PERGAMON

Vision Research 40 (2000) 3585–3597

Vision Research

www.elsevier.com/locate/visres

The properties of the motion-detecting mechanisms mediating perceived direction in stochastic displays

Timothy Ledgeway *, Robert F. Hess

McGill Vision Research Unit, 687 Pine Avenue West, Rm. H4-14, Montreal, Que., Canada H3A 1A1

Received 28 September 1999; received in revised form 14 June 2000

Abstract

Previous studies [e.g. Baker & Hess, 1998. *Vision Research*, *38*, 1211–1222] have shown that perceived direction in displays composed of multiple, limited-lifetime, Gabor micropatterns (G) is influenced by movement both at the fine spatial scale of the internal luminance modulation (first-order motion) and the coarse spatial scale of the Gaussian, contrast window (second-order motion). However it is presently indeterminate as to whether this pattern of results is indicative of the processes by which first-order and second-order motion signals interact within the visual system per se or those by which motion information, irrespective of how it is defined, is utilised across different spatial scales. To address this issue, and more generally the properties of the mechanisms that analyse motion in such displays, we employed stochastic motion sequences composed of either G, G added to a static carrier (G + C) or G multiplied with a carrier (G*C). Crucially G*C, unlike both G and G + C, micropatterns contain no net first-order motion and second-order motion only at the scale of the internal contrast modulation. For small displacements perceived direction in all cases showed a dependence on the internal sinusoidal spatial structure of the micropatterns and characteristic oscillations were typically observed, consistent with models in which first-order motion and second-order motion are encoded on the basis of similar low-level mechanisms. Importantly for larger displacements, and also when the internal spatial structure was randomised on successive exposures (so that motion at this spatial scale was unreliable), performance tended to be veridical for all types of micropattern, even though under these conditions displacements of the G*C micropatterns should have been invisible to current, low-level, motion-detecting schemes. This suggests that both low-level motion sensors and mechanisms utilising a different motion-detecting strategy such as high-level, attentive, feature-tracking may mediate perceptual judgements in stochastic displays. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: First-order motion; Second-order motion; Motion energy; Feature-tracking; Spatial scale

1. General introduction

Drifting contours in the retinal image may be defined by either local intensity or local wavelength differences (first-order motion) or by variations in higher-order, derived image properties such as local contrast, texture and flicker (second-order motion) (Chubb & Sperling, 1988; Cavanagh & Mather, 1989)[1]. Both first-order motion and second-order motion can be readily perceived and the processes by which this is achieved are of fundamental importance to our understanding of vision.

First-order, intensity-based, motion processing has been studied extensively (for reviews see Smith & Snowden, 1994) and computational schemes of the underlying mechanisms involved fall into two general classes. First, it has been suggested (e.g. Anstis, 1970; Braddick, 1974; Van Santen & Sperling, 1985) that motion is extracted directly from raw intensity variations in the retinal image using arrays of pre-attentive, low-level, hardwired sensors that have been identified with direction-selective cortical neurones (Hubel & Wiesel, 1968). Indeed first-order motion sensors have been modelled as motion energy detectors using receptive fields that

---

* Corresponding author.
E-mail address: tledgewa@vision.mcgill.ca (T. Ledgeway).

[1] In the present paper the terms 'first-order motion' and 'second-order motion' refer specifically to spatiotemporal variations in particular image statistics (Cavanagh & Mather, 1989), rather than the properties of the mechanisms that may mediate perception in each case, in order to avoid any potential ambiguity.

PII: S0042-6989(00)00202-9

are oriented and localised in space-time (Adelson & Bergen, 1985). Second, an alternative class of models exploit a different computational principle in which explicit image features (e.g. edges, boundaries and objects) are identified and then tracked and/or matched over time using post-attentive, high-level, cognitive processes (Ullman, 1979; Anstis, 1980; Braddick, 1980; Cavanagh, 1992). There is evidence that both strategies are used in human vision and the relative efficacy of each is influenced to some extent by the viewing conditions adopted such as the presence of an inter-stimulus-interval (ISI) and whether motion is presented monocularly or dichoptically (Anstis, 1980; Georgeson & Shackleton, 1989; Mather, Cavanagh, & Anstis, 1980; Georgeson & Harris, 1990).

There is emerging evidence that both low-level motion detectors (Chubb & Sperling, 1988; Cavanagh & Mather, 1989; Johnston, McOwan, & Buxton, 1992; Werkhoven, Sperling, & Chubb, 1993; Nishida, 1993) and high-level, feature-tracking (Cavanagh, 1991; Lu & Sperling, 1995b; Seiffert & Cavanagh, 1998; Derrington & Ukkonen, 1999) may also mediate second-order motion perception. In the case of low-level mechanisms, the weight of current evidence (for reviews see Smith, 1994a; Nishida, Ledgeway, & Edwards, 1997; Baker, 1999; Clifford & Vaina, 1999) favours the suggestion (e.g. Wilson, Ferrera, & Yo, 1992) that second-order motion detection has a similar basis to that proposed for first-order motion, but that each is encoded, at least initially, by distinct (separate) mechanisms. In order to make the second-order spatiotemporal image structure explicit, however, some form of gross nonlinear processing, prior to motion analysis, is typically deemed necessary. This is exemplified by the model of Wilson et al. (1992) in which second-order motion is extracted by means of a specialised processing pathway that utilises the 'filter-rectify-filter' (FRF) principle originally embodied in models of spatial texture segregation (e.g. Sutter, Beck, & Graham, 1989). Specifically the retinal image is convolved with an array of spatial-frequency-selective filters, subjected to a nonlinearity (such as full-wave rectification) and then secondary filtering at a lower (e.g. 1 octave) spatial frequency. Motion is detected in the resultant neural image using conventional, low-level motion sensors. First-order motion is encoded by a separate pathway but the two pathways subsequently converge to determine the net image motion. In the case of high-level feature-tracking, second-order motion may be detected on this basis, at least some of the time, particularly under viewing conditions analogous to those believed to favour feature-tracking of first-order motion (Smith, 1994b; Lu & Sperling, 1995a).

In natural vision motion information is likely to be conveyed by a mixture of first-order and second-order cues, rather than each occurring in isolation, and an important issue concerns the nature and extent to which the two varieties of motion interact perceptually when both are simultaneously present (Derrington, Badcock, & Holroyd, 1992; Stoner & Albright, 1992; Wilson et al., 1992; Yo & Wilson, 1992). In this respect Baker and colleagues (Boulton & Baker, 1993, 1994; Bex & Baker, 1997; Baker & Hess, 1998) and Clifford, Freedman, and Vaina (1998) have systematically investigated perceived direction, as a function of spatial displacement, in stochastic motion displays containing multiple, 'limited lifetime', Gabor micropatterns (patches of Gaussian-windowed sinusoidal luminance grating) which contain both first-order and second-order motion information. Baker and Hess (1998), for example, found that for small displacements perceived direction exhibited characteristic oscillations yoked to the centre spatial frequency of the luminance sinusoid (first-order motion) within each micropattern. Under these conditions performance was qualitatively consistent with the predictions of a low-level, first-order motion, energy model (Adelson & Bergen, 1985). For larger displacements, however, performance tended to be veridical, even when the orientation of the luminance sinusoid was orthogonal on successive exposures, consistent with the operation of a mechanism sensitive to second-order motion at the (coarse) scale of the Gaussian contrast envelope. Thus performance appeared to be an envelope of two qualitatively distinct types of motion perception, one governed by the internal first-order structure of the micropatterns and the other by the second-order structure of the Gaussian window.

Studies that have utilised displays containing multiple Gabor micropatterns have tended to assume that the mechanisms that detect second-order motion at the scale of the Gaussian envelopes are low-level. This is because the stimuli are stochastic, have a limited lifetime and are composed of dense arrays of elements, conditions believed to preclude the operation of high-level, feature-tracking mechanisms. However the validity of this assumption is indeterminate as several findings could equally be accommodated within a framework in which motion at the scale of the envelope is detected by an attentive, feature-tracking mechanism. For example Baker and Hess (1998) showed that performance, at least at large displacements, deteriorates rapidly when either the micropattern lifetime is reduced to a single displacement or when motion noise (incoherent random motion) is introduced. Both of these manipulations would be expected to prohibit attention-based tracking of individual Gabors, or Gabor clusters, over time and consequently have a detrimental effect on performance. Thus there are two opposing views on how second-order motion is processed and it remains a possibility that the second-order envelope motion in such displays could be detected on the basis of either low-level, pre-attentive motion sensors or reflect the involvement of a high-level, attentive, feature-tracking mechanism or even a combination of both processes.

To address this issue we compared direction-identification performance for arrays of either conventional Gabor micropatterns (G), Gabor micropatterns added to a static luminance carrier (G + C) or Gabor micropatterns multiplied with a carrier (G*C), over a range of displacements (see Fig. 1a). G and G + C micropatterns contain both first-order motion at the

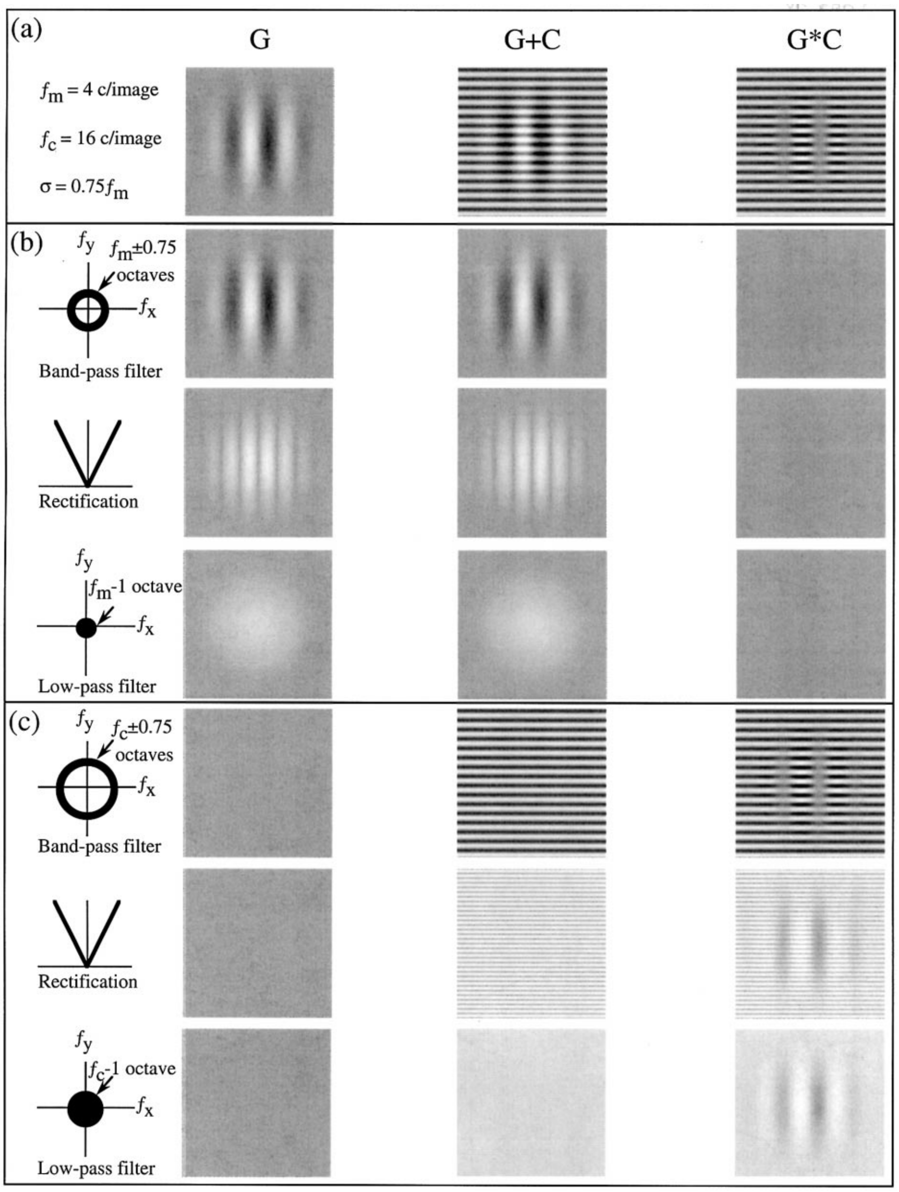

Fig. 1. (a) Isolated examples of the micropattern stimuli used in the present study. G micropatterns were composed of a patch of sinusoidal luminance grating (f_m), within a smooth 2-d Gaussian window. G + C micropatterns were constructed by adding a Gabor to a high spatial frequency, horizontally-oriented, sinusoidal luminance carrier (f_c) that filled the entire display area and G*C micropatterns were constructed by multiplying a Gabor with a static sinusoidal carrier. (b) and (c) illustrate for each type of micropattern the consequences of applying the 'filter-rectify-filter' (FRF) principle that is widely embodied in current computational models of motion detection (e.g. Wilson et al., 1992) in order to expose the second-order spatial (and temporal) structure to conventional low-level motion analysis. In (b) each image is first convolved with an array of bandpass spatial-frequency-selective filters centred on the modulation ($f_m \pm 0.75$ octaves), the outputs of which are then subjected to a nonlinearity (full-wave rectification) and secondary filtering at a lower (e.g. 1 octave below f_m) spatial frequency. For the G and G + C micropatterns this procedure produces a neural image in which the Gaussian window is made explicit. For the G*C micropatterns no spatial structure is evident in the resulting neural image (indicated by a homogenous grey field). In (c) the same FRF process is applied to each of the micropatterns but now the first spatial filtering stage is centred on the carrier ($f_c \pm 0.75$ octaves) and the second filtering stage at 1 octave below f_c. For both the G and G + C micropatterns the resulting neural images contain no spatial structure but for the G*C micropatterns the neural image is analogous to a simple Gabor pattern and conventional, low-level, motion sensors are sensitive only to the internal spatial structure of such a pattern and not the Gaussian envelope.

scale of the internal sinusoidal luminance modulation and second-order motion at the scale of the Gaussian envelope but G*C micropatterns contain second-order motion only at the scale of the internal sinusoidal contrast modulation and no net first-order motion. If first-order motion and second-order motion are indeed encoded predominantly, or exclusively, on the basis of qualitatively similar low-level mechanisms, as has been proposed, then for small displacements perceived direction in all cases (G, G + C and G*C displays) should show a similar dependence on the local internal spatial structure of the micropatterns and specifically exhibit characteristic oscillations yoked to their periodicity (cf. Baker & Hess, 1998). For a range of larger displacements, and also when the internal spatial structure is randomised on successive exposures, performance should tend to be veridical for both the G and G + C micropatterns as low-level, second-order motion sensors should still be able to extract the direction of the Gaussian contrast envelope (Fig. 1b). Importantly for the G*C micropatterns performance might be expected to be at chance under these conditions since current low-level, motion-detecting schemes will not be sufficient to expose motion at the coarse scale of the Gaussian envelope (Taub, Victor, & Conte, 1997). For example applying the ubiquitous FRF principle to an individual micropattern of this type (see Fig. 1c and legend for details) will yield a neural image that is analogous to a simple Gabor pattern and conventional, low-level motion sensors are sensitive only to the internal spatial structure of such a pattern and not the Gaussian envelope. In principle applying the FRF scheme twice in succession to the original G*C image could make the Gaussian envelope available to standard motion sensors but this would then render the internal spatial structure invisible. If, on the other hand, attentive, high-level, feature-tracking strategies are indeed utilised to some extent for detecting the displacements of micropatterns, irrespective of how those patterns are defined, then at least at large displacements performance should be qualitatively similar and veridical for G, G + C and G*C displays even when the internal spatial attributes of the G*C micropatterns are randomised on consecutive frames (providing unreliable signals for low-level mechanisms that encode second-order motion).

2. Experiment 1: direction-identification with displays composed of multiple, limited-lifetime, conventional Gabor (G) micropatterns

In order to obtain baseline measures of performance our principal objective, in Experiment 1, was to investigate how perceived direction varies as a function of displacement for stochastic displays composed of conventional Gabor (G) micropatterns (patches of sinusoidal grating windowed by a spatial Gaussian envelope) under conditions analogous to those used in previous studies. The purpose of this was 3-fold: (1) to establish the robustness of the results obtained in previous investigations that have employed stimuli of this type; (2) to provide a control condition against which the effects, if any, of the presence of the static carrier used in subsequent experiments could be assessed; and (3) to investigate further the consequences of changing the internal spatial structure of each micropattern on successive exposures. The latter is important because in the study of Baker and Hess (1998) the good performance found when the orientation of each micropattern was alternated on successive exposures could be due, at least in part, to observers perceiving first-order motion between the micropatterns presented on every other exposure (in which the orientations were identical) rather than second-order motion of the contrast envelope between micropatterns presented consecutively (those with orthogonal orientations).

2.1. Methods

2.1.1. Observers

Three observers participated in the experiment and each had normal or corrected-to-normal acuity. Observers TL and RFH were the authors and SOD was an experienced psychophysical observer who was naive to the purpose of the experiment.

2.1.2. Apparatus and stimuli

Motion stimuli were computer generated and displayed on a monochrome monitor (with a frame rate of 75.5 Hz) which was carefully gamma-corrected with the aid of internal look up tables. As an additional precaution psychophysical procedures were used to ensure that any residual luminance nonlinearities were minimised (see Ledgeway & Smith, 1994; Nishida et al., 1997). Stimuli were presented within a square display area subtending 20.4° vertically and horizontally at the viewing distance of 49 cm. The mean luminance of the remainder of the display (which was homogeneous) was ~ 24 cd/m^2. Viewing was binocular and a prominent fixation spot was located at the centre of the display in order to minimise eye movements.

Motion sequences were composed of ten frames, each of duration 106 ms, containing a pseudorandom array of 15 non-overlapping G micropatterns (see Fig. 1a, left panel) that were displaced by the same amount in the same direction (either leftwards or rightwards) on consecutive image updates. When a micropattern reached the boundary of the square display area it was 'wrapped-around' and immediately reappeared on the opposite edge of the display window. All micropatterns had a lifetime of four displacements. That is, at the beginning of a motion sequence each micropattern was initially assigned a random 'age', nominally between one and four displacements, that was independent of those assigned to the other micropatterns present. On each subsequent positional update this 'age' parameter was incremented by one and whenever this exceeded the limit of four displacements the micropattern was replotted randomly within the display area and its 'age' parameter was reset to 1. Micropatterns falling within a central exclusion zone (radius 1.61°) were not plotted in order to aid fixation and prevent ocular tracking of the motion stimuli.

Each G micropattern was composed of a patch of sinusoidal luminance grating, within a smooth two-dimensional (2-d) Gaussian window with a standard deviation (σ) of 0.73°, that was spatially truncated at $\pm 1.61°$. Prior to spatial windowing by the Gaussian envelope, the luminance modulation depth (Michelson contrast) of the sinusoidal waveform was 0.08. The spatial characteristics of the sinusoid were manipulated under two basic conditions: In the *standard condition*, for all micropatterns, the sinusoidal modulation was always vertically-oriented (0°), had a spatial frequency (f_m) of 1c/° and was in sine phase (0°) with respect to the centre of the Gaussian window. In the *random condition* all three parameters were independently randomised for each micropattern on every frame of the motion sequence and were uniformly assigned values in the range $\pm 90°$ for orientation, 0.5–2c/° for f_m and $\pm 180°$ for spatial phase with respect to the envelope. This was done in order to render the internal spatial characteristics of the micropatterns unreliable as signals for low-level, first-order motion detectors. As pointed out by Baker and Hess (1998) conventional linear motion models, such as that exemplified by the motion-energy model of Adelson and Bergen (1985), are selective to some extent to the local orientation and spatial frequency of luminance variations in the retinal image and also to the sign and magnitude (velocity) of temporal variations in spatial phase. Such a detector would be expected to give little, or no, consistent response to micropatterns in which these characteristics are uncorrelated over time.

2.1.3. Procedure

On each trial the observer was presented with a motion sequence in which all micropatterns were displaced by the same amount in the same direction on successive exposures. This direction was chosen at random at the beginning of each trial and could be either leftwards or rightwards with equal probability. The magnitude of the spatial displacement was varied from trial to trial, using the method of constant stimuli, and was chosen at random from a set of eight equally spaced values (selected on the basis of pilot studies) that ranged from either 0.04–0.89° (in order to adequately sample the range over which oscillations in perceived direction were expected) or 1.37–4.76° (to cover the broad range of displacements over which perception is veridical before reaching chance levels). The task of the observer was to indicate the perceived direction of motion using one of two response buttons. Observers completed four runs of 80 trials for each range of displacements, which were measured in separate randomised blocks of trials, for both the *standard condition* and the *random condition* and the order in which the runs were completed was randomised for each observer. Results are plotted as the percentage errors in direction-identification performance as a function of the spatial displacement of the G micropatterns (expressed as a fraction of the spatial period of f_m in the *standard condition*) for each condition.

2.2. Results and discussion

2.2.1. Standard condition

Fig. 2 shows the results obtained for three observers in the *standard condition* (i.e. when the orientation, spatial frequency and spatial phase of all G micropatterns were fixed). It is clear that for two of the observers (RFH and SOD) for relatively small displacements (typically < 1.5 spatial periods) direction-identification performance exhibits characteristic oscillatory behaviour linked to the periodicity of the windowed sinusoidal modulation. For displacements less than 0.5 of a period, performance is almost perfect (errors are close to 0% for both of these observers) and then rises rapidly to 90% errors or more for displacements close to 0.75 cycles of the modulation frequency. This indicates that for RFH and SOD under these conditions the reported direction of motion was consistently opposite that of the overall micropattern displacement. Between 0.75 and 1 spatial periods performance drops rapidly to chance levels (50% errors) but is generally veridical again for displacements around 1.5 spatial periods. Interestingly the results of observer TL do not follow this cyclical pattern and direction judgements always coincide with those of the micropatterns (0% errors) for this range of displacements. However for displacements beyond ~ 1.5 spatial periods all three observers show a very similar pattern of results in that error rates eventually rise, although performance is consistently better than chance levels (50%) for much of this range.

2.2.2. Random condition

When the spatial characteristics of the sinusoidal luminance modulation were randomised for each G micropattern over time (Fig. 3), the results obtained for each of the three observers were essentially identical. At the smallest displacements direction judgements were at chance and then error rates drop markedly to levels of

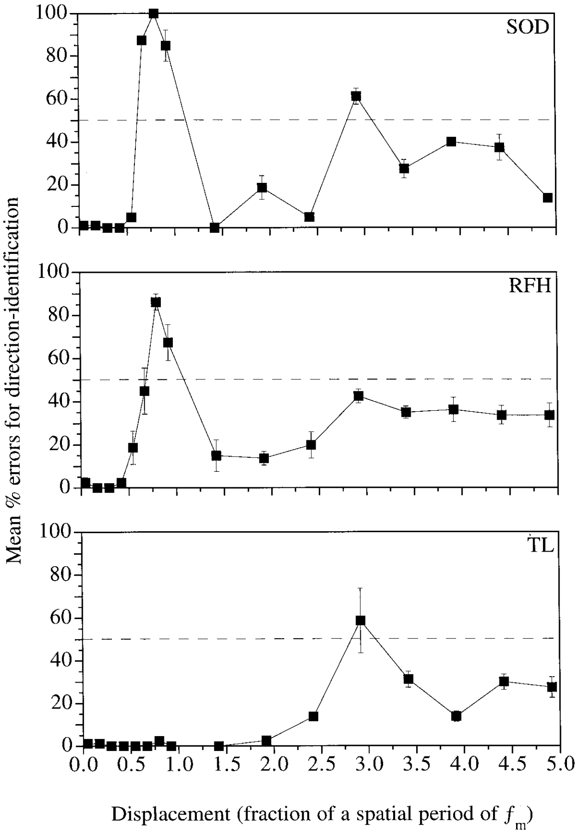

Fig. 2. Direction-identification performance for three observers as a function of the spatial displacement for motion stimuli containing multiple, limited-lifetime, G micropatterns in the *standard condition*. Each G micropattern was composed of a patch of sinusoidal luminance grating, within a smooth two-dimensional (2-d) Gaussian window with a standard deviation (σ) of 0.73°, that was spatially truncated at $\pm 1.61°$. Prior to spatial windowing by the Gaussian envelope, the luminance modulation depth (Michelson contrast) of the sinusoidal waveform was 0.08. The spatial characteristics of the sinusoid were fixed and for all micropatterns, the sinusoidal modulation was always vertically-oriented (0°), had a spatial frequency (f_m) of 1c/° and was in sine phase (0°) with respect to the centre of the Gaussian window. The vertical bars above and below each data point (where visible) represent ± 1 SEM based on variability between runs of trials.

between 0 and 13% for spatial displacements of the order of 0.8 periods (also equivalent to a spatial jump of 0.8°). Thereafter errors in direction-identification performance tend to increase with the magnitude of the displacement but once again performance is clearly much better than chance over a large range of displacements. Similar results were obtained (data not shown) when only one of the spatial characteristics of the sinusoidal luminance modulation (either the orientation, spatial frequency or relative phase) was randomised and the others were fixed.

Overall the results of this experiment are in good agreement with those of previous studies (e.g. Baker & Hess, 1998), that have examined direction perception with comparable stimuli, with the exception that TL's results in the *standard condition* do not exhibit any characteristic cyclical behaviour. If performance under this condition reflects an envelope of two qualitatively distinct types of motion perception, one governed by the internal first-order structure of the micropatterns and the other by the second-order structure of the Gaussian window, then it is apparent that for TL the overall perceived direction of drift is determined predominantly by the second-order motion present. This is also true for the other two observers, at least for relatively large spatial displacements, but as the displacement magnitude decreases first-order motion signals appear to have an increasingly dominant influence on their perceptual judgements (particularly for SOD) as evidenced by the marked oscillations in perceived direction found. As discussed previously (see Section 1 such oscillations are well predicted by low-level models of first-order motion detection, that operate directly on raw intensity variations in the retinal image, without the need to explicitly identify and track image features over time.

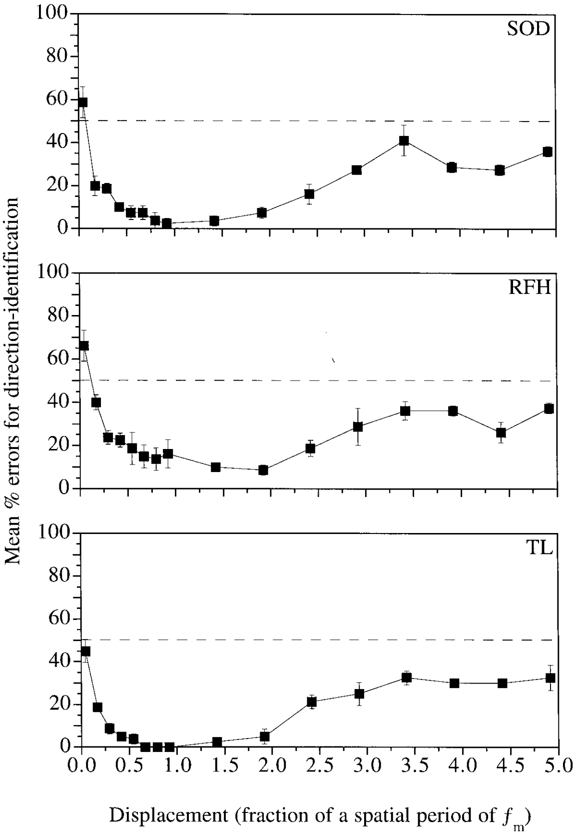

Fig. 3. Legend as for Fig. 2 with the exception that the spatial characteristics of the internal sinusoidal luminance modulation were independently randomised for each micropattern on every frame of the motion sequence and were uniformly assigned values in the range $\pm 90°$ for orientation, 0.5–2c/° for f_m and $\pm 180°$ for spatial phase with respect to the envelope. This was done, in the *random condition*, in order to render unreliable the internal spatial characteristics of the micropatterns as signals for low-level, first-order motion sensors.

In the *random condition* direction judgements were always governed by the second-order motion of the micropatterns, a result to be expected given the extensive measures taken to ensure that any first-order motion in the stimulus was rendered unreliable and noisy, even at the smallest displacements. At the very smallest displacements of between 0.04 and 0.5 periods performance for each of the observers was typically poor but was almost errorless in the *standard condition* for the same range of displacements. Importantly this implies that for all observers (including TL) the near perfect performance in the *standard condition* over this displacement range must have been based to some extent on the first-order motion of the micropatterns.

3. Experiment 2: direction-identification with displays composed of multiple, limited-lifetime, G + C and G*C micropatterns

Experiment 1 confirmed previous reports that perceived direction in displays composed of multiple G micropatterns can be influenced (to differing degrees) by movement both at the spatial scale of the internal sinusoidal modulation (in this case first-order motion) and at the coarse spatial scale of the Gaussian window (in this case second-order motion) and that the relative efficacy of each depends on the magnitude of the spatial displacement. Although the oscillations in performance found, under some conditions, are consistent with the properties of low-level, first-order, motion-sensors the nature of the mechanisms that extract second-order motion are still in question. If low-level mechanisms exist to extract second-order motion and these utilise qualitatively similar principles to those used to encode first-order motion, as has been suggested by several authors (see Section 1), then G*C micropatterns (which contain no net first-order motion and second-order motion only at the scale of the internal sinusoidal modulation) should also produce oscillations in perceived direction under conditions similar to those found with G micropatterns in Experiment 1. If, on the other hand second-order motion is encoded either predominantly, or exclusively, by mechanisms operating on different principles (e.g. attentive, high-level, feature tracking) then such oscillations in performance may not occur. Furthermore, if feature-based strategies do indeed play a role in detecting the displacements of micropatterns, irrespective of the particular image attributes that define the Gaussian envelope, then veridical motion perception may be possible even when the second-order motion signals in the stimulus are rendered unreliable and noisy (by randomisation of the internal spatial structure of the micropatterns over time). It is important to emphasise that under such circumstances the stimulus contains neither any net first-order motion nor any net second-order motion that the observers could use to reliably indicate the direction of motion. We sought, in Experiment 2, to investigate these possibilities and in order to control for any confounding effects of the presence of the static carrier on motion perception, performance for comparable G + C micropatterns was also measured.

3.1. Methods

3.1.1. Observers

Observers TL, RFH and SOD were the same observers that participated in Experiment 1.

3.1.2. Apparatus, stimuli and procedure

The apparatus, stimuli and procedure were identical to those used in Experiment 1 with the following exceptions. All motion sequences now contained a static (unless otherwise stated), sinusoidal luminance carrier that filled the entire display area (subtending 20.4° vertically and horizontally) to which multiple drifting Gabors were either added (G + C micropatterns) or multiplied (G*C micropatterns). In the case of the G + C micropatterns this produced motion sequences in which the local luminance of the static carrier was spatially modulated by each drifting Gabor waveform. For the G*C micropatterns the local contrast, rather than luminance, of the carrier was modulated in an analogous manner. For each motion sequence the orientation of the carrier on every frame was horizontal (90°), its spatial frequency (f_c) was always 4c/° (i.e. 2 octaves higher than f_m) and its luminance modulation depth (Michelson contrast) was 0.3.

For each G + C micropattern the modulation depth of the internal sinusoid was 0.08, prior to spatial windowing by the Gaussian waveform. As absolute sensitivity to second-order motion is ~ 10 times less (e.g. Smith, Hess, & Baker, 1994; Ledgeway & Smith, 1994, 1995; Nishida et al., 1997) than that to comparable first-order motion, in order to ensure that the sinusoidal contrast modulation in each G*C micropattern was clearly visible, its modulation depth was set to 0.8, prior to spatial windowing by the Gaussian envelope. Importantly for the present experiment pilot studies revealed that the particular choice of modulation depth values employed was not critical provided that the sinusoidal structure of the two types of micropattern was clearly visible. Furthermore, given that the G + C micropatterns contain both first-order motion at the scale of the internal sinusoidal luminance modulation and second-order motion at the scale of the Gaussian window and that the G*C micropatterns contain only second-order motion, it is evident that it was not possible to equate perfectly the two types of micropatterns for differences in motion sensitivity. In any case this would be unnecessary given that the purpose of this

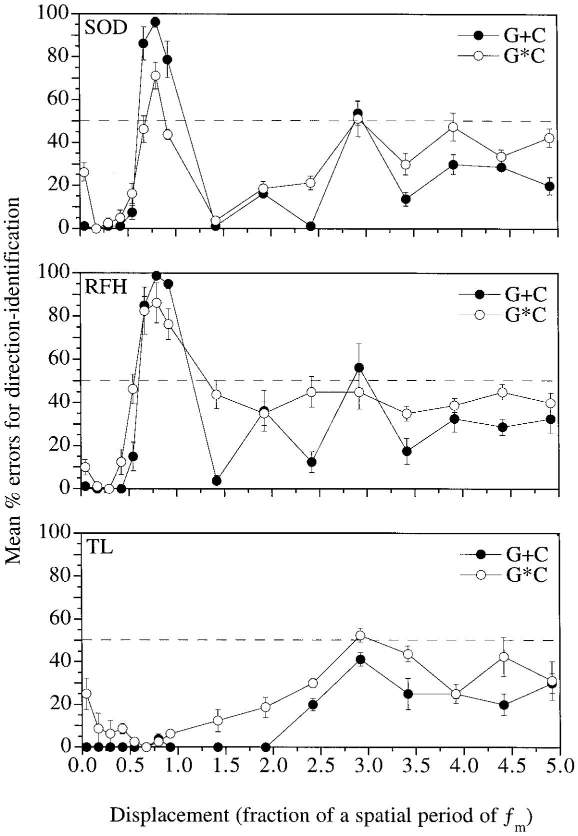

Fig. 4. Direction-identification performance for three observers as a function of the spatial displacement for motion stimuli containing either G + C micropatterns (filled circles) or G*C micropatterns (open circles) in the *standard condition*. All motion sequences contained a static, sinusoidal luminance carrier that filled the entire display area to which multiple drifting Gabors were either added (G + C micropatterns) or multiplied (G*C micropatterns). The drifting sinusoid modulation within each micropattern was always vertically-oriented (0°), had a spatial frequency (f_m) of 1c/° and was in sine phase (0°) with respect to the centre of the Gaussian window. The orientation of the carrier on every frame was horizontal (90°), its spatial frequency (f_c) was always 4c/° (i.e. 2 octaves higher than f_m) and its luminance modulation depth (Michelson contrast) was 0.3.The vertical bars above and below each data point (where visible) represent ± 1 SEM based on variability between runs of trials.

experiment was to compare qualitative, rather than quantitative, aspects of performance for the G*C and G + C micropatterns.

Again performance was measured under two basic conditions in which either the spatial parameters (orientation, spatial frequency and spatial phase) of all the G + C and G*C micropatterns were fixed (*standard condition*) or were independently randomised for each micropattern on every frame of the motion sequence (*random condition*).

3.2. Results and discussion

3.2.1. Standard condition

Fig. 4 shows the results for the three observers for motion sequences containing G + C micropatterns and those containing G*C micropatterns. It is evident that each observer's performance with the G + C micropatterns was extremely similar, both qualitatively and quantitatively, to that obtained with the G micropatterns shown in Fig. 2. This is important in that it demonstrates that the presence of the static, high spatial frequency grating carrier had little, or no influence, on the pattern of results obtained. Indeed this was also true when a vertically-oriented, rather than a horizontally-oriented, grating carrier was used (data not shown). For observers SOD and RFH perceived direction oscillates, with a periodicity governed by the internal sinusoidal luminance modulation, for micropattern displacements of less than ~ 1.5 to 2 spatial periods and then error rates generally approach levels that are better than chance (50%) for most of the remainder of the displacement range examined. Observer TL fails to show these characteristics oscillations in performance, as was the case in Fig. 2, and performance is veridical (almost a 0% error rate) for displacements less than two spatial periods of f_m. Thereafter it closely resembles that of the other two observers. This would imply, at first sight, that TL is more sensitive, in general, to the drifting second-order motion of the Gaussian envelope than the first-order motion of the internal structure of the micropatterns.

For the G*C micropatterns, in which the internal spatial structure of the Gaussian envelope is defined in terms of drifting second-order, rather than first-order, image attributes, direction-identification for each observer was very similar to that obtained with the G + C micropattern stimuli. That is, for an initial range of displacements, perceived direction oscillates for both SOD and RFH with a frequency consistent with these observers basing direction judgements on the internal contrast modulation present in the patterns. This provides compelling evidence that mechanisms do exist to extract second-order motion that have qualitatively similar properties to those that encode first-order motion and these properties are embodied in current low-level models of motion detection, typified by the Adelson and Bergen's (1985) motion energy model. At larger displacements these two observers are still able to identify the overall direction of micropattern motion but error rates (especially those for RFH) are close to chance levels. TL once again does not show oscillations in performance. This finding is interesting in that the results of Experiment 1, considered in isolation, would seem to imply that for this observer perceived direction for displays containing conventional G micropatterns is influenced by the second-order motion present in the stimulus to a greater extent than the first-order motion. However, the present results for the G*C micropatterns suggest that for TL motion at the spatial scale of the Gaussian envelope, irrespective of how it is defined, rather than second-order motion per se, predominantly governs his direction-identification performance over most, if not all, of the range of displacements examined. This in turn suggests that TL utilises alternative strategies for detecting motion in micropattern displays and is consistent with the proposal that both low-level motion sensors and high-level, attentive, feature-tracking strategies can be used to some extent, and perhaps to differing degrees by each observer, to mediate perceived direction in these stimuli.

That the perceived direction of motion of the G*C micropatterns, particularly over short displacements where oscillations in performance are typically observed, is not mediated by a low-level mechanism sensitive to the spatially localised first-order (luminance-defined) motion cues present along the horizontal axis of the grating carrier, is supported by the finding (Fig. 5) that performance was unchanged for each of the three observers, over all displacements tested, when the carrier spatial phase was randomised on each successive frame of the motion sequences (which should destroy any strategy based on this principle).

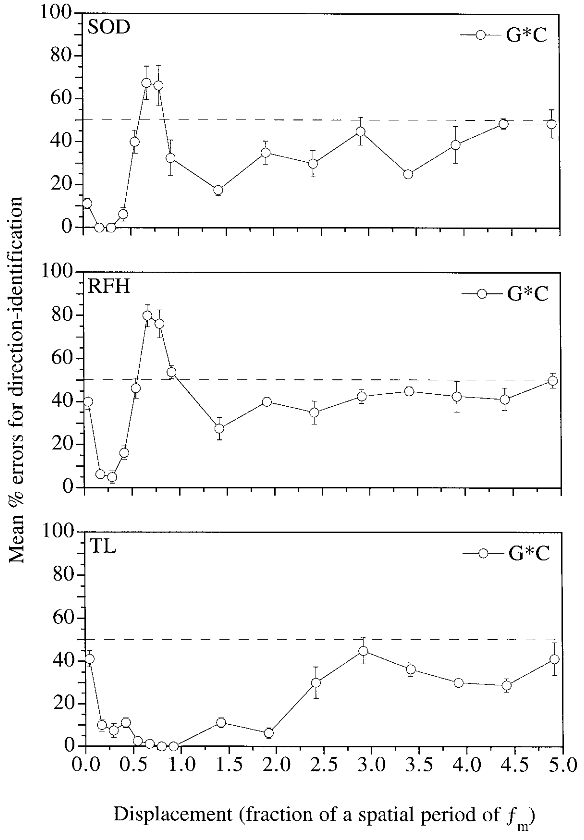

Fig. 5. Legend as for Fig. 4 with the exception that the absolute spatial phase of the horizontally-oriented, grating carrier was randomised for each successive frame of the motion sequence and was uniformly assigned values in the range ± 180°. This was done in order to render unreliable any spatially localised first-order motion cues, along the horizontal axis of the grating carrier, that could, in principle, be detected by low-level, first-order motion-detecting mechanisms operating over very restricted regions of the image.

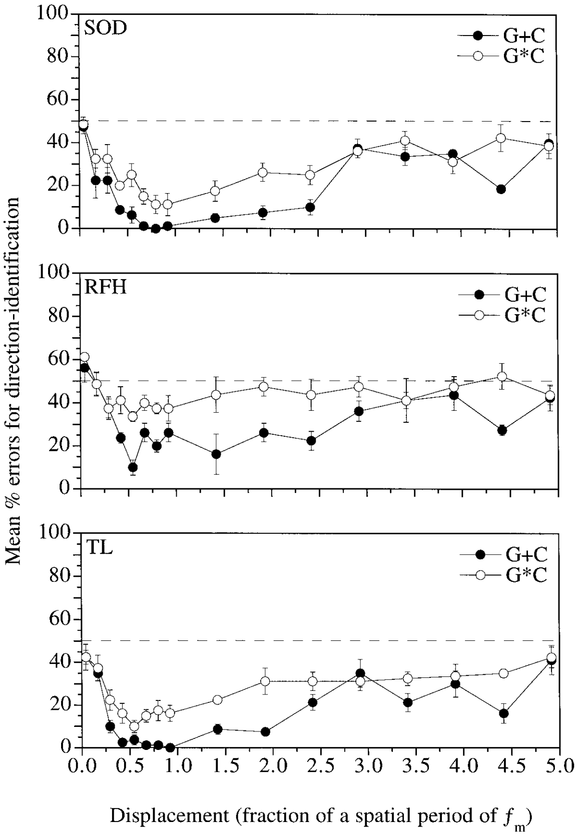

Fig. 6. Legend as for Fig. 4 with the exception that the spatial characteristics of the internal sinusoidal modulation in luminance (G + C micropatterns) or contrast (G*C micropatterns) were independently randomised for each micropattern on every frame of the motion sequence and were uniformly assigned values in the range ± 90° for orientation, 0.5–2c/° for f_m and ± 180° for spatial phase with respect to the envelope. This was done, in the *random condition*, in order to render unreliable the internal spatial characteristics of the micropatterns as signals for low-level motion sensors.

3.2.2. Random condition

In the *random condition*, motion signals present at the scale of the internal sinusoidal modulation in either luminance (G + C micropatterns) or contrast (G*C micropatterns) were made unreliable cues to the direction of motion (Fig. 6). It is clear that for the G + C micropatterns, which still contain coherent second-order motion at the coarse scale of Gaussian envelope, performance is typically good for all three observers over much of the displacement range and closely resemble those shown in Fig. 3, using G micropatterns without a superimposed carrier grating. For small displacements (typically < 0.5 spatial periods) performance was close to chance (50% errors) and then error rates drop rapidly to zero and remain low until displacements of the order of one to two spatial periods, when they rise gradually. Even at the very largest displacements, however, performance is still better than chance (~ 40% errors). For the G*C micropattern displays, the results are similar to those obtained with G + C micropatterns though performance is, if anything, generally less good, especially for observer RFH. Nevertheless it is evident that all three observers show better than chance performance over some of the displacement range and for observers SOD and TL error rates are remarkably low for much of the entire range examined (indicating that they perceived motion in the direction of the displacement even though the only reliable cue to this motion, at the coarse spatial scale of the Gaussian envelope, was defined by neither first-order nor second-order cues). That is, observers could consistently report the direction of micropattern motion (to differing degrees) even though this motion should be invisible to current first-order and second-order low-level, motion-detecting schemes. Thus, this clearly adds further support for the possibility of high-level, attentive, feature-tracking at the scale of the Gaussian envelope.

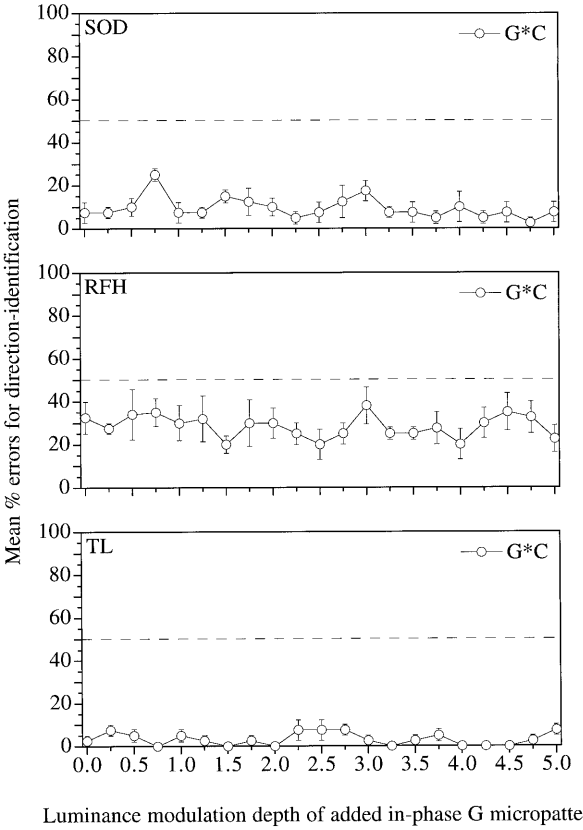

Fig. 7. Direction-identification performance in the *random condition* for three observers for G*C micropatterns as a function of the luminance modulation depth (prior to spatial windowing) of conventional Gabor (G) micropatterns that were added in-phase, in order to compensate for the possible effects of an early compressive, luminance nonlinearity occurring prior to the FRF scheme (depicted in Fig. 1) and low-level, second-order motion analysis. Observers completed four runs of 210 trials (ten at each of the 21 added luminance modulation depths) and the order in which trials were presented were randomised separately for each run and observer. The vertical bars above and below each data point (where visible) represent ± 1 SEM based on variability between runs of trials.

An intriguing possibility is that motion at the coarse spatial scale of the Gaussian envelope of individual G*C micropatterns could be made visible to low-level mechanisms that encode second-order motion if (say) a compressive, luminance nonlinearity were to precede the FRF scheme (see Fig. 1) implemented in the majority of current models (e.g. Wilson et al., 1992). An early nonlinearity of this form, fed by single cones and prior to band-pass spatial frequency filtering, has been identified in human vision (MacLeod, Williams, & Makous, 1992) and could account for some aspects of second-order motion perception. A compressive response to luminance would effectively transform a G*C micropattern into a conventional Gabor micropattern (G) such that subsequent FRF operations would render the structure of the Gaussian envelope available to low-level motion sensors [We are indebted to an anonymous referee for this suggestion]. In order to control for this possibility we employed the technique used previously by Scott-Samuel and Georgeson (1999), in which a sinusoidal luminance modulation is added in-phase with the sinusoidal contrast modulation (i.e. in antiphase with the putative compressive distortion) in order to null its effects. Although they were unable to abolish the perceived drift of second-order motion when drift speed was < 25°/s, some observers did exhibit a slight decline in performance when the added luminance modulation depth was increased from 0 to 0.06. We investigated direction-identification for G*C micropatterns, in the random condition, to each of which was added a conventional Gabor (G) micropattern (the orientations, spatial frequencies and spatial phases of the internal sinusoidal modulations were identical in each case) and the luminance modulation depth was varied from 0 to 0.05 in equal steps of 0.0025. All micropatterns were displaced by 0.5° on each successive positional update (equivalent to 0.5 spatial periods when f_m was 1c/°), a value chosen because it produced good performance for all three observers in Fig. 6. Contrary to the predictions of the proposal outlined above, direction-identification was not significantly affected [$F_{(20,180)} = 0.81$; $P = 0.6960$] by the addition of in-phase G micropatterns (Fig. 7), in that performance is veridical and importantly does not show a systematic decline towards chance levels (50% errors) as luminance modulation depth is increased. Indeed performance overall was significantly better than chance, even for observer RFH [$t_{(83)} = 14.23$; $P < 0.0001$]. It is unlikely therefore that a compressive luminance distortion coupled with FRF and subsequent low-level motion analysis is sufficient to account for the perception of motion at the scale of the Gaussian window for G*C micropatterns.

4. General discussion

The main objective of the present study was to investigate the nature of the mechanisms that mediate perceived motion in stochastic displays composed of multiple, local patches of drifting grating or texture. Previous research employing conventional Gabor micropatterns has shown that direction-identification performance appears to be indicative of two distinct types of motion perception, one governed by first-order motion (at the spatial scale of the internal sinusoid) and the other by second-order motion (at the coarse scale of the Gaussian envelope). These studies (Boulton & Baker, 1993, 1994; Bex & Baker, 1997; Baker & Hess, 1998; Clifford et al., 1998), however, have invariably attempted to interpret these results in terms of low-level mechanisms sensitive to either first-order motion or second-order motion and have neglected the possibility that alternative strategies (e.g. high-level, attentive, feature-tracking) may play a role in determining perceived motion even though there is much evidence for the existence of these processes. This is due in part to the assumption that the nature of the stimuli preclude feature-tracking strategies involving the identification and explicit tracking of individual micropatterns, or micropattern clusters, over time. The present results are important in that they clearly demonstrate that this assumption may not be entirely valid and that the use of stochastic motion patterns does not necessarily discourage observers from basing perceptual judgements on the behaviour of individual elements within the stimulus, as has been recognised previously (McKee & Watamaniuk, 1994). This raises the possibility that motion perception in such displays reflects an interplay between low-level, motion-detecting mechanisms and high-level, feature-based processes operating at a range of spatial scales. The resultant direction of perceived motion appears to be dependent not only the fidelity of the motion signals present at each spatial scale (i.e. whether they are noisy or reliable) but also on the individual observer, to some extent, as evidenced by the results of Experiments 1 and 2.

The present study also highlights the manner in which previous studies have tended to confound the interpretation of results, within the context of low-level motion mechanisms, by the fact that for Gabor micropatterns the first-order and second-order motion signals coexist at different spatial scales. Consequently the results may be indicative of either the processes by which first-order and second-order motion signals interact within the visual system (i.e. first-order motion dominates at small displacements and second-order motion at larger displacements and/or when first-order motion is noisy or unreliable) or reflect a more general scheme by which motion information is combined across different spatial scales, irrespective of how that motion is defined. In terms of the latter possibility several lines of evidence suggest that motion signals at coarse spatial scales (low spatial frequencies) can dominate or even override those at finer spatial scales (higher spatial frequencies) particularly when the latter are incoherent. For example $D_{\max}$ (the maximum displacement for reliable direction-identification of luminance-defined, random-dot-kinematograms) appears to be based on the lowest frequency information in the stimulus (Bex, Brady, Fredericksen, & Hess, 1995; Eagle, 1996, 1998) — the optimum strategy computationally since higher spatial frequencies typically exhibit aliasing (direction reversals). Furthermore Ramachandran and Inada (1985) and Ramachandran and Cavanagh (1987) have shown, using first-order motion stimuli, that a drifting low spatial frequency grating can 'capture' the motion of a higher spatial frequency grating or noise pattern such that overall the stimulus appears to move coherently in the direction of the low frequency grating. Similar interactions between first-order and second-order motion signals at different spatial frequencies have also been demonstrated (Mather & Murdoch, 1998) and the present results also have a bearing on this issue.

In Experiments 1 and 2 for all observers the perceived direction of G, G + C and G*C micropatterns tended to be consistently veridical for a range of large displacements (>2 spatial periods) and also for small displacements when motion signals at the relatively fine scale of the internal luminance or contrast modulation were made unreliable by randomising the spatial characteristics of that modulation over time. This demonstrates a clear reliance on using motion signals (however they might be defined) at the relatively coarse spatial scale of the Gaussian envelope, when those at finer spatial scales cannot be used. When coherent motion signals at the scale of the internal sinusoidal modulation were available (at small displacements when its characteristics were not randomised) direction judgements were influenced, albeit to different degrees for each observer, by this motion in a manner consistent with observers basing judgements on the outputs of low-level, mechanisms sensitive to motion at this spatial scale. This strongly suggests that the results of previous studies (e.g. Baker & Hess, 1998) in which performance for G micropatterns appears to be an envelope of two qualitatively distinct types of motion perception, are more indicative of the processes by which motion information is combined across different spatial scales rather than the processes by which first-order and second-order motion signals interact within the visual system per se.

In summary both low-level motion detectors and mechanisms utilising a different motion-detecting strategy such as high-level, feature-tracking may mediate, at least some of the time, perceptual judgements in stochastic motion displays. This finding is clearly relevant to studies of motion perception in general which typically assume that the mechanisms that detect the stimulus motion are predominantly (or exclusively) either pre-attentive and low-level or post-attentive and high-level. That both types of motion-detection appear to influence performance, even under conditions assumed to preclude high-level, feature-based mechanisms, highlights the difficulties associated with constructing stimuli to isolate each class of motion system. Furthermore future studies employing Gabor micropatterns need to take account not only of the first-order motion and second-order motion in the stimuli, but also the properties of the mechanisms that may detect each variety of motion and how the spatial scale over which each operates influences perceptual. judgements.

Acknowledgements

Supported by MRC and NSERC Grants #MT 10818 &OGPOO4652 (to RFH).

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. (1980). The perception of apparent movement. *Philosophical Transactions of the Royal Society of London*, *B290*, 153–168.

Baker, C. L., Jr. (1999). Central neural mechanisms for detecting second-order motion. *Current Opinion in Neurobiology*, *9*, 461–466.

Baker, C. L., Jr., & Hess, R. F. (1998). Two mechanisms underlie processing of stochastic motion stimuli. *Vision Research*, *38*, 1211–1222.

Bex, P. J., & Baker, C. L., Jr. (1997). The effects of distractor elements on direction discrimination in random Gabor kinematograms. *Vision Research*, *37*, 1761–1767.

Bex, P. J., Brady, N., Fredericksen, R. E., & Hess, R. F. (1995). Energetic motion detection. *Nature*, *378*, 670–672.

Boulton, J. C., & Baker, C. L., Jr. (1993). Dependence on stimulus onset asynchrony in apparent motion: evidence for two mechanisms. *Vision Research*, *33*, 2013–2019.

Boulton, J. C., & Baker Jr, C. L. (1994). Psychophysical evidence for both a 'quasi-linear' and a 'non-linear' mechanism for the detection of motion. In Lawton, T.B., *Computational Vision Based on Neurobiology*, *SPIE Proceedings*, 2054. Bellingham, SPIE.

Braddick, O. J. (1974). A short range process in apparent motion. *Vision Research*, *14*, 519–527.

Braddick, O. J. (1980). Low-level and high-level processes in apparent motion. *Philosophical Transactions of the Royal Society of London*, *B290*, 137–151.

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.

Clifford, C. W. G., Freedman, J. N., & Vaina, L. M. (1998). First- and second-order motion perception in Gabor micropattern stimuli: psychophysics and computational modelling. *Cognitive Brain Research*, *6*, 263–271.

Clifford, C. W. G., & Vaina, L. M. (1999). A computational model of selective deficits in first and second-order motion processing. *Vision Research*, *39*, 113–130.

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.

Derrington, A. M., & Ukkonen, O. I. (1999). Second-order motion discrimination by feature-tracking. *Vision Research*, *39*, 1465–1475.

Eagle, R. A. (1996). What determines the maximum displacement limit for spatially broadband kinematograms? *Vision Research*, *38*, 1775–1787.

Eagle, R. A. (1998). Upper displacement limits for spatially broadband patterns containing bandpass noise. *Journal of the Optical Society of America*, *A*, *13*, 408–418.

Georgeson, M. A., & Harris, M. G. (1990). The temporal range of motion sensing and motion perception. *Vision Research*, *30*, 615–619.

Georgeson, M. A., & Shackleton, T. M. (1989). Monocular motion sensing, binocular motion perception. *Vision Research*, *29*, 1511–1523.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, *195*, 215–243.

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*, *B250*, 297–306.

Ledgeway, T., & Smith, A. T. (1994). 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. (1995). The perceived speed of second-order motion and its dependence on stimulus contrast. *Vision Research*, *35*, 1421–1434.

Lu, Z-L., & Sperling, G. (1995a). Attention-generated apparent motion. *Nature*, *377*, 237–239.

Lu, Z-L., & Sperling, G. (1995b). The functional architecture of human visual motion perception. *Vision Research*, *35*, 2697–2722.

MacLeod, D. I. A., Williams, D. R., & Makous, W. (1992). A visual nonlinearity fed by single cones. *Vision Research*, *32*, 347–363.

Mather, G., Cavanagh, P., & Anstis, S. M. (1980). A moving display which opposes short-range and long range signals. *Perception*, *14*, 163–166.

Mather, G., & Murdoch, L. (1998). Evidence for global motion interactions between first-order and second-order stimuli. *Perception*, *27*, 761–767.

McKee, S. P., & Watamaniuk, S. N. J. (1994). The psychophysics of motion perception. In A. T. Smith, & R. J. Snowden, *Visual detection of motion*. London: Academic Press.

Nishida, S. (1993). Spatiotemporal properties of motion perception for random-check contrast modulations. *Vision Research*, *33*, 633–645.

Nishida, S., Ledgeway, T., & Edwards, M. (1997). Dual multiple-scale processing for motion in the human visual system. *Vision Research*, *37*, 2685–2698.

Ramachandran, V. S., & Inada, V. (1985). Spatial phase and frequency in motion capture of random-dot patterns. *Spatial Vision*, *1*, 57–67.

Ramachandran, V. S., & Cavanagh, P. (1987). Motion capture anisotropy. *Vision Research*, *27*, 97–106.

Scott-Samuel, N. E., & Georgeson, M. A. (1999). Does early non-linearity account for second-order motion? *Vision Research*, *39*, 2853–2865.

Seiffert, A. E., & Cavanagh, P. (1998). Position displacement, not velocity, is the cue to motion detection of second-order stimuli. *Vision Research*, *38*, 3569–3582.

Smith, A. T. (1994a). The detection of second-order motion. In A. T. Smith, & R. J. Snowden, *Visual detection of motion*. London: Academic Press.

Smith, A. T. (1994b). 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., Hess, R. F., & Baker, C. L., Jr. (1994). Direction identification thresholds for second-order motion in central and peripheral vision. *Journal of the Optical Society of America*, *A*, *11*, 506–514.

Smith, A. T., & Snowden, R. J. (Eds.) (1994). *Visual detection of motion*. London: Academic Press.

Stoner, G. R., & Albright, T. D. (1992). Motion coherency rules are form-cue invariant. *Vision Research*, *32*, 465–475.

Sutter, A., Beck, J., & Graham, N. (1989). Contrast and spatial variables in texture segregation: testing a simple spatial-frequency channels model. *Perception and Psychophysics*, *46*, 312–332.

Taub, E., Victor, J. D., & Conte, M. M. (1997). Nonlinear preprocessing in short-range motion. *Vision Research*, *37*, 1459–1477.

Ullman, S. (1979). *The interpretation of visual motion*. Cambridge: MIT Press.

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

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., 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.