Available online at www.sciencedirect.com
SCIENCE DIRECT®
Vision Research 44 (2004) 1499–1510
Vision Research
www.elsevier.com/locate/visres

Spatial frequency selective masking of first-order and second-order motion in the absence of off-frequency ‘looking’

Claire V. Hutchinson *, Timothy Ledgeway

School of Psychology, University of Nottingham, University Park NG7 2RD, UK

Received 16 October 2003; received in revised form 21 January 2004

Abstract

Converging evidence suggests that, at least initially, first-order (luminance defined) and second-order (e.g. contrast defined) motion are processed independently in human vision. However, adaptation studies suggest that second-order motion, like first-order motion, may be encoded by spatial frequency selective mechanisms each operating over a limited range of scales. Nonetheless, the precise properties of these mechanisms are indeterminate since the spatial frequency selectivity of adaptation aftereffects may not necessarily represent the frequency tuning of the underlying units [Vision Research 37 (1997) 2685]. To address this issue we used visual masking to investigate the spatial-frequency tuning of the mechanisms that encode motion. A dual-masking paradigm was employed to derive estimates of the spatial tuning of motion sensors, in the absence of off-frequency ‘looking’. Modulation-depth thresholds for identifying the direction of a sinusoidal test pattern were measured over a 4-octave range (0.125–2 c/deg) in both the absence and presence of two counterphasing masks, simultaneously positioned above and below the test frequency. For second-order motion, the resulting masking functions were spatially bandpass in character and remained relatively invariant with changes in test spatial frequency, masking pattern modulation depth and the temporal properties of the noise carrier. As expected, bandpass spatial frequency tuning was also found for first-order motion. This provides compelling evidence that the mechanisms responsible for encoding each variety of motion exhibit spatial frequency selectivity. Thus, although first-order and second-order motion may be encoded independently, they must utilise similar computational principles.

Keywords: Visual masking; Second-order motion; First-order motion; Spatial frequency; Off-frequency ‘looking’

1. Introduction

One of the fundamental tasks faced by the visual system is to encode the movements of objects in the world. Whenever objects move, the changing patterns of light that impinge on the retina convey information about that movement that can be broadly subdivided into two main categories: first-order motion (variations in luminance or colour) and second-order motion (variations in more complex textural properties such as contrast). There is now a wealth of evidence that observers can exploit both of these sources of visual information in order to perceive object movement (e.g. Baker, 1999; Cavanagh & Mather, 1989; Ledgeway & Hess, 2002; Ledgeway & Smith, 1994; Nishida, Ledgeway, & Edwards, 1997; Smith & Ledgeway, 1997, 1998).

The processes that mediate first-order motion detection have been studied extensively and several models of the underlying mechanisms have been proposed. For example the influential motion-energy model of Adelson and Bergen (1985) uses receptive fields that are oriented in space-time to detect moving luminance variations (Fourier energy) in the image. It is assumed that a population of such mechanisms exist and that each responds selectively to first-order motion in a particular direction and over a narrow range of spatial scales (frequencies). The precise principles by which second-order motion is encoded remains a fundamentally unresolved issue, but all current models assume that it has a computationally similar basis to first-order motion detection (though not necessarily utilising the same detectors). For example, it has been suggested that second-order motion is extracted by a specialised (separate)

---

[*] Corresponding author. Tel./fax: +44-115-95-15382.
E-mail address: lpxcvh@psychology.nottingham.ac.uk (C.V. Hutchinson).

0042-6989/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2004.01.014

visual pathway that applies a non-linearity (e.g. rectification or squaring) to the luminance profile of the image prior to motion-energy detection (Chubb & Sperling, 1988; Wilson, Ferrera, & Yo, 1992). Although alternatives to this general processing scheme have been developed (e.g. Johnston, McOwan, & Buxton, 1992), it is consistent with a large body of empirical evidence (e.g. see Baker, 1999).

Adaptation studies, in particular, have an important bearing on the issue of how second-order motion is processed in human vision. For example, Nishida et al. (1997) found that prior exposure to either first-order motion (luminance-defined gratings) or second-order motion (contrast-defined gratings) elevated thresholds for detecting the same variety of motion in a manner that was both direction selective and spatial frequency selective. This provides compelling evidence that first-order motion and second-order motion are encoded by mechanisms that are each selectively sensitive (tuned) to a particular range of spatial frequencies. Importantly, spatial frequency-selective aftereffects were not found under cross-adaptation conditions (i.e. when the adaptation and test stimuli were different types of motion), supporting the notion that the two varieties of motion are initially detected by separate spatial frequency mechanisms in human vision.

Although adaptation studies provide strong support for the existence of specialised mechanisms that respond to either first-order motion or second-order motion, they do not necessarily provide an accurate description of the actual spatial frequency tuning of those mechanisms (Blakemore, Carpenter, & Georgeson, 1970; Dealy & Tolhurst, 1974). As a result the spatial frequency selectivity of the aftereffects found in adaptation experiments should not be taken as the frequency tuning of the underlying units. This is further compounded by the fact that due to the time-consuming nature of adaptation experiments, studies primarily employ rapid, but inherently subjective, measurement methods (e.g. method of adjustment). Thus the results may be heavily influenced by the decision criteria and response biases adopted by individual observers. In addition adaptation does not always produce robust aftereffects when second-order motion stimuli are used (Cropper & Hammett, 1997; Culham et al., 1998) and so it may not be the most suitable technique for probing the characteristics of second-order motion detectors. However if spatial frequency tuned mechanisms do indeed exist for encoding second-order motion, as suggested by the results of adaptation studies (Nishida et al., 1997), then it should also be possible to measure their properties using an alternative psychophysical technique, such as visual masking.

Masking effects are typically spatial frequency selective, occurring only when the test and mask patterns have similar spatial frequencies, and have been used extensively to investigate the spatial frequency tuning of the channels that encode first-order spatial form and motion (Anderson & Burr, 1985; Anderson, Burr, & Morrone, 1991; Carter & Henning, 1971; Henning, 1988; Legge & Foley, 1980; Wilson, McFarlane, & Phillips, 1983).

Previous masking studies have generally presented a single mask stimulus in conjunction with the test pattern and have varied the test-mask spatial frequency difference in order to estimate the spatial frequency selectivity (bandwidth and channel shape) of visual mechanisms (e.g. Anderson & Burr, 1985; Legge & Foley, 1980). However potential problems can arise from what is commonly referred to as off-frequency 'looking'. This phenomenon was first identified in auditory masking and is termed off-frequency 'listening' in the auditory domain (Patterson, 1976; Patterson & Nimmo-Smith, 1980). It occurs when the listener is able to centre the auditory channel used for detection at a point other than the test frequency, thereby maximising the signal-to-noise ratio and increasing their ability to detect the test. This concomitant lowering of the listener's thresholds can distort measurements of channel shape and bandwidth leading to erroneous conclusions (typically the channel in question appears to be much more frequency selective than it really is). The same principle applies to any perceptual system that incorporates multiple mechanisms with overlapping passsbands.

It has been proposed (e.g. Losada & Mullen, 1994, 1995; Patterson, 1976; Patterson & Nimmo-Smith, 1980) that off-frequency detection can be minimised by using a dual-masking paradigm. This involves the simultaneous presentation of two masks, one lower than the test frequency and the other higher. Often in vision such studies employ a notched-noise mask, composed of the sum of spatially lowpass and highpass filtered noise with non-overlapping passbands, so that within the noise there is a spatial frequency notch. The test stimulus can then be placed within this notch thereby ensuring that the detection mechanism with the highest signal-to-noise ratio is always centred on the test frequency. However the use of two noise masks may not always be ideal. Instead, it may be advantageous to use pairs of masks composed of sinusoidal gratings, one positioned above and the other below the test frequency. One distinct advantage of sinusoidal masking patterns is that because each mask contains only one spatial frequency, all the energy is concentrated on that frequency thus increasing the effectiveness of the mask. In the case of noise masking, the relatively broad spectral composition of the masks means that much of the energy contained within the masks may be outside the visible range or passband of the system under scrutiny, thereby potentially reducing the effectiveness of the overall mask. This is an important consideration when using second-order motion patterns as stimuli, because absolute sensitivity and spatial acuity are much worse than they are for first-order motion (Smith, Hess, & Baker, 1994).

In the present study the spatial frequency tuning of the mechanisms responsible for encoding first-order (luminance-defined) motion and second-order (contrast-defined) motion was measured using a dual-masking paradigm and an objective (bias free) forced-choice task. This allowed us to explore further the characteristics of the mechanisms mediating motion detection in human vision, in the absence of off-frequency looking.

2. Methods

2.1. Observers

The observers were the two authors, CVH and TL. Both had corrected-to-normal acuity and had no history of any visual disorders.

2.2. Apparatus and stimuli

Stimuli were generated using *a Macintosh G4* computer and presented on a *Sony Trinitron Multiscan E530* monitor with an update rate of 75 Hz using custom software written in the C programming language. For precise control of luminance contrast the number of intensity levels available was increased from 8 to 12 bits by combining the outputs of the three digital-to-analog converters of the video card using a custom-built video attenuator (Pelli & Zhang, 1991). Images were presented in 'greyscale' on the colour monitor by amplifying the resulting 12-bit monochrome signal and sending this same signal to the red, green and blue guns of the display, allowing fine-grained control of the luminance levels in each stimulus. The mean luminance of the display was 25.3 cd/m^2. Images were viewed binocularly and in darkness at a distance of 69.5 cm. One pixel subtended 1.88 arcmin of visual angle resulting in a display that subtended 24 deg vertically and 32 deg horizontally.

To ensure that the second-order motion stimuli did not contain any luminance artifacts, the monitor was carefully gamma-corrected using a photometer and look-up-tables (LUT). As an additional precaution, the adequacy of the gamma-correction was also checked psychophysically using a sensitive motion-nulling task (Gurnsey, Fleet, & Potechin, 1998; Ledgeway & Smith, 1994; Lu & Sperling, 2001; Scott-Samuel & Georgeson, 1999).

Stimuli were presented for a total duration of 853 ms and were vertically oriented, grating patterns defined by sinusoidal variations in either luminance (first-order) or contrast (second-order). For first-order motion test stimuli, the luminance profile can be defined as

$$L_{(x,y,t)} = L_{\text{mean}}[1 + m\sin(2\pi f_x x + 2\pi f_t t + \phi)], \tag{1}$$

where $L_{(x,y,t)}$ represents the luminance at each point in the stimulus, L_{mean} is the mean luminance, m is the modulation depth (Michelson contrast) of the sinewave, f_x is the modulation spatial frequency (either 0.125, 0.25, 0.5, 1 or 2 c/deg), f_t is the modulation temporal frequency (2.34 Hz) and ϕ is the initial spatial phase of the modulation. The initial absolute (starting) phase of the test stimulus was randomised on each presentation and the direction of drift could be either leftwards or rightwards. First-order motion stimuli did not contain a carrier (e.g. noise) as it has been suggested that this in itself could act as a mask (Schofield & Georgeson, 2003) and potentially contaminate measurements of spatial tuning. Furthermore first-order motion stimuli were employed primarily as a control to verify the effectiveness and adequacy of our masking paradigm.

For second-order motion test stimuli, the luminance profile can be defined as

$$L_{(x,y,t)} = L_{\text{mean}}[1 + \langle\{1 + m\sin(2\pi f_x x + 2\pi f_t t + \phi)\}c_n R_{(x,y)}\rangle], \tag{2}$$

where $L_{(x,y,t)}$, L_{mean}, m, f_x, f_t and ϕ, refer to the same parameters as Eq. (1). C_n is the contrast (0.15) of the spatially two-dimensional (2-d) noise carrier $R_{(x,y)}$. In Eq. (2), $R_{(x,y)}$ refers to static noise but as an additional control thresholds for second-order motion were also measured at one spatial frequency (0.5 c/deg) using carriers composed of dynamic noise ($R_{(x,y,t)}$) to verify the robustness of the results to changes in carrier temporal properties. The noise carriers were generated by assigning individual screen pixels (1.88 arcmin) to be "black" or "white" with equal probability and there was no spatial variation in luminance within each noise pixel.

Modulation-depth thresholds were measured for test gratings in both the absence (unmasked thresholds) and presence of two counterphasing masks which were simultaneously positioned above and below the test frequency at varying test-mask separations (ranging from 1/32 to 3 octaves). Each mask was constructed by summing two vertical sinusoidal gratings (either both first-order or both second-order) drifting in opposite directions at 2.34 Hz to create a counterphasing grating. [1] A condition was also included where the test and a single counterphasing mask shared the same spatial frequency. In this condition, the modulation depth of the mask was adjusted such that its total power (when considered in the luminance domain or the contrast

---

[1] We used two counterphasing gratings, rather than drifting gratings, as masks because pilot studies revealed that when drifting gratings were used, observers' judgements of the test drift direction were often biased in the direction of the drifting masks. To remove this response bias we therefore utilised counterphasing mask patterns that have no net overall direction, similar to those used in other motion masking studies (Anderson & Burr, 1985).

domain, as appropriate) was equal to that of the dual masks used in all other conditions. This was done by equating the root-mean-square (RMS) modulation depths of the mask patterns. RMS contrast was used because unlike Michelson contrast which defines contrast using the maximum and minimum luminance levels of the stimulus, RMS contrast is based on the standard deviation of the luminance levels in the stimulus. Therefore, using RMS contrast allowed direct comparison between conditions containing one or two masks. In addition, whenever a single mask was employed, the test and the sinusoidal component of the mask drifting in the same direction as the test were added in phase in an attempt to ensure that any additive effects were uniform across trials. In all other conditions, when the mask and test patterns had different spatial frequencies, the phase relationship between the test and masks was randomised on each presentation.

2.3. Procedure

A single-interval forced-choice procedure was employed. On each trial observers were presented with a fixation cross followed by the presentation of the drifting sinusoidal test grating and masks (when present). After the presentation of the stimulus, observers were cued to respond with a key press and their task was to judge the direction of the test grating's motion (left or right). Feedback was given after trials in which the observer responded incorrectly. A direction-identification task, rather than a simple detection paradigm, was employed to ensure that observers' judgements were indeed based on the response properties of motion-encoding mechanisms, rather than those that mediate the encoding of spatial form per se. This is important because in the case of second-order motion stimuli, previous studies (Smith & Ledgeway, 1997) have shown that the absolute sensitivities of the mechanisms that extract spatial form and those that process motion direction are not the same, and thus it is vital to use a task that specifically probes motion mechanisms.

The modulation depth (strength) of the test stimulus was varied from trial to trial according to a modified 1-up 3-down staircase designed to converge on the modulation-depth corresponding to 79.4% correct performance (Cornsweet, 1962; Levitt, 1971; Wetherill & Levitt, 1965). At the beginning of each run of trials the modulation depth of the test grating was initially set to a suprathreshold level (typically ~6 dB above threshold) and the initial staircase step size was chosen to be half this value. On subsequent reversals the step size was halved and testing was terminated after a total of 16 reversals. Threshold estimates were taken as the mean of the last 4 reversals in each staircase. Each observer completed a minimum of 4 runs of trials (i.e. 4 staircases) for each condition and the order of testing was randomised. The mean threshold and the standard error of the mean were then calculated for each spatial frequency and stimulus type.

3. Results

3.1. Experiment 1: Unmasked (baseline) thresholds

Fig. 1 shows modulation-depth thresholds for identifying the test drift direction at each test frequency using first-order (FO) gratings and second-order (SO)

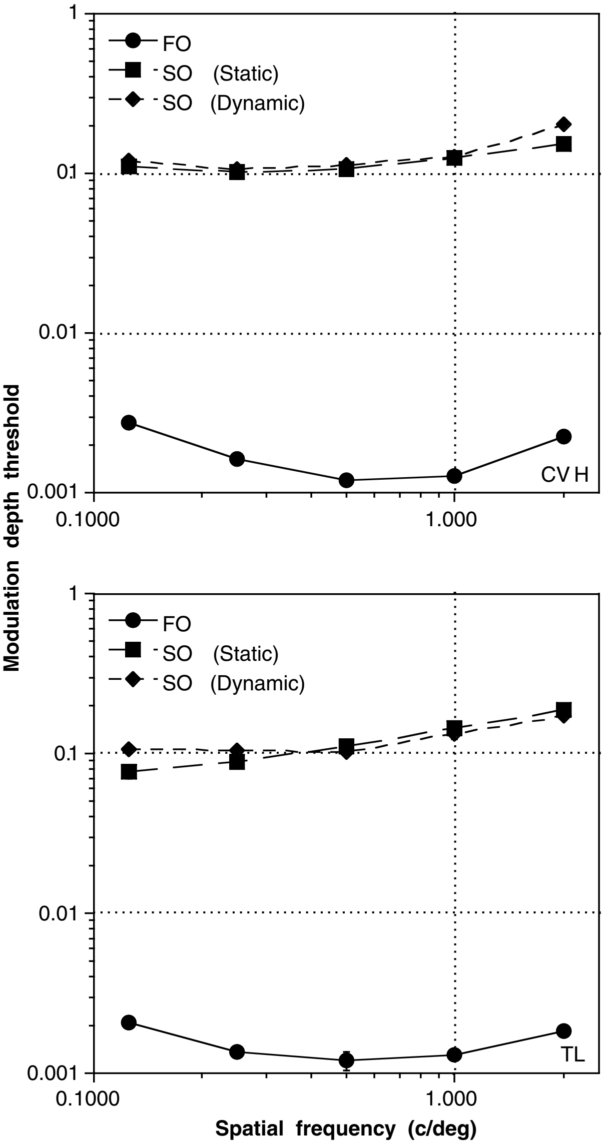

Fig. 1. Mean unmasked modulation-depth thresholds (as a function of spatial frequency) for two observers for identifying the drift direction of first-order (FO) and second-order (SO) drifting gratings. The second-order gratings contained either static or dynamic noise carriers, the contrast of which was modulated by a sinusoidal waveform. The vertical bars above and below each datum (where visible) represent ±1 SEM.

gratings (containing either static or dynamic noise carriers) in the absence of any mask. For both observers, absolute thresholds were considerably higher for the second-order motion stimuli than the first-order motion stimuli as found previously (e.g. Smith et al., 1994). Thresholds for second-order motion were generally spatially lowpass in character (irrespective of whether the noise carrier was static or dynamic) whilst those for first-order motion exhibited a characteristic bandpass profile. The results are in good agreement with those found in previous studies (e.g. Nishida et al., 1997) and ensure that, in the case of both first-order and second-order motion, testing was carried out within each system's visible range. Furthermore the similarity of the thresholds obtained with second-order stimuli, regardless of whether images contained a static or a dynamic noise carrier, also suggests that the these stimuli were isolating a 'true' second-order motion-detecting mechanism (Smith & Ledgeway, 1997, 1998). Whilst it is possible that the differences between the functions for first-order and second-order motion may have been due, in part, to the fact that first-order stimuli were not presented in conjunction with a noise carrier, recent work (Hutchinson & Ledgeway, in preparation) demonstrates that this is unlikely to be the case.

3.2. Experiment 2: The effects of mask spatial frequency and modulation depth on first-order motion detection

Although previous masking studies (e.g. Anderson & Burr, 1985) have established that first-order motion detectors exhibit some degree of spatial-frequency tuning, the current dual-masking paradigm has not been used previously to study motion. Therefore, testing was initially carried out with a 0.5 c/deg luminance-modulated test pattern to confirm the effectiveness of our protocol and to ascertain the effects of mask modulation depth (contrast) on first-order motion thresholds.

Thresholds were measured at five different test-mask separations (along the spatial frequency dimension) which were 0 (when the test and a single mask shared the same frequency), 0.031, 0.5, 1 and 1.5 octaves. A number of mask modulation depths ranging from 0.0012 to 0.1 were used and to aid comparison across the different conditions (i.e. those containing either one or two masks), these are expressed (in Fig. 2) as RMS values. Similarly thresholds are also expressed in RMS terms (equivalent to Michelson contrast * 0.7071 for single sinusoids).

The two observers show the same pattern of results. For stimuli defined by first-order motion, when the test and the mask had the same spatial frequency, subthreshold facilitation (a 'pedestal effect') was apparent at low mask modulation depths. That is, thresholds were lower in the presence of the counterphasing mask than when it was absent. Subthreshold facilitation has been reported previously for stationary first-order images (e.g. Legge & Foley, 1980) and for moving first-order stimuli and stationary pedestals (Zemany, Stromeyer, Chaparro, & Kronauer, 1998) and our results show that it also extends to scenarios in which both the test and mask patterns are in motion. Nonetheless it is evident that such facilitation effects disappear when masks of a sufficiently high modulation depth are used.

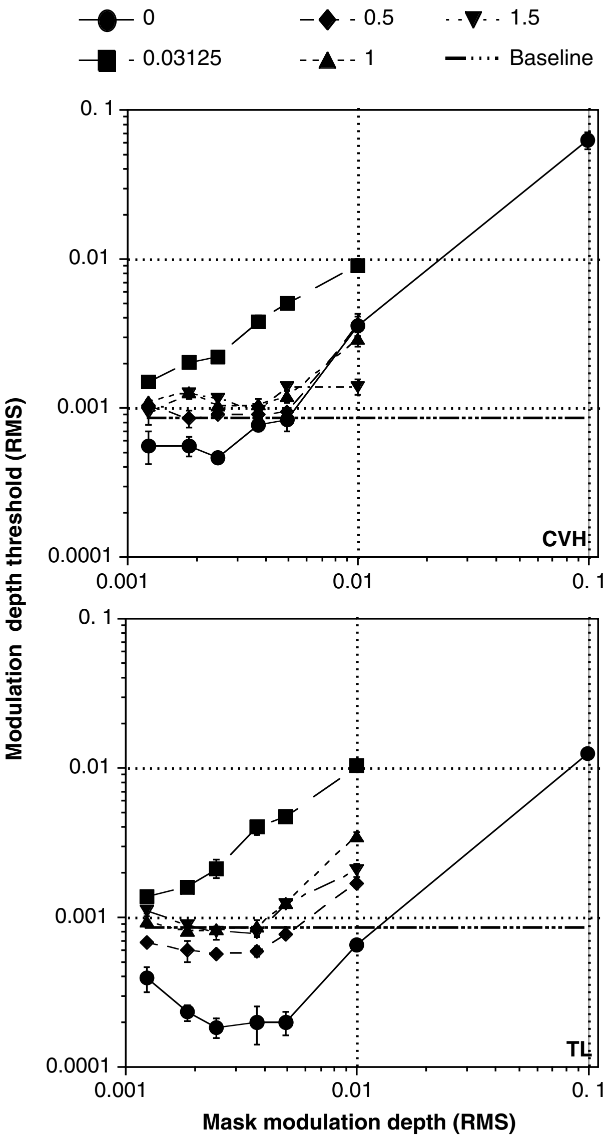

Fig. 2. First-order thresholds for two observers for identifying the direction of motion of a sinusoidal test pattern of 0.5 c/deg in the presence of two counterphasing mask patterns. The data plotted represents thresholds at each test-mask separation (indicated by the different symbols) as a function of the modulation depth of the masks. All values are expressed in terms of RMS contrast, rather than Michelson contrast, to aid comparison across the different conditions. Error bars represent ±1 SEM.

When the test and mask patterns differed in spatial frequency, the data for the two observers exhibited three main features: First, the degree of masking (threshold elevation relative to baseline) depends strongly on the spatial frequency difference between the test and mask patterns. The more similar the spatial frequency, the greater the magnitude of the masking effect. Second, for a given separation between the test and masks (e.g. 1 octave) the degree of masking observed is proportional to the modulation depth (strength) of the masks. This is most evident when the test-mask separation was 0.03125 octaves, but is also present at other test-mask separations. Third, the lowest modulation depth of the masks that first produces reliable threshold elevation varies with the spatial frequency separation between the test and the mask patterns. The closer the stimuli are in terms of spatial frequency, the lower the mask modulation depth that is needed to produce threshold elevation of the test stimulus. These patterns of performance are the characteristics expected of a bandpass, spatial-frequency-tuned motion mechanism. This is readily apparent in Fig. 3 which shows how thresholds vary as a function of the spatial frequency difference between the mask and test patterns, for each mask modulation depth tested. In Fig. 3, values are expressed in terms of conventional Michelson contrast rather than RMS contrast since all conditions contained two masks and it was not necessary to equate conditions containing single or dual masking patterns. The functions clearly exhibit spatial frequency selective masking and are invariably bandpass in character, except at the very lowest modulation depths tested where little or no masking could be measured.

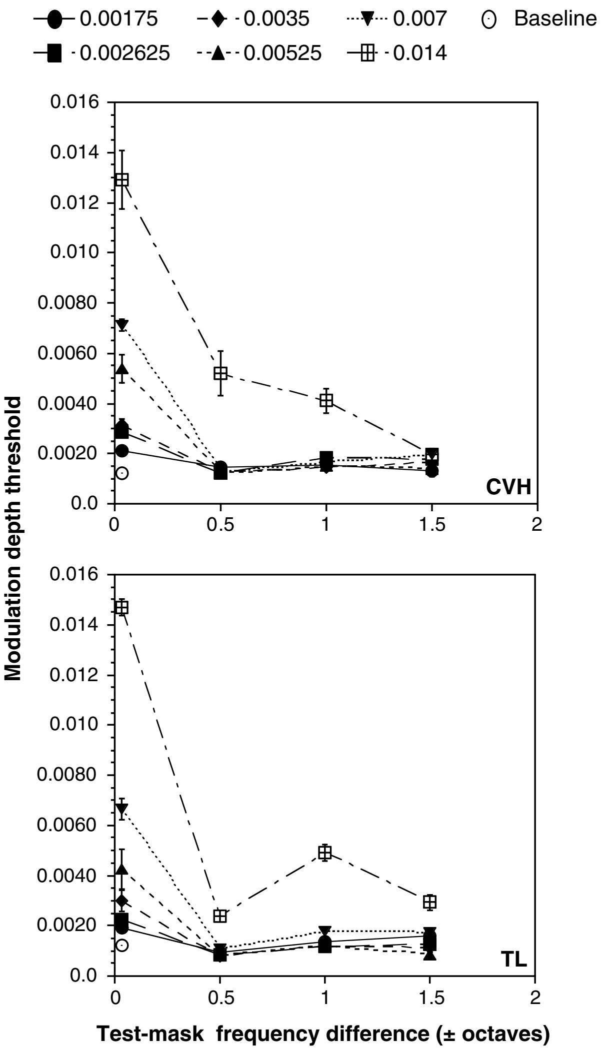

Fig. 3. First-order masking functions measured using different mask modulation depths (indicated by the different symbols). Masking levels (modulation-depth thresholds on a linear, rather than a logarithmic scale) are plotted as a function of test-mask separation. Masking levels achieved at a test-mask separation of 0.03125 octaves are represented by the left-most symbols. Error bars represent ±1 SEM.

3.3. Experiment 3: The effects of mask spatial frequency and modulation depth on second-order motion detection

We next sought to measure the effects of masking on thresholds for identifying the drift direction of second-order motion test stimuli. If spatial frequency tuned mechanisms exist for encoding second-order motion, as suggested by the results of adaptation studies (Nishida et al., 1997) and current models, then it should be possible to produce spatial frequency selective masking with second-order motion stimuli. Testing was carried out at 0.5 c/deg using contrast-modulated static noise patterns under equivalent conditions to those used for first-order motion. Once more, to aid comparison across the different conditions, mask modulation depths and test thresholds are expressed in terms of RMS values in Fig. 4. These were calculated on the basis of the second-order (contrast) information in the stimulus, rather than the first-order (luminance) profile.

Both observers showed some degree of subthreshold facilitation when the second-order mask and test gratings had the same spatial frequency. However, unlike the results found for first-order motion, this pedestal effect was still evident even at the highest available mask modulation depth that could be tested. This may be due to the fact that even at the maximum modulation depth available for the mask, the relatively poor absolute sensitivity of the visual system to second-order stimuli limits the perceptual effectiveness of the mask, which has insufficient power to combat the pedestal effect under these conditions.

When the test and mask patterns differed in spatial frequency, both observers showed a qualitatively similar pattern of results to those found using comparable first-order motion stimuli. That is, the magnitude of the masking effect was greatest when the spatial frequency difference between the masks and the test was small and declined as this difference increased. Furthermore

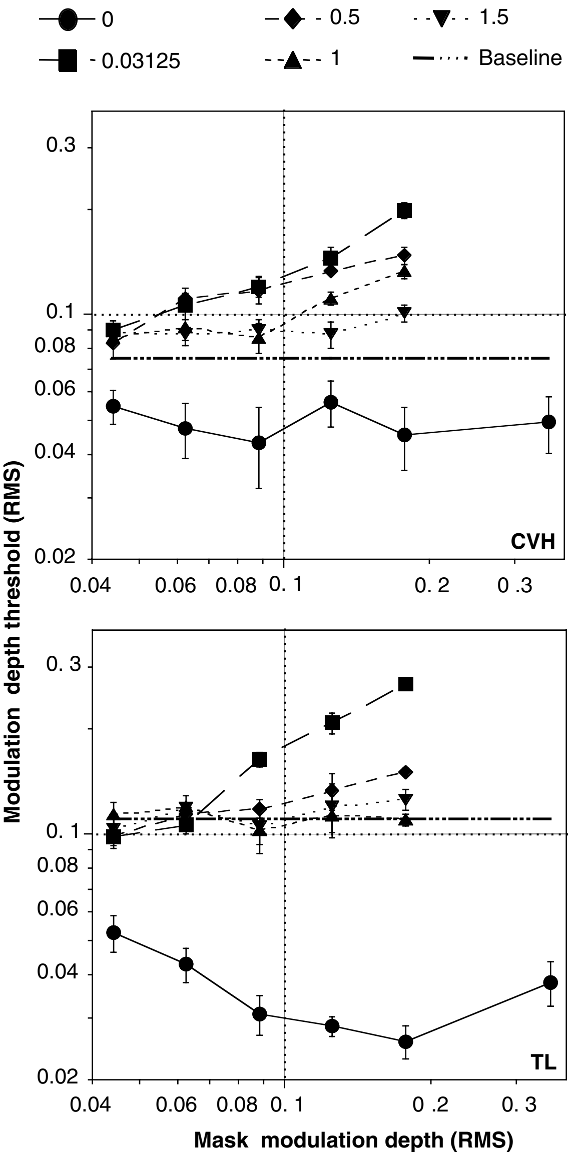

Fig. 4. Second-order thresholds for both observers at 0.5 c/deg measured using a range of mask modulation depths. The data plotted represent thresholds obtained at each signal-mask separation (indicated by the different symbols) as a function of mask modulation depth. All thresholds and modulation depth values are expressed in equivalent RMS terms (but computed from the second-order image statistics rather than the first-order, luminance information) and error bars represent ±1 SEM. (Note: a change in the scale of the x-axis compared with the data for first-order motion patterns shown in Fig. 2.)

thresholds generally increased as the modulation depth of the two masks increased (beyond some minimum value that depended on the proximity of the masks to the test in terms of spatial frequency). That the masking effects measured with a second-order test grating centred at 0.5 c/deg are also spatial frequency selective is readily evident in Fig. 5, which shows how thresholds vary as a function of the spatial frequency difference between the test and mask stimuli.

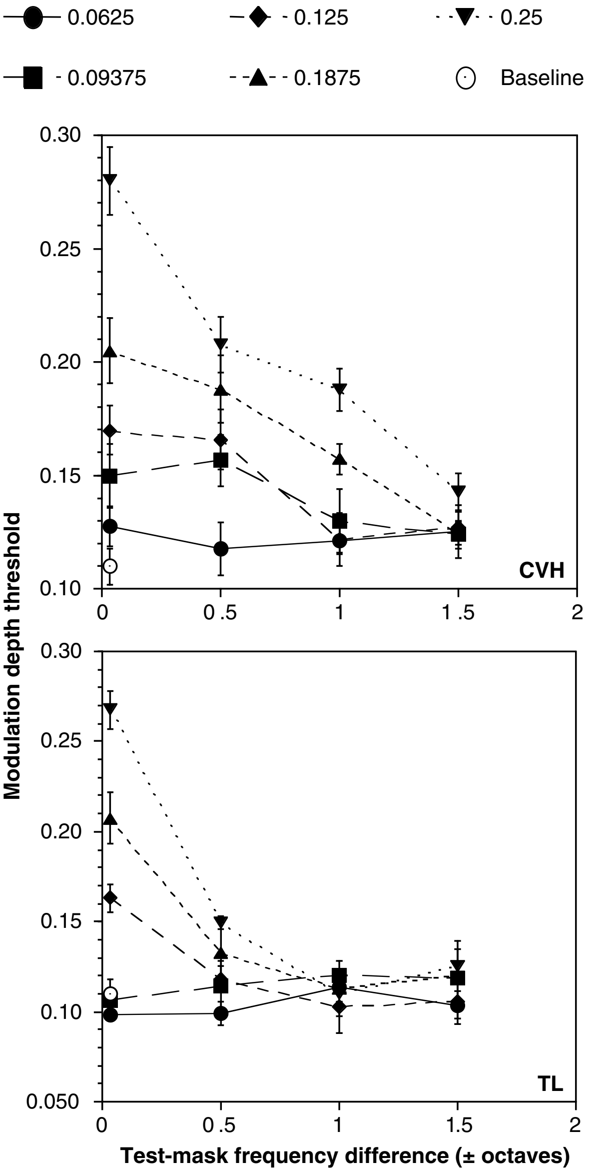

Fig. 5. Second-order masking functions measured using different mask modulation depths (indicated by the different symbols). Masking levels (modulation-depth thresholds on a linear ordinate) are plotted as a function of test-mask separation. Masking levels achieved at a test-mask separation of 0.03125 octaves are represented by the left-most symbols. Error bars represent ±1 SEM.

3.4. Experiment 4: Are there second-order motion sensors tuned to different spatial frequencies?

Although the results of the previous experiment provide compelling evidence that second-order motion detectors, like first-order motion sensors, are selectively sensitive to a narrow range of spatial frequencies, measurements were always centred on a test spatial frequency of 0.5 c/deg. In order to confirm the generality of this result, it is important to show that multiple band-pass mechanisms exist and that each encodes second-order motion over a different (but limited) range of spatial scales. To address this issue we measured spatial frequency selective masking using second-order motion test patterns (contrast-modulated static noise) spanning a 4-octave range that were centred on test spatial frequencies of either 0.125, 0.25, 0.5, 1 or 2 c/deg. Based on the results of the previous experiment, a modulation depth of 0.25 (c.f. Michelson contrast) was used for each mask since this produced reliable and maximal masking effects with second-order motion.

The masking functions obtained at each test frequency are depicted in Fig. 6(a) and (b), for observers CVH and TL respectively. The results for both observers are similar and clearly demonstrate the existence of mechanisms, centred on each test frequency, with spatially bandpass properties for encoding second-order motion. In general the maximum masking achieved was on average 2–3 times the unmasked threshold, which is comparable to the level of threshold elevation found with second-order motion in adaptation studies (Nishida et al., 1997). Additionally, in all cases, some degree of masking was still present as far away from the test spatial frequency as 3 octaves. To derive an approximate estimate of the spatial frequency bandwidths of the masking functions, the data obtained at each test spatial frequency were fitted with an exponential function using a least squares method. Exponential functions have been used extensively in previous studies to fit masking data and derive bandwidth estimates of spatial frequency tuning (e.g. Henning, 1988; Losada & Mullen, 1995). Although there was some marginal variation in the bandwidth estimates, both between the two observers and across the different test spatial frequencies used (especially 2 c/deg where observer CVH's thresholds were much less reliable than those at lower test frequencies), the average half-height, half-width bandwidth was 0.74 octaves (±0.21 octaves SEM).

3.5. Experiment 5: Effects of the noise carrier on second-order motion masking

All second-order motion stimuli must necessarily contain a carrier, such as static noise, some property of which (e.g. its contrast) is then modulated across space and over time to create movement. However there is always the possibility that a second-order image may

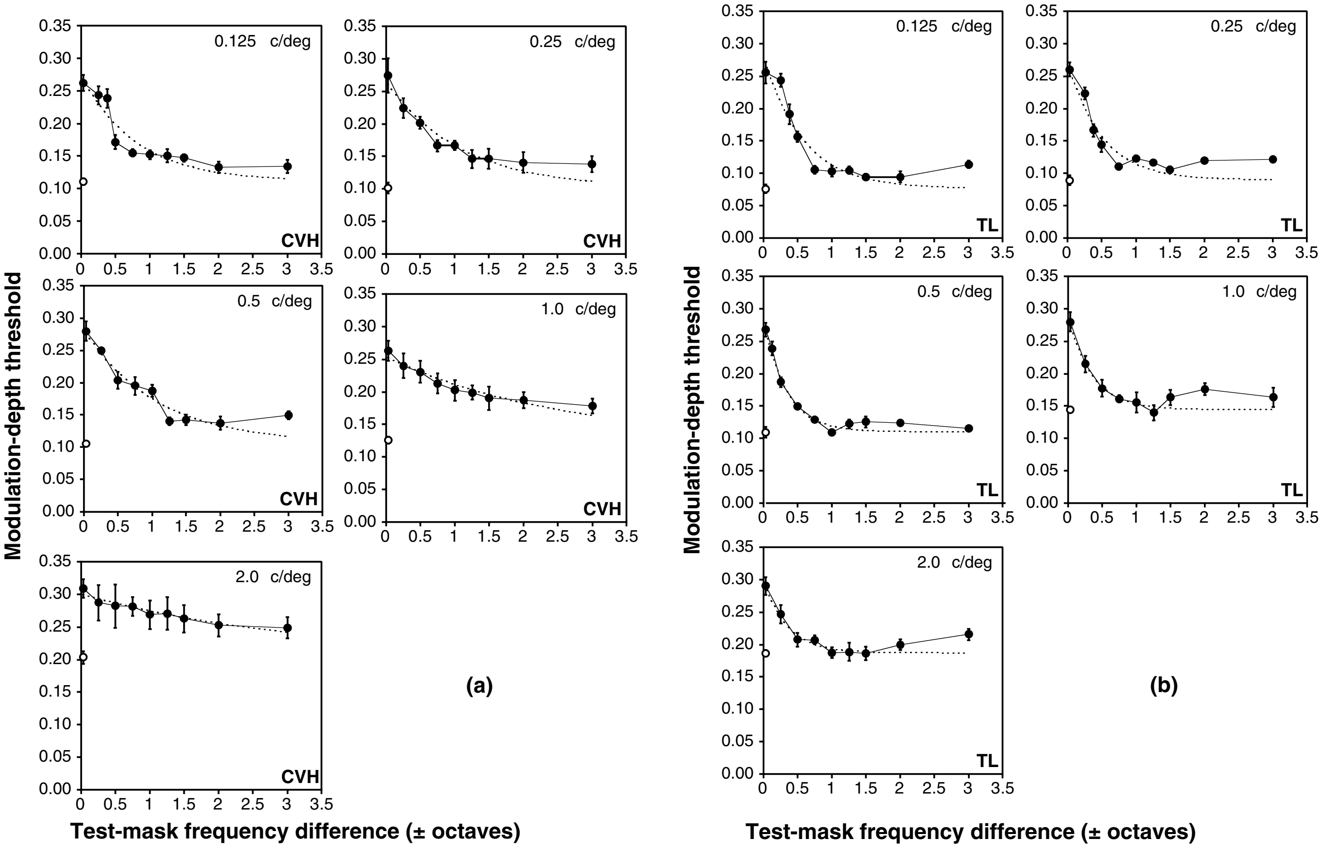

Fig. 6. Second-order masking functions measured for test patterns with spatial frequencies of either 0.125, 0.25, 0.5, 1 or 2 c/deg for observers CVH (a) and TL (b). Each plot indicates the degree of masking obtained (modulation-depth threshold) at each mask separation (in octaves) from the test frequency. The plots include the unmasked threshold (open circles) and the masked threshold data (filled circles) for each spatial frequency tested. Exponential functions have also been fitted to the data and are represented by the dotted lines. The vertical bars above and below each datum represent ±1 SEM.

inadvertently contain first-order (luminance) artifacts, upon which performance may be based. Such artifacts can potentially arise in contrast-modulated static noise patterns due to local imbalances in the proportion of 'light' and 'dark' noise pixels which lead to persistent spatial clumping of noise pixels (Smith & Ledgeway, 1997; but see Benton & Johnston, 1997). Although such aritifacts are minimal or absent when relatively small noise pixels are used and there is no spatial variation in luminance within each noise element (e.g. Ledgeway & Hess, 2002), it has been argued that a much more stringent test for their presence is to use a dynamic noise carrier. As the noise pixels vary over both space and time when a dynamic carrier is used, an observer's opportunity to utilise local first-order cues (due to clumping) within the carrier that might aid detection is even more limited (Smith & Ledgeway, 1998). In light of this controversy in a final control experiment we re-measured the masking function centred on a test frequency of 0.5 c/deg using contrast-modulated dynamic noise patterns.

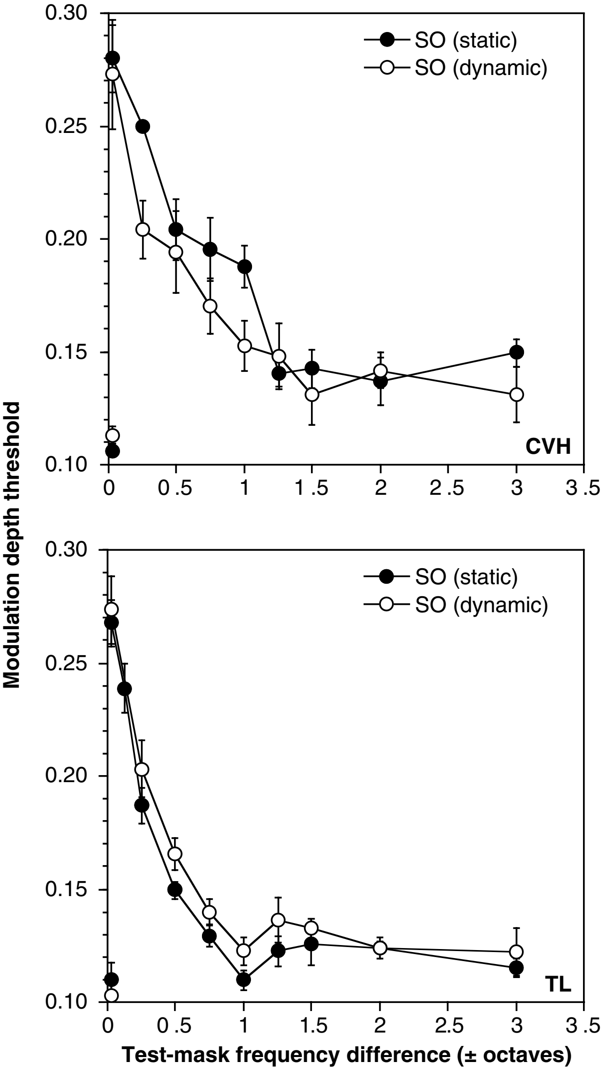

Fig. 7. Masking functions for both observers measured at 0.5 c/deg using second-order (SO) images containing carriers composed of either static (filled circles) or dynamic (open circles) noise. Error bars represent ±1 SEM.

Fig. 7 shows the masking functions for two observers measured with second-order motion patterns composed of either contrast-modulated static noise or contrast-modulated dynamic noise. It is evident that equivalent results were obtained regardless of whether the noise carrier was static or dynamic. Indeed, for observer TL the two masking functions are virtually indistinguishable. Therefore, based on these results we are confident that the static carriers used in the masking experiments did not give rise to any measurable luminance artefacts and, as such, the second-order images used were isolating second-order motion-detecting mechanisms. This conclusion is further bolstered by the fact that the unmasked (baseline) second-order motion thresholds measured previously using static and dynamic noise carriers (see Fig. 1) were also extremely similar over the entire range of spatial frequencies tested.

4. Discussion

The present study used a dual-masking paradigm in order to ascertain the spatial frequency tuning of the mechanisms that underlie the detection of first-order motion and second-order motion in human vision. This method sought to minimise the possibility of off-frequency looking and its rationale is directly analogous to the notched noise masking techniques that have been used previously to probe the spatial frequency tuning characteristics of first-order spatial vision. To our knowledge this technique has not been previously applied to study motion perception. Although the use of two masking patterns, equidistant in spatial frequency from the test pattern, implicitly assumes that the spatial frequency tuning of the underlying mechanisms is symmetrical (when plotted on an octave scale), the results of previous masking and adaptation studies are not inconsistent with this assumption (e.g. Anderson & Burr, 1985; Nishida et al., 1997).

The results of the present experiments clearly support the existence of motion-detecting mechanisms within the human visual system which are selectively sensitive to a range of spatial frequencies. Moreover it has demonstrated that both first-order motion and second-order motion are encoded by such mechanisms. The bandpass spatial frequency tuning found appears to be extremely robust and was relatively invariant with changes in the test spatial frequency, the masking pattern modulation depth and, in the case of second-order motion patterns, the temporal properties of the noise carrier. As such our results are in good agreement with those of Nishida et al. (1997) who used similar motion stimuli, but the completely different technique of selective adaptation, to demonstrate that second-order motion detection is mediated by multiple mechanisms that each respond to a limited range of modulation spatial frequencies.

Several key aspects of the results are worth considering in more detail. When a single mask pattern was employed (with the same spatial frequency as the test stimulus) we found that for a range of (low) mask modulation depths some degree of subthreshold facilitation, rather than masking, was evident. That is, thresholds for identifying the drift direction of either a first-order or second-order motion test pattern were somewhat lower in the presence of a single counterphasing mask than in its absence. Although similar effects have been reported previously for stationary test patterns (Legge & Foley, 1980) and for moving test patterns and stationary pedestals (Zemany et al., 1998), it is clear that they also apply to the moving test patterns used in the present study. Although the interpretation of this phenomenon has been the subject of much debate in the literature, Legge and Foley (1980) point out that when a test and a weak (barely perceptible) mask pattern share the same spatial frequency a simple additive effect occurs. Under these conditions the observer's task becomes somewhat different. In the context of the present study it changes to a simple modulation-depth (contrast) discrimination judgement between opposite directions of movement. Interestingly this facilitation effect disappeared when first-order motion stimuli were employed and the mask modulation depth was sufficiently high, but this was not the case when second-order motion stimuli were used. As mentioned previously it seems likely that this is because the relatively poor absolute sensitivity of the visual system to second-order motion (see Fig. 1) limits the perceptual efficacy of the mask pattern under these conditions (a second-order motion pattern with a modulation depth of unity is still only about 10 times above its own threshold). When the test and mask patterns differed in spatial frequency, there was little or no evidence of subthreshold facilitation and both varieties of motion exhibited spatial frequency selective masking (threshold elevation).

Although we have tentatively included bandwidth estimates of spatial frequency tuning derived from fitting exponential functions to the masking data, considerable variability exists in the literature concerning such quantitative estimates (although this is largely concerned with first-order masking studies). As such it may be unwise to place too much emphasis on the bandwidth measurements of the second-order masking functions in the present study (∼0.74 octaves at half-width and half-height), or indeed the bandwidth measurements in any masking study as there is still not a clear general consensus concerning their interpretation. Nevertheless, exponential functions have also been included for comparison. Whilst it has often been assumed that masking allows the direct quantification of the tuning of a mechanism that responds to a particular frequency (e.g. Henning, 1988), it may be that in some cases, there is more than one mechanism underlying masking (Ross & Speed, 1991). Indeed, it seems more likely that what is being measured is actually some form of mutual interaction (inhibition) between adjacent channels, a point noted by several previous authors (e.g. De Valois & De Valois, 1990). Nonetheless the fact that our masking functions are spatially bandpass in character, and extend over only a limited spatial frequency range, does suggest that multiple mechanisms encode second-order motion and that each has a much narrower tuning than the overall range of frequencies to which the visual system is sensitive.

It has been shown by Johnston et al. (1992) that, in principle, first-order motion and second-order motion could be detected by the same (common) mechanism. However, despite the evident computational similarities between the processing of first-order motion and second-order motion, there is converging evidence to suggest that, at least initially, the two varieties of motion are each encoded by separate populations of motion detectors. For example, as discussed previously, some of the most compelling psychophysical evidence to date has come from the finding that adaptation to one kind of stimulus does not affect responses to the other, at least at threshold stimulus levels (Nishida et al., 1997). Electrophysiological evidence for a distinction between second-order and first-order motion processing is evident from the behaviour of neurons found in the visual cortex of the cat which respond to a limited range of second-order envelope spatial frequencies. Although these neurons also respond to the motion of first-order gratings they do so over a very different range of spatial frequencies (Zhou & Baker, 1994, 1996). In addition, evidence has recently emerged both from neuropsychological patients (e.g. Greenlee & Smith, 1997; Vaina, Cowey, & Kennedy, 1999) and from fMRI studies (e.g. Dumoulin, Baker, Hess, & Evans, 2003; Nishida, Sasaki, Murakami, Watanabe, & Tootell, 2003; Seifert, Somers, Dale, & Tootell, 2003) demonstrating that the processing of first-order and second-order motion may be anatomically segregated.

In summary the results of the present masking study provide further compelling evidence that bandpass, spatial frequency tuned mechanisms exist for encoding both first-order motion and second-order motion in human vision. Thus, although it seems likely that luminance-defined motion and contrast-defined motion are detected independently, this detection appears to utilise similar computational principles.

Acknowledgements

A preliminary report of some of this work has been presented previously in abstract form (Hutchinson &

Ledgeway, 2003). This research was supported in part by a University of Nottingham Research Committee grant to TL.

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.

Anderson, S. J., & Burr, D. C. (1985). Spatial and temporal selectivity of the human motion detection system. *Vision Research, 25*, 1147–1154.

Anderson, S. J., Burr, D. C., & Morrone, M. C. (1991). Two-dimensional spatial and spatial-frequency selectivity of motion-sensitive mechanisms in human vision. *Journal of the Optical Society of America, 8*, 1340–1351.

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

Benton, C. P., & Johnston, A. (1997). First-order motion from contrast modulated noise? *Vision Research, 37*, 3073–3078.

Blakemore, C., Carpenter, R. H. S., & Georgeson, M. A. (1970). Lateral inhibition between orientation detectors in the human visual system. *Nature, 228*, 37–39.

Carter, B. E., & Henning, G. B. (1971). Detection of gratings in narrow-band visual noise. *Journal of Physiology (London), 219*, 355–365.

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.

Cornsweet, T. N. (1962). The staircase method in psychophysics. *American Journal of Psychology, 75*, 485–491.

Cropper, S. J., & Hammett, S. T. (1997). Adaptation to motion of a second-order pattern: the motion aftereffect is not a general result. *Vision Research, 37*, 2247–2259.

Culham, J. C., Nishida, S., Ledgeway, T., Cavanagh, P., von Grünau, M. W., Kwas, M., Alais, D., & Raymond, J. E. (1998). Higher order effects. In G. Mather, F. Verstraten, & S. Anstis (Eds.), *The motion after-effect: a modern perspective* (pp. 85–124). Cambridge: MIT Press.

Dealy, R. S., & Tolhurst, D. J. (1974). Is spatial adaptation an after-effect of prolonged inhibition? *Journal of Physiology, 241*, 261–270.

De Valois, R. L., & De Valois, K. K. (1990). *Spatial vision*. New York: Oxford University Press.

Dumoulin, S. O., Baker, C. L., Hess, R. F., & Evans, A. C. (2003). Cortical specialisation for processing first- and second-order motion. *Cerebral Cortex, 13*, 1375–1385.

Greenlee, M. W., & Smith, A. T. (1997). Detection and discrimination of first- and second-order motion in patients with unilateral brain damage. *Journal of Neuroscience, 17*, 804–818.

Gurnsey, R., Fleet, D., & Potechin, C. (1998). Second-order motions contribute to vection. *Vision Research, 38*, 2801–2816.

Henning, B. (1988). Spatial-frequency tuning as a function of temporal frequency and stimulus motion. *Journal of the Optical Society of America A, 5*, 1362–1373.

Hutchinson, C. V., & Ledgeway, T. (2003). Measurement of the spatial frequency tuning of second-order motion detectors with simultaneous masking in the absence of off-frequency 'looking'. *Perception, 32*(Suppl.), 101–102.

Hutchinson, C. V., & Ledgeway, T. (in preparation). Sensitivity to spatial and temporal modulations of first-order and second-order motion.

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., & Hess, R. F. (2002). Failure of direction-identification for briefly presented second-order motion stimuli: evidence for weak direction-selectivity of the mechanisms encoding motion. *Vision Research, 42*, 1739–1758.

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.

Legge, G. E., & Foley, J. M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America, 70*, 1458–1471.

Levitt, H. (1971). Transformed up–down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49*, 467–477.

Losada, M. A., & Mullen, K. T. (1994). The spatial tuning of chromatic mechanisms identified by simultaneous masking. *Vision Research, 34*, 331–341.

Losada, M. A., & Mullen, K. T. (1995). Colour and luminance spatial tuning estimated by noise masking in the absence of off-frequency looking. *Journal of the Optical Society of America A, 12*, 250–271.

Lu, Z.-L., & Sperling, G. (2001). Sensitive calibration and measurement procedures based on the amplification principle in motion perception. *Vision Research, 41*, 2355–2374.

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

Nishida, S., Sasaki, Y., Murakami, I., Watanabe, T., & Tootell, R. B. H. (2003). Neuroimaging of direction-selective mechanisms for second-order motion. *Journal of Neurophysiology, 90*, 3242–3254.

Patterson, R. D. (1976). Auditory filter shapes derived with noise stimuli. *Journal of Acoustical Society of America, 59*, 640–654.

Patterson, R. D., & Nimmo-Smith, I. (1980). Off-frequency listening and auditory-filter asymmetry. *Journal of the Acoustical Society of America, 67*, 229–245.

Pelli, D. G., & Zhang, L. (1991). Accurate control of contrast on microcomputer displays. *Vision Research, 31*, 1337–1350.

Ross, J., & Speed, H. D. (1991). Contrast adaptation and contrast masking in human vision. *Proceedings of the Royal Society of London B, 246*, 61–69.

Schofield, A. J., & Georgeson, M. A. (2003). Sensitivity to contrast modulation: the spatial frequency dependence of second-order vision. *Vision Research, 43*, 243–259.

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

Seifert, A. E., Somers, D. C., Dale, A. M., & Tootell, R. B. H. (2003). Functional MRI studies of human visual motion perception: texture, luminance, attention and after-effects. *Cerebral Cortex, 13*, 340–349.

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

Smith, A. T., & Ledgeway, T. (1997). Separate detection of moving luminance and contrast modulations: fact or artifact? *Vision Research, 37*, 45–62.

Smith, A. T., & Ledgeway, T. (1998). Sensitivity to second-order motion as a function of temporal frequency and eccentricity. *Vision Research, 38*, 403–410.

Vaina, L. M., Cowey, A., & Kennedy, D. (1999). Perception of first- and second-order motion: separable neurological mechanisms? *Human Brain Mapping, 7*, 67–77.

Wetherill, G. B., & Levitt, H. (1965). Sequential estimation of points on a psychometric function. *British Journal of Mathematical and Statistical Psychology, 18*, 1–10.

Wilson, H. R., Ferrera, V. P., & Yo, V. (1992). Psychophysically motivated model for two-dimensional motion perception. *Visual Neuroscience, 9*, 79–97.

Wilson, H. R., McFarlane, D. K., & Phillips, G. C. (1983). Spatial frequency tuning of orientation selective units estimated with oblique masking. *Vision Research, 23*, 873–882.

Zemany, L., Stromeyer, C. F., Chaparro, A., & Kronauer, R. E. (1998). Motion detection on flashed, stationary pedestal gratings: evidence for an opponent-motion mechanism. *Vision Research, 38*, 795–812.

Zhou, Y., & Baker, C. L., Jr. (1994). Envelope-responsive neurons in area 17 and 18 of cat. *Journal of Neurophysiology, 72*, 2134–2150.

Zhou, Y., & Baker, C. L., Jr. (1996). Spatial properties of envelope-responsive cells in area 17 and 18 of the cat. *Journal of Neurophysiology, 75*, 1038–1050.