on the perceived speed of second-order motion was examined. The contrast of the first-order motion stimulus was fixed and speed matches were obtained for second-order motion stimuli at several contrast modulation depths. The motion stimuli had the same spatial (1 or 4 c/deg) and temporal (5 or 20 Hz) frequencies. It was found that the perceived speed of second-order motion was approximately linearly related to log modulation depth. In agreement with previous studies we also confirmed that the perceived speed of first-order motion is similarly dependent on stimulus contrast (luminance modulation depth). The results are discussed in the context of current models of second-order motion perception.

First-order motion   Second-order motion   Perceived speed   Modulation depth

## GENERAL INTRODUCTION

Recent research has begun to focus on the proposed dichotomy between first-order motion processing and second-order motion processing. The terms "first-order motion" and "second-order motion" were introduced by Cavanagh and Mather (1989) to distinguish between the motion of luminance- or colour-defined contours in the retinal image and the motion of contours defined by variations in other spatial characteristics such as local contrast, texture and binocular disparity. Second-order motion has also been termed "non-Fourier" motion (e.g. Chubb & Sperling, 1988, 1989) in order to emphasize the point that the direction and speed of such motion is not directly conveyed by its Fourier spectrum in the luminance domain. Indeed, Chubb and Sperling (1988, 1989) have described a particular class of stimuli that give rise to second-order motion perception which they term "drift-balanced". Drift-balanced motion stimuli should be invisible to conventional low-level motion-detecting mechanisms, which operate by detecting motion energy (e.g. Adelson & Bergen, 1985), because any motion energy they contain is equal in opposite directions. Chubb and

Sperling have also shown that some members of this class of stimuli are "micro-balanced". That is, they remain drift-balanced even after any arbitrary spatiotemporal filtering imposed by the visual system. Such a stimulus can be constructed by modulating the contrast of two-dimensional (2-d), random noise with a drifting one-dimensional (1-d) sinusoidal waveform. However, much uncertainty remains as to the nature of the mechanisms that are involved in second-order motion perception in human vision.

Although some recent psychophysical evidence (e.g. Smith, 1994) suggests that second-order motion perception may be mediated, at least some of the time, by high-level feature-tracking mechanisms (e.g. Anstis, 1980; Cavanagh, 1991, 1992), all explicit models of second-order motion extraction are based on low-level motion-detecting schemes which negate the need explicitly to encode features. These models are supported by several lines of evidence. For example, observers can reliably perceive motion in dense, briefly presented random-dot kinematograms (RDKs) in which the dots are defined by second-order characteristics such as spatial variations in contrast (Cavanagh & Mather, 1989; Nishida, 1993). Such stimuli are assumed to preclude the involvement of feature-based motion-detecting strategies.

\*Department of Psychology, Royal Holloway, University of London, Egham, Surrey TW20 0EX, England.

These low-level models fall into two main classes. Firstly, several computational models suggest that second-order motion perception is mediated by motion detectors which are separate from those used for detecting first-order motion but which operate on qualitatively similar principles. In particular, Chubb and Sperling (1988, 1989) suggested that the outputs of spatiotemporal filters applied to the image are transformed in a nonlinear manner (e.g. rectified) so that motion energy is introduced into the neural representation of the image and is then detected conventionally. Werkhoven, Sperling and Chubb (1993) have recently developed a model of second-order motion perception based upon this principle. A somewhat similar idea has been proposed by Wilson, Ferrera and Yo (1992) and incorporated into a motion detection model that can successfully predict the perceived direction of texture boundaries and type II plaid patterns (Wilson & Mast, 1993). A second class of models propose that first- and second-order motion are detected by a single, common low-level mechanism. For example, Johnston, McOwan and Buxton (1992) have developed a computational model of first- and second-order motion detection based upon the spatiotemporal gradient scheme of Marr and Ullman (1981). Grzywacz (1992) has published a model in abstract form which detects motion using the method of Chubb and Sperling (1988, 1989) (i.e. band-pass spatial filtering followed by rectification and motion energy detection) but dispenses with the separate linear mechanism. Alternatively second-order motion may be detected by conventional, low-level detectors on the basis of an internally generated distortion product arising from a luminance nonlinearity early in the visual system. Distortion products at the level of the retina (e.g. Burton, 1973; MacLeod, Williams & Makous, 1992) and LGN (Derrington, 1987) have been proposed to account for the perception of spatial and moving "beats" (variations in contrast produced by summing two sine gratings of similar spatial frequency) which constitute a particular class of second-order motion stimuli, although some findings are inconsistent with such an explanation, at least for the perception of drifting beats and other contrast-defined patterns (e.g. Badcock & Derrington, 1989; Derrington & Badcock, 1985; Derrington, Badcock & Henning, 1993; Pantle & Turano, 1992). Thus, the idea that second-order motion is visible simply because distortion products stimulate a first-order motion-detection system may be ruled out.

Despite the parsimony of models that suggest a single, low-level motion mechanism for the detection of both first- and second-order motion, several lines of evidence support the existence of separate motion-detecting mechanisms for each type of motion. For example, Harris and Smith (1992) found that second-order motion fails to elicit optokinetic nystagmus (OKN) whilst first-order motion of the same visibility does, suggesting the existence of two distinct mechanisms, only one of which is able to drive OKN. Ledgeway and Smith (1994a) and Mather and West (1993) found that observers were unable to integrate frames in a motion sequence in which spatial structure is defined by first-order characteristics with

frames in which it is defined by second-order characteristics. Mather and West (1993) constructed two-frame RDKs in which the random-dot pattern on each frame was defined by either first-order characteristics (the dots were defined by luminance) or second-order characteristics (the dots were composed of random texture that differed from the background in contrast). Observers were consistently able to detect the direction of RDK displacement for pairs of frames that were defined by the same characteristics (i.e. both frames were composed of either first- or second-order dots). However, when the frames of the RDK were defined by different characteristics direction-identification performance was at chance, implying that the observers were unable to integrate the two frames of the RDK. Ledgeway and Smith (1994a) used a similar approach but employed sinusoidal gratings in place of RDKs, and obtained the same result. These results clearly support the existence of separate mechanisms for the processing of first- and second-order motion.

Although many experiments concerned with the processing of second-order motion have focused on the coding of direction information by the visual system, little attempt has been made to investigate the coding of the speed of second-order motion. This may be partly due to the fact that the principles underlying the coding of the speed of first-order (luminance-defined) motion are not as well established as those mediating the coding of direction information. For example, although direction-specificity is a fundamental property of visual neurones in mammalian visual cortex and direction appears to be systematically represented in an orderly array of cortical columns within area MT (Albright, Desimone & Gross, 1984), speed-selectivity does not appear to be a fundamental property of visual neurones (although see Newsome, Gizzi & Movshon, 1983). Motion-sensitive neurones tend to be tuned for temporal frequency rather than speed (e.g. Foster, Gaska, Nagler & Pollen, 1985; Holub & Morton-Gibson, 1981). Current computational models (Grzywacz & Yuille, 1990; Heeger, 1987) of first-order motion propose that local speed sensitivity emerges by a process of grouping the outputs of neurones that behave like spatiotemporal filters with receptive fields at the same spatial location in such a manner that neurones having similar velocity sensitivities are grouped together. For example, in Heeger's model the overall speed of an object moving in the visual field is then encoded as the peak in a distribution of the outputs of several such groupings having different speed sensitivities, and need not be encoded explicitly by individual neurones. This process of encoding speed by grouping the outputs of spatiotemporally tuned neurones is accomplished by speed-sensitive units which Heeger has tentatively identified with MT cells. However, the existence of true speed-tuned cells remains to be established.

Although some models of first-order motion detection confound the coding of speed information with contrast, several models attempt to disambiguate responses related to speed from those related to contrast. For example,

Watson and Ahumada's (1985) scalar motion sensor model signals local speed by the temporal frequency of the output of oriented spatiotemporal filters (a measure that is largely independent of stimulus contrast). Similarly, Heeger's (1987) model computes speed on the basis of the distribution of activity within sets of neurones (spatiotemporal filters) rather than their absolute firing rates which increase with contrast (e.g. Albrecht & Hamilton, 1982). Thus, such models predict that the perceived speed of first-order motion should be unaffected by changing stimulus contrast. This prediction, however, is not borne out by psychophysical evidence. For example, Thompson (1982) reported evidence using a speed-matching technique that at moderate contrast levels (at and below Michelson contrasts of 0.17) and drift temporal frequencies below 8 Hz, decreases in sine grating contrast lead to corresponding decreases in perceived speed. Similarly, Stone and Thompson (1992) reported that for simultaneously presented sine gratings across a wide range of contrasts (0.025–0.5) and temporal frequencies (3–10 Hz) speed-matching performance was dependent on the relative contrasts of the gratings. For example, a sine grating of contrast 0.7 had to be slowed by 35% in order to match the perceived speed of a sine grating of contrast 0.1 drifting at 2 deg/sec. There was little evidence of saturation in perceived speed as the contrasts of the sine gratings increased. However, the effect of contrast on perceived speed was impoverished when the comparison stimuli were presented sequentially, rather than simultaneously, although the effect was not completely abolished even with inter-stimulus intervals (ISIs) of several seconds.

Stone and Thompson (1992) tentatively identified the dependence of perceived speed on stimulus contrast for simultaneously presented sine gratings with a contrast-normalization process (e.g. Adelson & Bergen, 1986; Heeger, 1992). Within this scheme the motion energy of each grating is normalized with respect to the average (pooled) contrast of both gratings. As a result a low contrast stimulus is perceived to drift more slowly than a higher contrast stimulus. For sequentially presented stimuli, the motion energy of each stimulus is normalized with respect to its own average contrast and the effect of contrast on perceived speed should be diminished. Recent psychophysical evidence (Thompson, Stone, Swash & Stone, 1994) has suggested that such contrast normalization must be local, since manipulating the contrast of areas surrounding the motion stimuli has little impact on perceived speed.

Although much uncertainty exists as to the nature of first-order (luminance-defined) speed coding within the visual system, even greater uncertainty exists as to the coding of the speed of second-order motion. Although some MT cells have been identified which are sensitive to second-order motion (Albright, 1992; Olavarria, DeYoe, Knierim, Fox & Van Essen, 1992), no studies have systematically examined the spatiotemporal properties of such cells. From the standpoint of psychophysics, no studies have examined the perceived speed of second-order motion patterns.

In the present study we sought to address some of these issues by conducting speed-matching experiments using first- and second-order motion stimuli and a spatial two-alternative forced-choice (2AFC) task. Specifically, Expt 1 examined the relationship between the perceived speed of first-order motion and that of second-order motion and Expt 2 investigated whether or not the perceived speed of second-order motion is dependent on stimulus contrast (modulation depth) in a manner analogous to that reported for first-order (luminance-defined) motion. These issues are important because they offer valuable insights into the mechanisms responsible for encoding second-order motion and may provide useful constraints on current models of second-order motion perception.

## GENERAL METHODS

### *Observers*

Three observers participated in the study and all had normal or corrected-to-normal acuity. Observer TL was one of the authors and observers MW and TF were paid volunteers who were unaware of the purpose of the experiments.

### *Apparatus and stimuli*

Motion stimuli were generated under computer control by a Matrox IM-640 image processing system connected to a Manitron monitor and a Constable CRT image generator connected to Hewlett-Packard X-Y display. Both displays were monochrome and had white (P4) phosphor. The X-Y display had internal  $\gamma$ -correction and was approximately linear. As it was only used to display (first-order) sine gratings, no additional  $\gamma$ -correction was applied. In order to ensure that the Manitron monitor used to display the second-order motion stimuli did not introduce unwanted first-order motion components into the images as a result of inherent luminance nonlinearities (Henning, Hertz & Broadbent, 1975) it was carefully  $\gamma$ -corrected using a look-up table. Calibration was performed by measuring the local space-time luminance of second-order images similar to those employed in the experiment so that  $\gamma$ -correction was accurate with respect to such stimuli. We are confident that following this procedure any inherent residual luminance nonlinearities present in the display were too small to contribute to our results. This is supported by the results of previous studies that have used the same calibration procedure (e.g. Ledgeway & Smith, 1994a) in which any significant nonlinearities in the second-order images would lead to unambiguous motion percepts when alternated with first-order images of the same frequency. A result which was not observed.

The Constable image generator (which was used because it allowed very fine control of the drift speed of generated images) was used to generate a first-order motion stimulus. This was always a vertically-oriented sine grating with independently-controllable spatial frequency, contrast, drift direction and speed. Drift speed

was controlled by manipulating the magnitude of the spatial displacement of the sine grating on each image update with sub-pixel resolution. The update rate of the image was always 122 Hz. The sine grating was optically added to 2-d, broadband, static noise generated by the Matrox IM-640 image processing system to produce luminance-modulated noise. The luminance profile of such an image is shown schematically in Fig. 1(d). The perceived speed of this first-order motion stimulus (referred to as the match stimulus) was matched to that of either a similar first-order motion stimulus or a second-order motion stimulus (referred to as the standard stimulus) generated in real-time by the Matrox IM-640 image processing system. The second-order motion stimulus was composed of contrast-modulated noise produced by multiplying 2-d, static noise by a 1-d, raised sinusoid [see Fig. 1(c)] to give a vertically-oriented

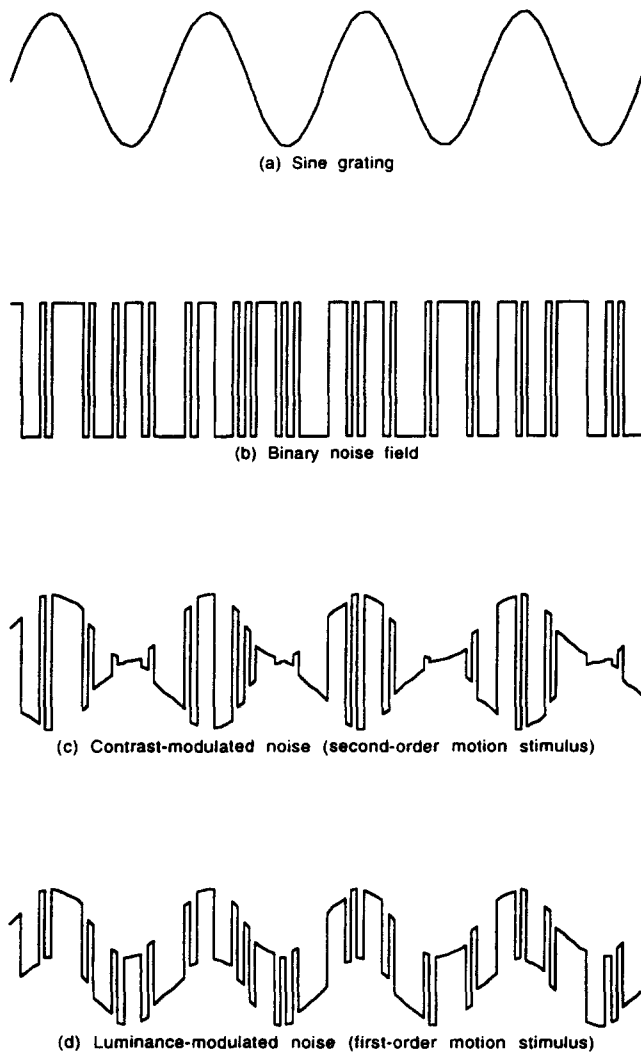


FIGURE 1. Luminance profiles of the first- and second-order images used. Each trace represents a horizontal section through an image and shows how luminance changes as a function of spatial position. Second-order images were composed of contrast-modulated noise (c) produced by multiplying a static, binary noise field (b) by a sine grating (a). For the purposes of multiplication the noise was signed (range -1 to +1) and the modulation signal was unsigned (range 0-1). First-order images were composed of luminance-modulated noise produced by summing a static, binary noise field (b) and a sine grating (a) and scaling.

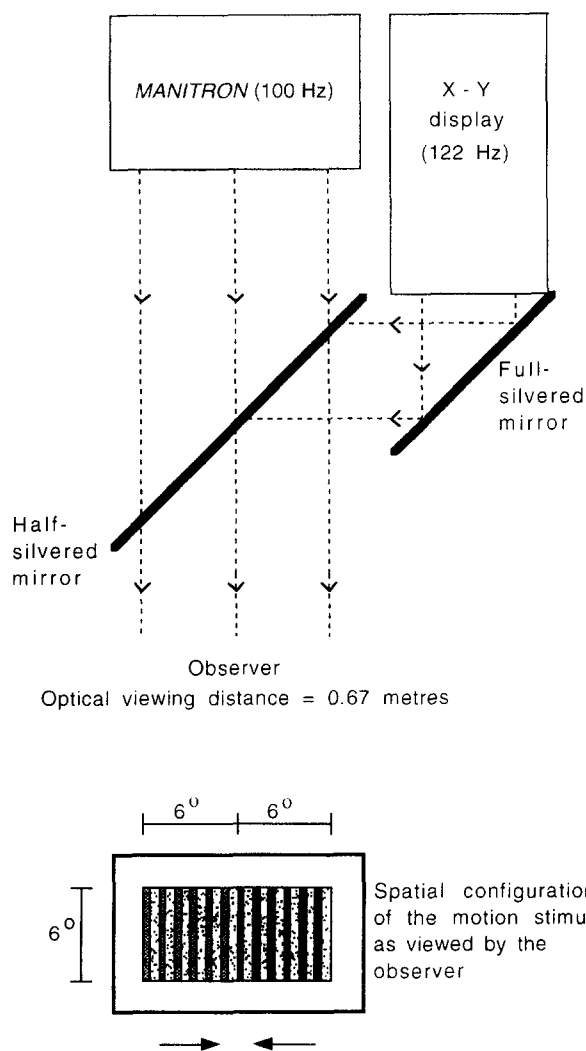


FIGURE 2. Schematic diagram of the apparatus used to display the first- and second-order motion stimuli employed in Expts 1 and 2. The X-Y display and the Manitron monitor were positioned so that they were horizontally adjacent to each other. Motion stimuli were displayed on them simultaneously and appeared side-by-side in the observer's field of view by the use of two mirrors (one was half-silvered and the other full-silvered).

second-order grating. Motion of the standard stimulus was achieved by displacing the sinusoid by some fraction of its spatial period prior to addition or multiplication with the noise (which remained static) and drift speed was controlled by manipulating the magnitude of the spatial displacement (always an integer number of pixels) and/or the image update rate between 50 and 100 Hz. The resulting luminance- or contrast-modulated noise had independently controllable spatial frequency, modulation depth, drift direction and speed.

A schematic representation of the spatial arrangement of the apparatus is shown in Fig. 2. The X-Y display and the monitor were positioned so that they were horizontally adjacent to each other. The motion stimuli displayed on them were presented simultaneously and appeared side by side in the observer's field of view by the use of two mirrors (one was half-silvered and the other full-silvered). Each motion stimulus subtended an angle of  $6 \times 6$  deg at the viewing distance of 0.67 m and had a

mean luminance of about 34 cd/m<sup>2</sup> when viewed through the half-silvered mirror. The luminance of the remainder of the display area (which was homogeneous) was also approx. 34 cd/m<sup>2</sup> when viewed through the half-silvered mirror. The motion stimuli were viewed binocularly and drifted in opposite directions towards a prominent fixation spot located at the centre of the observer's field of view (i.e. the direction of drift was centripetal) in order to facilitate fixation. Smith and Edgar (1990) found that this arrangement yielded less variance in speed-matching performance than other arrangements. A head rest was also used to minimize head movements and maintain the viewing distance of 0.67 m.

The image displayed on the Manitron monitor contained the standard motion stimulus together with a horizontally-adjacent, unmodulated noise field which was optically added to the sine grating (match stimulus) generated by the Constable image generator (see above). Thus, both the standard and match stimuli contained 2-d, static noise. This was in order to control for any possible confounding effects of the noise present in the second-order motion patterns on their perceived speed. (In fact, pilot studies revealed that similar speed matches were obtained when simple sine gratings rather than luminance-modulated noise served as the match stimuli.) The noise used for all stimuli was produced by randomly assigning elements (groups of screen pixels) to be "black" or "white" with probability 0.5. Each noise element subtended 3 × 3 arc min. For the second-order standard motion stimulus the mean contrast of the noise was half the maximum possible (Michelson contrast of 0.48) and the amplitude of the contrast modulation could be varied within the range 0.0–1.0 defined as:

$$\begin{aligned} \text{contrast modulation depth} \\ = (C_{\max} - C_{\min}) / (C_{\max} + C_{\min}) \end{aligned}$$

where  $C_{\max}$  and  $C_{\min}$  are the maximum and minimum local Michelson contrasts in the image, calculated over pairs of adjacent noise elements with opposite polarity.

For the first-order motion standard and match stimuli, the amplitude of the noise following addition with the sine grating was constant and was always equal to the mean amplitude of the noise in the second-order motion stimulus described previously. The amplitude of the luminance modulation (contrast or luminance modulation depth) could be varied within the range 0.0–0.5 defined as:

$$\begin{aligned} \text{luminance modulation depth} \\ = (L_{\max} - L_{\min}) / (L_{\max} + L_{\min}) \end{aligned}$$

where  $L_{\max}$  and  $L_{\min}$  are the maximum and minimum mean luminances averaged over pairs of adjacent noise elements with opposite polarity in the image. Note that this is not equal to the Michelson contrast of the image as a whole, but reflects only the contrast of the sinusoid.

In order approximately to equate the suprathreshold visibilities of the first- and second-order motion stimuli, the modulation depths of all stimuli were specified in terms of multiples of direction-identification threshold.

These were individually measured for each observer using the motion stimuli described above and the method of constant stimuli.

### Procedure

Speed matches were measured by means of a spatial 2AFC procedure involving simultaneous presentation of the standard and match stimuli to different parts of the observer's field of view. In each run of 50 trials, the standard stimulus was always presented at the same multiple of threshold and its spatial frequency and drift speed were also constant. The match stimulus had the same spatial frequency as the standard stimulus within each run of trials. The duration of each presentation of the standard and match stimuli was 1 sec and presentations were separated by a 3 sec interval containing a homogeneous field of mean luminance 34 cd/m<sup>2</sup>. At the beginning of each run of trials the drift speed of the match stimulus was chosen randomly from a set of possible values and the observer's task was to indicate using two response buttons which of the two motion stimuli appeared to be drifting faster. The drift speed of the match stimulus on subsequent trials was determined by a "best PEST" routine (Lieberman & Pentland, 1982; Pentland, 1980) which tracked the 50% performance level (the level at which the observer was unable to distinguish between the speeds of the standard and match stimuli). Observers completed four runs of trials for each condition examined. For two of the runs of trials, the standard stimulus was presented to the left of the match stimulus and for the remaining two runs of trials the positions of the stimuli were reversed. Thus, any hemifield differences in perceived speed (Smith & Hammond, 1986) were counterbalanced. The order in which each observer completed the runs of trials was randomized.

## EXPERIMENT 1—DO FIRST- AND SECOND-ORDER MOTION STIMULI HAVE THE SAME PERCEIVED SPEED WHEN THEY ARE EQUATED FOR VISIBILITY?

### Introduction and methodological details

The main aim of the present experiment was to investigate the relationship between the perceived speed of first- and second-order motion over a range of spatial frequencies, temporal frequencies and modulation depths.

The modulation spatial frequencies of the standard and match stimuli were always identical and were either 1 or 3 c/deg. Speed matches were measured at six modulation temporal frequencies of the standard stimulus ranging from 3.75 to 22.5 Hz spaced at equal intervals of 3.75 Hz.

The first-order motion match stimulus and the second-order motion standard stimulus were presented at the same multiple of direction-identification threshold. The thresholds for each observer (TL and MW), which were measured at both of the spatial frequencies and all six temporal frequencies examined, are shown in Fig. 3. In order to obtain as much information as possible

concerning the perceived speed of first- and second-order motion, two multiples of threshold were employed for each spatial frequency and drift temporal frequency examined. These were 2 and 13 times threshold (the latter being the maximum possible based on the highest thresholds obtained for each observer) when the spatial frequency of the standard and match stimuli was 1 c/deg and 2 and 3.9 times threshold (again, 3.9 was the maximum possible based on the highest thresholds obtained for each observer) when the spatial frequency of the standard and match stimuli was 3 c/deg.

### Results and discussion

Weibull (1951) functions were fitted to the resulting data and the speed match for each condition was taken as the 50% point on the function. Results are plotted as speed-matching functions for each spatial frequency and multiple of threshold condition examined.

(i) *Speed matches for the standard and match stimuli presented at low modulation depth.* Figure 4 shows speed-matching data for two observers when the spatial

frequencies of the standard and match stimuli were either both 1 c/deg or both 3 c/deg. It is apparent that, at least when the first-order match stimulus and the second-order standard stimulus are equated in terms of suprathreshold visibility, the speed-matching function is linear and has unity slope. That is, the perceived speeds of the first-order match stimulus and the second-order standard stimulus are equal for a given physical speed and spatial frequency. The data shown in Fig. 4 are fit well by straight lines for both of the observers.

(ii) *Speed matches for the standard and match stimuli presented at high modulation depth.* Figure 5 shows speed-matching data when the standard and match stimuli were both presented at either 13 times threshold (1 c/deg) or 3.9 times threshold (3 c/deg). Clearly the results are very similar to those obtained in the conditions discussed above (Fig. 4) suggesting that the absolute magnitude of the physical modulation depths of the standard and match stimuli does not affect the match provided that all stimuli are equated for suprathreshold visibility.

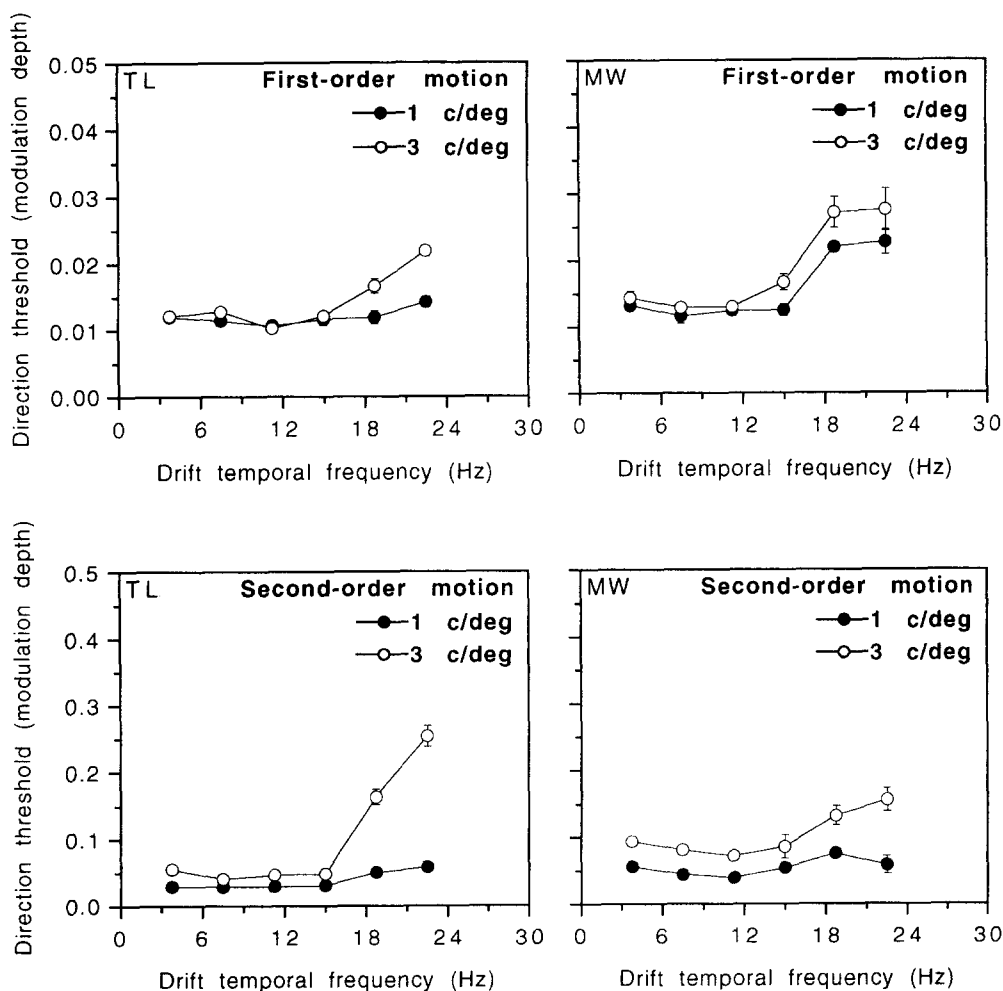


FIGURE 3. Direction-identification thresholds for two observers for first-order (top) and second-order (bottom) motion stimuli of the type used in Expt 1. These thresholds were used as the basis for equating the visibility of the stimuli in the main experiment. The first-order motion stimulus was composed of luminance-modulated noise and the second-order motion stimulus was composed of contrast-modulated noise. The spatial frequency of the modulation was either 1 c/deg (●) or 3 c/deg (○). Observers completed 6 runs of 50 trials for each condition examined and on each trial the motion stimulus was presented for 1 sec. The modulation depth of the stimulus on any one trial was determined by a PEST routine which tracked the 75% performance level.

The vertical lines above and below each data point (where visible) represent  $\pm 1$  SE.

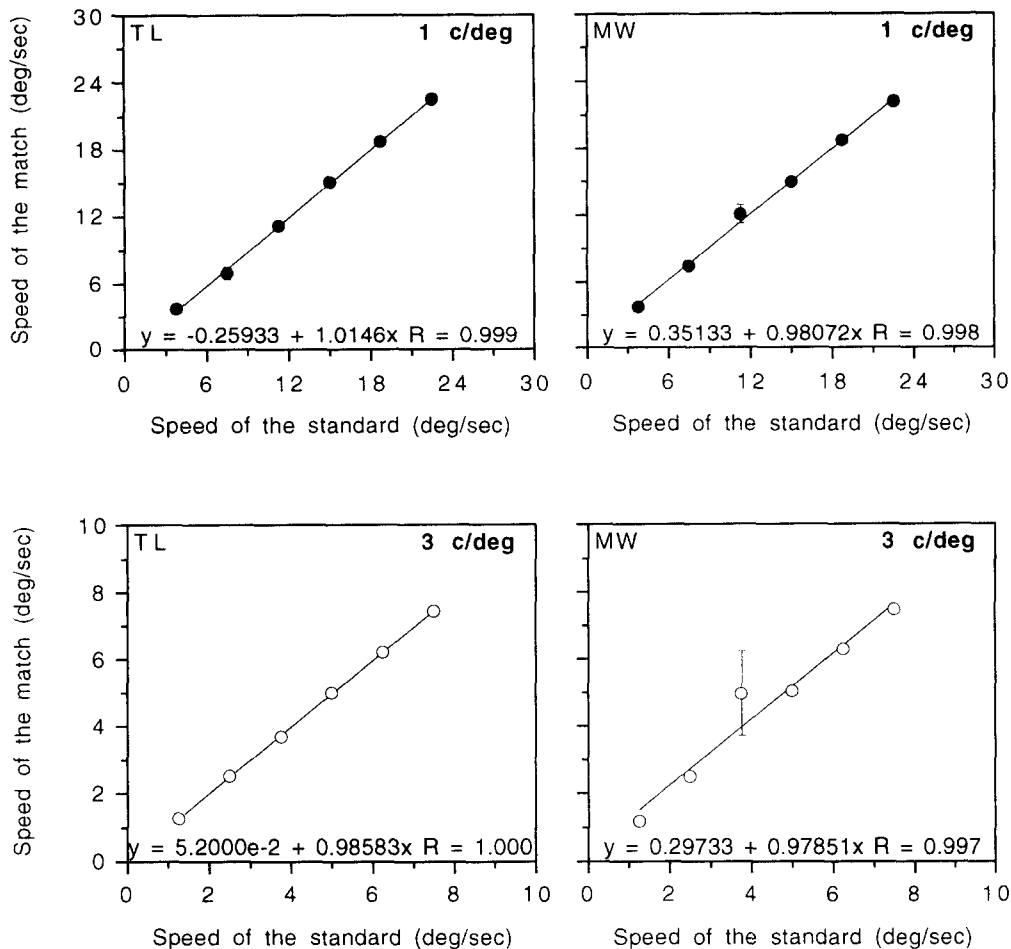


FIGURE 4. Speed matches for two observers. The first-order match stimulus was composed of luminance-modulated noise and the second-order standard stimulus was composed of contrast-modulated noise. The spatial frequency of the modulation was either 1 c/deg (●) or 3 c/deg (○) and all stimuli were presented at 2 times their respective direction-identification thresholds. The vertical lines above and below each data point (where visible) represent  $\pm 1$  SE.

Given that Stone and Thompson (1992) and Thompson (1982) have demonstrated that the perceived speed of luminance-defined (first-order) motion is dependent on physical modulation depth, the results imply that the perceived speed of contrast-defined (second-order) motion is also dependent on physical modulation depth. If this were not the case and the perceived speed of second-order motion were independent of physical modulation depth, then increasing the physical modulation depth of the stimuli would result in underestimation of the speed of the second-order stimulus. This was not the observed result. In Expt 2, the effect of modulation depth on the perceived speed of second-order motion was examined directly.

#### EXPERIMENT 2—THE EFFECT OF STIMULUS MODULATION DEPTH ON THE PERCEIVED SPEED OF FIRST- AND SECOND-ORDER MOTION

##### *Introduction and methodological details*

The main aim of Expt 2 was to investigate systematically the effect of modulation depth on

the perceived speeds of first- and second-order motion.

As for Expt 1, the match stimulus was always a first-order motion pattern composed of luminance-modulated noise. The standard stimulus could be either (first-order) luminance-modulated noise or (second-order) contrast-modulated noise. The modulation spatial frequencies of the standard and match stimuli were always identical and were either 1 or 4 c/deg. Speed matches were measured at one modulation temporal frequency of the standard stimulus which was 5 Hz.

In order to allow comparison of the results for the first- and second-order standard stimuli, the modulation depths of all stimuli were specified in terms of multiples of threshold. The thresholds for each observer (TL and TF), which were measured at both of the spatial frequencies examined, are shown in Fig. 6. The maximum multiple of threshold at which each motion stimulus could be presented was limited by the direction-identification thresholds obtained for the second-order motion stimulus at each spatial frequency. For the standard stimulus, nine multiples of threshold were employed for each spatial frequency. These ranged from 2 to 16 times threshold in

increments of 1.75, for observer TL and from 2 to 13.6 times threshold in increments of 1.45 for observer TF. For the match stimulus, one multiple of threshold was employed for each spatial frequency examined and this was always equal to the median multiple of threshold used for the standard stimulus. For observer TL the match stimulus was presented at 9 times threshold and for observer TF the match stimulus was presented at 7.8 times threshold. The same multiples of direction-identification threshold were used for the first- and second-order motion stimuli.

### Results and discussion

Weibull (1951) functions were fitted to the resulting data and the speed match was taken as the 50% point on the function. In line with studies that have examined the effect of stimulus modulation depth on the perceived speed of luminance-defined (first-order) motion (e.g. Stone & Thompson, 1992) the results are plotted as the speed of the match stimulus, expressed as a percentage of the speed of the standard stimulus, as a function of the modulation depth ratio in dB (i.e.  $20 \log_{10}$  of the ratio of the modulation depth of the standard stimulus to that of the match stimulus).

(i) *Speed matches obtained at 1 c/deg.* Figure 7 (top) shows speed-matching data for the two observers for first-order standard and match stimuli. The speed-matching data are very similar for both of the observers and fall close to a straight line. When the standard and match stimuli were presented at the same multiple of threshold (0 dB) they appeared to drift at the same rate, as expected. However, when the standard stimulus was presented at lower multiples of threshold than the match stimulus the latter had to be slowed down by as much as 27.2% for observer TL and 37.4% for observer TF in order to appear to drift at the same rate as the standard. In a similar manner, when the standard stimulus was presented at higher multiples of threshold than the match stimulus the latter had to be speeded up by as much as 30.2% for observer TL and 10.4% for observer TF in order to appear to drift at the same rate as the standard. The data are in broad agreement with those of Stone and Thompson (1992) in that the effect of stimulus modulation depth on the perceived speed of first-order motion appears to be approximately linear in log modulation depth (in this case luminance contrast). Indeed the data are fit well by straight lines. The maximum, physical modulation depth of the standard

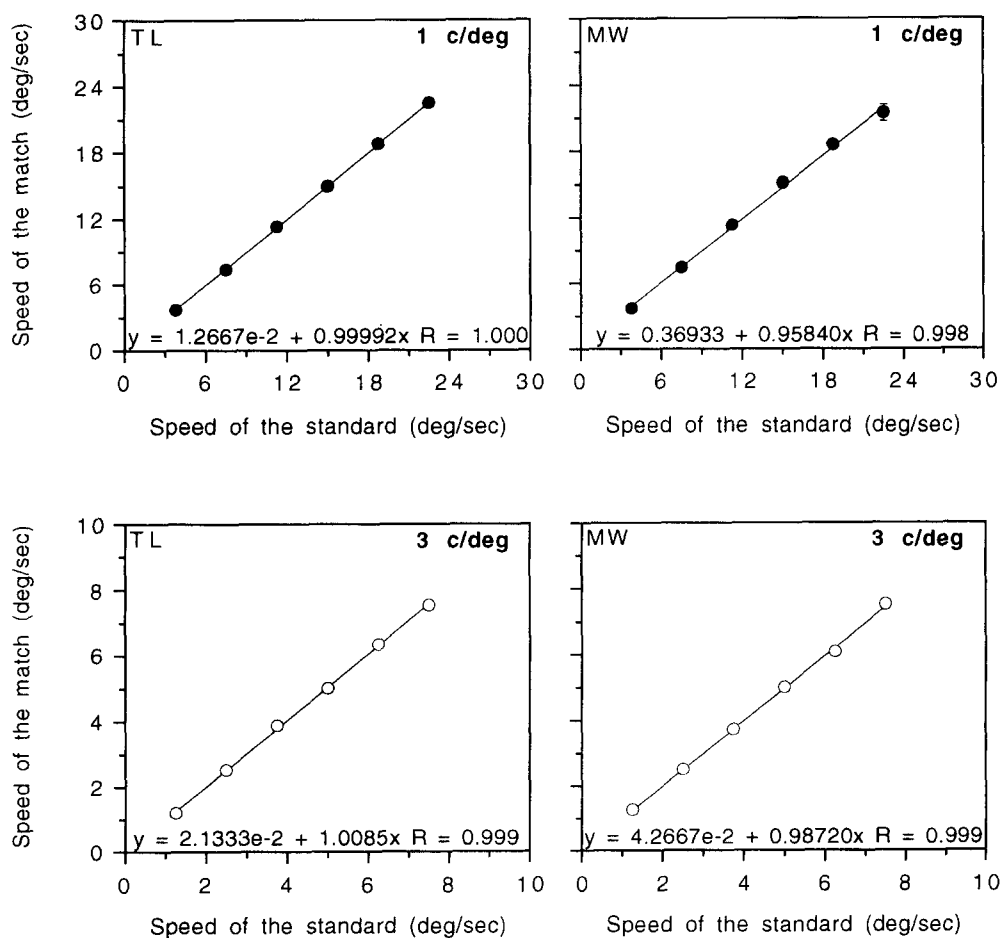


FIGURE 5. Same as Fig. 4 except that the stimuli were presented at either 13 times (●) or 3.9 times (○) their respective direction-identification thresholds.

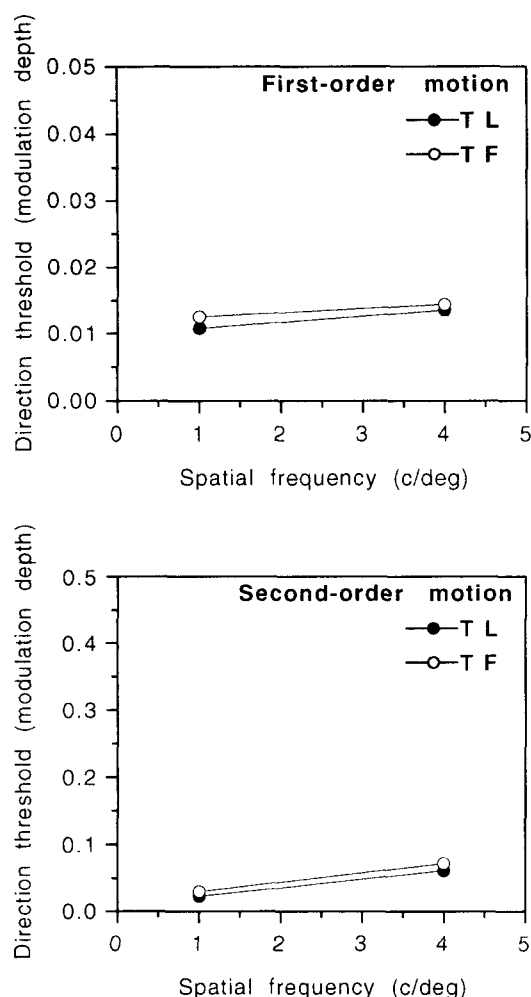


FIGURE 6. Direction-identification thresholds for two observers for the first-order (top) and second-order (bottom) motion stimuli used in Expt 2. The first-order motion stimulus was composed of luminance-modulated noise and the second-order motion stimulus was composed of contrast-modulated noise. The method was the same as that described in Fig. 3. The vertical lines above and below each data point (where visible) represent  $\pm 1$  SE.

was 0.17 for both observers and there was little evidence of saturation in perceived speed at these values.

Figure 7 (bottom) also shows speed matches obtained with a second-order standard stimulus and a first-order match stimulus. The pattern of results is similar to that for the first-order standard stimulus described above. The data for the two observers are reasonably well fit by straight lines that are comparable to the functions obtained with the first-order standard and match stimuli. Thus, the data clearly show that the perceived speed of second-order motion is dependent on physical modulation depth in a manner that is comparable to that of first-order motion.

(ii) *Speed matches obtained at 4 c/deg.* Figure 8 (top) shows speed-matching data for the two observers for first-order standard and match stimuli. The speed-matching data exhibit greater variability between observers than those for the 1 c/deg conditions, and are less well fit by linear functions. Nevertheless, there is a clear trend towards increased perceived speed as modulation depth is increased.

Figure 8 (bottom) also shows speed-matches obtained with a second-order standard stimulus and a first-order match stimulus. Although, again, a certain degree of variability is evident in the speed-matching data for the two observers, the data follow the same pattern as for the first-order standard stimulus. The variability of results between observers and the departure of the functions from linearity is probably attributable to the difficulty of the task. Despite extensive practice trials observer TF, in particular, expressed some difficulties in performing the speed-matching task, especially at high spatial and temporal frequencies, and this is reflected in the variability of the data obtained for this observer. Similarly, Stone and Thompson (1992) also reported a considerable degree of variability in the ability of individual observers to perform speed-matching tasks with first-order motion stimuli (e.g. see their Figs 2, 5 and 6). Nevertheless, the data once again show that the perceived speed of second-order motion is dependent on physical modulation depth. Stone and Thompson (1992) found that the effect of modulation depth on the perceived speed of first-order motion was robust and largely insensitive to a two-fold change in spatial frequency. The present results extend this finding to a four-fold change in spatial frequency and show that the same holds for second-order motion.

Although observer TF could not make reliable speed matches at drift temporal frequencies much greater than 5 Hz (data not shown), observer TL was able to perform the task at frequencies up to 20 Hz. Thompson (1982) found that observers were able to make speed matches at drift temporal frequencies of 16 Hz and other speed-matching studies (e.g. Smith & Edgar, 1990) have obtained reliable speed matches at drift temporal frequencies exceeding 20 Hz. However, Stone and Thompson (1992) reported that two of their observers could not make reliable speed matches at drift temporal frequencies exceeding 8.25 Hz whilst for one observer (PT) the corresponding limit was 10 Hz. Therefore, there appears to be considerable inter-observer variability in the highest drift temporal frequency at which speed-matching performance can be reliably measured when the relative modulation depths of the motion stimuli are manipulated.

Figure 9 shows speed-matching data for observer TL at 20 Hz when the spatial frequencies of the standard and match stimuli were either both 1 c/deg or both 4 c/deg. In general these data are similar to those obtained at the lower drift temporal frequency (5 Hz) and show that the effect of modulation depth on perceived speed is a robust phenomenon and is to all purposes largely unaffected by a four-fold change in drift temporal frequency. Stone and Thompson (1982) also found that a three-fold change in drift temporal frequency (from 3 to 10 Hz) had little effect on the shape of the function relating perceived speed to luminance modulation depth. However, two studies (Hawken, Gegenfurtner & Tang, 1994; Thompson, 1982) have reported that at drift temporal frequencies greater than about 8 Hz, the effect of modulation depth on the perceived speed of first-order motion declines and even

reverses. Further experimentation is required to clarify the circumstances in which this occurs. At present it is sufficient to conclude that the perceived speeds of first- and second-order motion show a similar dependence on stimulus modulation depth under the conditions we have employed.

### GENERAL DISCUSSION

The results of the present experiments show clearly that when luminance- (first-order) and contrast-defined (second-order) motion stimuli are equated in terms of suprathreshold visibility and the presence of noise, psychophysical performance for the two classes of stimuli can be remarkably similar, suggesting that the mechanisms responsible for detecting first- and second-order motion are either one and the same, or else are distinct but operate on similar principles. The results of Expt 1 demonstrate that when first- and second-order motion stimuli are presented at the same physical speed and multiple of threshold they are perceived to have the

same speed. From the results of Expt 2 it is apparent that the perceived speed of second-order motion is dependent on stimulus modulation depth in a manner that is similar to first-order motion.

The similarity between the speed-matching performance found for first- and second-order motion stimuli, over a range of stimulus parameters, has important implications concerning current models of second-order motion perception. Although the simplest interpretation of our results is that first- and second-order motion are both detected by the same (common) motion mechanism in human vision (e.g. Grzywacz, 1992), we also wish to demonstrate that models suggesting separate mechanisms for the processing of first- and second-order motion may also be able to account for our results. For example, Wilson *et al.* (1992) postulate the existence of separate motion pathways for the detection of first-order (Fourier) and second-order (non-Fourier) motion. The first-order pathway begins with band-pass spatial filtering followed by motion energy analysis. The second-order pathway is similar, except that motion energy analysis is preceded by

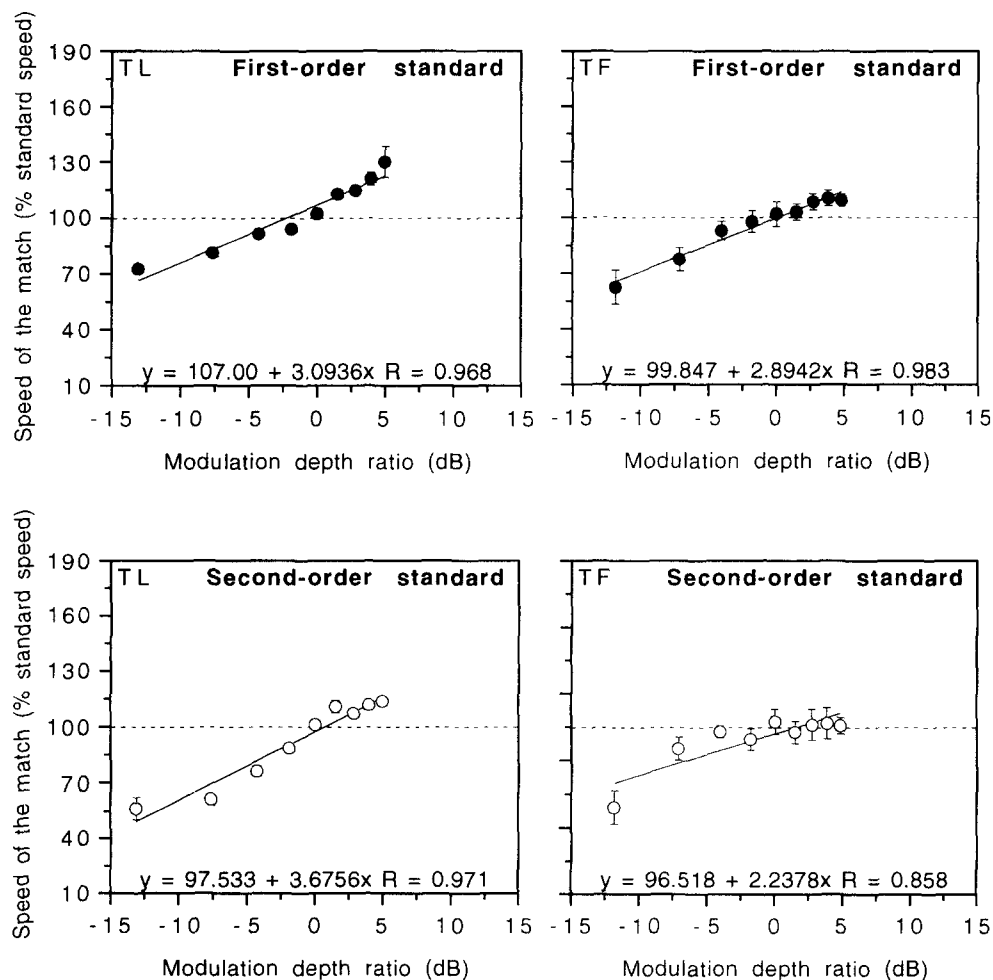


FIGURE 7. Speed matches for two observers as a function of the modulation depth ratio between the standard and match patterns. The match stimulus was always a first-order motion pattern composed of luminance-modulated noise and the standard stimulus was composed of either luminance-modulated noise (●) or contrast-modulated noise (○). The spatial frequency of the modulation was 1 c/deg and the drift speed of the standard was 5 deg/sec (5 Hz). The absolute physical modulation depths of the first-order motion standard stimulus ranged from 0.02 to 0.17 for TL and 0.03–0.17 for TF. The corresponding values for the second-order motion standard stimulus were 0.05–0.37 for TL and 0.06–0.41 for TF. The vertical lines above and below each data point (where visible) represent  $\pm 1$  SE.

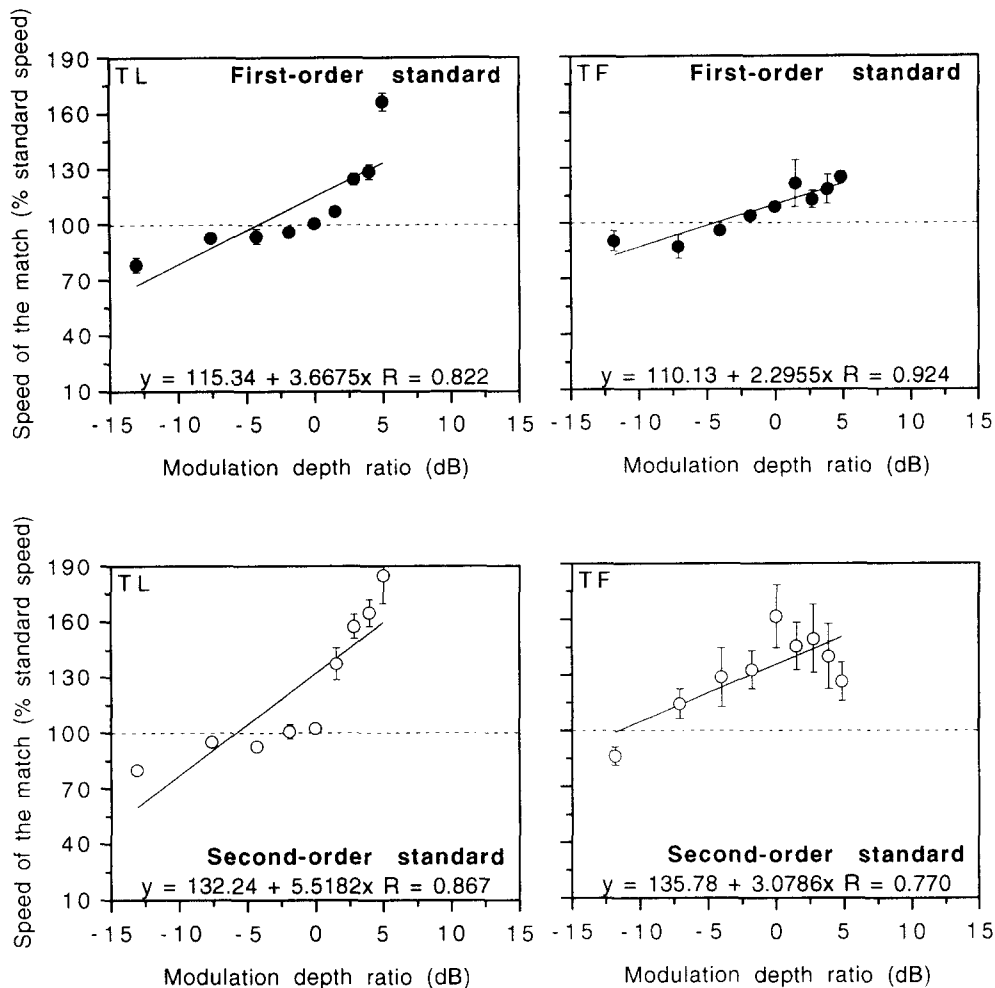


FIGURE 8. Same as Fig. 7 except that the spatial frequency of the modulation was 4 c/deg and the drift speed of the standard was 1.25 deg/sec (5 Hz). The absolute physical modulation depths of the first-order motion standard stimulus ranged from 0.03 to 0.22 for TL and 0.03–0.20 for TF. The corresponding values for the second-order motion standard stimulus were 0.12–0.99 for TL and 0.15–0.99 for TF.

spatial filtering, squaring and a second stage of filtering at a lower spatial frequency. The outputs of the first- and second-order pathways are then pooled in order to compute a local measure of the resultant direction (and implicitly speed) of motion at a given point in the retinal image. Although the model at present does not propose an explicit mechanism for speed extraction, it is evident that by appropriately grouping the outputs of motion energy detectors having similar velocity sensitivities so that they feed their outputs to a given unit in the pooling stage (c.f. Grzywacz & Yuille, 1990; Heeger, 1987) speed sensitivity could be achieved with little modification of the model's architecture (the exact implementation of such a process is unimportant for the present discussion, we aim simply to show that models exploiting separate first-order and second-order motion pathways may be able, in principle, to accommodate the results). The model also explicitly incorporates contrast gain-control mechanisms (i.e. normalization mechanisms) into the motion pathways. These function independently in each pathway but are functionally equivalent. That is, the outputs of the motion energy detectors in the first-order pathway are normalized with respect to the modulation depth

(luminance contrast) of the stimulus, as encoded in that pathway, by dividing the motion energy signal by the output of the spatial filtering stage. In the second-order pathway a similar mechanism divides the motion energy signal by the output of the second spatial filtering stage in this pathway. Although these gain-control mechanisms serve to minimize the effects of stimulus modulation depth on the motion energy detectors outputs, Wilson *et al.* (1992) acknowledge that modulation depth does have at least a modest effect on the perceived speed and direction (e.g. Stone & Thompson, 1992; Stone, Watson & Mulligan, 1990) of first-order motion patterns. Indeed the model's motion energy response is dependent on stimulus modulation depth, especially at modulation depths below 0.1–0.2, where the motion energy computation under-compensates for the effects of modulation depth. At higher modulation depths the response function rapidly asymptotes. Given that these normalization mechanisms do not completely disambiguate motion energy responses related to speed and modulation depth they may offer a common framework within which to interpret the results of the present experiments and those of Thompson (1982) and Stone and Thompson (1992).

In Expt 1, the first- and second-order motion stimuli were equated for visibility by presenting them at the same multiple of direction-identification threshold. If one plausibly assumes that this produces a quantitatively similar output from the spatial filtering stage that precedes motion energy detection in each pathway, then when the motion stimuli are presented simultaneously the outputs of the motion energy detectors in each pathway will be normalized with respect to equivalent first- and second-order modulation depth signals. As a result the two stimuli will have the same perceived speed. However, when the first- and second-order stimuli are presented at different multiples of threshold (Expt 2), the output of the spatial filtering stage preceding motion energy detection in each pathway will be different and this in turn will lead to different (normalized) motion energy outputs from each pathway. As a result, the motion stimulus presented at the higher multiple of threshold will be perceived to have the higher speed.

Although the results have been discussed within the framework of Wilson *et al.*'s (1992) model, as mentioned previously this is not meant to imply that our results could

not be successfully interpreted within other models of second-order motion. For example, the output of the model of Johnston *et al.* (1992) is largely modulation depth-invariant and provided that some assumptions are made concerning the normalization processes involved it may be able to account for the present data. However, the balance of current psychophysical and physiological evidence is consistent with the existence of separate mechanisms for the detection of first- and second-order motion. Psychophysical evidence bearing on this question was cited in the Introduction. A recent physiological study of the responses of motion-sensitive cells in areas 17 and 18 of feline cortex to luminance- and contrast-defined motion (Zhou & Baker, 1993) leads to the same conclusion. In that study, responses to drifting sine gratings and stationary high spatial frequency sine gratings with a drifting contrast modulation were compared. The spatial frequency tuning of cells responsive to both types of motion stimuli was typically lower for contrast modulations than for luminance modulations. This is consistent with the notion that second-order motion and first-order motion are detected

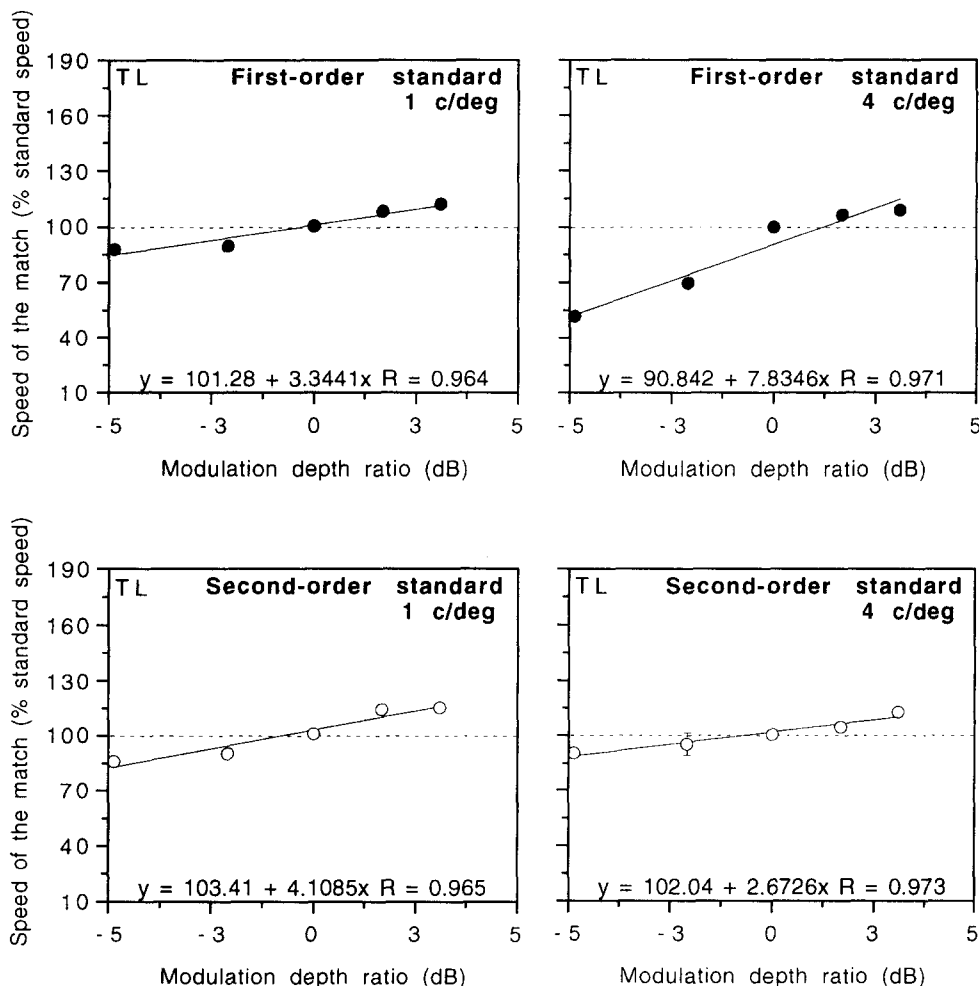


FIGURE 9. Same as Fig. 7 except that the drift speed of the modulation was 20 Hz. Data are shown for one observer only, for each of two spatial frequencies, 1 c/deg (20 deg/sec drift) and 4 c/deg (5 deg/sec drift). The absolute physical modulation depths of the first-order motion standard stimulus ranged from 0.03 to 0.06 when the modulation spatial frequency was 1 c/deg and 0.04–0.10 when the modulation spatial frequency was 4 c/deg. The corresponding values for the second-order motion standard stimulus were 0.09–0.22 (1 c/deg) and 0.36–0.90 (4 c/deg).

separately but that motion signals of the two types are then pooled, as suggested by Wilson *et al.* (1992).

A further point to be considered concerns the implications that our results have for stimuli that contain both first- and second-order motion signals. For example, the perceived direction of a plaid pattern composed of two first-order sinusoids of different contrasts has been shown to be biased in the direction of the higher contrast component (Stone *et al.*, 1990). This phenomenon has been attributed to a reduction in the perceived speed of the lower contrast component. There is also compelling evidence (e.g. Derrington, Badcock & Holroyd, 1992; Gorea & Lorenceau, 1989; Wilson *et al.*, 1992; Yo & Wilson, 1992) that the perceived direction of plaid patterns is not determined solely by the first-order components but also by the direction of motion of the second-order contrast variations ("blobs") formed at the intersections of the components. Our results (Expt 2) are consistent with the suggestion (Perrone & Stone, 1988) that the perceived direction of plaid stimuli may also be biased by altering the modulation depth of the second-order contrast variations relative to that of the first-order components. We are currently conducting experiments in order to address these issues.

## 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.
- Adelson, E. H. & Bergen, J. R. (1986). The extraction of spatiotemporal energy in human and machine vision. In *Workshop on motion: Representation and analysis* (pp. 151–155). Charleston, SC: IEEE Computer Society Press.
- Albrecht, D. G. & Hamilton, D. B. (1982). Striate cortex of monkey and man: Contrast response function. *Journal of Neurophysiology*, 48, 217–237.
- Albright, T. D. (1992). Form-cue invariant motion processing in primate visual cortex. *Science*, 255, 1141–1143.
- Albright, T. D., Desimone, R. & Gross, C. G. (1984). Columnar organization of directionally selective cells in visual area MT of the macaque. *Journal of Neurophysiology*, 51, 16–31.
- Anstis, S. M. (1980). The perception of apparent movement. *Philosophical Transactions of the Royal Society of London B*, 290, 153–168.
- Badcock, D. R. & Derrington, A. M. (1989). Detecting the displacement of spatial beats: No role for distortion products. *Vision Research*, 29, 731–739.
- Burton, G. J. (1973). Evidence for non-linear response processes in the human visual system from measurements on the thresholds of spatial beat frequencies. *Vision Research*, 13, 1211–1225.
- 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. & 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.
- Chubb, C. & Sperling, G. (1989). Second-order motion perception: Space/time separable mechanisms. In *Proceedings: IEEE workshop on visual motion* (pp. 126–138). Washington, DC: IEEE Computer Society Press.
- Derrington, A. M. (1987). Distortion products in geniculate X-cells: A physiological basis for masking by spatially modulated gratings? *Vision Research*, 27, 1377–1386.
- Derrington, A. M. & Badcock, D. R. (1985). Separate detectors for simple and complex grating patterns? *Vision Research*, 25, 1869–1878.
- Derrington, A. M., Badcock, D. R. & Henning, G. B. (1993). Discriminating the direction of second-order motion at short stimulus durations. *Vision Research*, 33, 1785–1794.
- Derrington, A. M., Badcock, D. R. & Holroyd, S. A. (1992). Analysis of the motion of 2-dimensional patterns: Evidence for a second-order process. *Vision Research*, 32, 699–707.
- Foster, K. H., Gaska, J. P., Nagler, M. & Pollen, D. A. (1985). Spatial and temporal frequency selectivity of neurones in visual cortical areas V1 and V2 of the macaque monkey. *Journal of Physiology, London*, 365, 331–363.
- Gorea, A. & Lorenceau, J. (1989). Motion perception in compound stimuli is 'blob'-dependent. *Investigative Ophthalmology and Visual Science (Suppl.)*, 30, 388.
- Grzywacz, N. M. (1992). One-path model for contrast-independent perception of Fourier and non-Fourier motions. *Investigative Ophthalmology and Visual Science (Suppl.)*, 33, 954.
- Grzywacz, N. M. & Yuille, A. L. (1990). A model for the estimate of local image velocity by cells in the visual cortex. *Proceedings of the Royal Society of London B*, 239, 129–161.
- Harris, L. R. & Smith, A. T. (1992). Motion defined by second-order characteristics does not evoke optokinetic nystagmus. *Visual Neuroscience*, 9, 565–570.
- Hawken, M. J., Gegenfurtner, K. R. & Tang, C. (1994). Contrast dependence of colour and luminance motion mechanisms in human vision. *Nature*, 367, 268–270.
- Heeger, D. J. (1987). Model for the extraction of image flow. *Journal of the Optical Society of America A*, 4, 1455–1471.
- Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience*, 9, 181–197.
- Henning, G. B., Hertz, B. G. & Broadbent, D. E. (1975). Some experiments bearing on the hypothesis that the visual system analyses spatial patterns in independent bands of spatial frequency. *Vision Research*, 15, 887–897.
- Holub, R. A. & Morton-Gibson, M. (1981). Response of visual cortical neurons of the cat to moving sinusoidal gratings: Response contrast functions and spatiotemporal interactions. *Journal of Neurophysiology*, 46, 1244–1259.
- Johnston, A., McOwan, P. W. & Buxton, H. (1992). A computational model of the analysis of some first-order and second-order motion patterns by simple and complex cells. *Proceedings of the Royal Society of London B*, 250, 297–306.
- Ledgeway, T. & Smith, A. T. (1994a). Evidence for separate motion-detecting mechanisms for first- and second-order motion in human vision. *Vision Research*, 34, 2727–2740.
- Ledgeway, T. & Smith, A. T. (1994b). The perceived speed of second-order motion patterns. *Investigative Ophthalmology and Visual Science (Suppl.)*, 35, 1405.
- Lieberman, H. R. & Pentland, A. P. (1982). Microcomputer-based estimation of psychophysical thresholds: The best PEST. *Behaviour Research Methods and Instrumentation*, 14, 21–25.
- MacLeod, D. I. A., Williams, D. R. & Makous, W. (1992). A visual nonlinearity fed by single cones. *Vision Research*, 32, 347–363.
- Marr, D. & Ullman, S. (1981). Directional selectivity and its use in early visual processing. *Proceedings of the Royal Society of London B*, 211, 151–180.
- Mather, G. & West, S. (1993). Evidence for second-order motion detectors. *Vision Research*, 33, 1109–1112.
- Newsome, W. T., Gizzi, M. S. & Movshon, J. A. (1983). Spatial and temporal properties of neurones in macaque MT. *Investigative Ophthalmology and Visual Science (Suppl.)*, 24, 106.
- Nishida, S. (1993). Spatiotemporal properties of motion perception for random-check contrast modulations. *Vision Research*, 33, 633–645.
- Olavarria, J. F., DeYoe, E. A., Knierim, J. J., Fox, J. M. & Van Essen, D. C. (1992). Neural responses to visual texture patterns in middle temporal area of the macaque monkey. *Journal of Neurophysiology*, 68, 164–181.
- Pantle, A. & Turano, K. (1992). Visual resolution of motion ambiguity

- with periodic luminance- and contrast-domain stimuli. *Vision Research*, 32, 2093–2106.
- Pentland, A. P. (1980). Maximum likelihood estimation: The best PEST. *Perception & Psychophysics*, 28, 377–379.
- Perrone, J. & Stone, L. S. (1988). Two-dimensional motion analysis: Information in the nodes of moving plaids. *Optical Society of America Technical Digest*, 11, 141.
- Smith, A. T. (1994). Correspondence-based and energy-based detection of second-order motion in human vision. *Journal of the Optical Society of America A*, 11, 1940–1948.
- Smith, A. T. & Edgar, G. K. (1990). The influence of spatial frequency on perceived temporal frequency and perceived speed. *Vision Research*, 30, 1467–1474.
- Smith, A. T. & Hammond, P. (1986). Hemifield differences in perceived velocity. *Perception*, 15, 111–117.
- Stone, L. S. & Thompson, P. (1992). Human speed perception is contrast dependent. *Vision Research*, 32, 1535–1549.
- Stone, L. S., Watson, A. B. & Mulligan, J. B. (1990). Effect of contrast on the perceived direction of a moving plaid. *Vision Research*, 30, 1049–1067.
- Thompson, P. (1982). Perceived rate of movement depends on contrast. *Vision Research*, 22, 377–380.
- Thompson, P., Stone, L. S., Swash, S. & Stone, R. (1994). Contrast dependence of speed perception: Effects of background contrast. *Investigative Ophthalmology and Visual Science (Suppl.)*, 35, 2077.
- Watson, A. B. & Ahumada, A. J. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A*, 2, 322–341.
- Weibull, W. (1951). A statistical distribution function of wide applicability. *Journal of Applied Mechanics*, 18, 292–297.
- Werkhoven, P., Sperling, G. & Chubb, C. (1993). The dimensionality of texture-defined motion: a single channel theory. *Vision Research*, 33, 463–485.
- Wilson, H. R. & Mast, R. (1993). Illusory motion of texture boundaries. *Vision Research*, 33, 1437–1446.
- Wilson, H. R., Ferrera, V. P. & Yo, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Visual Neuroscience*, 9, 79–97.
- Yo, C. & Wilson, H. R. (1992). Perceived direction of moving two-dimensional patterns depends on duration, contrast and eccentricity. *Vision Research*, 32, 135–147.
- Zhou, Y. & Baker, C. L. Jr (1993). A processing stream in mammalian visual cortex neurons for non-Fourier responses. *Science*, 261, 98–101.

---

*Acknowledgements*—The experiments described in this paper have previously been published in abstract form (Ledgeway & Smith, 1994b). The work was supported in part by a project grant to A. T. Smith from the Science and Engineering Research Council and was carried out at the School of Psychology, University of Wales, Cardiff, Wales.