Vision Research 50 (2010) 1766–1774

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Spatial summation of first-order and second-order motion in human vision

Claire V. Hutchinson [a,*], Timothy Ledgeway [b]

[a] School of Psychology, University of Leicester, Leicester LE1 9HN, UK
[b] Visual Neuroscience Group, School of Psychology, University of Nottingham, University Park, Nottingham NG7 2RD, UK

ARTICLE INFO

Article history:
Received 17 November 2009
Received in revised form 24 May 2010

Keywords:
First-order motion
Second-order motion
Spatial summation

ABSTRACT

This study assessed spatial summation of first-order (luminance-defined) and second-order (contrast-defined) motion. Thresholds were measured for identifying the drift direction of 1 c/deg., luminance-modulated and contrast-modulated dynamic noise drifting at temporal frequencies of 0.5, 2 and 8 Hz. Image size varied from 0.125° to 16°. The effects of increasing image size on thresholds for luminance-modulated noise were also compared to those for luminance-defined gratings. In all cases, performance improved as image size increased. The rate at which performance improved with increasing image size was similar for all stimuli employed although the slopes corresponding to the initial improvement were steeper for first-order compared to second-order motion. The image sizes at which performance for first-order motion asymptote were larger than for second-order motion. In addition, findings showed that the minimum image size required to support reliable identification of the direction of moving stimuli is greater for second-order than first-order motion. Thus, although first-order and second-order motion processing have a number of properties in common, the visual system's sensitivity to each type of motion as a function of image size is quite different.

1. Introduction

1.1. First-order and second-order motion

The human visual system is responsive to spatiotemporal information conveyed by a range of image properties. These are generally categorised as first-order (variations in luminance) or second-order (variations in more complex textural properties such as contrast) image statistics. A major unresolved debate in human vision concerns the issue of whether or not first-order motion and second-order motion are encoded by different low-level mechanisms. Although there is an abundance of evidence to suggest that first-order and second-order image properties are encoded separately in the mammalian visual system, at least in the initial processing stages (see Baker (1999), Smith (1994), Lu and Sperling (1995, 2001b) for reviews), a single mechanism could in principle handle both (Johnston, McOwan, & Buxton, 1992). For example in terms of the latter Benton and Johnston (2001) have shown mathematically that the motion of second-order contrast variations is available from conventional image spatiotemporal gradients in the lower contrast regions. Furthermore some phenomena such as the opposite motion induced in a static visual noise carrier, when its contrast is modulated by a moving waveform are difficult to explain if first-order and second-order motion are processed entirely separately in initial processing stages (Johnston, Benton, & McOwan, 1999). Thus the principles governing the perception of second-order image properties are still the subject of much controversy and warrant further study.

1.2. Spatial summation

Spatial summation in vision is a long-established phenomenon and, in short, refers to the fact that performance for detecting the presence of a visual stimulus improves as the size of that stimulus increases (Barlow, 1958; Campbell & Robson, 1968; Howell & Hess, 1978; Legge & Foley, 1980; Robson & Graham, 1981). Spatial summation functions are an important aspect of vision as they provide a behavioural measure of how visual information is integrated across retinal receptive fields (e.g. Anderson & Burr, 1991). A number of studies have investigated the nature of spatial summation for first-order and second-order information in the spatial domain (e.g. Schofield & Georgeson, 1999; Sukumar & Waugh, 2007; Wong & Levi, 2005). However these studies have not provided consistent results, suggesting that spatial summation functions may be heavily dependent on different stimulus parameters (stimulus type, frequency, etc.).

Schofield and Georgeson (1999) measured sensitivity to stationary first-order and second-order signals as a function of Gaussian blob width (defined as 2.5 times the standard deviation of the circularly symmetric Gaussian modulation function) and found similar effects for luminance-modulated and contrast-modulated noise. Sensitivity for detecting both types of stimuli increased as blob size increased and saturated at a similar blob size (~40 arc

* Corresponding author.
E-mail address: ch190@le.ac.uk (C.V. Hutchinson).

0042-6989/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2010.05.032

min). In addition, the sensitivity curves for luminance-modulated and contrast-modulated noise blobs were virtually parallel. These findings led Schofield and Georgeson (1999) to conclude that the similarity of the detection curves for first-order and second-order stimuli might reflect processing by the same or a functionally similar process. Landy and Oruc (2002) assessed the effects of spatial summation on second-order spatial processing using texture-quilts and found that performance plateaued at approximately 15°. Wong and Levi (2005) have also measured spatial summation areas for second-order stimuli in normal and amblyopic observers using static 1 c/deg. contrast-modulated noise Gabor patterns. They found that detection thresholds decreased at approximately the same rate in normal and amblyopic observers and saturated (flattened) at an image size of 6–8 cycles. In a control experiment, Wong and Levi (2005) compared their findings for second-order Gabors with spatial summation areas for static first-order Gabors in four normal observers. They found that, unlike second-order Gabors for which performance asympoted at an image size of around 6–8 cycles, for first-order Gabors performance failed to asymptote over the range of image sizes employed. Although Wong and Levi (2005) did not pursue this finding further, their results suggest that spatial summation areas are larger for first-order, compared to second-order, patterns.

Most recently, Sukumar and Waugh (2007) investigated spatial summation areas for static first-order and second-order patterns by measuring detection thresholds for luminance-modulated and contrast-modulated Gaussian noise blobs. They found that spatial summation areas were different for detecting luminance-defined and contrast-defined blobs. However they found that modulation thresholds saturated at smaller image sizes for luminance-defined than for contrast-defined stimuli. That is, spatial summation areas were larger for the contrast-defined (second-order) patterns, contrary to that found previously by Wong and Levi (2005).

The effects of spatial summation on the perception of second-order motion has received comparatively little attention. However, Zanker (1993) investigated the disruptive effect of uncorrelated visual noise on the ability to detect both first-order motion (a displaced rectangular region of random dots) and second-order motion (defined by either flicker or relative motion) across a limited range of image sizes. Performance was measured at a fixed image width (0.608°) but image height was varied in the range 0.076–4.864°. He found that when the height of the moving objects increased, sensitivity (percentage of noise superimposed on the image without destroying the perceived motion percept) continued to increase for all types of motion. In each case there was little evidence that a summation limit had been reached and changes in image size were restricted to a single spatial dimension. Thus the spatial integration area for second-order motion perception remains unclear and warrants further study. Therefore the present study investigated the effect of image size on thresholds for discriminating the direction of first-order (luminance-defined) and second-order (contrast-defined) motion.

2. Experiment 1: spatial summation of first-order and second-order motion signals

Experiment 1 investigated the effect of image size on performance for determining the drift direction of first-order (luminance-modulated dynamic noise) and second-order (contrast-modulated dynamic noise) motion.

2.1. Methods

2.1.1. Observers

Four observers (CVH, LS, MA and LA) took part in the study. CVH was an author and LS, MA, and LA were naïve observers. All had normal or corrected-to-normal visual acuity and had no history of any visual disorders.

2.1.2. Apparatus and stimuli

Stimuli were generated using a *Macintosh G5* and presented on a *Dell* monitor (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 14 bits using a Bits++ attenuator (*Cambridge Research Systems*). The mean luminance of the display was approximately 68 cd/m^2. Images were viewed binocularly and in darkness at a distance of 69.5 cm. 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, 2001a; Scott-Samuel & Georgeson, 1999).

Stimuli were vertically-oriented, 1 c/deg., luminance-modulated or contrast-modulated dynamic noise patterns, drifting at either 0.5, 2 or 8 Hz. The size of the image varied from 0.125° to 16°, horizontally and vertically. First-order motion was produced by adding a sinusoidal grating to a 1-bit, spatially 2-d, random noise carrier of 0.15 Michelson contrast. The noise carrier was generated by assigning individual (single) screen pixels (1.88 arc min) to be either "white" or "black" with equal probability and there was no spatial variation in luminance within each noise element. A new stochastic noise sample was used for each separate image in the motion sequence. The luminance profile of the first-order motion stimulus as a function of space and time, $L(x, y, t)$, can be described by the equation:

$$L(x,y,t) = L_{\text{mean}}[1 + m\cos\{2\pi(fx - \omega t) + \phi\} + cR(x,y,t)] \quad (1)$$

where L_{mean} is the mean luminance of the display (68 cd/m^2), f is spatial frequency (c/deg.), ω is temporal drift frequency (Hz), ϕ is the initial spatial phase (randomised at the beginning of each trial), m is the amplitude (modulation depth) and c is the contrast of the noise carrier $R(x, y, t)$ prior to modulation, chosen to be either −1 or +1 with probability 0.5. The modulation depth (m) of the sinusoidal luminance modulation could be varied in the range 0–1 according to the following equation:

$$m = (L_{\max} - L_{\min})/(L_{\max} + L_{\min}) \quad (2)$$

where $L_{\max}$ and $L_{\min}$ are the maximum and the minimum mean luminances in the image, respectively, averaged over adjacent noise elements with opposite polarity.

Second-order motion was produced by multiplying, rather than adding, a drifting sinusoidal waveform (unsigned for the purposes of multiplication) with a noise field:

$$L(x,y,t) = L_{\text{mean}}[1 + \langle 1 + m\cos\{2\pi(fx - \omega t) + \phi\}\rangle cR(x,y,t)] \quad (3)$$

where the parameters are identical to those in Eq. (1). The depth (m) of the contrast modulation could be varied in the range 0–1 according to the following equation:

$$m = (C_{\max} - C_{\min})/(C_{\max} + C_{\min}) \quad (4)$$

where $C_{\max}$ and $C_{\min}$ are the maximum and the minimum local Michelson contrasts in the image, respectively, computed over neighbouring noise elements with opposite polarity. Stimulus examples are shown in Fig. 1.

2.2. Procedure

A single-interval, two-alternative, forced-choice task was employed. On each trial, observers were presented with a fixation cross. Trials were self-paced. Observers initiated each trial by

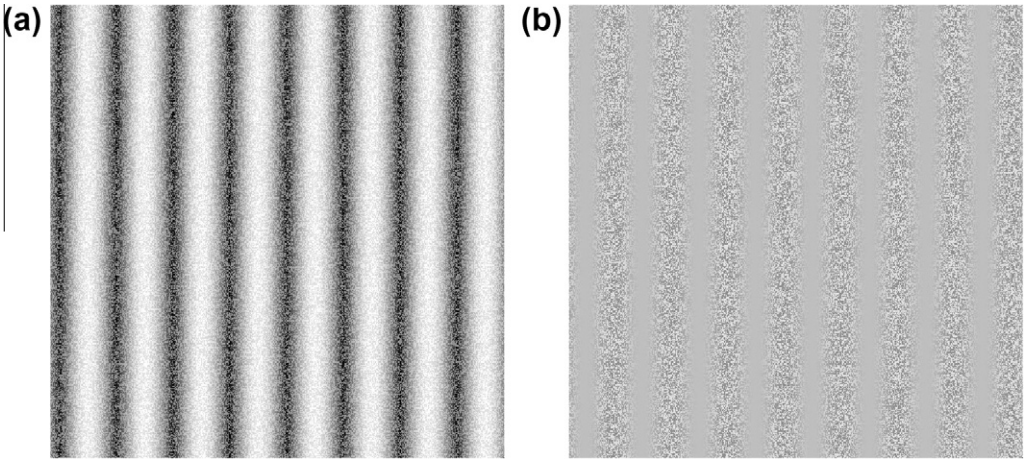

Fig. 1. Examples of the: (a) luminance-modulated (first-order) and (b) contrast-modulated (second-order) dynamic noise patterns.

pressing the spacebar, after which the stimulus was presented on-screen for 500 ms. The observers' task was to judge the motion direction of the pattern. This was always orthogonal to the pattern's orientation and was chosen to be left or right with equal probability.

On each trial the modulation depth of the moving pattern was varied according to a modified 1-up 3-down staircase designed to converge on the modulation depth corresponding to 79.4% correct performance (Levitt, 1971; Wetherill & Levitt, 1965). At the beginning of each run of trials the pattern modulation depth was initially set to a suprathreshold level (~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 four reversals in each staircase. Each observer completed a minimum of four runs of trials (i.e. staircases) for each condition and the order of testing was randomised. The mean threshold and SEM were then calculated for each condition.

3. Results

Fig. 2 shows thresholds for determining the direction of 1 c/deg. first-order and second-order motion as a function of overall stimulus size (identical in both spatial dimensions) at temporal frequencies of: (a) 0.5, (b) 2 and (c) 8 Hz.

To characterise spatial summation mechanisms for the two classes of motion, the data were fit with the following equation using a conventional least-squares procedure:

$$y = \left[\frac{(\mathrm{sgn}(a-x)+1)(\frac{x}{a})^{c} + \mathrm{sgn}(x-a)+1}{2}\right]b, \qquad (5)$$

where x is image size and a, b and c are constants. a is the image size after which there was no further performance improvement (corresponding to the knee-point of the function) and asymptoted at the lowest modulation depth threshold (b) and c determines the slope of the descending limb of the function (on log–log co-ordinates).

Note that sgn() is the signum function and is equal to either +1, 0 or −1 depending on whether the argument in parentheses is >0, 0 or <0, respectively. It is important to emphasise that the knee-point parameter of Eq. (5) was not chosen arbitrarily by eye, but was determined automatically (directly) by the fitting procedure. This ensured that the location of a knee-point, if indeed present in the data, was objective and unbiased. Curve fit values for each observer (CVH, LS, MA and LA) and stimulus type (luminance-modulated and contrast-modulated dynamic noise) are given in Table 1.

Unsurprisingly thresholds for second-order motion were markedly higher than for first-order motion. Observers could reliably determine first-order motion direction across all image sizes tested (with the exception of observer MA who could not accurately identify direction below an image size of 0.25° at 8 Hz). However second-order motion direction could not be determined below an image size of 0.25° at temporal frequencies of 0.5 and 2 Hz and 0.5° at 8 Hz (again with the exception of observer MA who was unable to determine the direction of second-order motion below an image size of 1° at 8 Hz). The general difference in asymptote and the increased variability between observers' performance at 8 Hz for second-order motion may be due, at least in part, to the fact that this temporal frequency is approaching the upper acuity limit for second-order motion stimuli (e.g. Hutchinson & Ledgeway, 2006).

In relation to spatial summation, thresholds for discriminating the direction of first-order and second-order motion decreased as image size increased. As is typical of spatial summation, after an initial rapid improvement in performance at relatively small image sizes, performance began to asymptote. Fig. 2 combined with the values derived from fitting Eq. (5) to the data revealed a number of important aspects of spatial summation for first-order and second-order motion. The slope of the initial performance improvement was steeper for first-order motion compared to second-order motion. Average (across observers) slopes for first-order motion were −1.178 (±0.07), −0.97 (±0.67) and −0.85 (±0.07) at temporal frequencies of 0.5, 2 and 8 Hz, respectively. The corresponding values for second-order motion were −0.80 (±0.16), −0.78 (±0.05) and −0.65 (±0.06), respectively. In addition, performance consistently reached asymptotic levels at smaller image sizes for second-order motion [2.54 (±0.52), 2.56 (±1.06) and 5.12 (±1.05) deg. at temporal frequencies of 0.5, 2 and 8 Hz, respectively], compared to first-order motion [3.48 (±0.46), 4.97 (±2.48) and 8.11 (±1.37) deg.]. Finally, in terms of the effects of temporal frequency, in general, slopes became shallower and summation areas became larger as temporal frequency increased. This was true of both first-order motion and second-order motion. Another point worth noting is that in a number of cases, estimates of the slope values were less reliable (errors were larger) for second-order than for first-order motion.

4. Experiment 2: luminance-defined gratings vs. luminance-modulated dynamic noise

Some of the findings of Experiment 1 are not in agreement with some previous studies that have examined spatial summation of

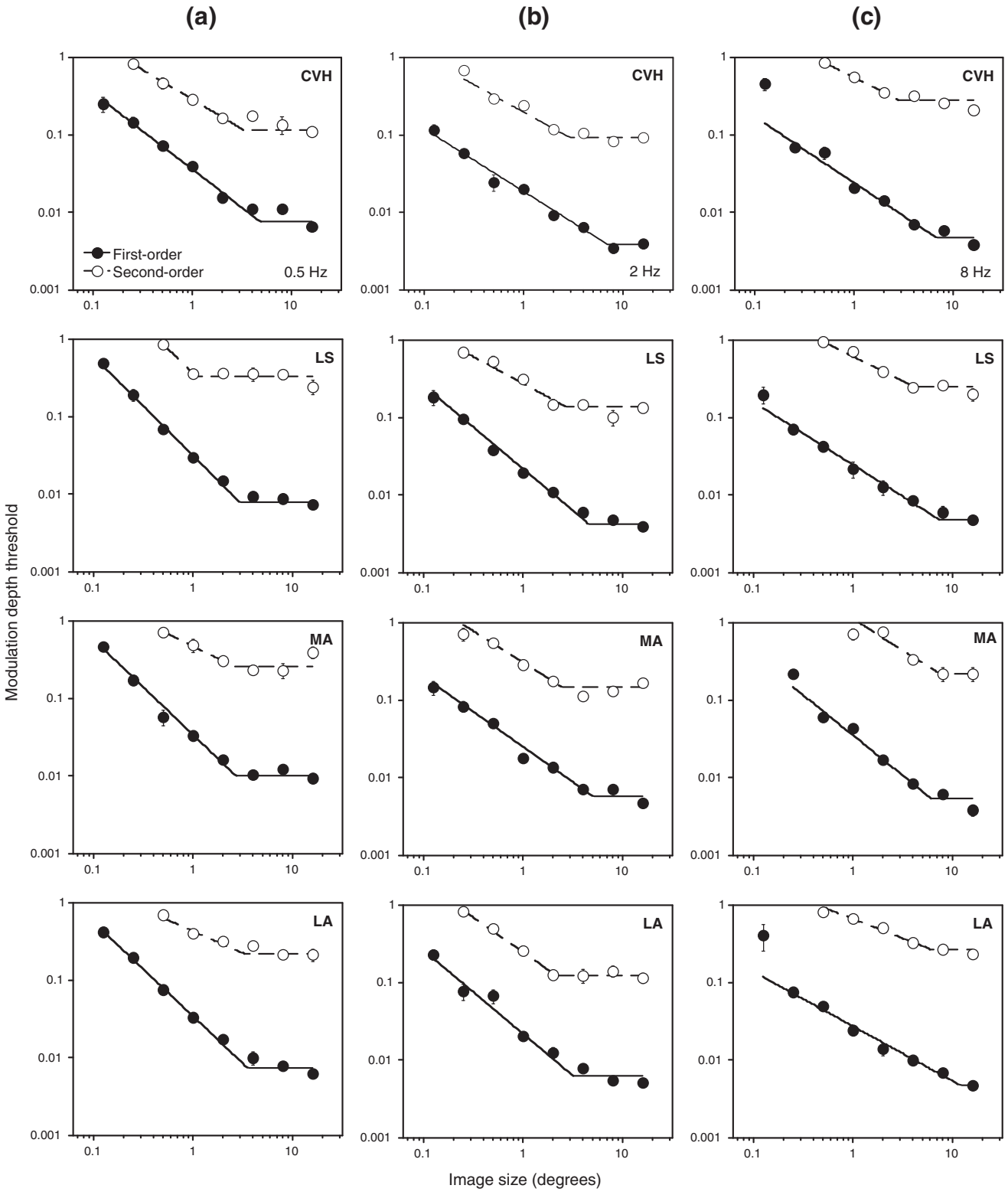

Fig. 2. Thresholds for discriminating the direction of first-order (closed circles) and second-order (open circles) motion as a function of stimulus size. Note that in this case the width and height of the image were always identical, so image size refers to the spatial extent in each dimension. Patterns had a spatial frequency of 1 c/deg. and drifted at temporal frequencies of either: (a) 0.5, (b) 2 or (c) 8 Hz. Data are shown for four observers (CVH, LS, MA and LA). Error bars represent ±1 SEM. Data have been fit with Eq. (5).

first-order and second-order signals in the spatial domain (e.g. Sukumar & Waugh, 2007). As such, the precise nature of spatial summation may be heavily dependent on the types of stimuli employed and, perhaps most importantly in this case, whether they are stationary or moving. In light of this, Experiment 2 repeated a previous experiment with first-order stimuli in the spatial domain (Schofield & Georgeson, 1999) but with moving stimuli.

Schofield and Georgeson (1999) have previously measured detection thresholds for stationary first-order Gaussian blobs, defined either by luminance or luminance-modulated noise, as a

Table 1
Curve fit parameters for each observer for luminance-modulated dynamic noise (first-order motion) and contrast-modulated dynamic noise (second-order motion). The change in threshold as a function of image size was, in most cases, well described by Eq. (5).

Observer	Temporal frequency (Hz)	Luminance-modulated dynamic noise (LMDN)				Contrast-modulated dynamic noise (CMDN)			
		a	*b*	*c*	R^2	*a*	*b*	*c*	R^2
CVH	0.5	4.775 (0.395)	0.0077 (0.0003)	−0.984 (0.047)	0.842	3.294 (0.474)	0.116 (0.006)	−0.764 (0.060)	0.964
	2	7.007 (0.589)	0.0039 (0.0001)	−0.807 (0.030)	0.974	2.951 (0.319)	0.092 (0.003)	−0.709 (0.044)	0.946
	8	6.859 (1.040)	0.0047 (0.0003)	−0.849 (0.052)	0.890	2.804 (0.497)	0.282 (0.006)	−0.649 (0.097)	0.548
LS	0.5	2.931 (0.288)	0.0079 (0.0004)	−1.280 (0.058)	0.969	1.068 (0.141)	0.333 (0.025)	−1.256 (0.204)	0.975
	2	4.590 (0.399)	0.0043 (0.0002)	−1.076 (0.028)	0.998	2.671 (0.503)	0.139 (0.005)	−0.692 (0.062)	0.969
	8	7.315 (1.363)	0.0048 (0.0002)	−0.813 (0.060)	0.968	4.021 (0.634)	0.253 (0.013)	−0.635 (0.043)	0.996
MA	0.5	2.686 (0.261)	0.0101 (0.0005)	−1.231 (0.043)	0.980	2.62 (0.445)	0.259 (0.025)	−0.624 (0.098)	0.819
	2	5.084 (0.434)	0.0057 (0.0003)	−0.906 (0.029)	0.943	2.477 (0.252)	0.148 (0.004)	−0.806 (0.054)	0.907
	8	6.116 (0.759)	0.0054 (0.0003)	−1.035 (0.054)	0.920	7.515 (1.506)	0.221 (0.029)	−0.802 (0.076)	0.822
LA	0.5	3.537 (0.318)	0.0074 (0.0002)	−1.217 (0.044)	0.983	3.212 (1.102)	0.223 (0.012)	−0.565 (0.135)	0.914
	2	3.196 (0.431)	0.0063 (0.0002)	−1.078 (0.075)	0.777	2.159 (0.327)	0.124 (0.012)	−0.897 (0.059)	0.991
	8	12.152 (1.829)	0.0048 (0.0004)	−0.702 (0.032)	0.963	6.127 (0.657)	0.269 (0.006)	−0.500 (0.031)	0.993

function of blob size. Interestingly they found that the range of spatial integration was different for these two first-order stimuli, in that sensitivity saturated at smaller blob sizes for luminance-defined blobs (~10 arc min) than for luminance-modulated noise blobs (~40 arc min). To investigate whether these differences in sensitivity to first-order stimuli in the spatial domain transfer to the perception of motion, we measured performance for discriminating the direction of conventional luminance-defined gratings as a function of stimulus size and compared performance to that for discriminating the direction of luminance-modulated dynamic noise.

4.1. Methods

All experimental methods were identical to those employed in Experiment 1 with the exception that stimuli were 1 c/deg. luminance-defined gratings (i.e. no noise carrier was present), drifting at 2 Hz.

4.2. Results

Fig. 3 shows thresholds for determining the direction of 1 c/deg. luminance-defined gratings. For comparison purposes the equivalent data for luminance-modulated dynamic noise have also been replotted from Fig. 2b. Curve fit values for luminance-defined gratings derived from fitting Eq. (5) to each observer's data are given in Table 2. In agreement with previous findings for static patterns (e.g. Schofield & Georgeson, 1999), thresholds for judging the direction of luminance-defined gratings were typically lower and summation areas were smaller than those for identifying the direction of luminance-modulated dynamic noise [3.94 (±0.49) and 4.97 (±2.48) deg., respectively]. However the initial slopes of the descending portion of the functions (i.e. prior to asymptote) for the two stimulus types were similar [−1.08 (0.05) and −0.97 (±0.67), respectively]. It may be worth noting that, in terms of the image size at which performance asymptoted, although the curve fitting procedure did converge on a knee-point, in some cases performance appeared, at least qualitatively, to still be improving at the largest image size employed (i.e. 16°). This was true of both luminance-defined gratings and luminance-modulated dynamic noise.

Also, at a given image size, thresholds for conventional luminance gratings were consistently lower than those for luminance-modulated dynamic noise. These findings are in agreement with Schofield and Georgeson (2003) who suggested that the addition of a noise carrier to a luminance-defined grating may mask the first-order signal and hence lead to poorer sensitivity.

5. Experiment 3: image height or number of cycles?

To separate the effects of curtailing image height and image width, in a final experiment we re-measured thresholds for determining the drift direction of first-order (luminance-modulated dynamic noise) and second-order (contrast-modulated dynamic noise) motion. However, rather than varying both the width *AND* the height of the stimuli to vary image size, in the present experiment we kept one dimension constant (either height or width) at 16° and varied the other from 0.125° to 16°.

5.1. Methods

The experimental methods were identical to those employed in Experiment 1, with the exception that image size only varied in one dimension, either horizontally or vertically, whilst the spatial extent in the other dimension remained at 16°. We measured thresholds for two observers, CVH (an author) and AA (a naïve observer), for first-order and second-order motion at 1 c/deg. and 2 Hz.

5.2. Results and discussion

Fig. 4a shows the results for two observers when only the image width (number of visible cycles) was varied. Fig. 4b shows the results for the same observers when the image height was varied, thereby maintaining the number of cycles. Curve fit values derived from fitting Eq. (5) to the data are given in Table 3.

Irrespective of whether image height or width was varied, performance for discriminating the direction of first-order motion became asymptotic at larger image sizes than performance for discriminating the direction of second-order motion, consistent with the results of Experiment 1. However there were some differences in the data shown in Fig. 4 depending on which dimension varied. Curtailing the image width (i.e. the number of grating cycles presented on-screen) produced results that most closely resembled those produced in Experiment 1 (see Fig. 2 & Table 1 for comparison). Varying image height, whilst image width remained at 16°, produced a greater difference between the image size at which performance for each variety of motion flattened (spatial summation areas), compared to when image width varied and height remained constant at 16°. The slopes representing the descending limb of the functions were also typically shallower.

6. General discussion

The findings of the present study have shown: (1) smaller spatial summation areas for second-order (contrast-modulated

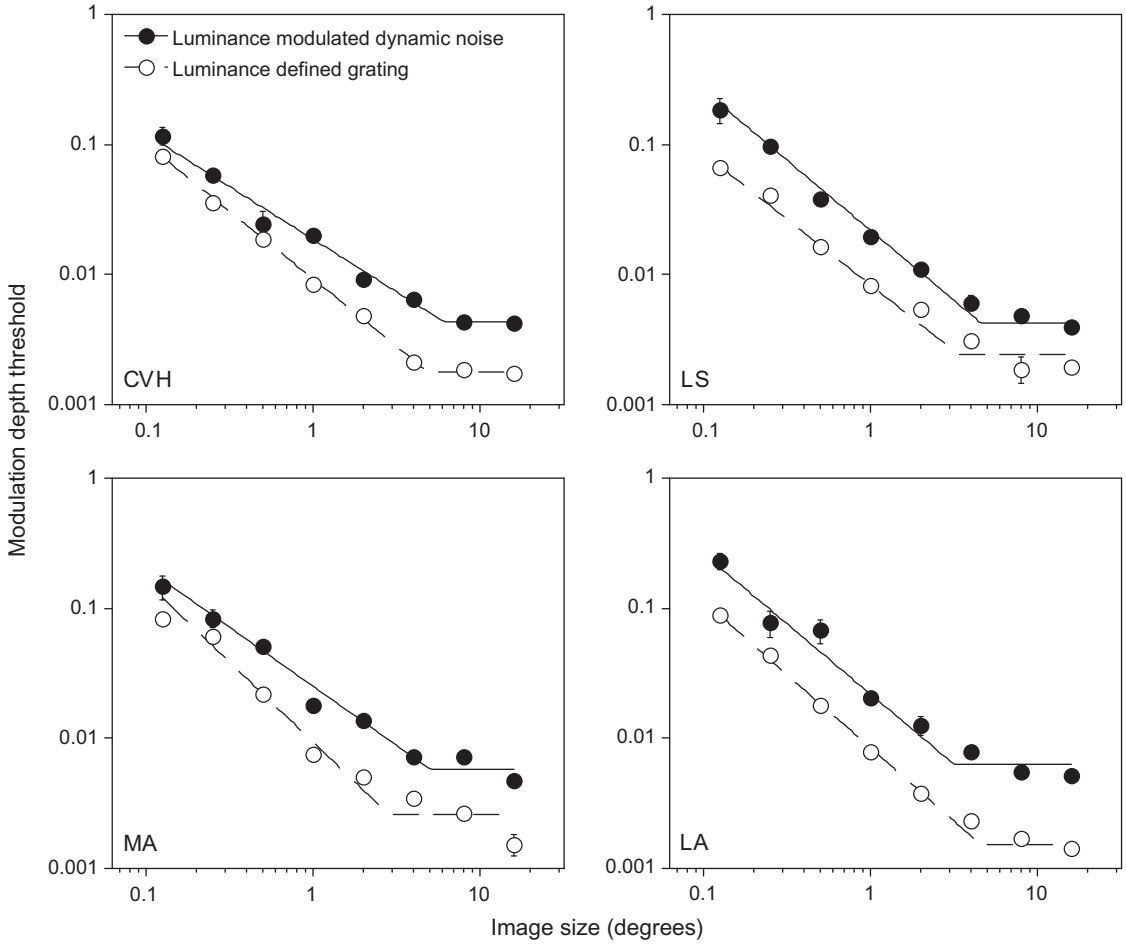

Fig. 3. Thresholds for discriminating the direction of luminance gratings (open circles) and luminance-modulated dynamic noise (closed circles) as a function of stimulus size. Patterns had a spatial frequency of 1 c/deg. and drifted at 2 Hz. Data are shown for four observers and error bars represent ±1 SEM. Data for luminance-modulated dynamic noise have been replotted from Fig. 2.

Table 2
Curve fit parameters for each observer for conventional luminance-defined gratings (first-order motion). The change in threshold as a function of image size was well described by Eq. (5).

Observer	*a*	*b*	*c*	R^2
CVH	4.964 (0.328)	0.0018 (0.00007)	−1.037 (0.021)	0.994
LS	3.344 (0.246)	0.0025 (0.00011)	−1.010 (0.036)	0.966
MA	2.870 (0.185)	0.0026 (0.00012)	−1.227 (0.030)	0.923
LA	4.568 (0.281)	0.0015 (0.00006)	−1.126 (0.025)	0.985

dynamic noise) compared to first-order (luminance-modulated dynamic noise) motion and (2) slightly smaller spatial summation areas for conventional luminance gratings compared to luminance-modulated dynamic noise.

Spatial summation ratios from Experiments 1 to 3 are shown in Fig. 5. Performance for determining the direction of first-order motion asymptoted at larger image sizes compared to second-order motion (Fig. 5a). This difference, averaged across observers, corresponded to factors of 1.580 (±0.399), 1.906 (±0.195) and 1.766 (±0.344) for 0.5, 2 and 8 Hz, respectively. Performance levels for determining the direction of luminance-modulated dynamic noise became asymptotic at larger image sizes compared to luminance-defined gratings. However this difference was much less pronounced (Fig. 5b), being in the region of around 1.25 (±0.224). Finally, when image width (number of cycles) varied and height remained constant at 16°, performance for determining the direction of first-order motion asymptoted at larger image sizes compared to second-order motion by a factor of 1.388 (±0.24) and when image height varied and width remained constant this difference was even greater. In this case (Fig. 5c), performance for determining the direction of first-order motion asymptoted at larger image sizes compared to second-order motion by a factor of 2.747 (±1.043).

The slopes derived from fitting Eq. (5) to the data produced average values (across conditions and observers) of −0.99 (±0.054) for first-order motion and −0.74 (±0.057) for second-order motion. The slopes corresponding to the initial performance improvement for first-order motion are in broad agreement with the linear systems-based model of Graham (1989), possibly reflecting summation within a single-channel or neuron receptive field. The slopes corresponding to the initial performance improvement for second-order motion lie somewhere between linear (−1) and probability (−0.5) summation, possibly reflecting some degree of summation both within (linear) and between (across) channels based on a pooling of each channel's independent output. That the slopes corresponding to the summation functions for second-order motion are typically shallower may reflect more between-channel pooling than for first-order motion. This may be due to broader tuning (less specificity and sensitivity) of second-order channels such that between-channel pooling provides a stronger signal of motion than within-channel summation from a less sensitive channel.

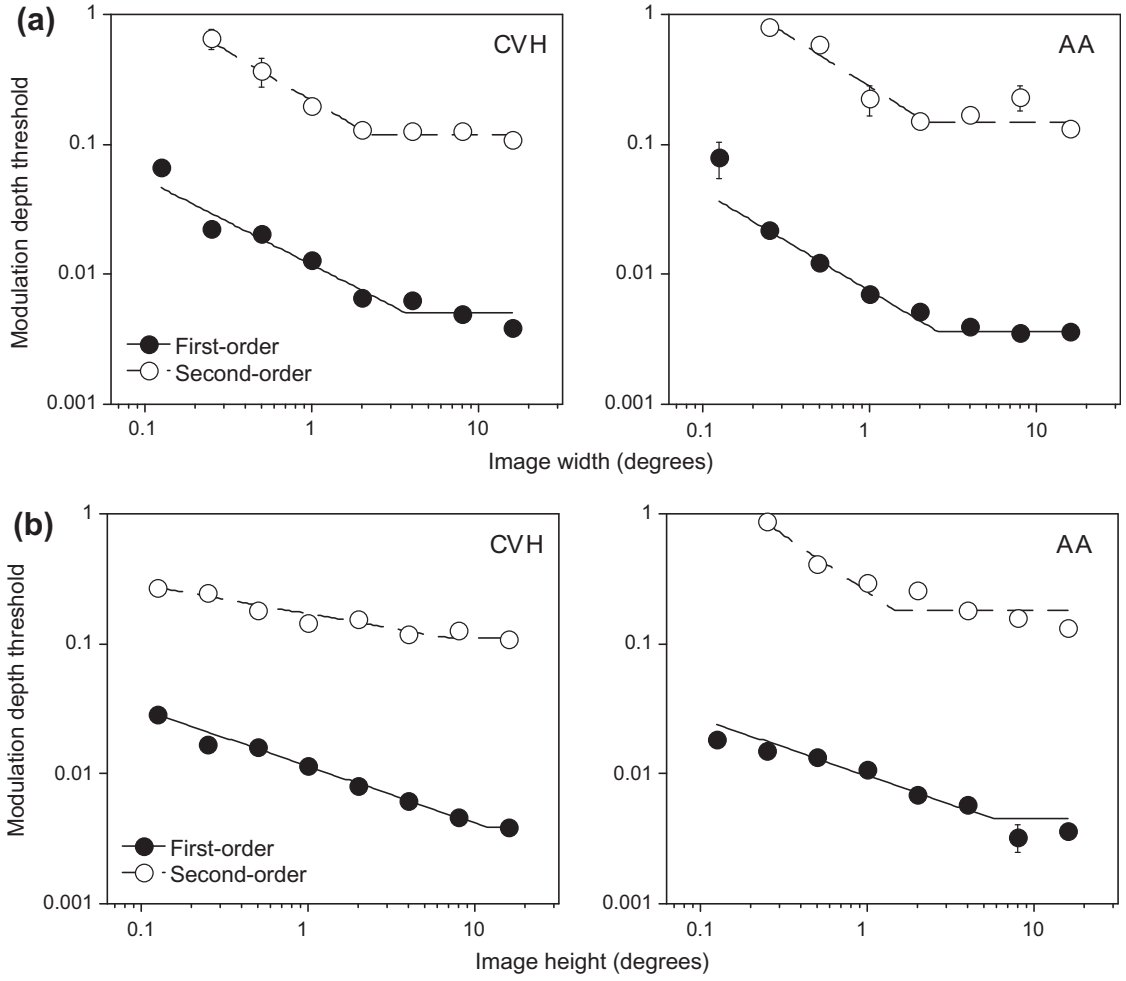

Fig. 4. Thresholds for discriminating the direction of first-order (closed circles) and second-order (open circles) motion as a function of (a) stimulus width (number of cycles) and (b) stimulus height. Patterns had a spatial frequency of 1 c/deg. and drifted at 2 Hz. Data is shown for two observers and error bars represent ±1 SEM.

Table 3
Curve fit parameters for each observer for luminance-modulated dynamic noise (first-order motion) and contrast-modulated dynamic noise (second-order motion) when the image width varied and its height remained constant at 16°, or vice versa.

Condition	Observer	Luminance-modulated dynamic noise (LMDN)				Contrast-modulated dynamic noise (CMDN)			
		a	*b*	*c*	R^2	*a*	*b*	*c*	R^2
Image width varies	CVH	3.652 (0.456)	0.0051 (0.0003)	−0.654 (0.037)	0.890	2.243 (0.184)	0.119 (0.001)	−0.752 (0.083)	0.487
	AA	2.544 (0.417)	0.0037 (0.00009)	−0.763 (0.073)	0.961	2.216 (0.203)	0.148 (0.011)	−0.803 (0.968)	0.968
Image height varies	CVH	11.99 (4.09)	0.0038 (0.0005)	−0.439 (0.002)	0.990	7.035 (1.236)	0.111 (0.002)	−0.221 (0.016)	0.945
	AA	5.534 (0.95)	0.0046 (0.0003)	−0.437 (0.019)	0.912	1.460 (0.279)	0.182 (0.005)	−0.866 (0.11)	0.653

Models of first-order and second-order motion propose that the image first undergoes orientation-selective filtering at a given spatial scale, after which two parallel pathways emerge, one for encoding first-order motion and the other for encoding second-order motion. After initial filtering, one pathway simply extracts the first-order motion energy whereas the other pathway incorporates rectification (or response squaring) followed by an additional orientation-selective filtering stage in order to make the second-order information visible to conventional 'first-order' motion detectors. It has been proposed that first-order motion undergoes simple luminance-based filtering in V1 whereas for second-order motion, first-order information is first extracted in V1, after which it is rectified, and the post-rectified outputs further filtered by V2 (e.g. Wilson, Ferrera, & Yo, 1992). The differences between spatial summation areas for the two types of motion may also reflect differences in the type and neural location of summation. Whereas for first-order motion it is likely that our findings tentatively reflect single-channel spatial summation mechanisms perhaps in V1, for second-order motion our findings may reflect single-channel summation of the second-order image by V1, followed by post-rectification summation of pooled outputs in say area V2, producing shallower slopes and different summation areas overall. However, this conclusion must remain speculative as comparatively little is known about the neural basis of second-order motion processing and pinpointing the exact anatomical locus where second-order motion sensitivity first emerges in visual cortex is fraught with difficulty (e.g. due to the extensive feedback connections between visual areas). Although evidence from human brain-imaging studies (fMRI) is somewhat equivocal with regard to the existence of grossly distinct cortical regions for encoding each type of motion

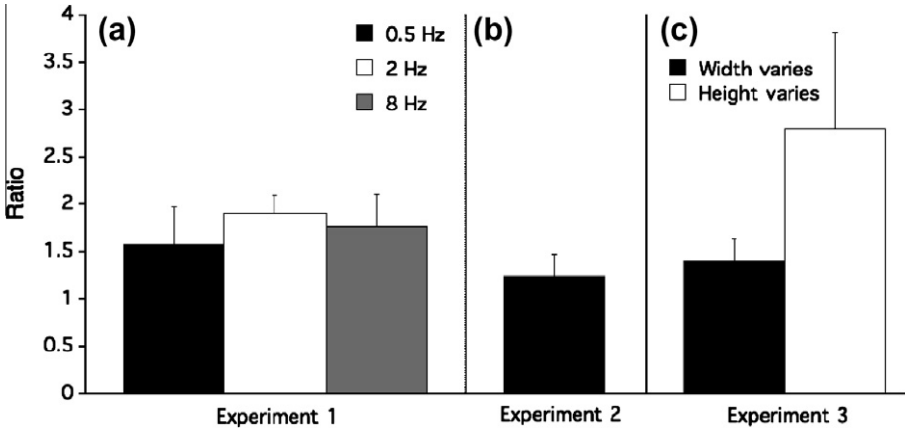

Fig. 5. Average spatial summation (image size at which performance became asymptotic) ratios for: (a) luminance-modulated (first-order) and contrast-modulated (second-order) dynamic noise in Experiment 1, (b) luminance-modulated dynamic noise and luminance-defined gratings in Experiment 2 and (c) luminance-modulated (first-order) and contrast-modulated (second-order) dynamic noise in Experiment 3. Error bars represent +1 SEM calculated across observers.

(Dumoulin, Baker, Hess, & Evans, 2003; Nishida, Sasaki, Murakami, Watanabe, & Tootell, 2003; Seifert, Somers, Dale, & Tootell, 2003)) a recent adaptation study (Ashida, Lingnau, Wall, & Smith, 2007) suggests that two separate, but co-localised, mechanisms may exist in extrastriate areas MT, MST and V3A. This is commensurate with electrophysiological research which has identified neurones sensitive to both first-order motion and second-order motion in many cortical areas of the mammalian visual system. These include areas 17 and 18 in the cat (Zhou & Baker, 1993, 1994, 1996), and primate areas V1, MT and MSTd (e.g. Albright, 1992; Chaudhuri & Albright, 1997; Geesaman & Andersen, 1996).

Our findings may also reflect differences in sensitivity to first-order and second-order motion. In the fovea the highest modulation depth possible for second-order motion stimuli is around 10 times threshold, whereas for an equivalent first-order stimulus the maximum modulation depth possible is around 100 times threshold. Such marked differences in sensitivity to first-order and second-order motion extend into the periphery where sensitivity to second-order motion is particularly poor. For low stimulus drift rates (e.g. 0.5 Hz), the direction of contrast-modulated gratings cannot even be identified by 8° eccentricity. Rather, contrast-modulated gratings appear to remain stationary (Pantle, 1992). A number of studies have suggested that sensitivity to second-order motion declines more rapidly with eccentricity than for first-order motion (Zanker, 1997). However, others suggest that sensitivity to the two motion types does in fact scale similarly with eccentricity (e.g. Smith, Hess, & Baker, 1994; Smith & Ledgeway, 1998; Solomon & Sperling, 1995). Regardless of the rate of sensitivity decline with increasing eccentricity, second-order motion perception is certainly much more restricted than first-order motion perception in peripheral vision. This may represent an important functional reason why second-order perception asymptotes at much smaller images sizes than first-order motion perception. In the case of first-order motion, although sensitivity is markedly poorer in the periphery compared to the fovea, it is still sufficient to lead to additive performance improvements as stimuli are spatially extended. However, for second-order motion, it may be that spatially extending stimuli into the periphery, where sensitivity is so poor, is of little further benefit.

Our findings are consistent with those of Wong and Levi (2005) who characterised spatial summation for static first-order and second-order patterns. They found that, whereas for second-order patterns performance asympoted at around six cycles, for first-order patterns performance failed to asymptote over the range of image sizes tested. Our findings for moving stimuli are, however, at odds with those of Sukumar and Waugh (2007) who found that summation areas for static images were larger for second-order compared to first-order stimuli. The disparity between the findings of Sukumar and Waugh (2007) and Wong and Levi's (2005) study also demonstrates that the nature and extent of spatial summation appears to be heavily dependent on the precise characteristics of a stimulus. For example, they may be due, in part, to the number of visible grating cycles.

In terms of stimulus size/number of cycles being the limiting factor in sensitivity to first-order and second-order motion direction, in agreement with previous observations by Cavanagh and Mather (1989), we found that observers required at least 0.25 cycles of grating (corresponding to 0.25° at a spatial frequency of 1 c/deg.) to be visible on-screen to accurately determine second-order motion direction. On the other hand, observers could accurately determine first-order motion direction with only 0.125 visible cycles. They may, in principle, have been able to successfully judge first-order motion direction with even fewer cycles. However, in the present study the smallest image size employed was 0.125° (i.e. 0.125 grating cycles). Interestingly this pattern of results closely mirrors the differences found between first-order motion and second-order motion on psychophysical tasks requiring temporal summation of direction information. For example, the direction of a sinusoidal contrast modulation, unlike a luminance modulation, cannot be identified when the stimulus exposure duration is less than about 200 ms (Derrington, Badcock, & Henning, 1993). One potential explanation of this phenomenon is that second-order motion detectors in human vision are simply less selective for stimulus direction than first-order motion sensors, and thus are more vulnerable to the disruptive effects of brief presentations, which introduce spurious motion in the opposite direction (see Fig. 1 of Ledgeway & Hess, 2002). It is also interesting to note that spatially restricting (windowing) a sinusoidal motion stimulus, such that the number of spatial cycles is severely curtailed, increases the spatial frequency bandwidth of the modulation signal and leads to an analogous smearing of the motion energy across the two directions. This inevitably introduces some degree of directional ambiguity into the stimulus. Thus the present finding that the direction of second-order motion, unlike first-order motion, could not be determined when image width was below 0.25°, is consistent with the notion that the mechanisms that encode each variety of motion differ in direction selectivity. Curtailing the image height (but not the width, so the number of cycles is constant) of a sinusoidal motion pattern introduces artifacts with energy at different orientations to the modulation signal,

but not spurious motion in the opposite motion (e.g. see Anderson & Burr, 1991).

In agreement with Schofield and Georgeson (1999) we found that, thresholds were typically higher and spatial summation areas slightly larger for luminance-modulated dynamic noise compared to luminance-defined gratings. However, as in Schofield and Georgeson's (1999) study, both stimulus types produced similar curves (initial slopes). This may be due, in part, to the 'temporal noise' in the luminance-modulated dynamic noise stimulus 'masking' the fist-order signal. However, in the present study, it was important that we used luminance-modulated dynamic noise so that our first-order and second-order (which also necessarily contained a noise carrier) patterns were as equivalent as possible.

In conclusion, the current literature on the nature of spatial summation of stationary first-order and second-order information in the human visual system is a matter for debate. This may be because the precise nature of all aspects of spatial summation (the slope of the initial performance improvement as stimulus size increases and the size at which performance asymptotes) appears to be critically dependent on the nature of the stimulus, and the manner in which data are expressed. The specific spatial characteristics of the stimuli that influence performance are unclear. In our case, the spatial summation functions in the present study support the findings of Wong and Levi (2005) and Schofield and Georgeson (1999). Moreover, in the case of moving patterns, stimulus drift rate does not appear to influence greatly these parameters. Finally, it is worth noting that, as is the case in spatial domain, spatial summation thresholds for first-order and second-order motion are not always clear. Consequently, it is difficult to draw firm conclusions about the functional basis of the differences between them or to develop a quantitative model at this stage, either for spatial or temporal vision.

Acknowledgments

The authors wish to thank observers LS, MA and LA for participating in the study and two anonymous observers for their comments.

References

Albright, T. D. (1992). Form-cue invariant motion processing in primate visual cortex. *Science, 255*, 1141–1143.
Anderson, S. J., & Burr, D. (1991). Spatial summation properties of directionally selective mechanisms in human vision. *Journal of the Optical Society of America A, 8*, 1330–1339.
Ashida, H., Lingnau, A., Wall, M. B., & Smith, A. T. (2007). FMRI adaptation reveals separate mechanisms for first-order and second-order motion. *Journal of Neurophysiology, 97*, 1319–1325.
Baker, C. L. Jr., (1999). Central neural mechanisms for detecting second-order motion. *Current Opinion in Neurobiology, 9*, 461–466.
Barlow, H. B. (1958). Temporal and spatial summation in human vision at different background intensities. *Journal of Physiology, 14*, 1337–1350.
Benton, C. P., & Johnston, A. (2001). A new approach to analysing texture-defined motion. *Proceedings of the Royal Society B, 268*, 2435–2443.
Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology (London), 197*, 551–566.
Cavanagh, P., & Mather, G. (1989). Motion: The long and short of it. *Spatial Vision, 4*, 103–129.
Chaudhuri, A., & Albright, T. D. (1997). A comparison of neuronal responses to edges defined by luminance or temporal texture in macaque area V1. *Visual Neuroscience, 14*, 949–962.
Derrington, A. M., Badcock, D., & Henning, B. (1993). Discriminating the direction of second-order motion at short stimulus durations. *Vision Research, 33*, 1785–1794.
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.
Geesaman, B. J., & Andersen, R. A. (1996). The analysis of complex motion patterns by form/cue invariant MSTd neurons. *Journal of Neuroscience, 16*, 4716–4732.
Graham, N. (1989). *Visual pattern analyzers*. Psychology Press: New York.
Gurnsey, R., Fleet, D., & Potechin, C. (1998). Second-order motions contribute to vection. *Vision Research, 38*, 2801–2816.
Howell, E. R., & Hess, R. F. (1978). The functional area for summation to threshold for sinusoidal gratings. *Vision Research, 18*, 369–374.
Hutchinson, C. V., & Ledgeway, T. (2006). Sensitivity to spatial and temporal modulations of first-order and second-order motion. *Vision Research, 46*, 324–335.
Johnston, A., Benton, C. P., & McOwan, P. W. (1999). Induced motion at texture-defined motion boundaries. *Proceedings of the Royal Society of London B, 266*, 2441–2450.
Johnston, A., McOwan, P. W., & Buxton, H. (1992). A computational model for 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.
Landy, M. S., & Oruc, I. (2002). Properties of second-order spatial frequency channels. *Vision Research, 42*, 2311–2329.
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.
Lu, Z.-L., & Sperling, G. (1995). The functional architecture of human visual motion perception. *Vision Research, 35*, 2697–2722.
Lu, Z.-L., & Sperling, G. (2001a). Sensitive calibration and measurement procedures based on the amplification principle in motion perception. *Vision Research, 41*, 2355–2374.
Lu, Z.-L., & Sperling, G. (2001b). Three-systems theory of human visual motion perception: Review and update. *Journal of the Optical Society of America A, 18*, 2331–2370.
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.
Pantle, A. (1992). Immobility of some second-order stimuli in human peripheral vision. *Journal of the Optical Society of America A, 9*, 863–867.
Robson, J. G., & Graham, N. (1981). Probability summation ad regional variation in contrast sensitivity across the visual field. *Vision Research, 21*, 409–418.
Schofield, A. J., & Georgeson, M. A. (1999). Sensitivity to modulations of luminance and contrast in visual white noise: Separate mechanisms with similar behaviour. *Vision Research, 39*, 2697–2716.
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. (1994). Correspondence-based and energy-based detection of second-order motion in human vision. *Journal of the Optical Society of America A, 11*, 1940–1948.
Smith, A. T., 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. (1998). Sensitivity to second-order motion as a function of temporal frequency and eccentricity. *Vision Research, 38*, 403–410.
Solomon, J. A., & Sperling, G. (1995). First-order and second-order motion and texture resolution in central and peripheral vision. *Vision Research, 35*, 59–64.
Sukumar, S., & Waugh, S. J. (2007). Separate first- and second-order processing is supported by spatial summation estimates at the fovea and eccentrically. *Vision Research, 47*, 581–596.
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.
Wong, E. H., & Levi, D. M. (2005). Second-order spatial summation in amblyopia. *Vision Research, 45*, 2799–2809.
Zanker, J. M. (1993). Theta motion: A paradoxical stimulus to explore higher order motion extraction. *Vision Research, 33*, 553–569.
Zanker, J. M. (1997). Second-order motion in the peripheral visual field. *Journal of the Optical Society of America A, 14*, 1385–1392.
Zhou, Y.-X., & Baker, C. L. (1993). A processing stream in mammalian visual cortex neurons for non-Fourier responses. *Science, 261*, 98–101.
Zhou, Y.-X., & Baker, C. L. (1994). Envelope-responsive neurons in areas 17 and 18 of cat. *Journal of Neurophysiology, 72*, 2134–2150.
Zhou, Y.-X., & Baker, C. L. (1996). Spatial properties of envelope-responsive cells in area 17 and 18 neurons of the cat. *Journal of Neurophysiology, 75*, 1038–1050.