Motion defined exclusively by second-order characteristics does not evoke optokinetic nystagmus

LAURENCE R. HARRIS 1 AND ANDREW T. SMITH 2
1Department of Psychology, York University, Toronto, Ontario, M3J 1P3, Canada
2Department of Psychology, University of Wales College of Cardiff, Cardiff, CF1 3YG, UK
(RECEIVED February 17, 1992; ACCEPTED May 15, 1992)

Abstract
We showed high-contrast, second-order motion stimuli to subjects whilst recording their horizontal eye movements. These stimuli were very poor at evoking optokinetic nystagmus. Smooth-pursuit eye movements and fixation were reduced by a masking band ±2.5 deg above and below an imaginary fixation point. First-order stimuli evoked vigorous optokinetic nystagmus (OKN) under identical conditions and also when matched for apparent contrast. These findings are discussed in terms of the site of detection of second-order motion.

Keywords: Motion perception, Optokinetic nystagmus, Second-order motion, Visual noise

Introduction
This paper addresses the question: at which stage of visual processing is motion information extracted from a visual stimulus? The answer to this question may depend on what it is that moves. The properties of a visual stimulus can be described in terms of a hierarchy with more elaborate properties being built up from more fundamental building blocks. The fundamental properties of any visual image are luminance and color. These properties have come to be called the first-order characteristics of the image. Higher-order attributes, such as depth or contrast, have subsequently to be derived from these first-order characteristics. Motion defined in terms of such derived characteristics might be detected at a higher level of processing than motion of first-order characteristics.

Information about form is normally contained in the distribution of first-order information over space. But it is possible in the laboratory to generate a stimulus in which information about form is not specified directly by the variations in first-order characteristics across the image. An example is a field of random visual noise whose contrast is modulated sinusoidally in one dimension across space. Such a pattern is readily visible as a grating, but the grating is defined only by variations in a second-order characteristic, in this case contrast. Its luminance (when averaged over an area containing a number of pixels of the noise) is the same at all points in the image. The same distinction may be made in the temporal domain. Normally motion is defined in terms of changes in first-order spatial characteristics over time. Not surprisingly, current computational models (Adelson & Bergen, 1985; van Santen & Sperling, 1985; Heeger, 1987; Grzywacz & Yuille, 1990) are mostly based on algorithms for detecting such first-order motion. But it is possible to generate patterns in which motion is defined exclusively by second-order characteristics. In our example, if the sinusoid used to modulate contrast is moved while the noise remains static, the motion, which is again readily perceived, is defined only by second-order characteristics. Such a stimulus has been described as “drift-balanced” (Chubb & Sperling, 1988) to denote the fact that Fourier analysis in the luminance domain reveals no net motion energy.

In primates, although not in some lower mammals, physiological evidence shows clearly that neither the retina nor the lateral geniculate nucleus of the thalamus (LGN) contains neurons sensitive to direction of motion (Hubel & Wiesel, 1961; Shapley & Lennie, 1985). Thus, the question posed at the beginning of the paper becomes: at which stage of cortical visual processing is motion information extracted from a visual stimulus? A partial answer is well established: direction sensitivity is found in neurons at the earliest stages of processing in the striate cortex (area V1: Shapley & Lennie, 1985). However, these neurons have almost invariably been tested with first-order motion patterns only. The site of detection of motion of second-order characteristics is unknown.

A simple way to progress towards locating the site of second-order motion detection is to find a way to tap motion signals at some intermediate level of processing and see whether second-order motion is represented at that level. Such an opportunity is afforded by the study of optokinetic nystagmus (OKN). Optokinetic eye movements have recently been revealing concerning the nature and site of motion processing. A drifting plaid (made out of two drifting gratings) can be arranged such that its coherent direction moves in one direction while its components move in other, distinguishable directions (Adelson &
Movshon, 1982; Movshon et al., 1986). The direction of OKN eye movements evoked by such plaids is close to that of the coherent (Harris et al., 1991). This direction could be obtained either from nonoriented detectors (i.e. the direct retinal pathway to the brain stem) or by reconstruction from the output of orientation-selective channels. The direct brain-stem pathway is completely crossed and therefore OKN eye movements relying on this route can only be nasally directed when viewing is monocular. OKN evoked by these plaids shows no nasal-temporal asymmetries when viewed monocularly (cf. cats, Smith & Harris, 1991), which suggests a minimal involvement of the direct retina to brain-stem pathway. OKN can thus be viewed in man as a cortical phenomenon (see also Hoffmann, 1989).

If second-order motion does not elicit optokinetic nystagmus, it implies that second-order motion is extracted from the image at a high level: after the overall direction of first-order motion of a pattern (which is certainly adequate to drive optokinetic nystagmus) has been deduced from its components. We therefore presented second-order motion to human subjects and recorded their eye movements under conditions when voluntary smooth pursuit and fixation were controlled. A preliminary report of these data has been presented (Harris & Smith, 1992).

Methods

Stimuli

The stimuli were generated using an image processing system (Matrox) and presented on a 17-in. monitor (IType) positioned 28.5 cm in front of the subjects. The image subtended 57° × 46 deg (31° × 24 cm). Second-order motion stimuli of three types could be generated. Each was produced by computing and displaying, in real time, the product of a horizontally drifting vertical sine grating and a stationary carrier. The image was updated at a rate of 33 Hz. The three types of pattern were as follows:

1. Contrast-modulated static noise. In this case, the carrier was a stationary field of two-dimensional static noise. Each noise element subtended approximately 0.48 deg for the main study; other element sizes gave the same result. To give the best chance of obtaining OKN, the modulation depth was 100% (i.e. local contrast varied across the image between zero and the maximum that the display could produce).

2. Contrast-modulated dynamic noise, produced as in (1) but using dynamic noise.

3. Amplitude-modulated (AM) gratings. In this case, the carrier was a stationary vertical grating of spatial frequency 1.4 cycle/deg.

There were no obvious differences in the eye movements evoked by the three stimulus types, and so our systematic study was carried out using dynamic random noise. For comparison purposes, a first-order motion stimulus was also used. This consisted of a simple drifting sine grating. The mean luminance of all patterns was 11.2 ft lamberts.

The spatial frequency for both conventional gratings and the second-order motion stimuli was kept constant at a value known to stimulate the OKN system optimally 0.12 cycle/deg (e.g. van den Berg & Collewijn, 1988; Schor, Narayan, 1981).

The patterns were drifted at speeds of 3.2, 12.8, and 22.4 deg/s. The contrast was 92% except where otherwise stated. The monitor was carefully calibrated and the image was accurately gamma-corrected to avoid the introduction of first-order components arising from brightness nonlinearities.

Subjects

Subjects were four graduate students of the York Vision Group and the authors. All had normal or corrected-to-normal vision at the viewing distance used. They were between 20 and 40 years of age.

Eye movement recording, calibration, and analysis

Horizontal eye movements were recorded by infrared reflection. Calibration trials were obtained by asking subjects to look alternately at points separated horizontally by 20 deg. Three subjects using rigidly mounted cental-acrylic bite bars provided the quantitative data. The observations were verified qualitatively in three further subjects without bite bars. Eye position was sampled and digitized every 12 ms and recorded on disc for off-line analysis.

Estimates of average eye velocity for each stimulus presentation were calculated as follows. For each record where nystagmus was clearly present, a number of periods showing the characteristic alternating fast and slow phases were identified. The average speed of each of a number of slow phases was then calculated from the total position change during the slow phase. These data were then averaged across a number of periods of nystagmus to give a large total number of slow phases (see Smith & Harris, 1991). For records where no clear nystagmus was evoked, the cumulative drift in position between saccades was measured and then divided by the duration to give an estimate of velocity.

Control of smooth pursuit and fixation

We were anxious that interpretation of our results might be complicated by either the presence of smooth pursuit (which might be mistaken for OKN) or fixation (which might suppress OKN). We controlled for smooth pursuit and fixation by two simultaneous means. First, for our systematic quantitative study we employed contrast-modulated dynamic visual noise. This reduced the chances of fixation considerably compared with modulated static noise or AM gratings, in both of which an obvious static structure was present that could readily be fixed. However, this measure alone did not stop pursuit, and fixation on the flickering, static borders of the pixels was still a possibility. Therefore we used a second technique of masking off the central 5 deg vertically of the stimulus, across its entire width. This has been shown to make fixation and pursuit very difficult, since they both rely on foveal information (Van Die & Collewijn, 1986; Howard & Ohni, 1984), whilst having almost no effect on optokinetic nystagmus (Murasegugi et al., 1986).

We positioned a 5-deg black band in front of the screen and instructed subjects to "try to fixate the center of the black band whilst keeping the image in focus." During control trials, in which the band was removed, we instructed subjects to "try to fixate the center of the screen whilst keeping the image in focus." The band had no fixation points and the edges were
Table 1. The average slow-phase eye velocity elicited by second order stimuli

<table>
<thead>
<tr>
<th>Stim velocity (deg/s)</th>
<th>Average eye velocity (deg/s)</th>
<th>Average gain (deg/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2</td>
<td>0.64 ± 0.29</td>
<td>0.20 ± 0.09</td>
</tr>
<tr>
<td>12.8</td>
<td>0.40 ± 0.29</td>
<td>0.03 ± 0.02</td>
</tr>
<tr>
<td>22.4</td>
<td>0.50 ± 0.18</td>
<td>0.02 ± 0.01</td>
</tr>
</tbody>
</table>

smooth and provided no landmarks. All recording was done with the room lights out to minimize reflections from the screen and the band.

Results

Eye movements evoked by second-order motion

The eye movements evoked by second-order motion were extremely weak. The overall average velocities of eye movements are summarized in Table 1. At no time did the average eye velocity evoked by a second-order stimulus in the presence of the central masking band (see Methods) exceed 1 deg/s. Moreover, eye velocity was independent of stimulus speed, although drift was always in the direction of the stimulus. Fig. 1 shows some examples of the vigorous optokinetic nystagmus evoked by a first-order drifting grating at 12.8 deg/s (traces A and B). The presence of the black band had no obvious effect suggesting that eye movements were optokinetic and not influenced by pursuit. In the case of the second-order motion, some slow-phases can be seen in the records obtained with free, full-field viewing (trace C). However, these were not present when the band was in place (trace D) and can therefore be regarded as voluntary smooth pursuit, dependent on the central retina.

Fig. 2 is a comparison of the response to first- and second-order stimuli over the range of velocities 3.2 to 22.4 deg/s. First-order drifting gratings evoked a vigorous OKN with a gain close to unity. Second-order patterns showed a much lower gain which was near zero at all but the lowest speeds (where pursuit would be easiest).

Comparison with eye movements elicited by drifting gratings

A possible interpretation of the absence of vigorous optokinetic eye movements in response to our second-order motion stimuli might be that the effective contrast was too low to elicit OKN. The contrast, as conventionally defined, was the same for the two stimuli (92.2%). However the critical variable in the second-order motion patterns is modulation depth (see Methods), not Michelson contrast. The latter relates the maximum and minimum luminance at any point in the image. This measure adequately reflects the amplitude of a first-order grating, but remains invariant as the modulation depth of our second-order patterns is manipulated. Modulation depth relates the maximum and minimum contrasts in the image, and defines the amplitude of the second-order patterns but is zero for our first-order gratings since their contrast is invariant across space. Clearly, it is not necessarily the case that patterns of the two types will elicit the same amplitude of neural signal when first-order Michelson contrast is numerically equated with second-order depth of contrast modulation. We addressed this problem
by measuring detection thresholds for each pattern, in terms of contrast for first-order and modulation depth for second-order motion, and then equated the suprathreshold patterns by presenting them at equal multiples of detection threshold. The subjective thresholds measured with the central ± 2.5 deg masked by the blank band, were between 0.7% and 1.5% contrast for first-order gratings and around 10% modulation depth (defined as the contrast of the drifting sine grating by which the carrier is multiplied) for the second-order stimulus. The variation of the response to first-order gratings with contrast is compared to the eye movements evoked by second-order stimuli in Fig. 3.

The average gain of the response to second-order motion of 12.8 deg/s was 0.03 ± 0.02. This is shown as a horizontal shaded band on Fig. 3 and overlaps with only the lowest contrast trials using first-order motion (less than 1%). The graph also allows comparison of the response to first- and second-order stimuli at the same multiple of subjective detection threshold. The second-order stimulus was presented at maximum modulation depth, which was approximately nine times threshold. The equivalent contrast of the first-order stimulus is therefore 10% (nine times the average threshold). The gains are 0.03 ± 0.02 (s.d.) for second-order and 0.46 ± 0.18 (s.d.) for first-order motion.

Discussion

We have shown that second-order motion stimuli, although producing a very strong percept of motion, are extremely poor at evoking optokinetic nystagmus. Such stimuli can be pursued (indeed voluntary effort can produce almost perfect tracking).

But when pursuit is disadvantaged (Fig. 1D), almost no optokinetic nystagmus is evident.

Implications of mechanisms of motion perception

Motion can be extracted from variations in the luminance (a first-order attribute) of an image over time by comparing the activity in two adjacent sensors and demanding a particular time delay between their activations (Barlow & Levick, 1965). There are numerous variants on this theme but motion detectors of this broad type are now known generically as Reichardt detectors after their detailed and successful application to the fly's visual system (Reichardt & Poggio, 1976). Cavanagh (1991) has recently pointed out that there is no reason to restrict the Reichardt principle to the processing only of first-order information. If the input to a Reichardt detector is from a system that extracts higher-order information, for example disparity or even motion itself, then such a detector would be able to signal movement between areas defined in terms of these higher-order attributes. In general, Reichardt detectors will indicate the same motion of a given object whether they base their decision on luminance or higher-order attributes. However, our second-order motion stimuli were arranged so that Reichardt detectors looking at first-order attributes would generate a very different outcome ("no motion") from detectors looking for movement by correlating higher-order attributes ("motion").

Drift-balanced stimuli with movement information contained entirely in the second-order attributes are indeed seen as moving. How then might such motion be detected? One possibility is that identification and tracking of the features defined by second-order properties occurs. This could be achieved by a high-level system with similarities to the long-range detection system (e.g. Anstis, 1980; Bradick, 1974, 1980) that is thought of as operating by identifying particular objects, features, or "tokens" (Ullman, 1979) and tracking their movements over time. An alternative hypothesis is that our sensitivity to second-order motion may result from grossly nonlinear transformations (such as rectification or squaring) of the image at an early stage in processing, before conventional intensity-based (first-order) motion analysis (Chubb & Sperling, 1988; Wilson, 1991).

A third possibility involves distortion products (Henning et al., 1975): a simple, compressive-brightness nonlinearity, of a type widely acknowledged to exist in the visual system, might create distortion products in the Fourier spectrum of the neural representation of the image. Such distortion could, in principle, be used to infer motion in many second-order stimuli. In the latter two cases, the neural representation of the image would then actually contain first-order motion information. If the distortion occurred early in visual processing (prior to the site of the "short-range" or energy-based process), the short-range system could detect them in the same way as real first-order motion. Direct physiological evidence exists for the existence of such distortion products in the lateral geniculate nucleus of the cat (Derrington, 1987), a site widely assumed to precede motion detection.

Since second-order motion stimuli evoke no OKN, the distortion product explanation appears to be ruled out. If our patterns were visible on the basis of distortion products detected by a single, energy-based system, then it is hard to see why they would not generate OKN when presented with a contrast of the same multiple of detection threshold that produces OKN with
conventional gratings. Our results can be reconciled with models in which two separate energy-based systems are proposed, one linear and one involving gross nonlinear transformations (e.g. Chubb & Sperling, 1988; Wilson, 1991). In such a system, it is possible in principle, to tap the output of only one of the two mechanisms to provide a drive for OKN. Models in which a single nonlinear energy-based system is proposed dealing with both first- and second-order motion (e.g. Grzywacz, 1992) are hard to reconcile with our results. In a single motion system, the cortical drive to OKN would reflect all types of motion that are visible to the motion-detection mechanism, including second-order motion. Other considerations also suggest that there are at least two motion-detection mechanisms and that they may be linked in series. For example, second-order motion is clearly visible when defined only in terms of movement between patches whose first-order motion attributes differ (in speed or direction). A single motion-detection system could not unambiguously signal both motions: two distinct, serially arranged stages are needed. Second-order motion could be deduced by higher-level detectors than those detecting first-order motion. These higher-order detectors work on the outputs of first-order attribute detectors, and also on higher-order detectors such as contrast-modulation detectors and disparity detectors. Cells responding to second-order motion have been recorded in cortical area MT (Albright, 1992). Higher-order detectors might even, as Cavanagh and Matin (1989) suggest, cross correlate the outputs of systems operating on different second-order characteristics. This would give the motion-detection system an extraordinary robustness when encoding the motion of objects, since the object could continue to be defined by a subset of its attributes if some became unreliable.

On the face of it, the lack of optokinetic eye movements when motion can be clearly seen in the image seems like a surprising result. But the argument above suggests that the second-order motion system may have evolved to enable the identification of motion of individual objects—possibly as part of object detection rather than an overt tracking system. Local motion does not, and should not, evoke optokinetic following movements which are to do with compensating for head rotation, not for following the motion of individual objects. Single objects evoke tracking (smooth pursuit) movements, whereas optokinetic nystagmus is generated by whole-field movement (Miyoisi, 1985). Furthermore, such “tracking” of individual objects does not necessarily need overt eye movements at all. Recent theories suggest that up to five objects can be “tracked” at once as they travel around within the visual field by an attention mechanism that is involved in the object defining process (Triesman, 1986; Pylyshyn & Storm, 1988).

As visual information passes through the brain, various analyses occur. At some stage after the overall visual direction of the scene has been calculated from its individual components, the information is used to generate OKN. Since OKN is not evoked by our stimuli, it can be deduced that such patterns do not generate a velocity signal for this purpose. So the percept must arise either from a higher stage or from a parallel stage, presumably diverging after the primary visual cortex.

Acknowledgments
This research is supported by NATO travel Grant CRG 910076 to A. T. Smith and L. R. Harris; and by NSERC Grant OGP0046271 to L. R. Harris. We would also like to thank the Institute for Space and Terrestrial Sciences for generous support.

References


